Use of Stochastic Calculus to Predict Stock Market Trends

Arnav Patri and Shashank Chidige

The aim of this project is to use differential equations to determine a method by which the behavior of stock prices can be modeled to gain financial insight. There are simply too many variables that have an impact on a stock's price, though, not all of which can be individually accounted for. The problem of optimizing buy and sell times of stocks is affected by the multitude of internal and external forces that affect the stock's price. In business, political, economic, social, and technological changes in the business environment can impact a stock's price. Rather than attempting to account for every individual variable, the aggregate result can be modeled stochastically.

Despite its complexity, an understanding of both statistics and differential equations being required, stochastic calculus enables traders to make more informed predictions regarding stock prices. Many firms already employ stochastic models for this reason. They stand to gain financially by increasing the accuracy of their prediction. In fact, many trading services directly incorporate stochastic models unbeknownst to those using them.

The model being employed is

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = \mu S_t + \sigma S_t \frac{\mathrm{d}W_t}{\mathrm{d}t}$$

where t is time (the independent variable), S_t is the stock price (as a function of time), μ is the (constant) expected return on the stock, σ is the (constant) volatility (standard deviation), and W_t is a standard Wiener process, with mean 0 and variance 1, as adapted from [2]:

The Wiener process is stochastic, meaning that its value changes over time randomly. As such, only the distribution of possible values is known at any given point in time. This distribution is defined by the mean and variance, in this case 0 and 1 respectively. The Wiener process is a series of randomly distributed normal variables with variances that increase over time, reflecting the increasing uncertainty in making predictions further into the future.

The change in the Wiener process can be found as

$$\Delta W = \varepsilon \sqrt{\Delta t}$$

where $\varepsilon \sim \mathcal{N}(0,1)$ (ε being a continuous random variable following the standard Normal distribution, but mean 0 and variance 1). This implies that the change in the Wiener process is itself a transformation of the standard Normal distribution, as $\sqrt{\Delta t}$ is simply some constant multiplier when parameter t is fixed.

As such, the mean of the Wiener process (its expected value $\mathbb{E}[W_t]$) can be found to be

$$\mathbb{E}[\Delta W] = \sqrt{\Delta t} \times \mathbb{E}[\varepsilon] = 0$$

and its variance

$$\mathrm{var}[x] = \left(\sqrt{\Delta t}\right)^2 \mathrm{var}[\varepsilon] = \Delta t$$

The Wiener process can therefore be denoted $\mathcal{N}(0, \Delta t)$.

The difference between the Wiener process at times T and 0 can be found as

$$W(T) - W(0) = \sum_{i=1}^{n} \varepsilon_i \sqrt{\Delta t} = \sum_{i=1}^{n} W_i \quad \text{where } n = \frac{T}{\Delta t}$$

As $W(T) - W(0)$ is simply a linear combination of n random normal variables W_i, it must itself also follow a Normal distribution. It should be noted that $W(0)$ must be fixed, as it is what defines the function's starting point. As such,

$$\mathbb{E}[W(0)] = \mathrm{var}[W(0)] = 0$$

It can then be seen that

$$\mathbb{E}[W(T) - W(0)] = \mathbb{E}[W(T)] - \mathbb{E}[W(0)] = \mathbb{E}[W(T)]$$

As the difference is a linear combination of ΔW_t n times,

$$\mathbb{E}[W(T)] = n \times \mathbb{E}[\Delta W_t] = 0$$

The distributions of the values Wiener process at times t_i and t_j are independent, so

$$\text{var}[W(T) - W(0)] = \text{var}[W(T)] + \text{var}[W(0)] = \text{var}[W(T)] = n \times \text{var}[\Delta W_t] = n\Delta t$$

In conclusion, $W(T) \sim \mathcal{N}(0, T)$.
As $n \to \infty$, Δt and ΔW_t become the differentials $\mathrm{d}t$ and $\mathrm{d}W_t$. The only conditions that the model must follow are that the stock price and time cannot fall below 0. Note the lack of an upper bound to the stock's price.

Table 1 shows information regarding Alphabet Inc., which is listed on the New York Stock Exchange as GOOG, as the company used to be known as Google. The data is taken from Yahoo Finance [1], which provides financial news and data for public use. The domain of the independent variable, time (t) is taken from 9/27/2022 to 10/15/2022.

Table 1	
Stock Price (USD)	Time (Days from start date of 9/27/22)
141.501007	0
136.184006	1
134.520996	2
133.265503	3
136.462494	4
133.764999	5
136.177002	6
137.354004	7
139.185501	8
140.056	9
138.847504	10
136.712997	11
137.899994	12
141.412003	13
141.675003	14
142.960495	15
143.822006	16
142.414993	17
142.780502	18
138.625	19
138.772995	20
139.671997	21
146.427505	22
146.128998	23
148.270493	24
143.774002	25
145.863007	26
146.789993	27
148.682999	28

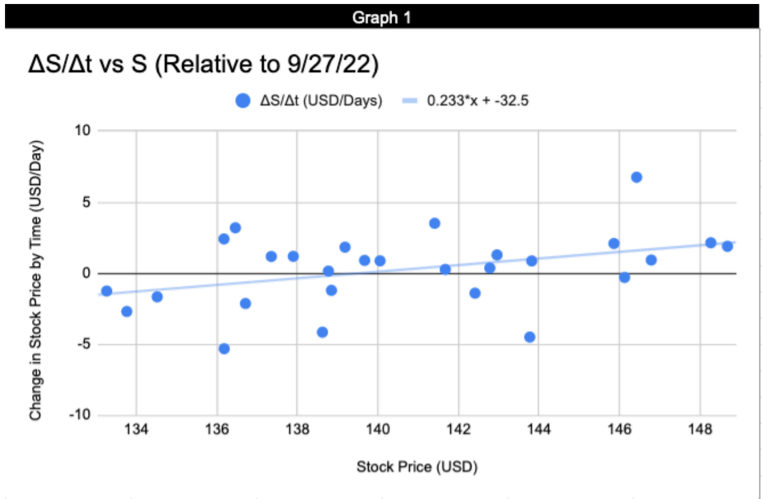
Table 2 shows the differential approach to deriving the differential equation model. As the stochastic part cannot be derived, only the derivation of the deterministic part is shown. Columns 1 and 2 are simply copied from Table 1. Equations 1 and 2 are used to derive the changes in stock price S_t and time t, put into columns 3 and 4 respectively. The derivative of S_t with respect to t can be approximated as $\Delta S_t/\Delta t$, as shown in column 5. The stock price S_t is then listed again, as the plot is of $\Delta S_t/\Delta t$ vs S_t.

Table 2					
Stock Price (USD)	Time (Days from start date of 9/27/22)	*ΔS** (USD)	*Δt*** (Days)	*ΔS/Δt* (USD/Days)	Stock Price (USD)
141.501007	0				141.501007
136.184006	1	-5.317001	1	-5.317001	136.184006
134.520996	2	-1.66301	1	-1.66301	134.520996
133.265503	3	-1.255493	1	-1.255493	133.265503
136.462494	4	3.196991	1	3.196991	136.462494
133.764999	5	-2.697495	1	-2.697495	133.764999
136.177002	6	2.412003	1	2.412003	136.177002
137.354004	7	1.177002	1	1.177002	137.354004
139.185501	8	1.831497	1	1.831497	139.185501
140.056	9	0.870499	1	0.870499	140.056
138.847504	10	-1.208496	1	-1.208496	138.847504
136.712997	11	-2.134507	1	-2.134507	136.712997
137.899994	12	1.186997	1	1.186997	137.899994
141.412003	13	3.512009	1	3.512009	141.412003
141.675003	14	0.263	1	0.263	141.675003
142.960495	15	1.285492	1	1.285492	142.960495
143.822006	16	0.861511	1	0.861511	143.822006
142.414993	17	-1.407013	1	-1.407013	142.414993
142.780502	18	0.365509	1	0.365509	142.780502
138.625	19	-4.155502	1	-4.155502	138.625
138.772995	20	0.147995	1	0.147995	138.772995
139.671997	21	0.899002	1	0.899002	139.671997
146.427505	22	6.755508	1	6.755508	146.427505
146.128998	23	-0.298507	1	-0.298507	146.128998
148.270493	24	2.141495	1	2.141495	148.270493
143.774002	25	-4.496491	1	-4.496491	143.774002
145.863007	26	2.089005	1	2.089005	145.863007
146.789993	27	0.926986	1	0.926986	146.789993
148.682999	28	1.893006	1	1.893006	148.682999

$$\Delta S_{t,i} = S_{t,i} - S_{t,i-1} \qquad \text{(Equation 1*)}$$

$$\Delta t_i = t_i - t_{i-1} \qquad \text{(Equation 2**)}$$

This is the graph of $\Delta S/\Delta t$ (in USD/day) vs S (in USD), using data from Table 2, with a linear regression performed.



It is clear that there is a linear relationships between the stock price and its derivative with respect to time. It can therefore be said that

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} \propto S_t \implies \frac{\mathrm{d}S_t}{\mathrm{d}t} = \mu S_t$$

where μ is a constant. Using information from [2], the Wiener process can be appended, adding the random motion to the model:

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = \mu S_t + \sigma S_t \frac{\mathrm{d}W_t}{\mathrm{d}t}$$

where S_t is stock price as a function of time t, μ is the (constant) drift, σ is the (constant) volatility, and W_t is a standard Wiener process.

The Wiener process is stochastic, meaning that its value changes over time randomly. As such, only the distribution of possible values is known at any given point in time. This distribution is defined by the mean and variance, in this case 0 and 1 respectively. The Wiener process is a series of randomly distributed normal variables with variances that increase over time, reflecting the increasing uncertainty in making predictions further into the future.

The dependent variable t is not present, making this an autonomous DE. The only information needed for the model are μ, σ, and the initial stock price S_0.

μ is the amount that $\mathbb{E}(S_t)$, the expected value of the stock price, changes per year, making it the coefficient of the linear regression of $\Delta S_t/\Delta t$ against S_t divided by 365, so $\mu \approx 0.00064$.

σ is simply the standard deviation of the stock price in the sample, so

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(S_{t,i} - \bar{S}_t)^2}{n-1}} \approx 4.356$$

where n is the number of days sampled, $S_{t,1\cdots n}$ are the particular stock prices, and $\bar{S}_t$ is the average stock price in the sample:

$$\bar{S}_t = \frac{\sum_{i=1}^{n} S_{t,i}}{n}$$

The model then becomes

$$\boxed{\frac{\mathrm{d}S_t}{\mathrm{d}t} = 0.00064 S_t + 4.356 S_t \frac{\mathrm{d}W_t}{\mathrm{d}t}}$$

The randomness term ensures that each trial of the model yields a different graph. One such trial is attached at the end of this document.

Itô's Lemma and Separation of Variables

The DE

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = \mu S_t + \sigma S_t \frac{\mathrm{d}W_t}{\mathrm{d}t}$$

where S_t is stock price (the dependent variable in USD) as a function of time t (the independent variable in days), μ is the (constant) drift, σ is the (constant) volatility, and W_t is a standard Wiener process (written in terms of the parameters solved for previously) can be rewritten as

$$\mathrm{d}S_t = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t$$

where $\mu_t = \mu S_t$ and $\sigma_t = \sigma S_t$. Integrating and reparameterizing μ_t, σ_t and W_t with s,

$$S_t = \int_0^t \mu_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s + C$$

where C is the constant of integration. This makes S_t an Itô process, a stochastic process expressible as the sum of two integrals, one with respect to a stochastic process and another with respect to time, and a constant.
The Taylor expansion of a twice-differentiable scalar function $f(t, s)$ is

$$\mathrm{d}f = \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial s}\,\mathrm{d}s + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}\,\mathrm{d}^2 s + \cdots$$

Substituting S_t for s and appropriately substituting for $\mathrm{d}s$ yields

$$\mathrm{d}f + \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial s}(\mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t) + \frac{\partial^2 f}{\partial s^2}\left(\mu_t^2 \,\mathrm{d}t^2 + 2\mu_t \sigma_t \,\mathrm{d}t \,\mathrm{d}W_t + \sigma_t^2 \,\mathrm{d}{W_t}^2\right) + \cdots$$

As $\mathrm{d}t$ approaches 0, $\mathrm{d}t^2$ and $\mathrm{d}t\,\mathrm{d}W_t$ tend to zero faster than $\mathrm{d}W_t^2$. Substituting 0 for $\mathrm{d}t^2$ and $\mathrm{d}t\,\mathrm{d}W_t$ and $\mathrm{d}t$ for $\mathrm{d}W_t^2$ yields

$$\mathrm{d}f = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial s} + \frac{\sigma_t^2}{2}\frac{\partial^2 f}{\partial s^2}\right)\mathrm{d}t + \sigma_t \frac{\partial f}{\partial s}\,\mathrm{d}W_t$$

This is itself an Itô process. Itô's lemma states that for any Itô process S_t and any twice-differentiable function $f(t, s)$, $f(t, S_t)$ is an Itô process.
Let $f(S_t) = \ln S_t$. Applying Itô's lemma,

$$\begin{aligned}
\mathrm{d}f &= f'(S_t)\,\mathrm{d}S_t + \frac{1}{2}f''(S_t)(\mathrm{d}S_t)^2 \\
&= \frac{1}{S_t}\,\mathrm{d}S_t - \frac{1}{2S_2^2}\left(S_t^2 \sigma^2 \,\mathrm{d}t\right) \\
&= \frac{1}{S_t}(\sigma S_1 \,\mathrm{d}W_t + \mu S_t \,\mathrm{d}t) - \frac{\sigma^2}{2}\,\mathrm{d}t \\
&= \sigma \,\mathrm{d}W_t + \left(\mu - \frac{\sigma^2}{2}\right)\mathrm{d}t
\end{aligned}$$

The Wiener process greatly limits the possibilities for solving this DE. It is inexact, nonlinear, and nonhomogenous. The integrating factor cannot be used either due to the Wiener process. As such, separation of variables is the only suitable method. Integrating the separable DE,

$$f(t) = \ln S_t = \ln C + \sigma W_t + \left(\mu - \frac{\sigma^2}{2}\right) \mathrm{d}t$$

This can finally be exponentiated, yielding S_t:

$$S_t = C\,\mathrm{e}^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t}$$

(This method of solution was largely adapted from [3] and [4])
Applying the initial condition $S_t(0) = S_0$, the initial stock price of \$141.501. Letting $t = 0$,

$$S_t(0) = Ce^{\sigma W_t}$$

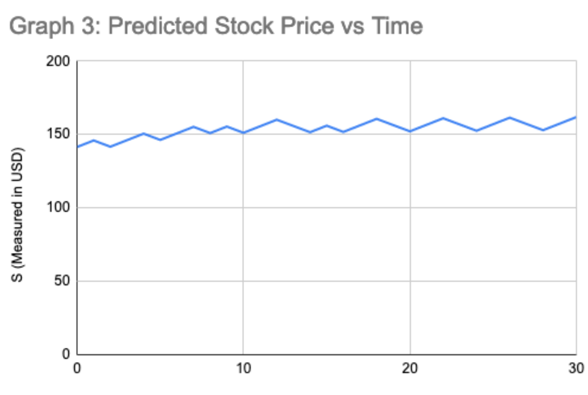
At time $t = 0$, the randomness from the Wiener process does not have any affect, is it only takes hold as t gets further from the initial time of 0. As such, it is equal to 0, so

$$S_t(0) = C = S_0 = \$141.501$$

Substituting for μ and σ,

$$\boxed{S_t = 141.501\,\mathrm{e}^{\left(0.00064 + \frac{4.356^2}{2}\right)t + 4.356W_t} = 141.501\,\mathrm{e}^{9.488t + 4.356W_t}}$$

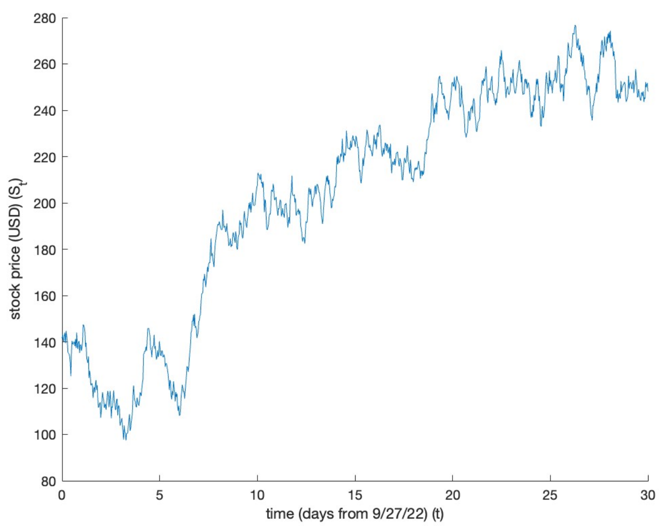
It should be noted that the Stock's price may not fall below 0, as the solution is an exponential with a positive coefficient.
Below is the graph of the solution in Excel:



the randomness term was simulated using the random walk approach (adapted from [5]), generating random numbers and compounding the change between them to create a random process with uncertainty that increases with time. The following are the MatLab code, solution (for the deterministic term), and graph of the model of S_t vs t:

```
Solution.m

close all;
clear all;

dom = [0, 30];
n = 1000;
dt = (dom(2) - dom(1))/50;
S0 = 141.501;

ndist = makedist('Normal', 0 , sqrt(dt));

mu = 0.00064;
sigma = 4.356;

tvals = linspace(dom(1), dom(2), n);
Svals = zeros(1, n);
Svals(1) = S0;
for i = 2:n
    t = dom(1) + (i - 1) .* dt;
    S = Svals(i - 1);
    dW = random(ndist);
    Svals(i) = S + mu .* dt + sigma .* dW
end
figure()
hold on;
plot(tvals, Svals);
xlabel('time (days from 9/27/22)')
ylabel('stock price (USD)')

% Solving for the non-stochastic term
syms s(t)
ode = diff(s, t) == s;
sol(t) = dsolve(ode);
```

```
sol
1x1 symfun

val(t) =

C1*exp(t)
```



Bibliography

[1] Yahoo Finance, *Alphabet Inc. (GOOG) Stock Price, News, Quote & history - Yahoo Finance*, New York, NY, 2022.

[2] Gregory F. Lawler, *Stochastic Calculus: An Introduction with Applications*, Chicago, IL, 2014.

[3] Wenyu Zhang, *Introduction to Itô's Lemma*, Ithaca, NY, May 6th, 2015

[4] Andrea Chello, *A Gentle Introduction to Geometric Brownian Motion in Finance*, October 30th, 2020

[5] Charles Zaiontz, *Random Walk*, 2021