Stochastically Modeling Stock Price using Differential Equations

Arnav Patri and Shashank Chidige

College of Art and Science, Lawrence Technological University

Introduction

It is incredibly advantageous to be able to make accurate predictions regarding the values of specific stocks. Doing so accurately is near impossible, though, due to the sheer number of external factors that simply cannot be accounted for realistically. The aggregate of all these variables can instead be considered. This resultant randomness can be modeled to some extent using a Wiener process [2].

The Wiener process is a stochastic process, meaning that it varies randomly with time. More specifically, it is comprised of the linear combination of infinite random variables that each follow the standard normal distribution; that is, random variables of mean 0 and variance 1. This creates a process that becomes increasingly unpredictable as time goes on, as randomness compounds, which is exactly what is needed for predicting stock price.

Let ε be a continuous random variable that follows a standard normal distribution; that is, $\varepsilon \sim \mathcal{N}(0,1)$. The change in the Wiener process can then be found as $\Delta W = \varepsilon\sqrt{\Delta t}$, where W is the Wiener process and t is time and Δ, when preceding a variable, denotes the change in the value of that variable. For any given Δt, then, ΔW is simply a constant multiple of ε, as $\sqrt{\Delta t}$ is rendered a constant. As such, the mean (or expected value) of the change in the Wiener process $\mathbb{E}[\Delta W]$ (where $\mathbb{E}$ is the expectation operator, returning the expected value of the argument) is simply $\sqrt{t} \times \mathbb{E}[\varepsilon]$. Recall, however, that ε is a standard Normal variable, making $\mathbb{E}[\varepsilon]$ and by proxy $\mathbb{E}[\Delta W] = 0$. The variance of the change in the Wiener process $\mathrm{var}[\Delta W]$ (where var returns the variance of the argument) can be found to be $\left(\sqrt{\Delta t}\right)^2 \times \mathrm{var}[\varepsilon]$. ε being a standard Normal variable means that $\mathrm{var}[\varepsilon]$ is 1, making $\mathrm{var}[\Delta W] = \Delta t$. The change in the Wiener process can therefore be denoted $\Delta W \sim \mathcal{N}(0, \Delta t)$.

The difference between the Wiener process at times T and 0, or $W(T) - W(0)$, can be found as $\sum_{i=1}^{n} \varepsilon_i\sqrt{\Delta t}$ where $n = T/\Delta t$ is the number of steps of size Δt being taken to reach T from 0. From this, it can be seen that $W(T) - W(0)$ is a linear combination of n random Normal variables, meaning that it must itself also follow a Normal distribution.

The only point at which the Wiener process is defined is the initial time 0. If it is fixed to 0, as in the case of the standard Wiener process used here, $\mathbb{E}[W(0)] = \mathrm{var}[W(0)] = 0$. 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 this difference is simply a linear combination of ΔW_i n times, $\mathbb{E}[W(T)] = n \times \mathbb{E}[\Delta W_i] = 0$.

The distributions of the values of the Wiener process at two distinct times are independent, so the variance of their difference is simply the sum of their variances. As such, $\mathrm{var}[W(T) - W(0)] = \mathrm{var}[W(T)] + \mathrm{var}[W(0)] = \mathrm{var}[W(T)]$. Using the fact that this difference is simply a combination, $\mathrm{var}[W(T)] = n \times \mathrm{var}[\Delta W_i] = n\Delta t$.

In summary, $W(T) \sim \mathcal{N}(0, T)$. As $n \to \infty$, $\Delta t \to 0$ and $\Delta W_i \to 0$, turning them into the differentials dt and dW_i respectively.

A baseline prediction can be made using recent data regarding the stock's value. This creates a sort of "through-line" about which the randomness of the Wiener process occurs. This through-line can be found by treating the model as exponential, disregarding the randomness term. The growth rate is then the drift μ of the stock, which is the change in the expected value of the stock price (which is different than the expected value of the change of the stock price).

As the Wiener process is inherently random, each simulation trial yields a different result. Examining multiple trials enables a more informed, accurate prediction to be made regarding the stock's future value.

An understanding of both statistics and differential equations is required for traders to be able to make educated predictions regarding a stocks' future pricing. Many firms already employ stochastic models such as that presented here in their own indicators. By using a more realistic model, traders are better able to make predictions, allowing investors to benefit more from financial gain.

Image 1: Example of Stock Market Volatility: Alphabet Inc. [1]

Image 2: A look at the floor of the New York Stock Exchange

Sources

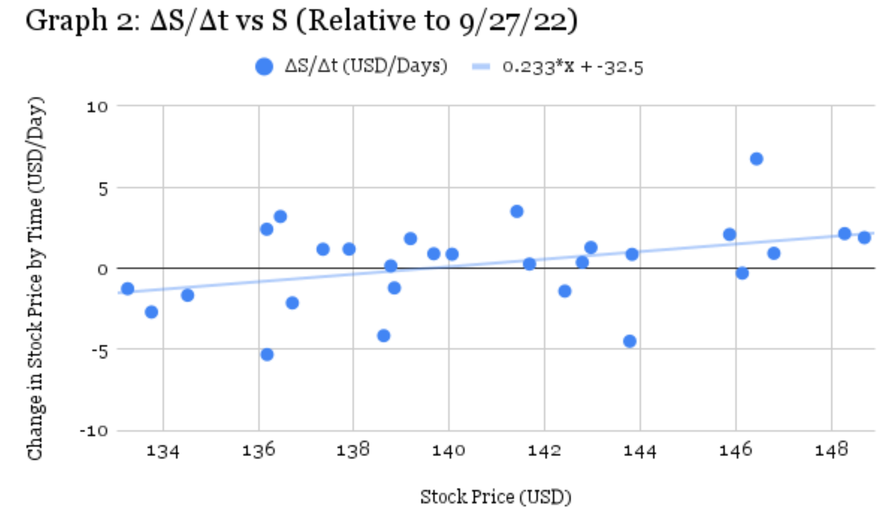
[1] Yahoo Finance, Alphabet Inc. (GOOG) Stock Price, News, Quote & history - Yahoo Finance, New York, NY, 2022.

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

[3] Wenyu Zhang, Introduction to It^o's Lemma, Ithaca, NY, May 6th, 2015

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

Graph 2: ΔS/Δt vs S (Relative to 9/27/22)



Statistical Data

The case study that the model was applied to is the change in GOOG's (Alphabet Inc.'s) stock price over the span of six months from 9/27/21 to 2/23/22. The reason for applying this model to model to Google is that it is one of the most formative blue chip stocks, making its trend is largely indicative of that of others. The data gathered from [1] is collected from its free API, which enabled the data to be transferred to Google Sheets.

Graph 1: Google Stock Price over Time

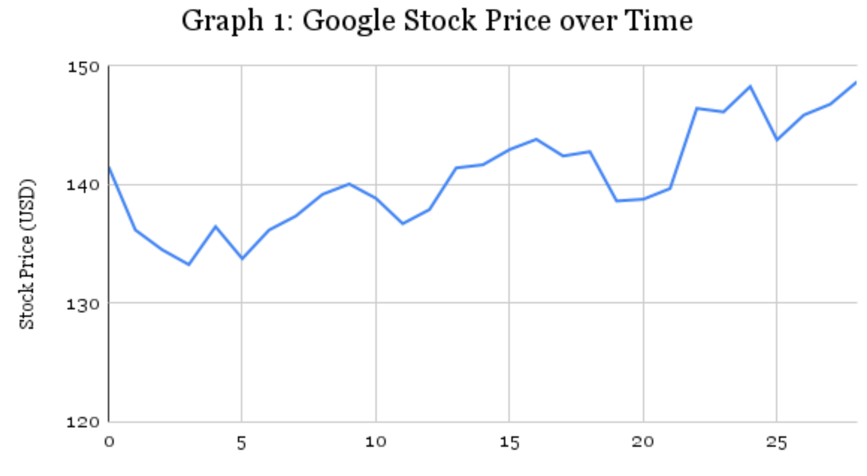


Table 1

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

Table 1 (cont.)

Stock Price (USD)	Time (Days from start date of 9/27/22)
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

Model Analysis

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

In order to derive the drift μ, the deterministic part of the model must be modeled as an exponential. Table 2 calculates the change in each variable. Performing a regression of ΔS vs t, performed on Graph 2, yields a slope of 0.233, but this must be divided by 365 to normalize the drift of the daily step, yielding

$$\mu \approx 0.00064$$

The volatility σ is simply the standard deviation of the stock price, making it

$$\sigma = \frac{\sum (S_{t,i} - \bar{S}_t)}{n-1} \text{ where } \bar{S}_t = \frac{\sum S_{t,i}}{n} \text{ and } n \text{ is the number of days tested}$$

which yields a result of $\sigma \approx 4.356$, making the model

$$\frac{dS_t}{dt} = 0.00064S_t + 4.356S_t\frac{dW_t}{dt}$$

Conclusion

This project yielded a model that can be used to predict the stock price of Alphabet Inc using data collected up to the date 10/25/2022, after which the stock price is to be predicted. This model can also be generalized to other stocks by applying the same differential equation model with a different set of time-series data of Stock Price (S_t) vs. Time (t) of that stock. Modeling stock prices allows more informed decisions to be made, taking into account more accurate predictions to make choices. The model is stochastic, so figuring out how to implement it required an understanding of stochastic calculus. Stochastic calculus is built on both calculus and probability, so learning about it forced us to better understand both. This project made clear the power of differential equations to obtain a family of functions which can be widely applied given parameters and boundary conditions. In this case, it was the initial stock price (S_0) that was the boundary condition for the initial value problem.

"Employ differential equations to make pragmatic predictions regarding something that is otherwise known to be quite random." - Arnav Patri

"Employ data to predict future outcomes and obtain quantifiable results by applying mathematics to a real-life problem." - Shashank Chidige

Alphabet Inc (GOOG) Price Model

Itô's Lemma and Separation of Variables

A stock's price S_t (in USD) can be predicted with respect to time t (in days from the starting time) using a growth function; that is, by relating the rate at which the stock price is changing to the current stock price as a proportion:

$$\frac{dS_t}{dt} \propto S_t$$

The constant of proportionality is the drift μ of the stock, which is the rate of change of the expected value of the stock price; that is,

$$\frac{dS_t}{dt} = \mu S_t \quad \text{where} \quad \mu = \frac{\Delta\mathbb{E}[S_t]}{\Delta t}$$

The stock market is constantly volatile, though, resulting in a stock's price changing in unpredictable ways. This element of randomness can be modeled by a randomness term, which is the product of the stock price and the rate of change of the standard Wiener process W_t with respect to time:

$$\text{rate of randomness} \propto S_t\frac{dW_t}{dt}$$

The proportionality constant is the volatility σ of the stock, which is simply its standard deviation over some timeframe over which data is collected:

$$\text{rate of randomness} = \sigma S_t\frac{dW_t}{dt} \quad \text{where}$$

$\sigma = \dfrac{\sum (S_{t,i} - \bar{S}_t)}{n-1}$ where $S_{t,i}$ is the stock's price at time i, n is the number of entries being used to calculate the volatility, and $\bar{S}_t = \dfrac{\sum S_{t,i}}{n}$

Laid out mathematically, the differential equation model is

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where t is time (the independent variable), S_t is the stock's price as a function of time (the dependent variable), μ is the stock's drift (the constant change in its expected value), σ is the stock's volatility (the constant standard deviation), and W_t is a standard Weiner process (with each random Normal variable having mean 0 and variance 1).

This differential equation can be rewritten as

$$dS_t = \mu_t dt + \sigma dW_t$$

where $\mu_t = \mu S_t$ and $\sigma_t = \sigma S_t$. Integrating from 0 to t after reparameterizing μ_t, σ_t, and W_t with s yields

$$S_t = \int_0^t \mu_s\, ds + \int_0^t \sigma_s\, dW_s + C$$

where C is the constant of integration (the initial stock price).

An Itô process is a stochastic process that is expressible as the sum of 2 integrals, one with respect to a stochastic process and another with respect to time, and a constant. The above expression for S_t is evidently an example of one.

The Taylor expansion of a twice-differentiable scalar function $f(t, s)$ is

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial s}ds + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}d^2 s + \cdots$$

Substituting S_t for s and appropriately substituting for ds yields

$$df + \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial s}(\mu_t dt + \sigma_t dW_t) + \frac{\partial^2 f}{\partial s^2}\left(\mu_t^2 dt^2 + 2\mu_t\sigma_t dt dW_t + \sigma_t^2 dW_t^2\right) + \cdots$$

As dt approaches 0, dt^2 and $dtdW_t$ tend to 0 faster than dW_t^2. Substituting 0 for dt^2 and $dtdW_t$ and dt for dW_t^2 yields

$$df = \left(\frac{\partial f}{\partial t} + \mu_t\frac{\partial f}{\partial s} + \frac{\sigma_t^2}{2}\frac{\partial^2 f}{\partial s^2}\right)dt + \sigma_t\frac{\partial f}{\partial s}dW_t$$

which 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.

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,

$$\frac{dS_t}{S_t} = \sigma dW_t + \left(\mu - \frac{\sigma^2}{2}\right)dt$$

Integrating yields

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

where C is the constant of integration. Exponentiating,

$$S_t = Ce^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t} \quad \text{where} \quad C = e^C$$

(This solution method was largely adapted from [3] and [4]) At time $t = 0$, S_t is equal to the initial stock price S_0, so

$$S_t(0) = S_0 = Ce^0 = C = 141.501$$

Substituting the values of the constants S_0, μ, and σ,

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

In MatLab,

```
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) (t)')
ylabel('stock price (USD) (S_t)')

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

Graph 3: MatLab Predicted Stock Price vs Time