STOCHASTIC CALCULUS AND THE BLACK-SCHOLES-MERTON MODEL

NOAH FISCHER

Abstract. We set up the mathematical background to construct the Black-Scholes-Merton differential equation and its solution. This includes Brownian motion and stochastic calculus. After the mathematical introduction, we show the derivation and solutions to the Black-Scholes-Merton differential equation from the canonical heat equation. Finally, we challenge the assumptions of the model using empirical data analysis of AAPL stock returns.

Contents

Introduction

First, we will establish a mathematical background for use in later sections. We will discuss Brownian motion and the building blocks of stochastic calculus. These topics are extremely useful in the modeling and understanding of random events, particularly those dependent on time. Further, these sections are requisite for the understanding and derivation of the Black-Scholes-Merton differential equation. After this derivation, we will solve the differential equation to yield the Black-Scholes-Merton options pricing model. We will then discuss the assumptions of the

Date: August 28, 2021.

model and prove through empirical data analysis of AAPL stock returns why the assumption of stock price as a geometric Brownian motion does not hold in reality.

1. Brownian Motion

We will define a **Brownian Motion (Wiener Process)** as follows.

Definition 1.1. A stochastic process B_t is a **Brownian Motion (Wiener Process)** with respect to time t if:

i) $B_0 = 0$
ii) for all $r \leq s < t$, the interval $B_t - B_s$ is independent of B_r
iii) for all $s < t$, the interval $B_t - B_s \sim N(\mu(t-s), \sigma^2(t-s))$
iv) B_t is a continuous path almost surely

Remark 1.2. Note that when the interval follows the standard normal distribution $N(0,1)$ we call the process a **Standard Brownian Motion**

We will now show the existence of Brownian motion following the proof in Lawler (2014) [1]. The sketch is as follows. First, we begin by showing that Brownian motions are defined on all times t in the set of dyadic rationals, which is a countable, dense set in the reals. Then, we show that the function mapping times t to values B_t is continuous on the dyadic rationals. Finally, we extend this continuity to the real line by properties of sequential continuity. We show the proof below.

Proposition 1.3. *Brownian motions are well defined functions mapping times t to values B_t.*

The proof of the previous proposition is beyond the scope of this paper, however an interested reader can find it in Lawler (2014) p45. Now, let us consider some basic properties of Brownian motion.

Observation 1.4. A Brownian motion has the following properties:

i) It is almost surely nowhere differentiable.
ii) It exhibits the **Strong Markov Property**.
iii) It exhibits the **Reflection Principle**.

Let us consider each property. Property i) states that with probability one, the derivative of a Brownian motion does not exist at t for any t. We will further define Property ii) and iii). First, let us consider the definition and uses of stopping times.

Definition 1.5. A stopping time is a positive time value t at which a stochastic process exhibits a certain property.

Example 1.6. Let X_t be a stochastic process with state space K. We call $\tau = \inf\{t \mid X_t = k\}$ for some $k \in K$ a stopping time for state k.

Property 1.7. Let B_t be a Brownian motion. Let T be a stopping time with respect to B_t such that $T < \infty$ with probability one and let

$$Y_t = B_{T+t} - B_T$$

The **Strong Markov Property** states that Y_t is a standard Brownian motion independent of $\{B_r \mid 0 \leq r \leq T\}$.

Property 1.8. Let B_t be a standard Brownian Motion with $B_0 = 0$ and define $\Phi(t) = \mathbb{P}\{Z \leq t\}$ for some $Z \sim N(0,1)$. The **Reflection Principle** states that for all $a > 0$,

$$P\left\{\max_{0\leq s\leq t} B_S \geq a\right\} = 2\mathbb{P}\{B_t > a\} = 2\left(1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right)$$

An interested reader can find the proof for Property 1.7 in Lawler (2014) p60. Further discussion on stopping times and their properties are useful in the explanation of Brownian motion.

2. Stochastic Calculus

We now introduce background for stochastic integration. We begin the discussion with a comparison between Riemann and Ito integrals. First, consider the standard Riemann integral. A one dimensional integral of an integrable function f over an interval $[a,b] \in \mathbb{R}$ begins by partitioning the interval into subintervals indexed for instance as $a = t_0 < \cdots < t_i < \cdots < t_n = b$. Then, we consider step function approximations at each subinterval $f_i(t) = f(s_i)$ for $s_i, t \in [t_{i-1}, t_i]$. Then, we define the integral as

$$\int_a^b f(t)dt = \lim_{n\to\infty} \sum_{i=1}^{n} f(s_i)(t_i - t_{i-1})$$

We know from analysis that this limit exists and is independent of our choice of partition.

Now, let us discuss the stochastic analogue of the Riemann integral. Let us partition our time interval such that

$$0 = t_0 < \cdots < t_i < \cdots < t_n = T < \infty$$

Consider a collection of random variables Y_j for $j = 0, 1, \ldots, n$ that are $\mathcal{F}_{t_j}$ measurable. The stochastic analogue to a step function is a simple process. We call A_t a simple process if $A_t = Y_j$ for $t_j \leq t < t_{j+1}$. We let $t_{n+1} = \infty$. By the $\mathcal{F}_{t_j}$ measurabiltity of each Y_j, we then have that A_t is $\mathcal{F}_t$ measurable. We assume that $\mathbb{E}[Y_j^2] < \infty$ for all j. Now, we define the stochastic integral with respect to an infinitesimal change in a Brownian motion dB_s as

$$Z_t = \int_0^t A_s dB_s$$

Then, we define

$$Z_{t_j} = \sum_{i=0}^{j-1} Y_j(B_t - B_{t_j})$$

and finally

$$Z_t = Z_{t_j} + Y_j(B_t - B_{t_j})$$

$$\int_r^t A_s dB_s = Z_t - Z_r$$

The stochastic integral has a number of useful properties that we will define below. In the following properties, let B_t denote a standard Brownian motion and let A_t and C_t be simple processes.

Property 2.1.

- **Linearity:** For constants a, b, we have that
$$\int_0^t (aA_s + bC_s)dB_s = a\int_0^t A_s dB_s + b\int_0^t C_s dB_s$$
and moreover for $0 < q < t$ we have that
$$\int_0^t A_s dB_s = \int_0^q A_s dB_s + \int_q^t A_s dB_s$$
- **Martingale Property:** Z_t as defined above is a Martingale with respect to filtration $\mathcal{F}_t$.
- **Variance Property:** Z_t is such that Z_t^2 is integrable (square integrable) and
$$\text{Var}[Z_t] = \mathbb{E}[Z_t^2] = \int_0^t \mathbb{E}[A_s^2]ds$$
- **Continuity:** The function mapping times to Z_t is almost surely continuous for all t.

Interested readers can find the proof of the above properties in Lawler (2014) p87.
We now have sufficient background to state and prove the following two theorems called **Ito's Lemma**.

Theorem 2.2. ***(Ito's Lemma 1)*** *Let f be a C^2 function and let B_t be a standard Brownian motion. Then, for all t we have that*
$$f(B_t) = f(B_0) + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds$$
or in differential form
$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt$$

We will prove this theorem using the proof from Lawler (2014) p104.

Proof.
$$f(x+y) = f(x) + f'(x)y + \frac{1}{2}f''(x)y^2 + o(y^2)$$
where $o(y^2)$ represents lower order terms such that $\lim_{y\to 0}\frac{o(y^2)}{y} = 0$. Consider
$$f(B_1) - f(B_0) = \sum_{j=1}^{n}[f(B_{j/n} - f(B_{(j-1)/n}]$$
Now, we will use the Taylor approximation of each interval to get
$$f(B_{j/n}) = f(B_{(j-1)/n}) + f'(B_{(j-1)/n})(B_{j/n} - B_{(j-1)/n}) + \frac{1}{2}f''(B_{(j-1)/n})(B_{j/n} - B_{(j-1)/n})^2 + o((B_{j/n} - B_{(j-1)/n})^2)$$

where $o((B_{j/n}-B_{(j-1)/n})^2)$ vanishes as above. We can now see that $f(B_1)-f(B_0)$ is the sum of three limits.

$$f(B_1) - f(B_0) = \tag{2.3}$$

$$\lim_{n\to\infty}\sum_{j=1}^{n} f'(B_{(j-1)/n})[B_{j/n} - B_{(j-1)/n}] \tag{2.4}$$

$$+\lim_{n\to\infty}\frac{1}{2}\sum_{j=1}^{n} f''(B_{(j-1)/n})[B_{j/n} - B_{(j-1)/n}]^2 \tag{2.5}$$

$$+\lim_{n\to\infty}\sum_{j=1}^{n} o([B_{j/n} - B_{(j-1)/n}]^2) \tag{2.6}$$

First, we see that $[B_{j/n}-B_{(j-1)/n}]^2 \approx \frac{1}{n}$ by properties discussed above. Considering (2.6), we see that this limit is the sum of n intervals of order less than $\frac{1}{n}$. Thus, (2.6) equals zero. Now, consider (2.4). This is the simple process approximation of a stochastic integral, as discussed above. Thus, this limit equals

$$\int_0^1 f'(B_t)dB_t$$

Limit (2.5) requires a bit more work. Let $h(t) = f''(B_t)$. We know that $h(t)$ is continuous. Thus, for all $\epsilon > 0$ there exists a step function $h_\epsilon(t)$ such that $|h(t) - h_\epsilon(t)| < \epsilon$ for every t. Fix $\epsilon > 0$. Now, we can consider each interval on which the step function is constant. We see that

$$\lim_{n\to\infty}\sum_{j=1}^{n} h_\epsilon(t)[B_{j/n} - B_{(j-1)/n}]^2 = \int_0^1 h_\epsilon(t)dt$$

We also see that

$$\left|\sum_{j=1}^{n}[h(t) - h_\epsilon(t)][B_{j/n} - B_{(j-1)/n}]^2\right| \le \epsilon\sum_{j=1}^{n}[B_{j/n} - B_{(j-1)/n}]^2 \to \epsilon$$

Thus, we see that (2.6) is equivalent to

$$\lim_{\epsilon\to 0}\frac{1}{2}\int_0^1 h_\epsilon(t)dt = \frac{1}{2}\int_0^1 h(t)dt = \frac{1}{2}\int_0^1 f''(B_t)dt$$

Finally, we conclude that

$$f(B_1) - f(B_0) = \int_0^1 f'(B_t)dB_t + \frac{1}{2}\int_0^1 f''(B_t)dt$$

or more generally and in the form of the theorem,

$$f(B_t) = f(B_0) + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds$$

□

Let us consider a toy example to show the use of Ito's Lemma.

Example 2.7. Let $f(x) = x^2$. The first two derivatives are $f'(x) = 2x$ and $f''(x) = 2$. Applying Ito's Lemma, we see that

$$\begin{aligned} B_t^2 &= B_0^2 + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds \\ &= 2\int_0^t B_s dB_s + t \end{aligned}$$

and thus

$$\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$$

We will now state, but not prove, the higher dimensional version of Ito's Lemma.

Theorem 2.8. ***(Ito's Lemma 2)*** *Let $f(t, x)$ be a C^1 function in t and a C^2 function in x and let B_t be a standard Brownian motion. Then,*

$$f(t, B_t) = f(0, B_0) + \int_0^t \partial_x f(s, B_s)dB_s + \int_0^t \partial_s f(s, B_s) + \frac{1}{2}\partial_{xx} f(s, B_s)ds$$

or in differential form

$$df(t, B_t) = \partial_x f(t, B_t)dB_t + [\partial_t f(t, B_t) + \frac{1}{2}\partial_{xx} f(t, B_t)dt]$$

This is a higher dimensional version of Ito's Lemma. An important application of stochastic calculus is in the stochastic differential equation representing a **geometric Brownian motion** which we define below.

Definition 2.9. A stochastic process X_t is a geometric Brownian motion with drift m and volatility σ if it satisfies the SDE

$$\begin{aligned} dX_t &= mX_t dt + \sigma X_t dB_t \\ &= X_t[mdt + \sigma dB_t] \end{aligned}$$

where B_t is a standard Brownian motion.

The applications of geometric Brownian motions will be discussed below, particularly their theoretical ability to model stock prices.

3. Black-Scholes-Merton Differential Equation

Perhaps the most well known application of stochastic calculus in finance is in the **Black-Scholes-Merton European Option Pricing Model**, developed in the 1970's by Fischer Black, Myron Scholes, and Bill Merton. Their paper earned a Nobel prize in economics in 1997. We will derive the model here. First, we will establish some basic financial knowledge.

A **financial derivative** is a financial instrument which derives value from another financial instrument. A class of financial derivatives called **equity options** are contracts deriving their value from the price of an underlying stock. Options are separated into two contract types. The first type is a **call option**, which gives the holder the right but not the obligation to buy a predetermined number of shares at a predetermined price. Conversely, a **put option** gives the holder the right but not the obligation to sell a predetermined number of shares at a predetermined price. Additionally, we will define **arbitrage** as the ability to simultaneously buy and sell an asset across different markets to generate a riskless profit.

Let us derive the Black-Scholes model now following the proof Hull (2018) [2].

Theorem 3.1. *Let S be a stock, r be the risk-free interest rate, f be the price of a European option, and let t be the time until expiration of the option. We make the following assumptions.*

1. *S follows a geometric Brownian motion.*
2. *Short selling at any quantity is possible.*
3. *There are no transaction costs and fractional share buying is possible.*
4. *S pays no dividends.*
5. *The market is arbitrage free.*
6. *Trading is time continuous.*
7. *r is a constant over the life of the derivative.*

Then, the ***Black-Scholes-Merton differential equation***

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf$$

holds.

Proof. First, we assume that stock price S follows the following geometric Brownian motion

$$dS = \mu S dt + \sigma S dB_t$$

where B_t is a standard Brownian motion, μ is the constant drift rate, and σ is the constant volatility. Consider f, a derivative whose value depends on S like a European option. Applying Ito's Lemma, we see that

$$df = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt + \frac{\partial f}{\partial S}\sigma S dB_t$$

Now, we will seek to eliminate the Brownian motion term. To do so, we construct a portfolio of stock shares and derivative contracts. We see that this portfolio is

$$\begin{aligned} &-1 \text{ contracts} \\ &+\frac{\partial f}{\partial S} \text{ shares} \end{aligned}$$

In other words, the portfolio is short one derivative and long $\frac{\partial f}{\partial S}$ shares of the underlying stock. Let Π be the portfolio value. Thus,

$$\Pi = -f + \frac{\partial f}{\partial S}S$$

and

$$d\Pi = -df + \frac{\partial f}{\partial S}dS$$

We substitute the geometric Brownian motion for dS and the differential equation for df to get

$$d\Pi = \left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt$$

Note that we have no dependence on dB_t and thus we say that the portfolio is **riskless** for an infinitesimal time period dt. The assumption of no arbitrage implies that the portfolio must yield the same risk-free rate of return r as other short-term risk-free assets. Thus, we have that

$$d\Pi = r\Pi dt$$

Substituting again, we see that

$$\left(\frac{\partial f}{\partial t}+\frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt = r\left(f-\frac{\partial f}{\partial S}\right)dt$$
$$\implies \frac{\partial f}{\partial t}+rS\frac{\partial f}{\partial S}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} = rf$$

Thus we have arrived at the Black-Scholes-Merton differential equation.

□

4. Solutions to the Black-Scholes-Merton Differential Equation

In this section we present the canonical solutions to the differential equation. First, we must discuss the **heat equation** and its relevance to the problem.
The heat equation represents the dissipation of heat within a metal rod. This model is known as a forward parabolic equation. The differential equation takes the form

$$\frac{\partial u}{\partial \tau}=\frac{\partial^2 u}{\partial x^2}$$

for finite x and $\tau > 0$ and where $u(x, \tau)$ measures the temperature of at location x and time τ. We can solve the heat equation once we impose boundary conditions on x and τ. Since the equation is first order in τ and second order in x we impose one and two boundary conditions respectively.

Let us apply this framework to the Black-Scholes-Merton model. We will focus on European call options, since the European put option derivation is parallel. Let us consider the boundary conditions. We must impose two conditions on S and one condition on t based on the differential equation. First, we impose our condition on t. We know that at expiry or when $t = T$, the value of a call is equal to its **intrinsic value**, that is

$$C(S, T) = \max(S - K, 0)$$

where K is the strike price and the call price $C(S, T)$ is a function of underlying price and time. Now, we impose our conditions on S. First, we consider when $S = 0$. By the definition of the stock as a geometric Brownian motion, this implies that $dS = 0$ and thus the stock price is a constant zero and call price is identically zero. Thus,

$$C(0, t) = 0$$

at all t. We now consider taking S to ∞. As stock price increases to infinity, the probability that the option will be exercised approaches certainty as well. Thus, the call value is equivalent to the difference between the equity value and the discounted value of the price paid for the shares. Mathematically,

$$C(S, t) = S - Ke^{-r(T-t)}$$

as S goes to ∞.

Now, we have our options pricing model in the same form as the heat equation. Solutions to the heat equation are well documented and an interested reader can find a complete mathematical derivation of the solutions in Hull (2018). We see through this proof that a solution to the Black-Scholes-Merton differential equation is as follows in the subsequent theorem.

Theorem 4.1. *For a European call option, the solution to the Black-Scholes-Merton differential equation is*

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

where $N(x)$ is the standard normal CDF and

$$d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = \frac{\ln(\frac{S}{K}) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

*We call this equation the **Black-Scholes-Merton Options Pricing Formula**.*

Remark 4.2. An analogous statement holds for European put options.

Remark 4.3. An important clarification here is that σ here represents the **implied volatility**. Rather than the observed historical standard deviation, or historical volatility, the implied volatility is a value indicating the market's predicted move of the underlying equity. Often, the Black-Scholes-Merton options pricing formula is inverted to solve for this value, since a liquid market will always provide an observable notion of price.

Definition 4.4. We define the **delta, gamma, theta, vega** and **rho** as partial derivatives of Π as follows. Together, they are known as **the Greeks**.

$$\Delta = \frac{\partial \Pi}{\partial S} \qquad \Gamma = \frac{\partial^2 \Pi}{\partial S^2}$$
$$\Theta = -\frac{\partial \Pi}{\partial t} \qquad v = \frac{\partial \Pi}{\partial \sigma}$$
$$\rho = \frac{\partial \Pi}{\partial r}$$

The Greeks allow traders to understand contract sensitivity to various risk factors.

5. Challenging the Assumptions of the Black-Scholes-Merton Model

The Black-Scholes-Merton model relies on a very stringent set of assumptions. In this section, we will discuss why these assumptions do not reflect reality. We refer to the assumptions by the numbers in Theorem 3.1.

5.1. **Assumption 1: Stock Price as a GBM.** Geometric Brownian Motion is a popular model for stock behavior due to its ease of calculation, assumption of only positive values, and its independence between stock price and expected returns. We can show empirically[1] that stock price does not follow a geometric Brownian motion in reality. First, we consider the price series for AAPL [3]. We choose AAPL because it is extremely liquid and we can get close price data at the 1 second resolution.

We train $\hat{\mu}$ and $\hat{\sigma}$ on one trading day of close data. We use the normal distribution MLE to train these values, namely sample mean and sample standard deviation. Recall that a geometric Brownian motion follows the process $S = \mu S\, dt + \sigma S\, dB_t$. To simulate the Brownian motion, we sample one day of data from the distribution

[1] Full code for this section is available in the Appendix.

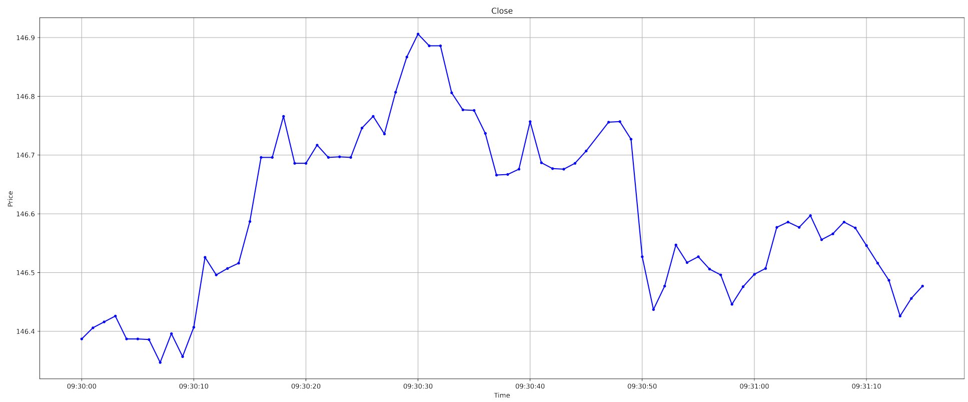

FIGURE 1. AAPL at 1 Second Resolution on 2/8/2021

$N(\hat{\mu}, \hat{\sigma})$. These values represent the secondly returns of the equity. Then, we simulate an equity that began at the start of the trading day at the same price and its movement. Now, we repeat this process 2,000 times to visualize likely paths of

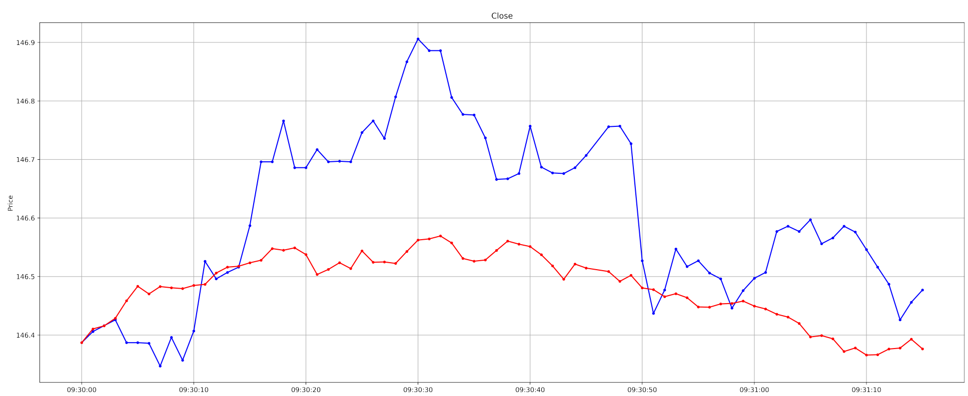

FIGURE 2. AAPL and a Geometric Brownian Motion

geometric Brownian motions.

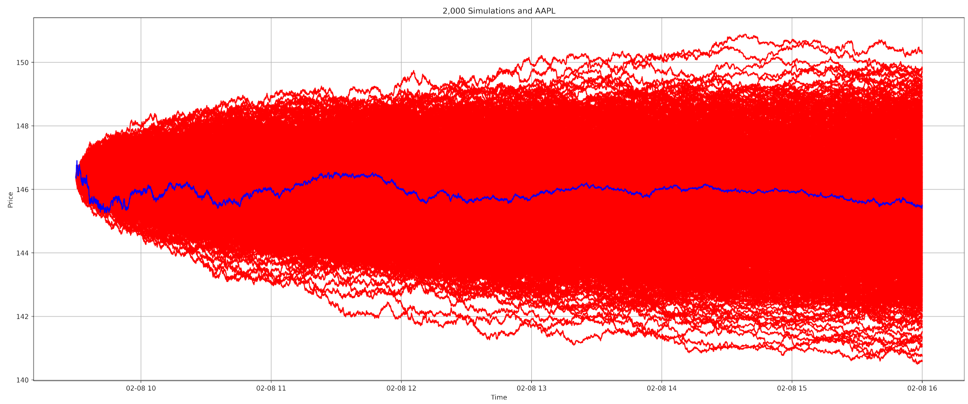

FIGURE 3. Actual Stock with Simulations

On inspection, AAPL seems to fit within the geometric Brownian motion framework. However, we can zoom into the first several minutes of the time frame and disprove this visually. As we can see, it is unlikely for AAPL to be a geometric

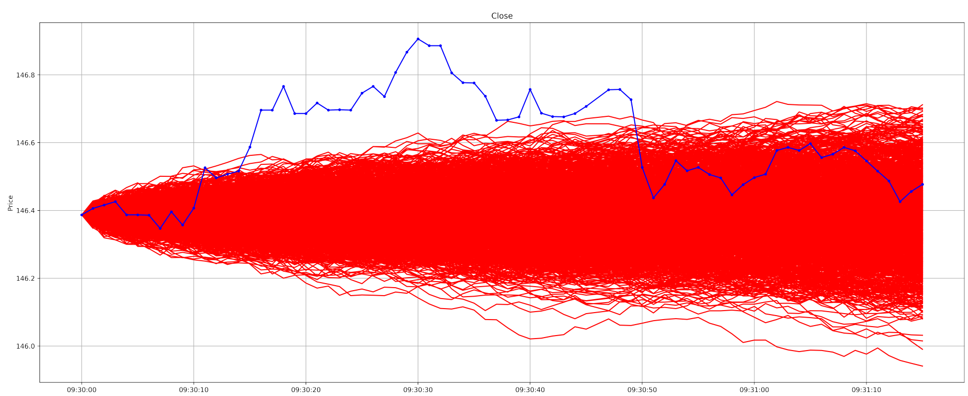

FIGURE 4. First Minutes of Simulation

Brownian motion. We can prove this statistically by examining the distribution of log returns of AAPL at the one second resolution in comparison to a $N(\hat{\mu}, \hat{\sigma})$ distribution. To show that the AAPL log returns are not drawn from a normal distribution, we use the Kolmogorov-Smirnov goodness-of-fit test. Below, we plot both the normal distribution and the empirical log return distribution for AAPL.

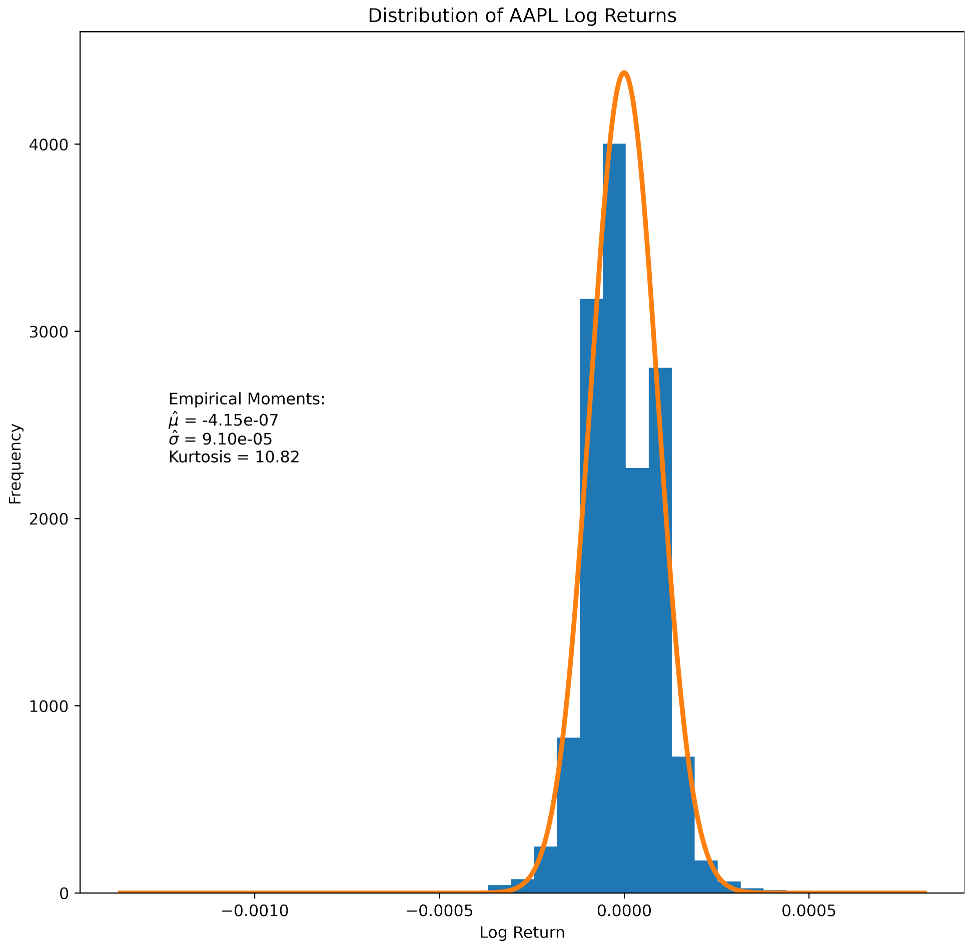

FIGURE 5. AAPL Distribution of Secondly Log Returns

We get a Komogorov-Smirnov test statistic of .168 and a corresponding p-value of almost identically 0. This proves that the distribution is almost certainly not normal. Thus, we have shown empirically that equity price is almost certainly not pursuant to a geometric Brownian motion. This points to a large flaw in the Black-Scholes-Merton options pricing formula.

5.2. **Assumption 2: Short Selling.** The assumption of easy short selling is largely true, but not of all equities. Less liquid assets have limited availability of shares for shorting. These assets are known as **hard to borrow** due to the mechanics of short selling. Further, in times of financial crisis government regulation can prevent short selling.

5.3. **Assumption 3: Transaction Costs.** This assumption is not accurate. For retail traders with small amounts of capital, there is in fact commission free trading, however costs are implicit in the price at which the equity is provided to the trader. For institutional traders, there are commissions for their trades. Additionally, fractional share buying is typcially not available.

5.4. **Assumption 4: Dividends.** This assumption is not accurate, however it is fixable. Dividends are payed by many equities, however adjusting the model by subtracting the dividend from the share price at the ex-dividend date accounts for this.

5.5. **Assumption 5: Arbitrage Free Market.** Arbitrage opportunities are present in equity and options markets. However, in the era of high frequency trading, they are often in existence for milliseconds or less as algorithmic traders edge out these market inefficiencies. Thus, this assumption holds largely true if we consider a long enough time increment.

5.6. **Assumption 6: Continuous Trading.** This assumption is false. Even with some popular assets entering 24/5 trading, weekends and off hours make most equity trading continuous but with frequent large time jumps over night.

5.7. **Assumption 7: Constant r.** First, the definition of a risk-free rate is ambiguous. Assuming we use the Fama-French definition of the risk free rate, we see daily movements of this value. Thus, risk free rate is not constant over the life of an option. The same conclusion can be drawn independent of your choice of definition for the risk free rate.

5.8. **Conclusion.** Several of the assumptions of this model do not prove realistic. Namely, the assumption that equity price follows a geometric Brownian motion. The question of what distribution is most accurate for equity returns is a difficult one, and is far beyond the scope of this paper.

6. Appendix

This section gives the code for challenging Assumption 1. First, note that we use closing price at the one second resolution in this analysis. This value is the price at which the last completed trade occurred during the 1 second interval. We import the following libraries.

```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import scipy as sp
```

First, we estimate $\hat{\mu}$ and $\hat{\sigma}$.

```
mu_hat = np.mean(np.log(data['Close']) -
    np.log(data['Close'].shift(1)))
sigma_hat = np.std(np.log(data['Close']) -
    np.log(data['Close'].shift(1)))
```

Then, we generate a series of simulated secondly returns. We then cumulatively multiply this starting from the initial price. This creates one simulated path.

```
np.random.seed(0)
sim_rets = np.array([0])
sim_rets = np.append(sim_rets,
np.random.normal(loc=mu_hat, scale=sigma_hat,
```

```
    size=len(data.index)-1))
sim_prices = data.iloc[0]['Close']*(1+sim_rets).cumprod()
```

Now, we plot this series with the real equity.

```
fig, ax = plt.subplots(figsize=(25, 10))
ax.grid()

ax.plot(data.index[:75], data['Close'][:75], color='blue')
ax.plot(data.index[:75], data['Close'][:75], marker='o',
    color='blue', markersize=3)
ax.plot(data.index[:75], sim_prices[:75], color='red')
ax.plot(data.index[:75], sim_prices[:75], marker='o',
    color='red', markersize=3)

plt.title('Close')
plt.xlabel('Time')
plt.ylabel('Price')
```

We repeat this process for the multi-simulation plot

```
for i in range(0,2000):
    sim_rets = np.array([0])
    sim_rets = np.append(sim_rets, np.random.normal(loc=mu_hat,
        scale=sigma_hat, size=len(data.index)-1))
    sim_prices = data.iloc[0]['Close']*(1+sim_rets).cumprod()
    sim_list.append(sim_prices)

fig, ax = plt.subplots(figsize=(25, 10))
ax.grid()

for sim in sim_list:
    ax.plot(data.index, sim, color='red')
ax.plot(data.index, data['Close'], color='blue')

plt.title('2,000 Simulations and AAPL')
plt.xlabel('Time')
plt.ylabel('Price')
```

and for the zoomed in version of the previous plot.

```
sim_list = []

fig, ax = plt.subplots(figsize=(25, 10))
ax.grid()

for sim in sim_list:
    ax.plot(data.index[:75], sim[:75], color='red')
ax.plot(data.index[:75], data['Close'][:75], color='blue')
ax.plot(data.index[:75], data['Close'][:75], marker='o',
    color='blue', markersize=3)
```

```
plt.title('Close')
plt.xlabel('Time')
plt.ylabel('Price')
```

We now plot our log return distribution and overlay the $N(\hat{\mu}, \hat{\sigma})$ density plot.

```
data['log_ret'] = (np.log(data['Close']) -
    np.log(data['Close'].shift(1)))

min_x = min(data['log_ret'][1:])
max_x = max(data['log_ret'][1:])

fig, ax = plt.subplots(figsize=(10, 10))

counts, bins, bars = plt.hist(data['log_ret'], bins=35)
plt.title('Distribution of AAPL Log Returns')
plt.xlabel('Log Return')
plt.ylabel('Frequency')

plt.plot(np.linspace(min_x, max_x, len(data.index)),
sp.stats.norm.pdf(np.linspace(min_x, max_x,
len(data.index)), mu_hat, sigma_hat), linewidth=3)
plt.annotate('Empirical Moments: \n'+'$\hat{\mu}$ = ' +
    str('{:0.2e}'.format(mu_hat)) + '\n$\hat{\sigma}$ = ' +
    str('{:0.2e}'.format(sigma_hat)) +
    '\nKurtosis = ' +
    str(round(data['log_ret'].kurtosis(),2)),
    xy=(.1, .5), xycoords='axes fraction')
```

Finally, we perform our Kolmogorov-Smirnov test.

```
dist = getattr(sp.stats, 'norm')
params = dist.fit(data['log_ret'][1:])

test_stat = sp.stats.kstest(data['log_ret'][1:], 'norm', params)
```

Acknowledgments

I would like to thank Stephen Yearwood for being my mentor throughout the REU and for his help on this paper. I would also like to thank Dr. Greg Lawler for leading the probability group and Dr. Peter May for organizing the REU.

References

[1] Lawler, Gregory. Stochastic Calculus: An Introduction with Applications. 2014.
[2] Hull, John. Options, Futures, and Other Derivatives. 10th ed., Pearson, 2018.
[3] AAPL Historical Data Feed. Dukascopy Swiss Banking Group., https://www.dukascopy.com/swiss/english/marketwatch/historical/