An exploration of Kelly's Criterion

RQ: What proportion of one's total asset should one invest in stocks for optimum growth?

Introduction

Recently, a stock named Gamestop has gained much attention due to it's dramatic price swings. This piqued my interest in stock markets, where I started reading into the basics of trading. Due to the fact that the US market opens at a inconvenient time (for Singapore), it is impossible for me to make decisions and trade real time. Therefore, I decided to code a simple bot that will buy and sell based on a mathematical formula which maximizes growth on the long term. One thing we wanted to avoid is to lose all our money, and to do this, I must only invest a proportion of our total asset each time when we invest in a stock.

This lead to me thinking about the following question: What percentage of our total asset should we invest, so that growth is maximized when we trade an arbitrarily large number of times?

This idea of an ideal or growth optimal investment fraction comes from the concept of Kelly fraction, which is primarily used in betting games. In this exploration, I will extend this concept to determine the most optimum fraction we should invest in a stock. To investigate this, I will first pick the stock of Gamestop and model the probability of price change under a discrete time frame, and use this as the basis for the calculation of the optimum proportion. Then, I will explore the problem when multiple stocks are available for investment, and investigate the significance of covariance between two stocks in the optimum investing fraction. Finally, I will look at a way to model stocks under continuous time frame by considering the mean and variance.

Modeling stock price under discrete time

We shall begin the investigation through the modeling of the stock. In order to calculate the optimum proportion for investment, we must first make an estimation of the profitability of the stock, where we need to find out the probability of how often the price of the stock will go up.

To start off, we will look at stocks under a discrete time frame, where assume that the price "jump" from one point to another every t seconds, where nothing happens within the t seconds. At time = 0, we let the value of the stock to be S_0. After t seconds, the price "jumps", with probability of p to multiply by u times and a probability of $1 - p$ to multiply by d times. Hence, we can also express the price multiplier for every t seconds as a random variable X, where

$$X = \begin{cases} u & P(X) = p \\ d & P(X) = 1 - p \end{cases} \tag{1}$$

The following diagram illustrates the probability tree of the price after Nt seconds.

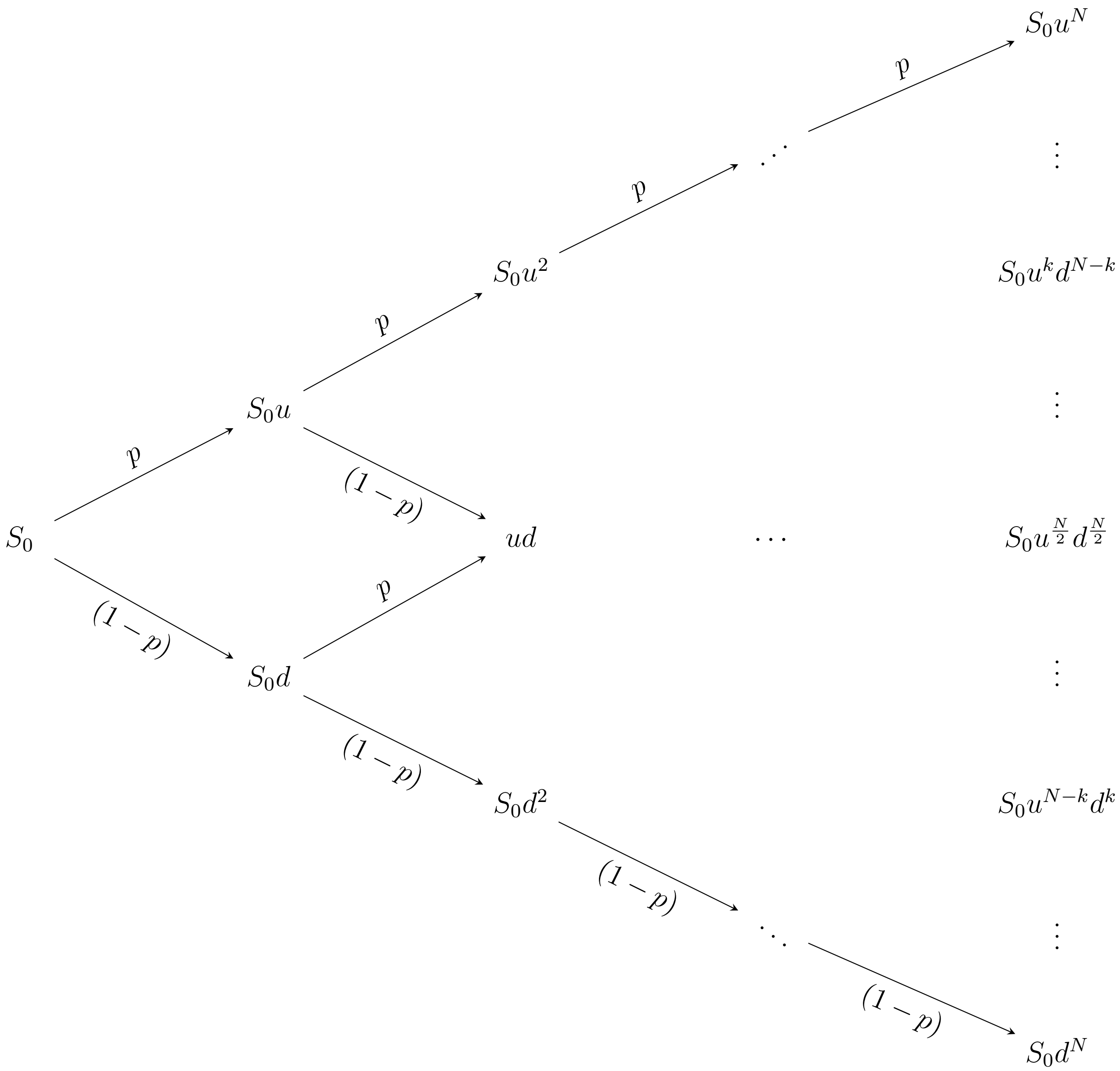


Fig. 1

The movement of the price is similar as doing repeated coin flips with a biased coin. The best way to understand this is to imagine flipping a coin every t seconds, where a heads means the price is u times higher, and a tail means d times lower. The limitation is that in reality, the motion of the stock price isn't absolutely random, and is depended on the supply and demand of the market. Therefore, this is only an approximation of the real price movement. Furthermore, Because of the way we defined u to be an increase and d to be an decrease, we have $u > 1$ and $d < 1$.

This is formally known as the Cox-Ross-Rubinstein Model[Ortiz-Latorre]. One limitation with this model is that the initial price is fixed. Because market data is real time and after the time period Nt, the price will have changed, making the initial price different. Therefore, we can modify the model by dividing by S_0 at every point in the tree, which gives us something like the following:

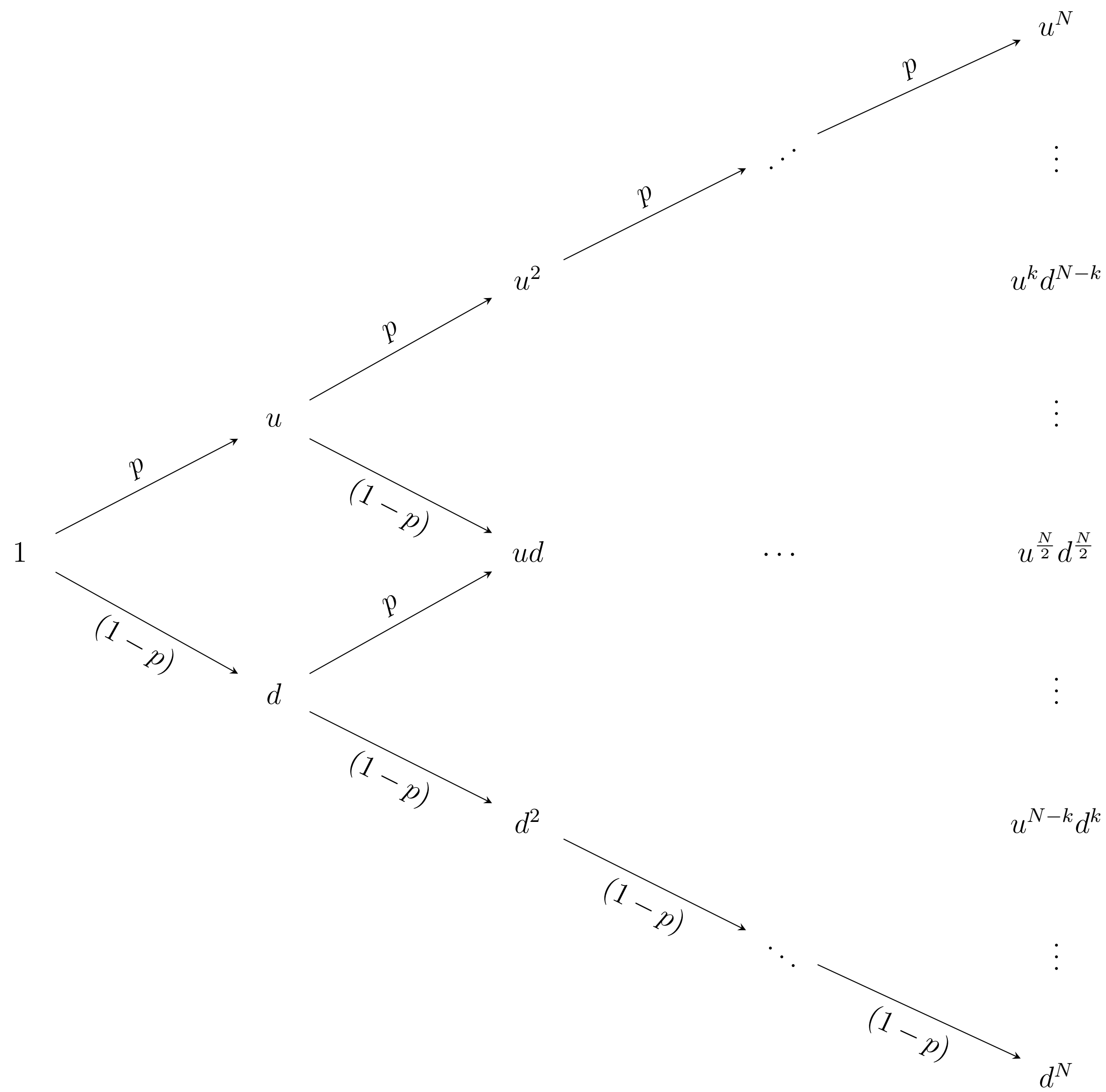
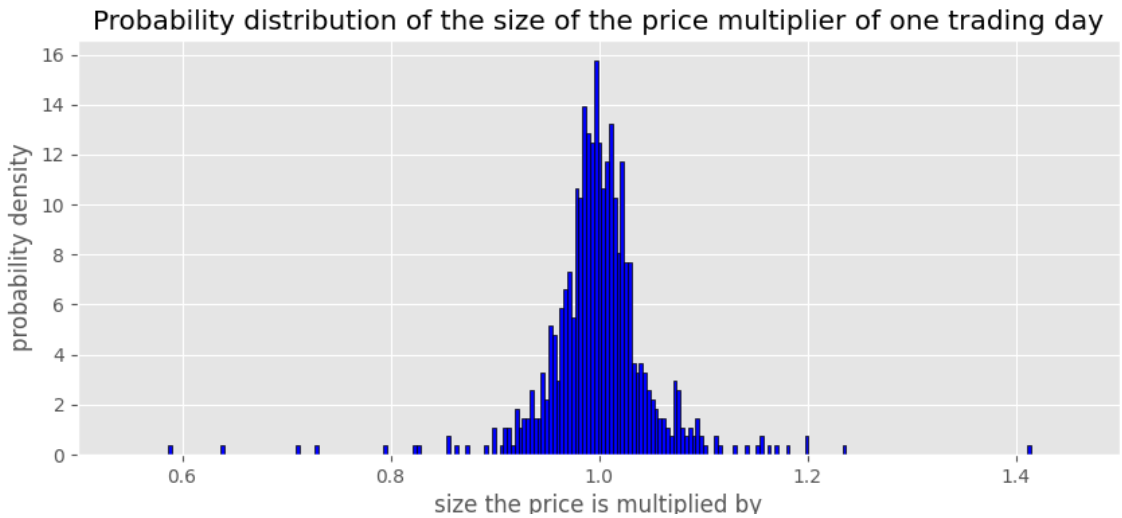


Fig. 2

Here, we take S_0 out of the equation, and only look at the price multiplier from $T = 0$ to $T = Nt$. We are essentially disregarding what value the initial price was at, and only look at the new price as a multiple of whatever the initial price was. Furthermore, we can treat the probability for the price to multiply by a specific value, $u^k d^{N-k}$ in a period of Nt, very similarly to a binomial distribution, where

$$P(u^k d^{N-k}) = \binom{N}{k} \cdot p^k (1-p)^{N-k} \tag{2}$$

This can be seen from $Fig.2$, where k u's(with the probability p^k) and $N-k$ d's (with the probability $(1-p)^{N-k}$)leads to the price of $u^k d^{N-k}$, and there are $\binom{N}{k}$ combinations of the paths. This modification to the model makes it possible to generate a histogram for the probability distribution of the price multiplier after Nt seconds has elapsed. Here, we assume Nt seconds to be the length of one trading day, and sampled the size of the price multiplier from Gamestop over 700 trading days in the past 3 years. The probability distribution of the price multiplier is in the chart below:



Date	Open	Close	Multiplier
2018-02-09	15.98000	16.17000	1.011890
2018-02-12	16.29000	15.80000	0.969920
2018-02-13	15.76000	15.58000	0.988579
2018-02-14	15.58000	16.21999	1.041078
2018-02-15	16.29999	16.41000	1.006749

. . .

Date	Open	Close	Multiplier
2021-02-02	140.75999	90.00000	0.639386
2021-02-03	112.01000	92.41000	0.825016
2021-02-04	91.19000	53.50000	0.586687
2021-02-05	54.04000	63.77000	1.180052
2021-02-08	72.41000	60.00000	0.828615

The data is obtained from Yahoo Finance [Yahoo Finance]. Only the first and last 5 data points are shown, and the rest are omitted due to the size of dataset, but all specific data points can be found in the code repository linked in the Appendix. From the raw data in the table, the price multiplier in the time period of a day will simply be calculated as $\frac{\text{Open price}}{\text{Close price}}$.

We can see how the price has a high probability of multiplying by a small amount(high density when the multiplier is between 0.9 and 1.1), and a low probability of multiplying by a large amount(low density when the multiplier is either really large or really small). Here, we are tempted to transform the data points, that is, to derive a function $t(x)$ that does the following:

$$t(u^N d^0) = N \qquad t(u^0 d^N) = 0 \qquad t(u^k d^{N-k}) = k \tag{3}$$

Where instead of having price multiplier ranging from $u^0 d^N$ to $u^N d^0$ (see $Fig.2$), it ranges from 0 to N. This treats every upward movement in price as a "success", and every downward movement in price as a "failure". What $t(x)$ does is it makes the data from the chart above distribute exactly like a binomial distribution. Mathematically, we have

$$P(t(u^k d^{N-k})) = P(k) = \binom{N}{k} \cdot p^k (1-p)^{N-k} \tag{4}$$

Which is derived from (2) and (3). The following is my derivation of $t(x)$ that will transform the data.

First, the set of price multipliers S is the following:

$$S = \{d^N, d^{N-1}u, d^{N-2}u^2, \cdots, du^{N-1}, u^N\} \tag{5}$$

Note that these are what we have at the end of the binomial tree in $Fig.2$. Now, we first take ln of all elements in the set, obtaining the following:

$$\ln S = \{N \ln d, (N-1)\ln d + \ln u, \cdots, \ln d + (N-1)\ln u, N \ln u\} \tag{6}$$

From here, we get some interesting result by subtracting $N \ln d$ from every term:

$$\ln S - N \ln d = \{0, \ln u - \ln d, \cdots, (N-1)(\ln u - \ln d), N(\ln u - \ln d)\} \tag{7}$$

Where every term is a multiple of $\ln u - \ln d$. Therefore, we can divide every element by $\ln u - \ln d$ to get the following:

$$\frac{\ln S - N \ln d}{\ln u - \ln d} = \{0, 1, 2, \cdots, N-1, N\} \tag{8}$$

We have obtained the function $t(x)$, where

$$t(x) = \frac{\ln x - N \ln d}{\ln u - \ln d} \tag{9}$$

All we did here is to transform the histogram to the standard binomial distribution, where the price multiplier is transformed to a Bernoulli variable in the following form:

$$X = \begin{cases} u & P(X) = p \\ d & P(X) = 1-p \end{cases} \qquad \rightarrow \qquad X' = \begin{cases} 1 & P(X) = p \\ 0 & P(X) = 1-p \end{cases} \tag{10}$$

This may feel unmotivated, by this transformation makes it possible to analyze the data as a Bernoulli variable, where we have ample knowledge on its statistical properties.

However, there are still a few extra steps before we can map the variable. First, there are a few estimations we need to make for the value of N, d, and u. We assume that the discrete time period t is one minute long, and in 1 trading day, there are 390 minutes, so $Nt = 390$ minutes $\iff N = 390$. Next, for the values for u and d, We can assume that the smallest price multiplier is d^N and the largest price multiplier to be u^N. This comes directly from $Fig.2$. There exist one limitation to the approximation, because the maximum and minimum price changes can be the result of forces that are outside of the market itself, such as government intervention, or are result of rare events, for example, the 2008 economic crisis, and may not be intrinsic to the stock. From our data, the smallest price change is 0.58668 and the largest price change is 1.67095 [Yahoo Finance]. Hence, we have the following:

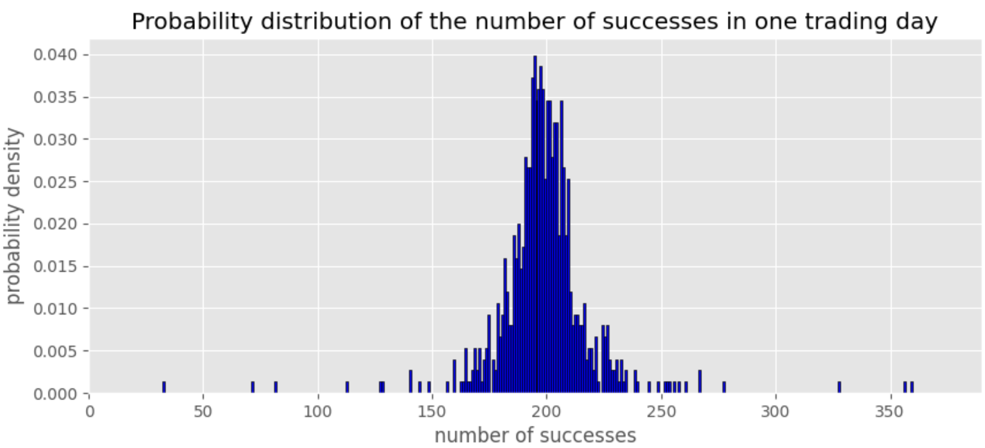
$$0.58668 = d^{390} \iff d = 0.99863 \tag{11}$$

$$1.67095 = u^{390} \iff u = 1.00131 \tag{12}$$

Then,

$$t(x) = \frac{\ln x - 390 \ln 0.99863}{\ln 1.00131 - \ln 0.99863} \tag{13}$$

Is the function that we will use to transform our raw data into a binomial probability density distribution.



The above chart is the transformed histogram, with a mean μ_n of 198.63(calculation of the mean done electronically by code in the Appendix, and will be omitted here due to the size of the dataset). The value of the mean will become extremely useful after we demonstrate that μ_n is related to the probability p by proving that the mean of a Binomial distribution is Np. by definition, $\mu_n = \mathbb{E}(Y)$, We have the following:

$$
\begin{aligned}
\mu = \mathbb{E}(Y) =& \sum YP(Y) \\
=& \sum_{x=1}^{N} x \cdot \binom{N}{x} p^x (1-p)^{N-x} \\
=& \sum_{x=1}^{N} x \cdot \frac{N!}{x!(N-x)!} p^x (1-p)^{N-x} \\
=& \sum_{x=1}^{N} N \cdot \frac{(N-1)!}{(x-1)!(N-x)!} p^x (1-p)^{N-x} \\
=& \sum_{x=1}^{N} Np \cdot \binom{N-1}{x-1} p^{x-1} (1-p)^{N-x} \\
=& Np \cdot \sum_{x=1}^{N} \binom{N-1}{x-1} p^{x-1} (1-p)^{(N-1)-(x-1)} \\
=& Np \cdot \sum_{x=0}^{N-1} \binom{N-1}{x} p^{x} (1-p)^{(N-1)-x}
\end{aligned} \tag{14}
$$

By the binomial theorem[Khan], we have

$$
(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \tag{15}
$$

Therefore, when $x = p$, $y = p - 1$, and $n = N - 1$, we have

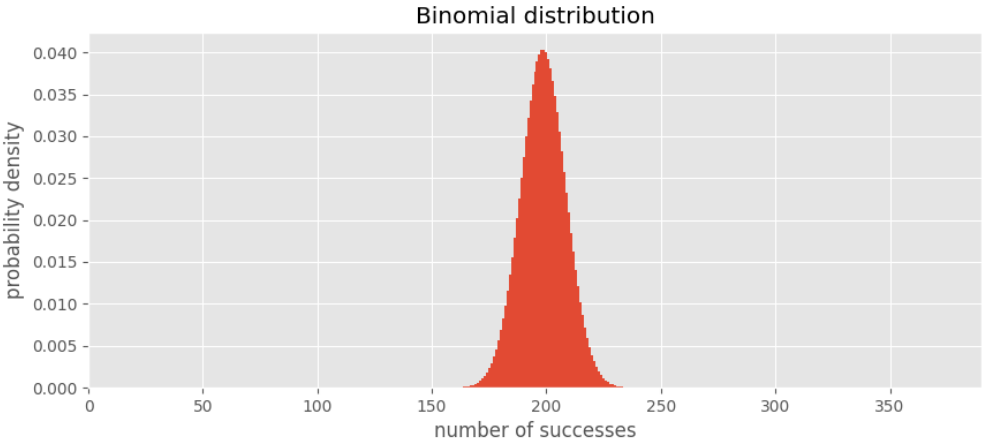
$$
\sum_{x=0}^{N-1} \binom{N-1}{x} p^x (1-p)^{(N-1)-x} = (p + 1 - p)^{N-1} = 1 \tag{16}
$$

Substituting our result from (16) into (14), we have

$$
\mu = Np \cdot \sum_{x=0}^{N-1} \binom{N-1}{x} p^x (1-p)^{(N-1)-x} = Np \cdot 1 = Np \tag{17}
$$

Then, $\mu_n = Np$. This was the reason why we transformed the raw data using $t(x)$; We can utilize our knowledge of the mean of a binomial distribution to calculate the probability p. Solving this equation, we have $p = \frac{\mu_n}{N} = \frac{198.63}{390} = 0.50931$. Hence, price multiplier X can be described as the following:

$$
X = \begin{cases} 1.00131 & P(X) = 0.50931 \\ 0.99863 & P(X) = 0.49069 \end{cases} \tag{18}
$$



Comparing this to the probability density of the binomial distribution $B(0.50931, 390)$ generated by computer, we also see that our distribution matches it closely, which gives us confidence that the price multiplier(success and failures) of the stock does follow a binomial distribution. The p value is above 0.5, which indicates that the stock is more likely to rise than to fall. However, we also need to take into account of u and d, because the magnitude of the change in price also needs to be accounted. Therefore in the next section, we will consider the value of u, d, and p together to optimize our investing proportion.

Optimization of investment

Modelling with binomial distribution

After the analysis of data, we will now begin our optimization step. From our previous chapter, we know that price change can be approximated by the discrete random variable X, defined in (1).

Now, let $g(f)$ be the function that takes in the variable f, which is the fraction of our total asset, and returns the value of our asset after the discrete period t seconds, Since X is the price change, if we only invest the fraction f in the stock, only f will experience the price change. Hence, the following will represent the outcome of the investment, where

$$g(f) = 1 - f + Xf \tag{19}$$

To express this with more clarity, we have the following expression:

$$P(g(f)) = \begin{cases} p & g(f) = 1 - f + uf \\ 1 - p & g(f) = 1 - f + df \end{cases} \tag{20}$$

Which directly follows from the probability of (8). Hence, the expected value of $g(f)$ will be the following:

$$E(g(f)) = \sum g(f) \cdot P(g(f)) = p \cdot (1 - f + uf) + (1 - p) \cdot (1 - f + df) \tag{21}$$

representing the the amount we make per investment on average. Clearly, when this is maximized, we make the most money per investment. Therefore, it seems like our goal is to maximize $E(g(f))$. However, this isn't the case. This is because utility, or satisfaction of wealth, plays a huge role in how the investment should be optimized. If we assume that utility is linear, i.e. People's satisfaction grows proportionally with their wealth, it is not realistic, because in reality, utility is always diminishing, where the same amount of money leads to less satisfaction. We can see utility as a function of wealth, where

$$h(g(f)) = \ln(g(f)) \tag{22}$$

To better model reality, we are incentivized to maximize the expected value of $ln(g(f))$ instead. Moreover, maximize $\ln(g(f))$ doesn't change the position of maxima. This is because $\ln(x)$ is monotone increasing, where when $g(f)$ is highest, $\ln(g(f))$ is also highest.

Therefore, we will attempt to find an f that maximizes the expected value of $\ln g(f)$.

Similar to (20), the probability of $\ln(g(f))$can be expressed as the following:

$$P(\ln g(f)) = \begin{cases} p & \ln g(f) = \ln(1 - f + uf) \\ 1 - p & \ln g(f) = \ln(1 - f + df) \end{cases} \tag{23}$$

Where the expected log wealth E will simply be the following:

$$E(\ln g(f)) = p \ln (1 - f + uf) + (1 - p) \ln (1 - f + df) \tag{24}$$

An alternative way to arrive at this equation is to use the law of large numbers. Let $V(f)$ be the value of the asset after N rounds of investments given the optimum fraction f.

When N is arbitrarily large, pN of the time, the price will go up, and $(1 - p)N$ of the time, the time will fall down. This is a result of the law of large numbers, where the average numbers of ups and downs(successes and failures) in the price converges to the expected numbers of ups and downs.

Therefore, the value of our asset after N rounds is

$$V(f) = V_0(1 - f + uf)^{pN}(1 - f + df)^{(1-p)N} \tag{25}$$

Where V_0 is the value of the initial asset.

Dividing V_0 to the left hand side, and taking the N^{th} root, we obtain an equation for the change of asset value per investment(remember that $V(f)$ is for N rounds of investment, and we only need to maximize the return for one round), where

$$g(f) = \sqrt[N]{\frac{V_f}{V_0}} = (1 - f + uf)^p (1 - f + df)^{1-p} \tag{26}$$

Here, we are also maximizing $\ln g(f)$ instead, because like we mentioned before, ln is monotone increasing and will not change the position of the maximum, as we explained earlier, while making the equation easier to differentiate.

Let $h(f) = \ln(g(f))$, then

$$h(f) = p \ln(1 - f + uf) + (1 - p) \ln(1 - f + df) \tag{27}$$

This alternative method gives us the same equation. To find the maximum, we simply take $\frac{dE}{df} = 0$, where

$$\frac{dE}{df} = \frac{p(u-1)}{1 - f + uf} + \frac{(1-p)(d-1)}{1 - f + df} = 0 \tag{28}$$

Simplifying the equation, we get an expression for the f that maximizes $h(f)$, where

$$f = \frac{p}{1-d} + \frac{1-p}{1-u} \tag{29}$$

Which is surprisingly elegant. We also know this must be a maxima because ln is concave, The derivation is modified from [Kelly, Thorp].At first, it seems that the result only makes sense of $1 \geq f \geq 0$, since f represents a fraction of your total asset. However, there are a few implications when value of f is outside this range.

When $f < 0$, this means it is not profitable to invest into the stock. However, this means it is profitable to short the stock, which is a technique where investors make money when the price of the stock falls.

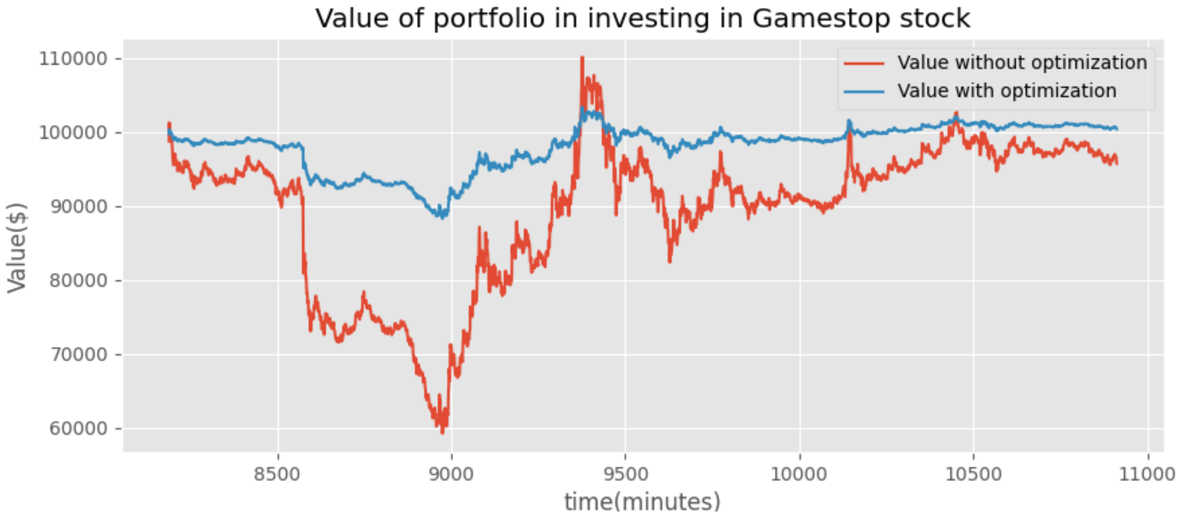
When $f > 1$, this means it is even profitable to borrow money you didn't own in the first place and invest in the stock. This is a technique called leveraging. This means we can still use f to guide our investment even if it is outside the normal range.

Now, we can substitute in the values for p, u, and d from (18),

$$f = \frac{p}{1-d} + \frac{1-p}{1-u} = \frac{0.50931}{0.00137} - \frac{0.49069}{0.00131} = 0.24466950 \tag{30}$$

This gives us the optimum fraction of 24.46%. This means after every minute, we should adjust our portfolio to ensure that only 24.46% of our total asset is invested in Gamestop. This in fact follows the intuition of buying low and selling high: When the price rise, the value of the shares increase, making our invested asset exceed 21.46%, making us sell some of the shares, while the reverse is also true.

This strategy is tested, where I simulated trading using this strategy, and compared it to buying and holding:



Where the price data is gathered from the week between 23-3-2021 to 31-3-2021. The axis is adjusted to minutes to illustrates the numbers of trades the optimizing strategy executed. We can see that at the end of the period, the method of optimization has outperformed simply buying and not doing anything. Moreover, we observe less volatility in our asset value, where the fluctuations are a lot less severe. One limitation with this is that it doesn't take into account of commission fees, which is the fee payed to brokers for the execution of every trade. Therefore, this is only viable when no commission fees are charged.

In the following sections, we will extend this idea and explore it in more complex situations.

Extension

Multiple stocks for investment

Another question we can ask ourselves is, how does the situation change if multiple stocks are up for investment? What proportion should we divide our total asset, which in this case is our portfolio, to maximize long term gain? Here, I decide to also widen the investment options to also include cryptocurrencies, which is also gaining traction recently. Here, I picked Litecoin as the specific investment option we will look into.

If we were to invest in both Gamestop and Litecoin, the optimum fraction should be f_1 percent of asset to Gamestop and f_2 percent of asset to Litecoin. Assume we have the following two random variables, X, for the price multiplier of Gamestop and Y for the price multiplier of Litecoin, where

$$P(X) = \begin{cases} p_1 & X = u_1 \\ 1 - p_1 & X = d_1 \end{cases} \qquad P(Y) = \begin{cases} p_2 & Y = u_2 \\ 1 - p_2 & Y = d_2 \end{cases} \tag{31}$$

Let us first consider the case that that X and Y are independent, where $P(X, Y) = P(X) \cdot P(Y)$. Then the joint probability will be as the following:

$$P(X, Y) = \begin{cases} p_1 p_2 & X = u_1, Y = u_2 \\ 1 - p_1 p_2 & X = d_1, Y = u_2 \\ p_1(1 - p_2) & X = u_1, Y = d_2 \\ 1 - p_1(1 - p_2) & X = d_1, Y = d_2 \end{cases} \tag{32}$$

Now, similarly to when we have only one stock to consider, for two stocks, the value of our asset after one discrete time period t will be $g(f_1, f_2)$, expressed as the following:

$$g(f_1, f_2) = 1 - f_1 + X f_1 - f_2 + Y f_2 \tag{33}$$

Similar to what we did before, we take the expected value of the log wealth, or $E(\ln g(f_1, f_2))$, which is expressed as the following:

$$\begin{aligned} E(\ln g(f_1, f_2)) = p_1 p_2 \ln\left[1 + (u_1 - 1) f_1 + (u_2 - 1) f_2\right] \\ +(1 - p_1) p_2 \ln\left[1 + (d_1 - 1) f_1 + (u_2 - 1) f_2\right] \\ +p_1(1 - p_2) \ln\left[1 + (u_1 - 1) f_1 + (d_2 - 1) f_2\right] \\ +(1 - p_1)(1 - p_2) \ln\left[1 + (d_1 - 1) f_1 + (d_2 - 1) f_2\right] \end{aligned} \tag{34}$$

Here, each row of the equation corresponds to the combination of successes and failures of the price multiplier of the two stocks as we summarized in (32).

When dealing with multivariate functions, to find the stationary points, we cannot differentiate with respect to one variable only and set the derivative to 0. This is because a function is stationary in one dimension doesn't imply it is stationary in all dimensions. Hence, we differentiate with respect to every variable separately, obtaining what is known as a partial derivative. Do note that when taking the partial derivative of single variable, every other variable is treated as a constant.

Here, we need to employ partial differentiation to find out the maximum of $g(f_1, f_2)$, by considering the partial derivative of each variable.

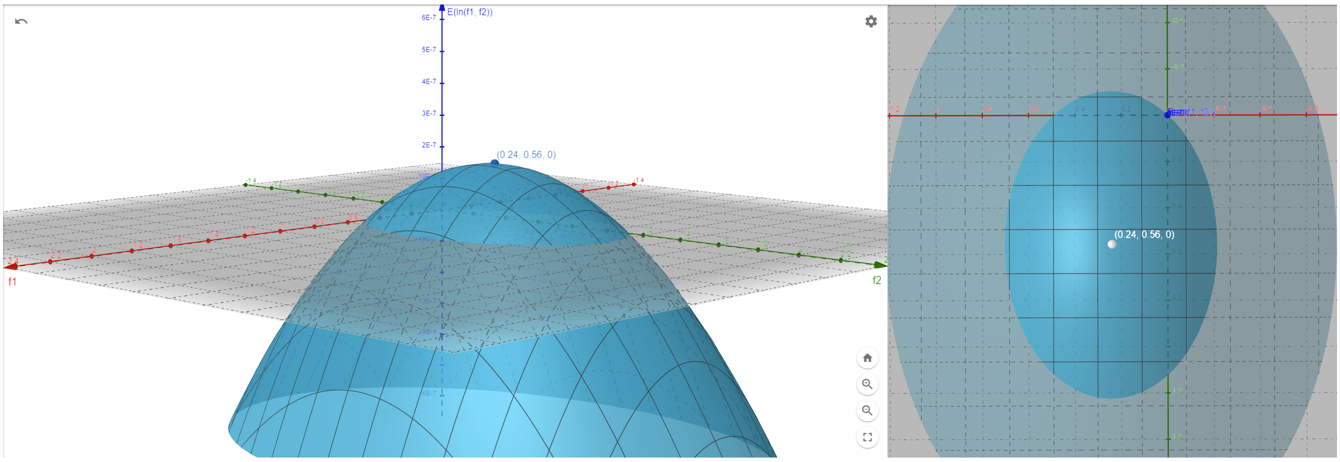
Differentiating with respect to x and setting the partial derivative to 0, we obtain

$$\begin{aligned}\frac{\partial g}{\partial f_1} &= \frac{p_1 p_2 (u_1 - 1)}{1 + (u_1 - 1) f_1 + (u_2 - 1) f_2} + \frac{(1 - p_1) p_2 (d_1 - 1)}{1 + (d_1 - 1) f_1 + (u_2 - 1) f_2} \\ &+ \frac{(1 - p_2) p_1 (u_1 - 1)}{1 + (u_1 - 1) f_1 + (d_2 - 1) f_2} + \frac{(1 - p_1)(1 - p_2)(d_1 - 1)}{1 + (d_1 - 1) f_1 + (d_2 - 1) f_2} = 0\end{aligned} \tag{35}$$

Differentiating with respect to f_2, we obtain

$$\begin{aligned}\frac{\partial g}{\partial f_2} &= \frac{p_1 p_2 (u_2 - 1)}{1 + (u_1 - 1) f_1 + (u_2 - 1) f_2} + \frac{(1 - p_1) p_2 (u_2 - 1)}{1 + (d_1 - 1) f_1 + (u_2 - 1) f_2} \\ &+ \frac{(1 - p_2) p_1 (d_2 - 1)}{1 + (u_1 - 1) f_1 + (d_2 - 1) f_2} + \frac{(1 - p_1)(1 - p_2)(d_2 - 1)}{1 + (d_1 - 1) f_1 + (d_2 - 1) f_2} = 0\end{aligned} \tag{36}$$

Instead of solving this explicitly, we substitute in the respective values for u_i, d_i and p_i, and attempts to solve it numerically. From the previous section, we calculated $u_1 = 1.00131$, $d_1 = 0.99863$, $p_1 = 0.50931$ which corresponds to the Gamestop. The same steps were carried out to calculate $u_2 = 1.00074$, $d_2 = 0.99884$, $p_2 = 0.60711$ for Litecoin(Data snippets can be found in Appendix). After solving using a graphical calculator, we obtain the following pair for f_1 and f_2: $(0.24466949, 0.55996150)$.



This means we should invest 19.02% of our total asset in Gamestop and 55.99% of our total asset in Litecoin. We see that even taking 6 decimal places, f_1 calculated here is the same as the f we calculated in (30). Then for investing in only two stocks, it is a good approximation to simply calculate the optimum fraction separately and treat them as seperate investments. The fact that we can treat these as two seperate investments actually fits with our assumption, that the price of two stocks are independent. If we look back at our method of obtaining the maxima, one can do this systemically for more stocks using partial differential equations, providing a pragmatic method to find the optimum fraction when considering even more stocks.

Covariance between stock prices

The limitation of our previous calculations is that we assumed Gamestop and Litecoin to be independent. In real life however, stocks are all correlated in someway their price cannot be seen as independent from each other. In economics, this ties closely with the concept of substitute and compliments. Here, we will analyze the covariance between the price of Gamestop and Litecoin, and examine how this will impact the Kelly fraction.

First of all, we are no longer assuming independence, where we treat X and Y, the price multipliers are correlated variables. Hence, the expected value of $\ln g(f_1, f_2)$ is modified from (34) and is now the following:

$$\begin{aligned} E(\ln g(f_1, f_2)) = & a \ln\left[1 + (u_1 - 1)f_1 + (u_2 - 1)f_2\right] \\ & + b \ln\left[1 + (d_1 - 1)f_1 + (u_2 - 1)f_2\right] \\ & + c \ln\left[1 + (u_1 - 1)f_1 + (d_2 - 1)f_2\right] \\ & + d \ln\left[1 + (d_1 - 1)f_1 + (d_2 - 1)f_2\right] \end{aligned} \tag{37}$$

To find the values of a, b, c and d by examining the joint probability distribution of $P(X', Y')$. The reason we are using X', Y', the Bernoulli version of X and Y, is because we defined them to have the same joint probability as X and Y, while they will simplify our calculation later on(see (10) for the mapping process). Here, we have their joint probability as the following:

$$P(X', Y') = \begin{cases} a & X' = 1, Y' = 1 \\ b & X' = 0, Y' = 1 \\ c & X' = 1, Y' = 0 \\ d & X' = 0, Y' = 0 \end{cases} \tag{38}$$

Our goal now is to find the specific values of a, b, c, and d . To do this, we will first look at a metric that measures the relationship between two variables called covariance. Covariance is defined as the following, where

$$cov(X, Y) = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{N - 1} \tag{39}$$

where, $\bar{X}$ and $\bar{Y}$ are the mean for the two datasets, and N is the number of data points in both datasets. Covariance is useful, because if two variables are independent, the covariance will be 0, and if the two variables have a perfect linear relationship, covariance will be positive, and vice versa.

Here, we will derive an alternative expression of covariance in the form of expected values, where

$$cov(X', Y') = \mathbb{E}[(X' - \mathbb{E}[Y'])(Y' - \mathbb{E}[X'])] \tag{40}$$

By the linearity of expectation, we have:

$$\begin{aligned} cov(X', Y') &= \mathbb{E}[(X' - \mathbb{E}[Y'])(Y' - \mathbb{E}[X'])] \\ &= \mathbb{E}[X'Y'] - \mathbb{E}[X']\mathbb{E}[Y'] - \mathbb{E}[X']\mathbb{E}[Y'] + \mathbb{E}[X']\mathbb{E}[Y'] \\ &= \mathbb{E}[X'Y'] - \mathbb{E}[X']\mathbb{E}[Y'] \end{aligned} \tag{41}$$

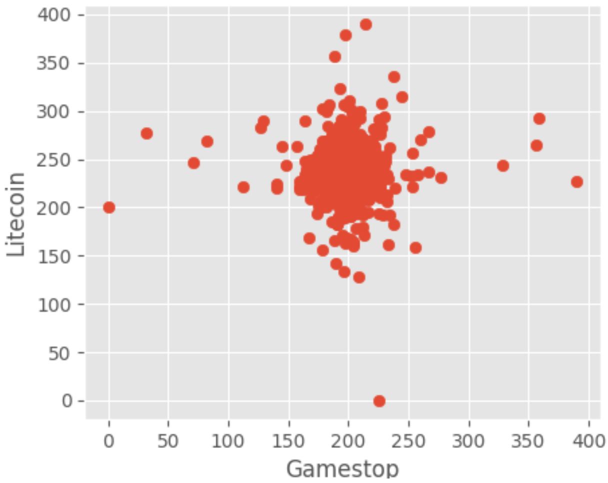
This is useful because we can express the expected values in terms of the probabilities a, b, c and d. First of all, we have

$$\begin{aligned} \mathbb{E}[X'Y'] &= P(X' = 1, Y' = 1) \cdot (1 \cdot 1) \\ &+ P(X' = 1, Y' = 0) \cdot (1 \cdot 0) \\ &+ P(X' = 0, Y' = 1) \cdot (0 \cdot 1) \\ &+ P(X' = 0, Y' = 0) \cdot (0 \cdot 0) \\ &= a \end{aligned} \tag{42}$$

Substituting (42) into (41), we obtain the following equation for a, where

$$cov(X', Y') = a - \mathbb{E}[X']\mathbb{E}[Y'] \Leftrightarrow a = cov(X', Y') + \mathbb{E}[X]\mathbb{E}[Y'] \tag{43}$$

Moreover, we can obtain numerical values for $cov(X', Y')$, $\mathbb{E}[X']$, $\mathbb{E}[Y']$. We can calculate covariance directly from our data using (39). The follow chart is the joint distribution of X' and Y':



Where $cov(X', Y') = 0.000101663$. The calculation is done electronically using code that can be found in the Appendix. Just a remark, we see that the value is very close to zero, which means our previous assumption that they are independent is very good. Given by the fact that one is a stock in the retail industry and the other is a cryptocurrency, it is understandable that they are not closely related.

Furthermore, we know that $\mathbb{E}[X'] = 0.50931$, $\mathbb{E}[Y'] = 0.55996$, because they can be expressed as probabilities in the following way, where

$$\begin{aligned} \mathbb{E}[X'] &= P(X = 1) \cdot 1 + P(X = 0) \cdot 0 = a + c = p_1 = 0.50931 \\ \mathbb{E}[Y'] &= P(Y' = 1) \cdot 1 + P(Y' = 0) \cdot 0 = a + b = p_2 = 0.60711 \end{aligned} \tag{44}$$

Hence, substituting these values into (43), we have

$$a = 0.00010166 + 0.50931 \cdot 0.60711 = 0.30931 \tag{45}$$

Using the value of a, we can find the value of b and c also from (44), where

$$\begin{aligned} c &= \mathbb{E}[X'] - a = 0.20000 \\ b &= \mathbb{E}[Y'] - a = 0.29779 \end{aligned} \tag{46}$$

Furthermore, since $\sum P(X', Y') = 1$, we also have $1 = a + b + c + d$, which gives

$$d = 1 - 0.30931 - 0.29779 - 0.20000 = 0.19288$$

We can substitute the probabilities back into (37), and following the similar steps from (35) and (36), where we take the partial derivatives with respect to f_1 and f_2, we obtain the following results, where

$$f_1 = 0.2445086 \qquad f_2 = 0.5598149 \tag{47}$$

The entire calculation process is omitted because it is the same from the last section in terms of the methodology. We see that the values are both slightly lower. This makes sense because the more related two investments are, the more similar their movement will be, which makes it more dangerous to invest more in both. This is also inline with the intuitive method of diversifying portfolios to reduce risk, where stocks that are more dependent are more risky to invest at the same time.

Extended extension:

Modeling stock price under continuous time

We can look make our approximation better by looking at the price multiplier under a continuous time frame. Here, we will reconstruct our price model from scatch. Instead of writing the price multiplier as (1), we will express it as the following,

$$x = \begin{cases} (\mu + \sigma) & P(x) = 0.5 \\ (\mu - \sigma) & P(x) = 0.5 \end{cases} \tag{48}$$

where X has a mean of μ and a variance of σ^2. The probability distribution of x can be graphically represented as the following:

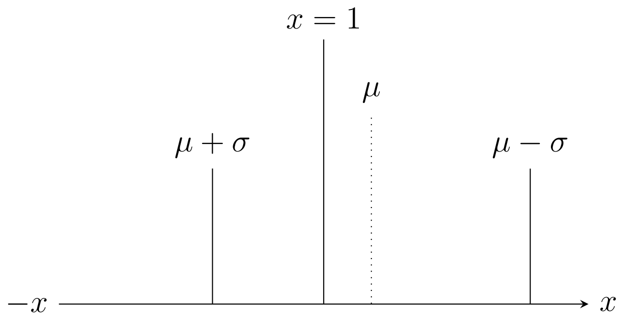


Fig. 3

Note that x is still a discrete random variable. What we are doing differently here is that we assume x has equal probabilities of leading to a higher price and leading to a lower price; However, it's mean and variance will determine how much the multiplier favors an increase in price or decrease in price. Note that in $Fig.3$, the mean is greater than 1, where x is favoring an increase in price. This is essentially a more succinct and useful way of expressing the price multiplier, because things like mean and variance can be applied to a lot of other types of distributions.

After a round in investment, the value of the asset, $g(f)$, will be

$$g(f) = 1 - f + fx = 1 + f(x-1) \tag{49}$$

which is exactly the same as (19), where f is the optimum fraction of investment. Note how $g(f) = 1 + f(x-1)$, and we can simplify the expression of $g(f)$ by defining a new variable X where $X = x - 1$. Here, we will translate x to the left by 1, where

$$x = \begin{cases} (\mu+\sigma) & P(x) = 0.5 \\ (\mu-\sigma) & P(x) = 0.5 \end{cases} \quad \rightarrow \quad X = \begin{cases} (\mu-1)+\sigma & P(X) = 0.5 \\ (\mu-1)-\sigma & P(X) = 0.5 \end{cases} \tag{50}$$

Then, we have

$$g(f) = 1 + fX \tag{51}$$

This transformation here simplifies the expression of $g(f)$, and will make the calculations less tedious later on. Here, X can be represented as the following

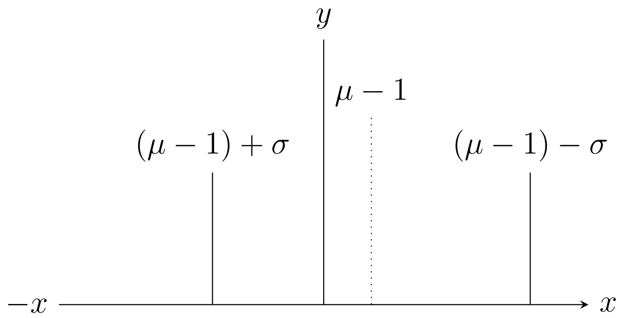


Fig. 3

Where instead of having an axis at $x = 1$, we have it at the y-axis itself. This transformatoin is not arbitrary, since it also has some implications. This new variable X is no longer a "price multiplier", but closer to what is commonly known as the "return" [Finance formula]. It essentially represents how much the price has changed as a fraction of the original price. Note how X can now be negative, which represents a downward change in price.

Now, what we will do differently than the discrete case, is to assume that under the discrete time in which the price multiplied by X, there were actually n sub steps, where in each sub step, the price was multiplied by a smaller multiplier X_n, defined as the following:

$$X_n = \begin{cases} (\frac{\mu-1}{n} + \frac{\sigma}{\sqrt{n}}) & P(X_n) = 0.5 \\ (\frac{\mu-1}{n} - \frac{\sigma}{\sqrt{n}}) & P(X_n) = 0.5 \end{cases} \tag{52}$$

The distribution of X_n can be visually represented in the following diagram:

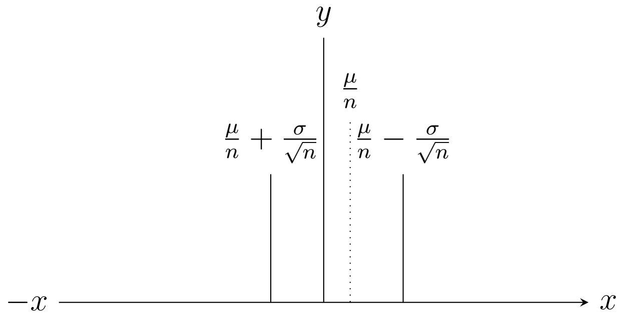


Fig. 4

Then, the value of the asset after n sub step, $V_n(f)$, will be expressed as the following

$$g_n(f) = (1 + fX_1)(1 + fX_2)\cdots(1 + fX_{n-1})(1 + fX_n) \tag{53}$$

which is modified from (51), where $g_n(f)$ overall amount the stock is multiplied by after every single discrete time, expressed in the form of n sub movements. Interestingly, even though X_n is a discrete random variable, $g_n(f)$, the overall return, can take on a wide range of values, and becomes more continuous the more sub movements we consider. Just like how a Binomial distribution approaches to a normal distribution, the idea of considering sub steps makes $g_n(f)$ behave like a continuous random variable. Like previously, we will try to maximize the expected value of $\ln(g_n(f))$. Let $h_n(f) = E(\ln(g_n(f)))$, then

$$h_n(f) = E(\ln(g_n(f))) = \sum_{i=1}^{n} E(\ln(1 + fX_n)) \tag{54}$$

Here, the expected value of the expression can be found using (52), where

$$\begin{aligned} E(\ln(1 + fX_n)) =& \sum \ln(1 + fX_i)P(X_i) \\ =& \frac{1}{2}\ln(1 + f(\frac{\mu-1}{n} + \frac{\sigma}{\sqrt{n}})) + \frac{1}{2}\ln(1 + f(\frac{\mu-1}{n} - \frac{\sigma}{\sqrt{n}})) \\ =& \frac{1}{2}\ln[f^2((\frac{\mu-1}{n})^2 - \frac{\sigma^2}{n}) + 2f(\frac{\mu-1}{n}) + 1] \end{aligned} \tag{55}$$

Substitute (55) into (54), we have

$$\begin{aligned} h_n(f) =& \sum_{i=1}^{n} \frac{1}{2} \ln[f^2((\frac{\mu-1}{n})^2 - \frac{\sigma^2}{n}) + 2f(\frac{\mu-1}{n}) + 1] \\ =& \frac{n}{2} \ln[f^2((\frac{\mu-1}{n})^2 - \frac{\sigma^2}{n}) + 2f(\frac{\mu-1}{n}) + 1] \end{aligned} \tag{56}$$

Furthermore, we can use the Taylor series and expand the ln part at $f = 0$. The Taylor series for $h_n(f)$ is defined as the following, where

$$h_n(f) = h_n(0) + \frac{f}{1!}\frac{dh_n(0)}{df} + \frac{f^2}{2!}\frac{d^2h_n(0)}{df^2} + \cdots \tag{57}$$

For a good approximation, we can simply take the terms of the Taylor series up until the second degree. The derivatives are calculated as the following, where

$$h_n(0) = \frac{n}{2}\ln(1) = 0 \tag{58}$$

$$\frac{dh_n(0)}{df} = \frac{n}{2} \cdot 2(\frac{\mu-1}{n}) = \mu - 1 \tag{59}$$

$$\begin{aligned} \frac{d^2h_n(0)}{df^2} &= n \cdot [(\frac{\mu-1}{n})^2 - \frac{\sigma^2}{n} - 2(\frac{\mu-1}{n})^2] \\ &= -\sigma^2 - (\frac{(\mu-1)^2}{n}) \end{aligned} \tag{60}$$

This results in the following expression:

$$h_n(f) = f(\mu-1) - f^2\frac{\sigma^2}{2} - \frac{(\mu-1)^2}{n} \tag{61}$$

Here, when $\lim_{n\to\infty}$, $\frac{(\mu-1)^2}{n} \to 0$, leaving us with just

$$h_\infty(f) = f(\mu-1) - f^2\frac{\sigma^2}{2} \tag{62}$$

This derivation is modified from [Thorp, Hung]. Now, to find the maximum of this function, we take it's derivative and equate it to 0,

$$\frac{dh_\infty(f)}{df} = \mu - 1 - f\sigma^2 = 0 \tag{63}$$

Which gives us the optimum fraction f at

$$f = \frac{\mu-1}{\sigma^2} \tag{64}$$

Substituting in the values, with $\sigma^2 = 0.004113$ and $\mu = 1.001915$(generated from the dataset by code in the Appendix), we have the optimum investing fraction $f = \dfrac{\mu - 1}{\sigma^2} = \dfrac{1.001915 - 1}{0.004113} = 46.5\%$.The continuous time model gives a larger investment fraction, which means if we could trade arbitarily fast, it is more viable to invest a larger proportion of our asset into Gamestop. The continuous time model looks at the variance, which takes into consideration of all the data points, while previously in (11) and (12), we used two data points to generate the probability value and the investment fraction. Therefore, we can conclude that this method gives a more accurate investing fraction. However, one thing we didn't take into account is that realistically, we can never trade fast enough for one stock, so the continuous approximation may not be as good as our previous discrete time model, even though theoretically it is more accurate. Therefore, this method also has its limitations.

Conclusion

Overall, we achieved our aim of finding an optimum fraction for stock investment, and derived many methods using different models of the price. However, one major assumption we made was that we can trade a very large number of times. How likely will the value of our assets deviate from the expected value in the short run, where the law of large numbers don't apply? When N is small enough, the likelihood of deviation becomes large, because we cannot trade an arbitrarily large numbers of times. Will it still be optimal to invest at f? In fact, a lot of other papers, such as [Thorp], suggests investing half the optimum, or even less to reduce the risk in the short run.

Moreover, there are large impacts if the modeling wasn't accurate. We only have a few hundred data points to construct the probability histogram, which is not enough for the distribution to be accurate. More data is needed for a closer and better approximation. Furthermore, the money we make is quite sensitive to the investing fraction. The inaccuracy in our modeling could potentially lead to significant over or under investing, making it far away from optimum. To improve, we must gather finer data at smaller time frames.

Lastly, another assumption we made was, the price at time kt is independent from the price at $(k+1)t$, but this is unlikely to be true. In fact, when a price drops sharply, human emotions such as fear will incentive more people to sell, so the price at one point in time is definitely dependent on the price preceding that. Therefore, there are a lot of areas that can be refined and improved upon.

Appendix

A Litecoin data snippets

Date	Open	Close	Multiplier
2018-02-09	149.72999	163.94999	1.094971
2018-02-12	150.10404	161.56599	1.076360
2018-02-13	161.77499	159.55400	0.986271
2018-02-14	159.57899	213.35899	1.337012
2018-02-15	212.34700	225.42599	1.061593

...

Date	Open	Close	Multiplier
2021-02-02	131.95291	142.50697	1.079984
2021-02-03	142.47673	155.61203	1.092193
2021-02-04	155.65647	145.14819	0.932491
2021-02-05	145.15063	154.85221	1.066838
2021-02-08	151.10751	167.21905	1.106623

B Source code for graphs and data

`https://github.com/BienBienBuen/stock-market-bot`

References

[Finance formula] "Total Stock Return - Formula (with Calculator)." Finance Formula, `financeformulas.net/Total-Stock-Return.html`. Accessed 2 April 2021.

[Hung] Hung, Jane. "Betting with the Kelly criterion." Massachusetts Institute of Technology (2010).

[Thorp] Thorp, Edward O. (June 1997). "The Kelly criterion in blackjack, sports betting, and the stock market" (PDF). 10th International Conference on Gambling and Risk Taking. Montreal.

[Khan] Khan, Sal. "Intro to the Binomial Theorem." Khan Academy, `www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:polynomials/x9e81a4f98389efdf:binomial/v/binomial-theorem`. Accessed 16 Apr. 2021.

[Kelly] Kelly, J. L. (1956). "A New Interpretation of Information Rate" (PDF). Bell System Technical Journal. 35 (4): 917–926.

[Ortiz-Latorre] Ortiz-Latorre, Salvador. "Cox-Ross-Rubinstein and Black-Scholes Models." University of Olso, (2018).

[Yahoo Finance] Yahoo Finance. Verizon Media, `https://finance.yahoo.com/`. Accessed 14 Feb. 2021.