Pricing European Options in Continuous-Time: Theory and Applications

Tallis Bowers

University of Notre Dame

Undergraduate Mathematics Student

Acknowledgements

I first must thank Professors Alex Himonas and Tom Cosimano so much for their time and generosity. Without their weekly guidance, this paper would never have come together. I also need to thank Professor Jeffrey Diller for his support throughout my undergraduate tenure and for encouraging me to pursue this thesis. Many other people offered their time and invaluable critiques, and I would like to thank particularly Professor Andrei Jorza, Professor Peter Kelly, and Ed Stivers. Finally, I must thank my friend and colleague Sean Kent; his weekly assistance helped to coherently structure this thesis.

Abstract

This paper develops several common mathematical models used in finance to describe stocks and European options. These models include the binomial model of stock growth and options pricing, the log-normal model of stock growth, and the Black-Scholes-Merton model of options pricing. Binomial models are used to understand the financial instruments of interest to this paper in discrete time. Most of this paper focuses on understanding these instruments in continuous-time, and the ultimate result is the use of stochastic calculus to derive the famed Black-Scholes-Merton equation. Finally, applications of this equation in the financial industry are discussed briefly.

Contents

1. Introduction 1
2. Discrete Time 2
2.1. One-Period Binomial Stock Movement 2
2.2. European Call Options 4
2.3. The Multi-period Binomial Model 6
3. Continuous Randomness 9
3.1. Random Walks 9
3.2. Statistical Detour 11
3.3. Brownian Motion 13
4. Continuous-Time 16
4.1. Log-Normal Stock Growth 16
4.2. BSM Formula: First Derivation 19
5. Stochastic Calculus 21
5.1. Simple Stochastic Processes 21
5.2. General Itô Integral 23
5.3. Itô's Formula 25
6. BSM Formula Revisited 27
6.1. BSM Partial Differential Equation 27
6.2. Solving the PDE 28
7. Applications 32
7.1. Implied Volatility Surfaces 32
7.2. Heston's Model 34
8. References 36

1. Introduction

The financial world was changed significantly in 1973 when a pair of economists named Fischer Black and Myron Scholes published a paper entitled “The Pricing of Options and Corporate Liabilities.”(1) Their work provided a formula that could be used to determine the theoretical price of a European option given that arbitrage should never exist as a result of a portfolio combining stocks and options. Given the age of this work today, it is not surprising that other, more sophisticated, models have been developed and are currently more common in industry. We will briefly introduce one such model at the end of this paper, but we focus on the model created by these two men because of its historical importance and because of the interesting mathematics necessary for its derivation.

Often the formula that will be the climax result of this paper is referred to as the Black-Scholes formula, but we will refer to it as the Black-Scholes-Merton formula in recognition of the work of Robert C. Merton. In 1990, Merton, another renowned economist, published a book entitled *Continuous-Time Finance*(7) in which he provided even more mathematical rigor for the financial model developed by Black and Scholes, and he greatly advanced numerous concepts in continuous-time finance. It is his work that inspires our use of stochastic calculus in the later sections of this paper.

To help make clear the ideas of Black, Scholes, and Merton, as well as to introduce rigorous mathematics in a financial context, Alex Himonas, Professor of Mathematics at the University of Notre Dame, and Tom Cosimano, Professor Emeritus of Economics at the University of Notre Dame, have worked for many years on a series of notes, soon to be published as a book, entitled *Mathematical Methods in Financial Economics*(3). This paper most closely follows this wonderful work.

The ultimate goal of this paper is to allow the reader an understanding of the Black-Scholes-Merton formula and the math behind its derivation in a fairly brief account. In order to do this, we must assume some prior knowledge of basic mathematics. Many of the most basic concepts of calculus, measure theory, and statistics are assume. Whenever possible, definitions of mathematical concepts and statements of theorems are given, but whenever an unexplained concept is foreign to the reader, I recommend the following sources (numbers found in References section) for the respective fields of mathematics: (8) - calculus, (4) - measure theory, and (13) - statistics. All economic concepts should be explained well enough that no prior knowledge of economics is necessary on the part of the reader.

2. Discrete Time

This section considers the implications of a very simple model economy, in which time elapses in discrete increments over which a stock can only take only one of two values. Such a model is referred to as binomial, and we will move from the one-period binomial model to the multi-period binomial model.

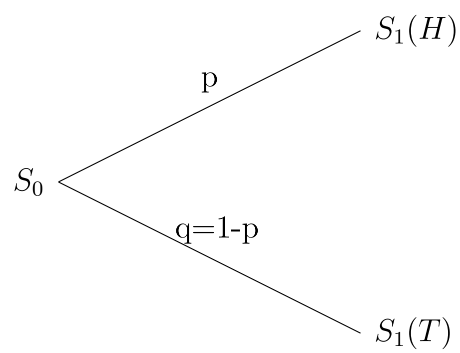
2.1. **One-Period Binomial Stock Movement.** To begin the one-period bionomial model, we consider the price of a stock at time zero S_0 and the price at time one S_1. The future price S_1 shall be dependent upon a random event with two possible outcomes. We let this event be a coin toss and call the outcomes H and T for heads and tails. We will make the assumption that the probability p of heads lies in the open interval $(0, 1)$ so that we can say the same for the probability of tails q, i.e. $0 < q = 1 - p < 1$.

Definition 2.1. In the context of the one-period binomial stock model, we define the **up** and **down factors** by

$$u = \frac{S_1(H)}{S_0} \text{ and } d = \frac{S_1(T)}{S_0},$$

respectively.

Figure 1. A tree representing stock growth in the one-period binomial model.



In order to say anything interesting, we must have a second investment alternative to stocks in our model economy. For this, we will imagine investing in a bank that guarantees a fixed return. This **risk-free interest rate** is assumed to be the same for money invested in the bank as for money borrowed from the bank. Letting r be the risk-free interest rate, we also assume $d < u$, and further that

$$0 < d < 1 + r < u \ . \tag{2.1}$$

In order to explain why it is reasonable to accept this assumption, the concept of arbitrage must be introduced.

Definition 2.2. An investment opportunity is considered to be an **arbitrage** if it has the following properties:

(1) The investment requires no initial wealth.
(2) The probability of losing money as a result of the investment is zero.
(3) The probability of making money from the investment is greater than zero.

Essentially an arbitrage presents an investor with a chance to make money without taking any risk. It is reasonable to assume such scenarios do not occur frequently in an economy because any such opportunity would be desired by every investor in an economy, which would quickly eliminate any arbitrage. With this assumption, we can justify our earlier inequality regarding the up and down factors with the following theorem.

Theorem 2.3. *The equation 2.1 holds if and only if there is no arbitrage in our one-period binomial model.*

Proof. First, if $d \geq 1+r$, then we can take out a loan at interest rate r to buy stock at price S_0. From this stock, we would be certain of making at least $dS_0 \geq (1+r)S_0$ at time one, and we would have a chance of making $uS_0 > dS_0 \geq (1+r)S_0$. This is an arbitrage because we can be certain of our ability to pay back the loan, meaning we will not be losing money, and we have a positive probability of making money since $uS_0 - (1+r)S_0 > 0$.

Similarly, if $u \leq 1+r$, we can employ an investing strategy to bring about arbitrage. By selling a stock short at the price S_0 and investing the money at the risk-free rate r, we would at most have to pay $uS_0 \leq (1+r)S_0$ at time one, which would insure that we do not lose money. We would also have a positive probability of making $(1+r)S_0 - dS_0 > 0$ since $d < u$. These two investing strategies prove that equation 2.1 follows from the assumption of no arbitrage.

Now we assume $d < 1+r < u$. We want to show that this precludes arbitrage. At time zero we will buy Δ_0 shares of stock with $\Delta_0 S_0$ borrowed (or invested if Δ_0 is negtive, and hence we have shorted the stock) money so that your initial investment is $X_0 = 0$. Then our return is given by

$$X_1 = \Delta_0 S_1 - (1+r)\Delta_0 S_0. \tag{2.2}$$

Since $d < 1+r$, when $\Delta_0 > 0$ we have with probability $q > 0$

$$X_1(T) = \Delta_0 S_1(T) - (1+r)\Delta_0 S_0 = (d-(1+r))\Delta_0 S_0 < 0.$$

Similarly since $u > 1+r$, when $\Delta_0 < 0$ we have with probability $p > 0$

$$X_1(H) = \Delta_0 S_1(H) - (1+r)\Delta_0 S_0 = (u-(1+r))\Delta_0 S_0 < 0.$$

Here we have shown that any investing stragtegy consisting of stocks and a risk-free return has a positive chance of losing money in both the case of buying the stock ($\Delta_0 > 0$) and selling the stock short ($\Delta_0 < 0$). This is all that is necessary to say that arbitrge does not exist in our model economy. $\square$

The material of this section was inspired by an is elaborated upon by Shreve (2004)[9].

2.2. **European Call Options.** To this point, we have discussed the mechanics behind stock price movement in the one-period binomial model. We now introduce the financial instrument which we hope to study throughout this paper.

Definition 2.4. A **European call option (ECO)** is a contract between two parties. The party selling the contract gives the party buying the contract the right to buy a stock at a stated **strike price** K on a specified future date, known as the **maturity date** T.

Because the ECO only gives the buyer the *right* to purchase a stock in the future, the contract will go unexercised on the maturity date if the price of the stock is less than or equal to the strike price of the ECO. In the one-period binomial model, the value of an ECO at time one is given by $(S_1 - K)^+$, where the notation $(...)^+$ means $max(S_1 - K, 0)$. The ultimate goal in this section is to derive a formula for the price of an ECO at time zero in a way that avoids arbitrage. For this, we will replicate the option using a porfolio of initial wealth X_0 invested partially in Δ_0 shares of stock and partially at a risk-free interest rate r. For the purpose of notation, let V_0 and V_1 denote the value of the ECO at times zero and one, respectively. The next theorem explicitly proves the formula we seek.

Theorem 2.5. *The value of a European call option at time zero in the one-period binomial pricing model is given by the formula*

$$V_0 = \frac{1}{1+r}\left[\tilde{p}V_1(H) + \tilde{q}V_1(T)\right]$$

where $\tilde{p} = \frac{1+r-d}{u-d}$ *and* $\tilde{q} = 1 - \tilde{p} = \frac{u-1-r}{u-d}$.

Proof. Using our replicating portfolio, we begin by observing that the value of the total portfolio at time one is given by

$$X_1 = \Delta_0 S_1 + (1+r)(X_0 - \Delta_0 S_0).$$

Since we wish to replicate the call option, we will let $V_1(H) = X_1(H)$ and $V_1(T) = X_1(T)$ and solve for V_0. It is also important to note here that because the option has payoffs equal to those of the portfolio, it will also have the same initial value, i.e. $X_0 = V_0$. Now the equations we get for V_1 are

$$V_1(H) = \Delta_0 S_1(H) + (1+r)(X_0 - \Delta_0 S_0) \tag{2.3}$$

and

$$V_1(T) = \Delta_0 S_1(T) + (1+r)(X_0 - \Delta_0 S_0). \tag{2.4}$$

Subtracting equation 2.3 from equation 2.4 gives the following equation, which can be solved for Δ_0

$$V_1(H) - V_1(T) = \Delta_0(S_1(H) - S_1(T)) \implies \boxed{\Delta_0 = \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)}.}$$

The boxed equation above is often referred to as the *delta hedging formula*, and it represents the amount of stock that needs to be purchased in order to replicate the option's payoffs.

Now, we divide equation 2.3 and equation 2.4 by $(1+r)$ to get

$$\frac{1}{1+r} V_1(H) = X_0 + \Delta_0 \left(\frac{1}{1+r} S_1(H) - S_0 \right) \tag{2.5}$$

and

$$\frac{1}{1+r} V_1(T) = X_0 + \Delta_0 \left(\frac{1}{1+r} S_1(T) - S_0 \right). \tag{2.6}$$

Defining $\tilde{p}$ and $\tilde{q}$ as in the theorem, we multiply equation 2.5 by $\tilde{p}$ and equation 2.6 by $\tilde{q}$ and add the products to get

$$\begin{aligned} \frac{1}{1+r} \left(\tilde{p} V_1(H) + \tilde{q} V_1(T) \right) = \\ (\tilde{p} + \tilde{q}) X_0 + \Delta_0 \left(\frac{1}{1+r} \left[\tilde{p} S_1(H) + \tilde{q} S_1(T) \right] - (\tilde{p} + \tilde{q}) S_0 \right). \end{aligned} \tag{2.7}$$

Next, we look at the middle term of equation 2.7 and observe

$$\begin{aligned} & \frac{1}{1+r} \left[\tilde{p} S_1(H) + \tilde{q} S_1(T) \right] \\ &= \frac{(1+r)S_1(H) - dS_1(H) + uS_1(T) - (1+r)S_1(T)}{(1+r)(u-d)} \\ &= \frac{S_1(H) - S_1(T)}{u-d} + \frac{uS_1(T) - dS_1(H)}{(1+r)(u-d)} = S_0. \end{aligned}$$

Finally, because of the way $\tilde{p}$ and $\tilde{q}$ were defined, we can reduce equation 2.7 to

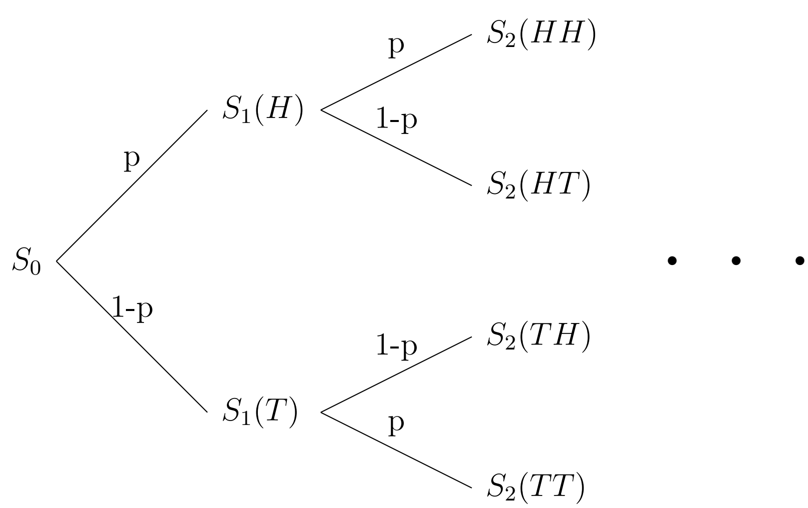
$$\boxed{\frac{1}{1+r} \left(\tilde{p} V_1(H) + \tilde{q} V_1(T) \right) = X_0 = V_0.}$$

We have succesfully replicated the option with a stock/risk-free portfolio because of the delta hedging formula. □

In deriving the above equation, we made many assumptions that created a very simple model economy. It is important to think about what these assumptions are and how they limit the applicability of this model to the real world. First, we assumed shares of stock could be arbitrarily subdivided (i.e. Δ_0 need not be a whole number). This assumption is not too concerning as most investors work with many shares of stock, and thus don't care about having to round Δ_0 to a whole number. Two more assumption to acknowledge involve the lack of an ask-bid spread and the lack of a difference between the borrowing and investing rates at a bank. In reality banks charge more interest to borrow money than they give for investing, and brokers sell stocks at slightly higher prices than they are willing to buy them at. These assumptions would certainly affect the payoffs considered above, but again, if one is investing large enough quantities, both spreads will be small. Finally, the most important assumption of the one-period binomial model, and the reason we will eventually develop a continuous-time analog, is that a stock can only take two possible values in any given time step. The methods we used in this section followed those of Shreve (2004)[9].

2.3. **The Multi-period Binomial Model.** The one-period binomial model can easily be extended to a model in which the underlying stock price changes over any number of periods. In this case, we wish to find the value of an option whose payoffs after N periods are

Figure 2. Here is a tree representing two iterations of multi-period, binomial stock growth.



determined by the price of a stock which exhibits multi-period binomial growth. To do this, we must find the value of the option at every time $t = 1, 2, ..., N-1$ by using the formula from theorem 2.5 recursively backwards. The next theorem makes this idea rigorous.

Theorem 2.6. *Let $\tilde{p}$ and $\tilde{q}$ be defined as in the previous theorem and V_N be the payoff of a European call option expiring at time N, which is a random variable dependent upon the coin tosses $\omega_1\omega_2...\omega_{N-1}$. We define recursively backward the value of the option at times*

$n = 0, 1, ..., N-1$ by

$$V_n(\omega_1\omega_2...\omega_n) = \frac{1}{1+r}\left[\tilde{p}V_{n+1}(\omega_1\omega_2...\omega_n H) + \tilde{q}V_{n+1}(\omega_1\omega_2...\omega_n T)\right]$$

according to the one-period model. The option with given payoffs at time N can be replicated by a portfolio containing

$$\Delta_n = \frac{V_{n+1}(\omega_1...\omega_n H) - V_{n+1}(\omega_1...\omega_n T)}{S_{n+1}(\omega_1...\omega_n H) - S_{n+1}(\omega_1...\omega_n T)}. \tag{2.8}$$

shares of stock and initial wealth $X_0 = V_0$.

Proof. This proof will depend upon induction on n. In a very similar fashion to the proof of the one-period case, we assume that we have

$$X_n(\omega_1...\omega_n) = V_n(\omega_1...\omega_n)$$

where X is the value of the replicating portfolio. We will show that

$$X_{n+1}(\omega_1...\omega_n\omega_{n+1}) = V_{n+1}(\omega_1...\omega_n\omega_{n+1}).$$

For ease of notation, let $\omega = \omega_1...\omega_n$. We can see that the value of the replicating portfolio in the case that $\omega_{n+1} = H$ is

$$\begin{aligned} X_{n+1}(\omega H) &= \Delta_n S_{n+1}(\omega H) + (1+r)\left(X_n(\omega) - \Delta_n S_n(\omega)\right) \\ &= (1+r)X_n(\omega) + \Delta_n\left[S_{n+1}(\omega H) - (1+r)S_n(\omega)\right]. \end{aligned}$$

Now, using equation 2.8 we can see that

$$\begin{aligned} X_{n+1}(\omega H) &= (1+r)X_n(\omega) + \frac{V_{n+1}(\omega H) - V_{n+1}(\omega T)}{S_{n+1}(\omega H) - S_{n+1}(\omega T)}\left[S_{n+1}(\omega H) - (1+r)S_n(\omega)\right] \\ &= (1+r)X_n(\omega) + \frac{V_{n+1}(\omega H) - V_{n+1}(\omega T)}{(u-d)S_n(\omega)}\left[S_{n+1}(\omega H) - (1+r)S_n(\omega)\right] \\ &= (1+r)X_n(\omega) + \frac{V_{n+1}(\omega H) - V_{n+1}(\omega T)}{(u-d)}\left[u - (1+r)\right] \\ &= (1+r)X_n(\omega) + \tilde{q}V_{n+1}(\omega H) - \tilde{q}V_{n+1}(\omega T). \end{aligned}$$

Because of the induction assumption, we can replace this value by

$$X_{n+1}(\omega H) = (1+r)V_n(\omega) + \tilde{q}V_{n+1}(\omega H) - \tilde{q}V_{n+1}(\omega T).$$

Now, by the one-period theorem, we can replace $(1+r)V_n$ as follows

$$X_{n+1}(\omega H) = \tilde{p}V_{n+1}(\omega H) + \tilde{q}V_{n+1}(\omega T) + \tilde{q}V_{n+1}(\omega H) - \tilde{q}V_{n+1}(\omega T) = V_{n+1}(\omega H).$$

This is what we want to show, and the case for $\omega_{n+1} = T$ is very similar, so we will leave it out of this discussion. This proof comes from Cosimano and Himonas (2017)[3]. $\square$

What we have shown is that we can replicate a call option with a portfolio containing a stock and risk-free investment at every time in the multi-period binomial process by readjusting the shares of stock held at each time. The most important thing to realize about this result is that the one-period binomial pricing model is the key to replicating an option in a multi-period binomial economy.

3. Continuous Randomness

So far we have considered a discrete model of stock growth and shown how to price European options within this model. Now we will look at a continuous-time model of stock growth, and we will develop the famed Black-Scholes-Merton equation using two methods. Before we can look at either of these methods; however, we must learn how to add randomness to a continuous model. The continuous stochastic process we will develop in this section is called Brownian motion, and much of the framework of this section is inspired by Cosimano and Himonas (2017)[3] and Shreve (2004)[10]. We begin our construction by considering random walks in one dimesion.

3.1. **Random Walks.** In order to visualize this process, it is best to consider a person walking along a number line. This person randomly chooses whether to move one unit right or one unit left at discrete intervals of time. To create randomness we will allow the toss of a fair coin (infinitely many times) to determine the direction of each step, and we denote the outcome of the j^{th} coin toss by ω_j. Then we let the set of all possible sequences of coin tosses be a sample space denoted by

$$\Omega = \{\omega_0\omega_1\omega_2... : \omega_i = \text{ H or T, for all } i = 0, 1, 2...\}.$$

Next, we define a set of random variables over this sample space given by

$$X_j = \begin{cases} 1 \text{ when } \omega_j = \text{H} \\ -1 \text{ when } \omega_j = \text{T}. \end{cases}$$

We will also denote the distance covered in each"step" by Δx and each discrete amount of time between coin tosses by Δt. Finally, we say that the walker's movement speed is constant so that the path is continuous everywhere. In this context, the position of the walker along the number line after k coin tosses is given by the random variables

$$M_{k\Delta t} = \sum_{i=1}^{k} X_i \Delta x. \tag{3.1}$$

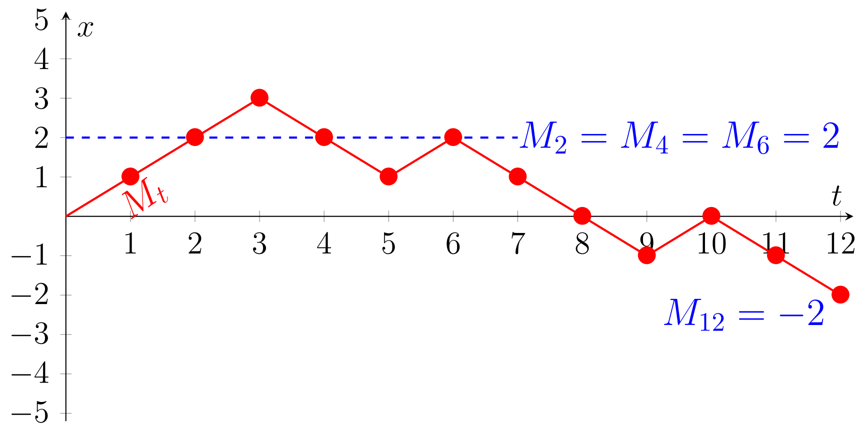
Definition 3.1. A **random walk** is a process $\{M_t\}_{t\geq 0}$ defined at all times t by

$$M_t = \left(k + 1 - \frac{t}{\Delta t}\right) M_{k\Delta t} + \left(\frac{t}{\Delta t} - k\right) M_{(k+1)\Delta t} \tag{3.2}$$

where k and $k+1$ are such that $k\Delta t \leq t \leq (k+1)\Delta t$. Also, we say $M_0 = 0$ and $M_{j\Delta t}$ are random variables as defined above for $j = 0, 1, 2, ...$

This method of continuously connecting each of our discrete points with lines is known as linear interpolation, and we will see that continuity of random walks is very important later in this paper.

FIGURE 3. In this graph of one possible random walk, we see the position of the walker on the number line x versus time t with parameters $\Delta x = \Delta t = 1$. Note the continuiuty of the graph, which is a result of linear interpolation.



Definition 3.2. Given a random walk $\{M_j\}_{j\geq 0}$ we define an **increment** of that walk to be

$$M_n - M_m = \sum_{i=m}^{n} X_i \Delta x$$

where $0 \leq m < n$ and n, m are both multiples of Δt.

Now that we have defined random walks and their increments, we will prove a few important properties about them. The next three theorems show that random walks are unbiased in their expected direction, their disjoint increments are independent, and they are continuous.

Theorem 3.3. *Let $\{M_j\}_{j\geq 0}$ be a random walk, and let $M_n - M_m$ be an increment of this walk. Then the expectation and variance of the increment are given by*

$$E(M_n - M_m) = 0 \tag{3.3}$$

and

$$V(M_n - M_m) = (\Delta x)^2 (n - m), \tag{3.4}$$

respectively.

Proof. First, we give our increment the notation

$$M_n - M_m = \sum_{i=m}^{n} X_i \Delta x.$$

Then, because the X_i are determined by different coin tosses, each is independent, i.e. $E(X_i) = 0$ for all $i = 0, 1, 2, ...$ Thus we have

$$E(M_n - M_m) = E\left(\sum_{i=m}^{n} X_i \Delta x\right) = \sum_{i=m}^{n} E(X_i)\Delta x = 0.$$

Now, we can calculate the variance of X_i by observing

$$V(X_i) = (1)^2 \frac{1}{2} + (-1)^2 \frac{1}{2} = 1.$$

Finally, using the independence of all X_i again, we obtain

$$V(M_n - M_m) = V\left(\Delta x \sum_{i=m}^{n} X_i\right) = (\Delta x)^2 \sum_{i=m}^{n} V(X_i) = (\Delta x)^2 (n - m).$$

□

Theorem 3.4. *Let $\{M_j\}_{j \geq 0}$ be a random walk, and let $M_{n_1} - M_{m_1}$ and $M_{n_2} - M_{m_2}$ be two increments of this walk such that $0 \leq m_1 < n_1 < m_2 < n_2$. Then these two increments are independent.*

Proof. Since $M_{n_1} - M_{m_1} = \sum_{i=m_1}^{n_1} X_i \Delta x$ and $M_{n_2} - M_{m_2} = \sum_{i=m_2}^{n_2} X_i \Delta x$ share no random variables X_i, the independence of the increments follows from that of the X_i. □

Theorem 3.5. *Let $\{M_j\}_{j \geq 0}$ be a random walk. Then for any $\omega \in \Omega$ we have $M_0(\omega) = 0$ and the map $t \longmapsto M_t(\omega)$ is continuous.*

Proof. This result follows from our construction of random walks. We said M_0 is identically 0, and we defined M_t where $k\Delta t < t < (k+1)\Delta t$ by connecting $M_{k\Delta t}$ and $M_{(k+1)\Delta t}$ using linear interpolation. □

3.2. **Statistical Detour.** In this section, we inroduce some basic concepts from statistics that will be necessary throughout this paper. First, we discuss standard normal random variables, then we look at the well-known Central Limit Theorem.

Definition 3.6. We say that Z is a **standard normal random variable** if it has the density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \ \ x \in \mathbb{R}$$

The next theorem proves that f is a valid density function.

Theorem 3.7. *If f is defined as above,*

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Proof. Let I denote the integral we wish to calculate and use polar coordinates to observe

$$I^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-x^2/2}dx \int_{-\infty}^{\infty} e^{-y^2/2}dy = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}dydx$$

$$\overset{r=\sqrt{x^2+y^2},\ \theta=\arctan\frac{y}{x}}{=} \frac{1}{2\pi}\int_0^{\infty}\int_0^{2\pi} re^{-r^2/2}d\theta dr = \int_0^{\infty} re^{-r^2/2}dr = \left[-e^{-r^2/2}\right]_0^{\infty} = 1.$$

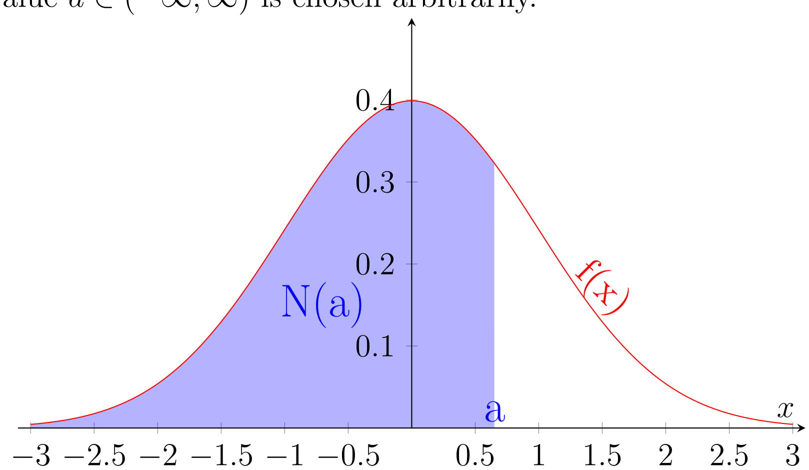
Thus, since $f > 0$ everywhere, we must have $I = 1$. □

Now that we have shown f is a good density function, we can define its associated cumulative distribution function, which we will use in our final options pricing formula.

Definition 3.8. The **cumulative distrbution function of a standard normal random variable** is denoted by $N(a)$. It can be written as

$$N(a) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-x^2/2}dx.$$

FIGURE 4. This is a graphical representation of the density function f and the cumulative distribution function N of a standard normal random variable Z. The value $a \in (-\infty, \infty)$ is chosen arbitrarily.



Another important idea from statistics is the Central Limit Theorem. It will be used several times in this paper and is stated below.

Theorem 3.9. *Let $Y_1, Y_2, ..., Y_n$ be indepent random variables, all with the same distribution, so that $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$ for all $i = 1, 2, ..., n$. Also let U_n be defined by*

$$U_n = \frac{\Sigma_{i=1}^{n} Y_i - n\mu}{\sigma\sqrt{n}}.$$

Then the distribution of U_n converges to the standard normal distribution, and we say $U_n \longrightarrow Z$ as $n \longrightarrow \infty$.

The proof of this theorem can be found in many statistics texts, including Wakerly, Mendenhall, and Scheaffer (2008)[(13)].

3.3. **Brownian Motion.** Now we can begin the construction of Brownian motion as the limit of random walks. For this, we want to find a stochastic process $\{B_t\}_{t\geq 0}$ which is a function $B(t,\omega) : [0,\infty) \times \Omega \longrightarrow \mathbb{R}$ over time and the probability space Ω. This function can also be denoted $B_t(\omega)$ and must have the following properties:

1. $B_0(\omega) = 0$ for all $\omega \in \Omega$.
2. The map $t \longmapsto B_t(\omega)$ is continuous for all $\omega \in \Omega$.
3. The increment $B_t - B_s$, where $0 \leq s < t$, is normally distributed with mean 0 and variance $t - s$. In other words, the increment has the distribution function
$$P\left(a \leq B_t - B_s \leq b\right) = \frac{1}{\sqrt{2\pi(t-s)}} \int_a^b e^{-\frac{1}{2(t-s)}x^2} dx.$$
4. Any two non-overlapping increments are independent.

It is important to note that all of these assumptions except (3) have been shown for random walks.

To begin our construction, consider the family of random walks $\{M_j^n\}_{j\geq 0, n\in\mathbb{N}}$ such that for any positive integer n, we have a random walk with the time and position intervals

$$\Delta t = \frac{1}{n} \text{ and } \Delta x = \frac{1}{\sqrt{n}},$$

respectively. We want to explicitly define M_t^n using equation 3.2. For this, let k be the largest integer such that $k\Delta t = \frac{k}{n} \leq t$. Then k will be the floor value $\lfloor nt \rfloor$, and we have

$$\begin{aligned}
M_t^n &= \left(\lfloor nt \rfloor + 1 - nt\right) M^n_{\lfloor nt \rfloor \cdot \frac{1}{n}} + \left(nt - \lfloor nt \rfloor\right) M^n_{(\lfloor nt \rfloor + 1)\cdot \frac{1}{n}} \\
&= M^n_{\lfloor nt \rfloor \cdot \frac{1}{n}} + \left(nt - \lfloor nt \rfloor\right)\left(M^n_{(\lfloor nt \rfloor + 1)\cdot \frac{1}{n}} - M^n_{\lfloor nt \rfloor \cdot \frac{1}{n}}\right) \\
&= \sum_{i=1}^{\lfloor nt \rfloor} X_i \frac{1}{\sqrt{n}} + \frac{\left(nt - \lfloor nt \rfloor\right)}{\sqrt{n}} X_{\lfloor nt \rfloor + 1} \\
&= \frac{\sqrt{\lfloor nt \rfloor}}{\sqrt{n}} \cdot \frac{1}{\sqrt{\lfloor nt \rfloor}} \sum_{i=1}^{\lfloor nt \rfloor} X_i + \frac{\left(nt - \lfloor nt \rfloor\right)}{\sqrt{n}} X_{\lfloor nt \rfloor + 1}.
\end{aligned}$$

Now we can observe the simple limit

$$\lim_{n\to\infty} \frac{\sqrt{\lfloor nt \rfloor}}{\sqrt{n}} = \sqrt{t},$$

and since $\mid X_i \mid$ is bounded by 1 for all i, we can also see that

$$\lim_{n\to\infty} \frac{\left(nt - \lfloor nt \rfloor\right)}{\sqrt{n}} X_{\lfloor nt \rfloor + 1} = 0.$$

Recall from the proof of theorem 3.3 that the expectation and variance of $\sum_{i=1}^{\lfloor nt \rfloor} X_i$ are given by 0 and $\sqrt{n}$, respectively. Thus by the Central Limit Theorem, we can say that

$$\frac{1}{\sqrt{\lfloor nt \rfloor}} \sum_{i=1}^{\lfloor nt \rfloor} X_i \longrightarrow Z \text{ as } n \to \infty$$

where Z is a standard normal random variable. Now that we have determined the limit of each term, we can define B_t by

$$M_t^n \longrightarrow \sqrt{t}Z =: B_t \text{ as } n \to \infty.$$

Theorem 3.10. *The stochastic process B_t has all four properties we require of Brownian motion.*

Proof. Let $\omega \in \Omega$ be arbitrary. To check property (1), we observe that

$$B_0(\omega) = \lim_{n\to\infty} M_0^n(\omega) = \lim_{n\to\infty} 0 = 0.$$

Now let $\epsilon > 0$ be arbitrary. Then, since M_t^n converges to B_t pointwise at every ω and B_t is measurable, we can find a $\delta > 0$ such that this convergence is uniform in the open ball of radius δ around t. So there is an $N \in \mathbb{N}$ such that $\mid B_s(\omega) - M_s^n(\omega) \mid < \epsilon/3$ for $s \in B(\delta, t)$ and $n > N$. Also, by the continuity of M_t^n, we can pick s, t sufficiently close so that $\mid M_t^n(\omega) - M_s^n(\omega) \mid < \epsilon/3$ and $s \in B(\delta, t)$. Thus we see

$$\begin{aligned} \mid B_t(\omega) - B_s(\omega) \mid &= \mid B_t(\omega) - M_t^n(\omega) + M_t^n(\omega) - M_s^n(\omega) + M_s^n(\omega) - B_s(\omega) \mid \\ &\leq \mid B_t(\omega) - M_t^n(\omega) \mid + \mid M_t^n(\omega) - M_s^n(\omega) \mid + \mid M_s^n(\omega) - B_s(\omega) \mid < \epsilon \end{aligned}$$

shows property (2).

For property (3), consider the increment $B_t - B_s$ where $0 \leq s < t$. We can observe,

$$\begin{aligned} B_t - B_s &= \lim_{n\to\infty} (M_t^n - M_s^n) \\ &= \lim_{n\to\infty} \left(M^n_{\lfloor nt \rfloor \cdot \frac{1}{n}} - M^n_{\lfloor ns \rfloor \cdot \frac{1}{n}} + \frac{nt - \lfloor nt \rfloor}{\sqrt{n}} X_{\lfloor nt \rfloor} - \frac{ns - \lfloor ns \rfloor}{\sqrt{n}} X_{\lfloor ns \rfloor} \right) \\ &= \lim_{n\to\infty} \left(\frac{\sqrt{\lfloor nt \rfloor - \lfloor ns \rfloor}}{\sqrt{n}} \cdot \frac{1}{\sqrt{\lfloor nt \rfloor - \lfloor ns \rfloor}} \sum_{i=\lfloor ns \rfloor}^{\lfloor nt \rfloor} X_i \right) = \sqrt{t-s}Z. \end{aligned}$$

So it is clear that the increments are normally distributed with expectation 0 and variance $t - s$. Finally, we need that any two non-overlapping increments are independent. Since increments of B_t are the limit of increments of M_t^n; however, the independence of non-overlapping random walk increments implies the same for non-overlapping increments of B_t. ∎

The stochastic process given by B_t fulfills our four properties, and so it will be our definition of **Brownian motion** moving forward.

4. Continuous-Time

Now that we have seen how to price an ECO given a discrete model of stock growth, we will use Brownian motion to construct a continuous-time model stock growth. This model is known as the log-normal model, and it is one of the most commonly used stock models.

4.1. **Log-Normal Stock Growth.** Intuitively, we would like our model stock to grow on average but be subject to some randomness. To capture this, we introduce the following assumptions:

(A1) Stock prices grow at the risk-free rate r on average, i.e

$$E(S(t)) = S_0 e^{rt}. \tag{4.1}$$

(A2) The randomness within stock prices is introduced through Brownian motion.

First, we want to look at assumption (A1). It says that, on average, a stock grows at the fixed rate r, like money invested in a bank or Treasury bond. This means we want

$$\begin{cases} \frac{dS}{dt} = rS \\ S(0) = S_0. \end{cases} \tag{4.2}$$

The second equation in this system is simply and initial condition, and this starting price S_0 must be positive.

Theorem 4.1. *A solution to the initial value ordinary differential equation (ODE) 4.2 is*

$$S(t) = S_0 e^{rt}. \tag{4.3}$$

Proof. We start by rearranging the ODE and integrating both sides to get

$$\frac{dS}{S} = rdt \Longrightarrow \int \frac{dS}{S} = \int rdt$$

$$\Longrightarrow ln(S) = rt + C \Longrightarrow S(t) = Ce^{rt}.$$

All solutions of the ODE will be of this form, but we only need that equations of this form are solutions for our purposes. Now, using the initial data, we can see that a solution which satisfies equation 4.2 completely is

$$S(0) = Ce^0 = C = S_0 \Longrightarrow S(t) = S_0 e^{rt}.$$

□

Since our original assumption was that the *average* price of the stock would follow the system 4.2, we know that it is actually the expectation of S that is given by equation 4.3. We now move on to consider (A2). For this assumption, the standard normal random varible

Z is entered into the log-linear form of equation 4.3 to give

$$ln(S(t)) = ln(S_0) + rt + cZ.$$

We use the log-linear form of equation 4.3 because it allows us to insure that the stock price never becomes negative. Also, c is a constant that will be determined later to insure our model immitates the volatility of a stock. Introducing randomness in this way fulfills (A2), however, since we showed in the previous section that $B_{c^2} = cZ$ is Brownian motion. If we solve for $S(t)$ again, we get the equation

$$S(t) = S_0 e^{rt+cZ}.$$

We will proceed by showing that the expectation of this equation is not $S_0 e^{rt}$ as we want, and we will adjust the random variable slightly. The randomness is then introduced at every moment to find c.

We begin with a basic theorem from statistics.

Theorem 4.2. *Let X be a random variable with density funciton $f(x)$, and let $g : \mathbb{R} \longrightarrow \mathbb{R}$ be integrable with respect to $f(x)$. Then we have*

$$E\left[g(X)\right] = \int_{\mathbb{R}} g(x) f(x) dx.$$

Proof. First, we let $G \subset \mathbb{R}$ be the range of the function g. Then we see $g(Y)$ has a distribution function given by

$$F_{g(Y)}(u) = P(g(Y) < u) = \int_{y \ s.t. \ g(y)<u} f(y) dy. \tag{4.4}$$

It is clear that the expectation of $g(Y)$ is given by

$$\int_{y \ s.t. \ g(y) \in G} g(y) f(y) dy = \int_{\mathbb{R}} g(y) f(y) dy.$$

This proof follows a similar result in Wackerly, Mendenhall, and Scheaffer (2008)(13). □

Using this theorem, we can calculate the expectation of $S(t)$ to be

$$E\left[S_0 e^{rt} e^{cZ}\right] = S_0 e^{rt} E(e^{cZ}),$$

and

$$E(e^{cZ}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{cx} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-c)^2 + \frac{1}{2}c^2} dx = e^{\frac{1}{2}c^2}.$$

The average value of $S(t)$ now contains the extra term $e^{\frac{1}{2}c^2}$, which we can eliminate by redefining $S(t)$ as

$$S(t) = S_0 e^{rt} e^{c(Z - \frac{1}{2}c)}. \tag{4.5}$$

Now, $S(t)$ satisfies (A1) and (A2), and so we only need to find c. To do this, we will consider adding a standard normal variable at finitely many moments of time. Imagine that we split the interval $[0, T]$ into k intervals of length $\Delta t = T/k$. Then it is easy to see that we can obtain the following formulas for the price of the stock at times Δt, $2\Delta t$, and $j\Delta t$ where $1 \leq j \leq k$.

$$S(\Delta t) = S_0 e^{r\Delta t} e^{c(Z_1 - \frac{1}{2}c)}$$
$$S(2\Delta t) = S_0 e^{2r\Delta t} e^{c(Z_1 + Z_2 - c)}$$
$$\vdots$$
$$S(j\Delta t) = S_0 e^{jr\Delta t} e^{c(Z_1 + Z_2 + \cdots + Z_j - \frac{j}{2}c)} \tag{4.6}$$

For ease of notation, we will let $W_j := Z_1 + Z_2 + \cdots + Z_j$ for the remainder of this discussion. Now that we have introduced randomness at k moments, we want to define c so that the volatility of our model matches that of the stock we are modeling. We don't worry about how to determine the volatility of a stock, but rather we assume the stock's price has a given variance σ^2. With this variance, c can be solved for as follows

$$\sigma^2 j\Delta t = V(cW_j) = c^2 V(Z_1 + Z_2 + \cdots + Z_j) = jc^2$$
$$\Longrightarrow c = \sigma\sqrt{\Delta t} = \sigma\sqrt{\frac{T}{k}}.$$

Using this value for c in equation 4.6 gives

$$S(j\Delta t) = S_0 e^{jr\Delta t} e^{\sigma\sqrt{\frac{jT}{k}}\frac{W_j}{\sqrt{j}}} e^{-\frac{j}{2}\sigma^2\Delta t}.$$

Since we have come this far, it is worth nothing that one can apply the Central Limit Theoremto see

$$\frac{W_k}{\sqrt{k}} \longrightarrow Z \text{ as } k \longrightarrow \infty,$$

and so, letting k go to ∞ in our equation for $S(j\Delta t)$, our discrete times $j\Delta t$ can be replaced with the continuous-time variable t, and we get

$$S(t) = S_0 e^{rt} e^{\sigma\sqrt{t}Z} e^{-\frac{1}{2}\sigma^2 t} = S_0 e^{(r - \frac{1}{2}\sigma^2)t} e^{\sigma\sqrt{t}Z}.$$

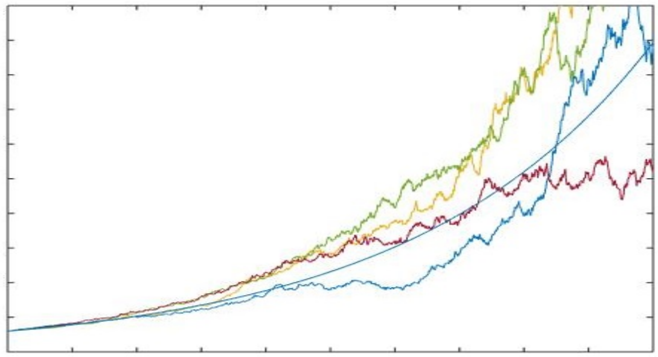
This is also where we would arrive if we had replaced c in equation 4.5. The ideas of this section are stated concisely in the following definition.

Definition 4.3. For any $t > 0$, let $S(t)$ denote the price of a stock at time t. Let S_0 denote the initial price of the stock. Then, in the log-normal model, $S(t)$ is given by

$$S(t) = S_0 e^{(r - \frac{1}{2}\sigma^2)t} e^{\sigma\sqrt{t}Z}$$

where Z is a standard normal random variable and σ is the standard deviation of the stock.

Figure 5. This is a look at how stocks might move in the log-normal model. The smooth blue line is the expected growth of the stock, and the other lines are actual stock movements in this model.



4.2. **BSM Formula: First Derivation.** We now have a way to price stocks in continuous-time, and we will use the log-normal model, as well as the concept of replicating portfolios, to price ECOs. The following theorem for the discussion considered in this section.

Theorem 4.4. *If the motion of a stock is described by the log-normal model, then the price V_0 of an option with uncertain payoff V_T at a fixed time T is given by*

$$V_0 = e^{-rT} E\left(V_T\right) \tag{4.7}$$

This result can be found in Cosimano and Himonas (2017)$^{(3)}$. It is not easy to prove without sotchastic calculus, which we will look at in later sections. It is a somewhat intuitive result given our model of stocks, though. It simply says that the value of an option at any time is the present value of its expected payoff. This is exactly what we would find for the present value of a stock if we solved for S_0 in equation 4.1. This symmetry is a result of the fact that we are working over the risk-neutral probability space, which is a concept beyond the scope of this paper.

Equation 4.7 is important because it tells us how to price any option with known payoffs at a known time T. To arrive at the BSM formul, it is necessary to apply equation 4.7 to the payoffs of a European call option. The following corollary derives the formula we seek.

Corollary 4.5. *The value of a European call option with strike price K, and whose underlying stock follows the log-normal model is given by*

$$V_0 = S_0 N(d_1) + e^{-rT} K N(d_2), \tag{4.8}$$

where r is the risk-free rate, N is the normal distribution function, and

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \text{ and } d_2 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.$$

Proof. First, recall that the payoff of an ECO is given by $(S_T - K)^+$. Applying equation 4.7 we see that we must solve

$$V_0 = e^{-rT} E\left((S_T - K)^+\right).$$

Using theorem 4.2, we can calculate the expectation in the above equation by solving the integral

$$E\left((S_T - K)^+\right) = \int_{-\infty}^{\infty} (S_T - K)^+ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx.$$

Now, we want to find where $S_T - K \geq 0$ in order to understand what we are integrating. For this, the log-normal equation is substituted in place of S_T, and we find

$$S_0 e^{(r-\frac{1}{2}\sigma^2)T} e^{\sigma\sqrt{T}x} - K \geq 0 \implies x \geq \frac{\ln\left(\frac{K}{S_0}\right) - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.$$

It is clear that this value is $-d_2$ as defined in the corollary, and so the integral we wish to solve is given by

$$\frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} \left(S_0 e^{(r-\frac{1}{2}\sigma^2)T} e^{\sigma\sqrt{T}x} - K\right) e^{-\frac{1}{2}x^2} dx$$
$$= e^{rT} \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} S_0 e^{-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}x - \frac{1}{2}x^2} dx - K \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} e^{-\frac{1}{2}x^2} dx.$$

It is clear that the right integral above is $K(1 - N(-d_2))$ or $KN(d_2)$. We can complete the squares within the exponent of the left integral to get the expression

$$-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}x - \frac{1}{2}x^2 = -\frac{1}{2}(x - \sigma\sqrt{T})^2.$$

With this expression, we can solve the left integral by using substituion

$$e^{rT} \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} S_0 e^{-\frac{1}{2}(x-\sigma\sqrt{T})^2} dx = e^{rT} S_0 \frac{1}{\sqrt{2\pi}} \int_{-d_2-\sigma\sqrt{T}}^{\infty} e^{-\frac{1}{2}y^2} dy = e^{rT} S_0 N(d_1).$$

Finally, we use the solutions of these integrals to solve for value of the option

$$V_0 = e^{-rT}\left[e^{rT} S_0 N(d_1) - KN(d_2)\right] \implies \boxed{V_0 = S_0 N(d_1) - e^{-rT} KN(d_2).}$$

This proof follows a similar result from Cosimano and Himonas (2017)[(3)]. □

With corollary 4.5 , we have reached our first derivation of the BSM equation. From here, we will introduce stochastic calculus in order to derive the BSM partial differential equation.

5. Stochastic Calculus

In order to derive the partial differential equation that will lead to the BSM formula, we need to look at some ideas from stochastic calculus originally developed by Itô. We first consider simple processes, and then we will look at how they can allow us to build into general stochastic processes. We then construct the Itô integral and conclude with Itô's formula.

5.1. Simple Stochastic Processes.

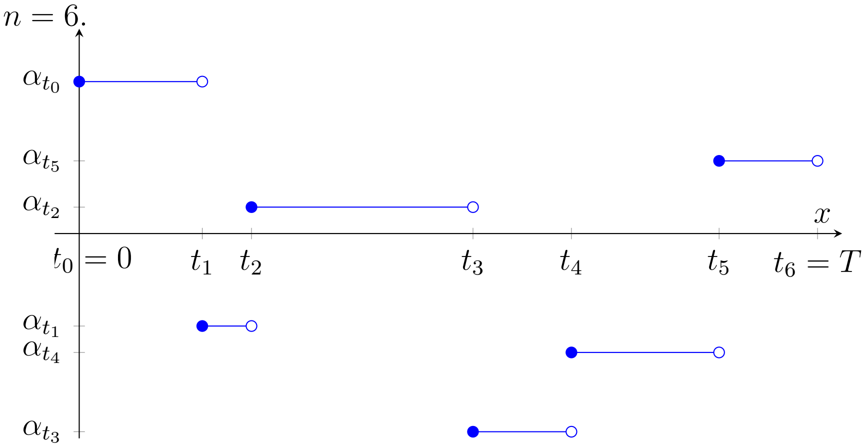
We let $B_t : \Omega \to \mathbb{R}$ denote a Brownian motion stochastic process for the remainder of this discussion, and we will need to recall the properties of Brownian motion from section 3.3. Also, consider the following restricted class of stochastic processes.

Definition 5.1. A **simple stochastic process** is a function of the form

$$f(t, \omega) = \sum_{j=0}^{n-1} \alpha_j(\omega) \mathbb{1}_{[t_j, t_{j+1})}(t),$$

where $\mathbb{1}$ is the characteristic function, $0 = t_0 < t_1 < \cdots < t_n = T$, and $\alpha_j(\omega)$ are random variables dependent upon the sequence of coin tosses ω.

Figure 6. This is a graphical representation of a simple stochastic process in which $n = 6$.



This definition is helpful because it cuts the time interval $[0, T]$ into finitely many half-intervals on which the function f is constant. We can now simply define the Itô integral of a simple stochastic process with respect to B_t by

$$\int_0^T f(t, \omega) dB_t = \sum_{j=0}^{n-1} \alpha_j(\omega)(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)). \tag{5.1}$$

The sum here looks a lot like that of the Riemann integral, except in Riemann integration, we would continually partition $[0, T]$ into smaller intervals to find the limiting sum. The important distinction here is that on each interval, we only consider the value of the function at the left-most point. In fact, at the right-most point on each interval, f has the value α_{j+1} instead of α_j. This is necessary, because B_t is not differentiable at any point (t, ω), and so there are general stochastic processes f for which a left-sum and right-sum would not converge with respect to any partition. For clarification, we consider the following proposition.

Proposition 5.2. *If B_t is a Brownian process, define the following sums*

$$S_n^L = \sum_{j=0}^{n-1} B_{t_j}(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)) \text{ and } S_n^R = \sum_{j=0}^{n-1} B_{t_{j+1}}(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)).$$

Then, $E(S_n^L) \neq E(S_n^R)$ for any n.

Proof. First, we use the fact that B_{t_j} and $(B_{t_{j+1}}(\omega) - B_{t_j}(\omega))$ are independent. This can be seen intuitively since the position B_{t_j} precedes the step, or increment, $(B_{t_{j+1}}(\omega) - B_{t_j}(\omega))$. We also have from our properties of brownian motion that the expectation of each is zero. Thus, we can say

$$E(S_n^L) = E\left(\sum_{j=0}^{n-1} B_{t_j}(B_{t_{j+1}}(\omega) - B_{t_j}(\omega))\right) = \sum_{j=0}^{n-1} E(B_{t_j}) E\left(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)\right) = 0. \tag{5.2}$$

Now, we want to consider the following expectation

$$\begin{aligned} E\left[\left(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)\right)^2\right] &= \frac{1}{\sqrt{2\pi(t_{j+1} - t_j)}} \int_{-\infty}^{\infty} x^2 e^{\frac{-x^2}{2(t_{j+1}-t_j)}} dx \\ &= \frac{\sqrt{t_{j+1} - t_j}}{\sqrt{2\pi}} \left(\left[-x e^{\frac{-x^2}{2(t_{j+1}-t_j)}} \right]_{x=-\infty}^{x=\infty} + \int_{-\infty}^{\infty} e^{\frac{-x^2}{2(t_{j+1}-t_j)}} dx \right) \\ &= \sqrt{t_{j+1} - t_j} \left(0 + \sqrt{t_{j+1} - t_j}\right) = t_{j+1} - t_j. \end{aligned} \tag{5.3}$$

This solution follows from the fact that $B_{t_{j+1}}(\omega) - B_{t_j}(\omega)$ is normally distributed with variance $t_{j+1} - t_j$. We can also use this solution to see that

$$\begin{aligned} E\left(S_n^R - S_n^L\right) &= E\left(\sum_{j=0}^{n-1} B_{t_{j+1}}(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)) - \sum_{j=0}^{n-1} B_{t_j}(B_{t_{j+1}}(\omega) - B_{t_j}(\omega))\right) \\ &\sum_{j=0}^{n-1} E\left[(B_{t_{j+1}}(\omega) - B_{t_j}(\omega))^2\right] = \sum_{j=0}^{n-1} (t_{j+1} - t_j) = T. \end{aligned}$$

Thus, by equation 5.2, we have that $S_n^R = T$. This proof was inspired by Cosimano and Himonas (2017)[3]. ☐

The importance of this proposition is that no matter how we partition the interval $[0, T]$, the sums S_n^L and S_n^R will not converge. This is why we must define the Itô integral as in equation 5.1, using the left-most point of every interval.

5.2. **General Itô Integral.** Having defined the Itô integral for simple stochastic processes, we now seek to develop a definition for general stochastic processes. To this end, we introduce some concepts from measure theory, and A deeper explanation of these concepts can be found in Folland (1999)[4]. The space of $L^2(\omega)$ functions is important for to the construction to come.

Definition 5.3. We say that a function $f(t,\omega)$ belongs to the **space of $L^2(\omega)$ functions** if

$$E\left[\int_0^T f^2(t,\omega)dt\right] < \infty.$$

The most important property of this space is that simple stochastic processes are dense in it. This means that we can approximate any stochastic process with a sequence of simple processes in the sense of the following theorem.

Theorem 5.4. *If $f(t,\omega)$ is a stochastic process such that $f \in L^2(\omega)$, then there exists a sequence of simple stochastic processes $f_n(t,\omega)$ such that*

$$E\left[\int_0^T (f(t,\omega) - f_n(t,\omega))^2 dt\right] \to 0 \text{ as } n \to \infty.$$

The proof of this theorem is a major piece of any text on measure theory and can be found in Folland (1999)[4]. This theorem can be utilized once we consider the next result, often known as the Itô Isometry.

Theorem 5.5. *For any simple stochastic process $f(t,\omega)$, we have that*

$$E\left[\left(\int_0^T f(t,\omega)dB_t\right)^2\right] = E\left(\int_0^T f^2(t,\omega)dt\right).$$

Proof. First, let f be a simple stochastic process, and use the previously developed definition for the Itô integral for a simple stochastic process to observe (here we neglect the dependence of variables on ω for brevity of notation)

$$E\left[\left(\int_0^T f(t,\omega)dB_t\right)^2\right] = E\left[\left(\sum_{j=0}^{n-1} \alpha_j (B_{t_{j+1}} - B_{t_j})\right)^2\right]$$

$$= \sum_{j=0}^{n-1} E\left(\alpha_j^2\right) E\left[(B_{t_{j+1}} - B_{t_j})^2\right] + \sum_{j=0,k=0,j\neq k}^{n-1} E(\alpha_j)E(\alpha_k)E\left(B_{t_{j+1}} - B_{t_j}\right) E\left(B_{t_{k+1}} - B_{t_k}\right).$$

In the right sum, we can separate the two increments since they are non-overlapping, and thus independent. This sum then disappears as the expectation of each increment is zero.

Finally, we can rewrite the entire expectation as

$$\sum_{j=0}^{n-1} E\left(\alpha_j^2\right)(t_{j+1}-t_j) = E\left(\sum_{j=0}^{n-1} \alpha_j^2 (t_{j+1}-t_j)\right) = E\left(\int_0^T f^2(t,\omega)dt\right).$$

□

With this resul,t we can move from the definite differential dt to the stochastic differential dB_t, which will eventually allow us to use theorem 5.4 in a way that relates to Itô's integral. The next step in our construction, however, is to introduce a type of convergence.

Definition 5.6. Let $I_n(\omega)$ be a sequence of random variables over Ω, and let $I(\omega)$ be another such random variable. We say that I_n **converges in** $L^2(\omega)$ to I if

$$E\left[(I(\omega) - I_n(\omega))^2\right] \to 0 \text{ as } n \to \infty.$$

This convergence can also be referenced using the notation $I \stackrel{L^2(\omega)}{=} \lim_{n\to\infty} I_n$. We will define the Itô integral in terms of such a limit of simple stochastic processes. Our next result will show that this construction works for all $f \in L^2(\omega)$.

Theorem 5.7. *Let $f : [0,T]\times\Omega \to \mathbb{R}$ be an $L^2(\omega)$ stochastic process, and let $f_n : [0,T]\times\Omega \to \mathbb{R}$ be a sequence of simple stochastic processes which converge to f as in theorem 5.4. There must exist a random variable $I(\omega)$ such that*

$$I(\omega) \stackrel{L^2(\omega)}{=} \lim_{n\to\infty} \int_0^T f_n(t,\omega)dB_t(\omega).$$

Proof. Since $L^2(\omega)$ is a complete space, we need only show that the sequence of integrals of f_n is Cauchy. For this, let $n \neq m$, and observe by theorem 5.5 that

$$E\left[\left(\int_0^T (f_m(t,\omega) - f_n(t,\omega))dB_t\right)^2\right] = E\left[\int_0^T (f_m(t,\omega) - f_n(t,\omega))^2 dt\right].$$

Now we can say

$$\begin{aligned}(f_m(t,\omega) - f_n(t,\omega))^2 &= (f_m(t,\omega) - f(t,\omega) + f(t,\omega) - f_n(t,\omega))^2 \\ &\leq 2(f_m(t,\omega) - f(t,\omega))^2 + 2(f(t,\omega) - f_n(t,\omega))^2.\end{aligned}$$

This implies that our desired expectation is less than or equal to

$$2\int_0^T (f_m(t,\omega) - f(t,\omega))^2 dt + 2\int_0^T (f(t,\omega) - f_n(t,\omega))^2 dt.$$

By construction of the sequence f_n, these integrals both go to 0 as $n \to \infty$. Therefore the sequence f_n is Cauchy, and it has a limit $I(\omega)$ in $L^2(\omega)$. □

It might be guessed that I is the integral value we seek, but a concise definition follows to make this point clear.

Definition 5.8. Let $f : [0,T] \times \Omega \to \mathbb{R}$ be an $L^2(\omega)$ stochastic process, and let $f_n : [0,T] \times \Omega \to \mathbb{R}$ be simple stochastic processes such that

$$E\left[\int_0^T (f(t,\omega) - f_n(t,\omega))^2 dt\right] \to 0 \text{ as } n \to \infty.$$

We define the **Itô integral of** f to be

$$\int_0^T f(t,\omega)dB_t \stackrel{L^2(\omega)}{=} \lim_{n\to\infty} \int_0^T f_n(t,\omega)dB_t(\omega). \tag{5.4}$$

As a final note, it is important to realize we have shown only that such an integral exists. Finding a numerical value for the Itô integral of a specific stochastic process requires finding a converging sequence of simple stochastic processes.

5.3. **Itô's Formula.** So far, we have constructed the Itô integral using the density of simple processes in $L^2(\omega)$. In order to derive the partial differential equation we seek, another result from stochastic calculus is needed. This section will culminate in the result, known as Itô's Formula. We begin by introducing the following definition.

Definition 5.9. A stochastic process $X_t : [0,T] \times \Omega \to \mathbb{R}$ is known as a **stochastic integral** if it satisfies

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s,$$

where X_0 is fixed, and $v \in L^2(\omega)$, i.e. $E\left[\int_0^T v^2(s,\omega)ds\right] < \infty$. Another form for a stochastic integral is that of a **stochastic differential equation**, which is given by

$$dX_t = udt + vdB_t.$$

With this definition, we wish to introduce a quick set of rules. In the finaly result, we will need to calculate $(dX_t)^2$, and for this we note that $(dt)^2 = dtdB_t = dB_tdt = 0$, and $(dB_t)^2 = dt$. We are now ready to consider Itô's Formula, which is stated below.

Theorem 5.10. *Let X_t be a stochastic integral and $g(t,x) : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function. Also, let $Y_t = g(t,X_t)$ be another stochastic process. In this case, Y_t is a stochastic integral with*

$$dY_t = \frac{\partial g}{\partial t}(t,X_t)\,dt + \frac{\partial g}{\partial x}(t,X_t)\,dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\,(dX_t)^2. \tag{5.5}$$

In order to better understand Itô's formula, we would like to write equation 5.5 in integral form. We begin by observing that we can simplify the final term as follows

$$\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\,(dX_t)^2 = \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\,(udt + vdB_t)^2$$

$$= \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\left(u^2(dt)^2 + uvdtdB_t + vudB_tdt + v^2(dB_t)^2\right) = \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\,v^2dt.$$

We can also write the middle term of equation 5.5 as follows

$$\frac{\partial g}{\partial x}(t, X_t)\, dX_t = \frac{\partial g}{\partial x}(t, X_t)\, udt + \frac{\partial g}{\partial x}(t, X_t)\, vdB_t.$$

Thus, we can rewrite equation 5.5 as

$$dY_t = \left(\frac{\partial g}{\partial t}(t, X_t) + \frac{\partial g}{\partial x}(t, X_t)\, u(t,\omega) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t)\, v^2(t,\omega)\right) dt + \frac{\partial g}{\partial x}(t, X_t)\, v(t,\omega) dB_t.$$

Now, since $Y_0 = g(0, X_0)$, we get the integral form to be

$$Y_t = g(t, X_t) = g(0, X_0) + \int_0^t \left(\frac{\partial g}{\partial s}(s, X_s) + \frac{\partial g}{\partial x}(s, X_s)\, u(s,\omega) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(s, X_s)\, v^2(s,\omega)\right) ds$$
$$+ \int_0^t \frac{\partial g}{\partial x}(s, X_s)\, v(s,\omega) dB_s.$$

This concludes the stochastic calculus that we will need to derive the Black-Scholes-Merton PDE. An elaboration on Itô's Formula can be found in Cosimano and Himonas (2017)[3].

6. BSM Formula Revisited

This section provides a more rigorous derivation of the BSM formula shown earlier. It is inspired by methods used in Merton (1990)[7], and its structure follows that of Cosimano and Himonas (2017)[3]. The brilliant change of variables utilized in section 6.2 is so important to the following discussion that we note it is the product of Merton's work.

6.1. **BSM Partial Differential Equation.** To help us derive the Black-Scholes-Merton PDE, we will use the concept of a replicating portfolio again. We want the value of the replicating portfolio X_t at any time t to be equal to that of the call option V_t. In order to do this, we will say that our porfolio consists of Δ_t shares of stock and $X_t - \Delta_t S_t$ dollars invested at the risk-free rate r at any time t. From this, we obtain the differential

$$dX_t = \Delta_t dS_t + r(X_t - \Delta_t S_t)dt, \tag{6.1}$$

which describes how the value of our portfolio will change with time. Now, since we have $X_t = V(t, S_t)$ by assumption, we see that

$$d\left(e^{-rt}X_t\right) = d\left(e^{-rt}V(t, S_t)\right). \tag{6.2}$$

This important equation comes from Cosimano and Himonas (2017)[3]. To calculate the left differential in equation 6.2, we must apply Itô's Formula with $g(t, x) = e^{-rt}x$ to get

$$d(e^{-rt}X_t) = -re^{-rt}X_t dt + e^{-rt}dX_t.$$

Because we assume that stocks grow log-normally at a given rate r_s (which we will eventually equate to the risk-free rate), we have the stochastic differential

$$dS_t = r_s S_t dt + \sigma S_t dB_t.$$

Applying this fact and equation 6.1, we see that

$$\begin{aligned} d(e^{-rt}X_t) &= -re^{-rt}X_t dt + e^{-rt}\left(\Delta_t dS_t + r(X_t - \Delta_t S_t)dt\right) \\ &= \Delta_t e^{-rt}\left(dS_t - rS_t dt\right) \\ &= \Delta_t e^{-rt}\left(r_s S_t dt + \sigma S_t dB_t - rS_t dt\right) \\ &= \Delta_t e^{-rt}\left((r_s - r)S_t dt + \sigma S_t dB_t\right). \end{aligned} \tag{6.3}$$

Now, going back to the differential on the right side of equation 6.2, we will use Itô's Formula again with $g(t, x) = e^{-rt}V(t, x)$. For clarity, we list the relevant partials

$$\frac{\partial g}{\partial t}(t, x) = -re^{-rt}V(t, x) + e^{-rt}\frac{\partial V}{\partial t}(t, x)$$

$$\text{and } \frac{\partial g}{\partial x}(t, x) = e^{-rt}\frac{\partial V}{\partial x}(t, x) \text{ and } \frac{\partial^2 g}{\partial x^2}(t, x) = e^{-rt}\frac{\partial^2 V}{\partial x^2}(t, x).$$

Thus we can solve for the desired differential

(6.4)

$$d\left(e^{-rt}V(t,S_t)\right) = \left(-re^{-rt}V(t,S_t) + e^{-rt}\frac{\partial V}{\partial t}(t,S_t)\right)dt + e^{-rt}\frac{\partial V}{\partial x}(t,S_t)dS_t + e^{-rt}\frac{\partial^2 V}{\partial x^2}(t,S_t)d(S_t)^2.$$

Recall that since S_t is a stochastic integral, we have $(dS_t) = \sigma^2 S_t^2 dt$. This implies that the last two terms in equation 6.4 can be written as

$$e^{-rt}\frac{\partial V}{\partial x}(t,S_t)(r_s S_t dt + \sigma S_t dB_t) + \sigma^2 S_t^2 e^{-rt}\frac{\partial^2 V}{\partial x^2}(t,S_t)dt.$$

Finally, we can solve as follows

$$d\left(e^{-rt}V(t,S_t)\right) = e^{-rt}\left(-rV(t,S_t) + \frac{\partial V}{\partial t}(t,S_t) + r_s S_t\frac{\partial V}{\partial x}(t,S_t) + \sigma^2 S_t^2\frac{\partial^2 V}{\partial x^2}(t,S_t)\right)dt$$

$$+\, e^{-rt}\sigma S_t\frac{\partial V}{\partial x}(t,S_t)dB_t. \tag{6.5}$$

Now both the differentials 6.3 and 6.5, which have been calculated, contain a dt component and a dB_t component. Both of these must be equal according to equation 6.2. So we first set the dB_t components equal to get

$$\Delta_t e^{-rt}\sigma S_t = e^{-rt}\sigma S_t\frac{\partial V}{\partial x}(t,S_t) \implies \Delta_t = \frac{\partial V}{\partial x}(t,S_t).$$

This result is the hedging formula, or the shares of stock that would need to be purchased at any time t to replicate the option. Now, we can equate the dt components to get

$$\Delta_t e^{-rt}(r_s - r)S_t = e^{-rt}\left(-rV + \frac{\partial V}{\partial t} + r_s S_t\frac{\partial V}{\partial x} + \sigma^2 S_t^2\frac{\partial^2 V}{\partial x^2}\right).$$

Next, substituting the value of Δ_t into the above equation and combining terms gives

$$\frac{\partial V}{\partial x}e^{-rt}(r_s - r)S_t = e^{-rt}\left(-rV + \frac{\partial V}{\partial t} + r_s S_t\frac{\partial V}{\partial x} + \sigma^2 S_t^2\frac{\partial^2 V}{\partial x^2}\right).$$

When setting up this model, we assumed that stocks grow at the risk-free rate on average. This means that $r_s = r$, and thus

$$\boxed{\frac{\partial V}{\partial t} + rs\frac{\partial V}{\partial s} + \sigma^2 s^2\frac{\partial^2 V}{\partial s^2} - rV = 0.} \tag{6.6}$$

This is the Black-Scholes-Merton partial differential equation.

6.2. **Solving the PDE.** Many differential equations are solved by transformation into a simpler equation with a known solution or solutions. We shall solve the Black-Scholes-Merton PDE by using Merton's change of variables to turn it into the well-known diffusion equation. The diffussion initial value problem (IVP) is the partial differential equation given by

$$
(6.7) \qquad \begin{cases} u_{xx} - u_t = 0, & -\infty < x < \infty \text{ and } t > 0 \\ u(x,0) = u_0. \end{cases}
$$

Theorem 6.1. *The system 6.7 has*

$$
(6.8) \qquad u(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4t}} u_0(y)dy
$$

as a solution.

Proof. It can be shown that 6.8 is in fact the unique solution of the diffusion IVP, but for our purposes, the fact that it is a solution is sufficient. Showing this only requires checking the derivatives of u and its initial condition.

Beginning with the derivatives, we can see that

$$
\frac{\partial u}{\partial x} = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} \frac{\partial}{\partial x} e^{-\frac{(x-y)^2}{4t}} u_0(y)dy = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} \left(-\frac{x-y}{2t}\right) e^{-\frac{(x-y)^2}{4t}} u_0(y)dy
$$

$$
\text{and } \frac{\partial^2 u}{\partial x^2} = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} \left(\frac{(x-y)^2}{4t^2} - \frac{1}{2t}\right) e^{-\frac{(x-y)^2}{4t}} u_0(y)dy.
$$

We also observe

$$
\frac{\partial u}{\partial t} = -\frac{1}{2t\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4t}} u_0(y)dy + \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} \left(\frac{(x-y)^2}{4t^2}\right) e^{-\frac{(x-y)^2}{4t}} u_0(y)dy.
$$

From these derivative calculations, it is clear that $u_{xx} - u_t = 0$. To calculate the value of u at $t = 0$, we will need to pass to the limit, which requires integration by parts. First, however, we will define the function Q as

$$
Q(x,t) = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{x/\sqrt{4t}} e^{-p^2} dp.
$$

The convenience of this definition comes from the facts that

$$
\frac{\partial Q}{\partial x} = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x}{4t}}
$$

$$
\text{and } \lim_{t\to 0} Q(x,t) = \begin{cases} \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-p^2} dp = 1, \text{ if } x > 0 \\ \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{-\infty} e^{-p^2} dp = 0, \text{ if } x < 0. \end{cases}
$$

Now, we say that

$$
\lim_{t\to 0} u(x,t) = \lim_{t\to 0} \int_{-\infty}^{\infty} \frac{\partial Q}{\partial x}(x-y,t) u_0(y)dy = \lim_{t\to 0} -\int_{-\infty}^{\infty} \frac{\partial}{\partial y}[Q(x-y,t)] u_0(y)dy
$$

$$
= \lim_{t\to 0} \left[\int_{-\infty}^{\infty} Q(x-y,t)u_0'(y)dy - Q(x-y,t)u_0(y) \Big|_{y=-\infty}^{y=\infty} \right].
$$

From this point, we assume that the second term vanishes because of limiting conditions of u_0 at $y = \pm\infty$. Finally, we can see that our limit becomes

$$= \int_{-\infty}^{\infty} \lim_{t \to 0} Q(x-y,t) u_0'(y) dy = \int_{-\infty}^{x} u_0'(y) dy = u_0(x).$$

This proof follows the method found in Strauss (2008)[(12)]. $\square$

Now, let us consider the Black-Scholes-Merton PDE given by equation 6.6. We know that our function V depends on the variables $s > 0$ and $0 < t < T$. We can make the following change of variables (where $\tau = T - t$ and $\gamma = r - \frac{\sigma^2}{2}$)

$$\begin{cases} G(x,y) = e^{r\tau} V(s,t) \\ x = ln(s) + \gamma\tau \\ y = \frac{\sigma^2 \tau}{2}. \end{cases}$$

From this change of varibales, we observe

$$V_t = \frac{\partial}{\partial t}\left(e^{-r\tau}\right) G + \frac{\partial G}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial G}{\partial y}\frac{\partial y}{\partial t} = re^{-r\tau} G - \gamma G_x - \frac{\sigma^2}{2} G_y$$

$$\text{and } V_s = \frac{\partial G}{\partial x}\frac{\partial x}{\partial s} = \frac{1}{s} G_x \Longrightarrow V_{ss} = -\frac{1}{s^2} G_x + \frac{1}{s^2} G_{xx}.$$

Now, substituting into equation 6.6, we get

$$re^{-r\tau} G - \frac{\sigma^2}{2} G_x + \frac{\sigma^2}{2} G_{xx} - (r - \frac{\sigma^2}{2}) G_x - \frac{\sigma^2}{2} G_y + rG_x - re^{-r\tau} G = 0$$

$$\Longleftrightarrow G_{xx} - G_y = 0. \tag{6.9}$$

We need an intial condition for this diffusion equation. We know the terminal value of the option at time T is given by $[S_T - K]^+$. Thus we can see that $V(s,T) = [s - K]^+$. But, then

$$x(s,T) = ln(s) - \gamma(T - T) = ln(s)$$

$$\Longrightarrow V(s,T) = G(x,0) = [e^x - K]^+.$$

We can use this as our initial condition, and we will look to solve the diffusion IVP

$$\begin{cases} G_{xx} - G_y = 0 \\ G(x,0) = [e^x - K]^+. \end{cases} \tag{6.10}$$

Using theorem 6.1, we can say

$$G(x,y) = \frac{1}{\sqrt{4\pi y}} \int_{-\infty}^{\infty} e^{-\frac{1}{4y}(x-z)^2} [e^z - K]^+ dz$$

$$= \frac{1}{\sqrt{4\pi y}} \int_{ln(K)}^{\infty} e^{-\frac{1}{4y}(x-z)^2} (e^z - K) dz$$

is a solution of IVP 6.10. From here, we work with the exponent on the first term to reduce it to

$$\frac{x^2 - 2xz + z^2 - 4yz}{4y} = \frac{(z-x-2y)^2}{4y} - (x+y).$$

The first term then becomes

$$\frac{1}{\sqrt{4\pi y}} \int_{ln(K)}^{\infty} e^{-\frac{(z-x-2y)^2}{4y}} e^{x+y} dz = \frac{e^{x+y}}{\sqrt{2\pi}} \int_{\frac{ln(K)-x-2y}{\sqrt{2y}}}^{\infty} e^{-\frac{p^2}{2}} dp$$

$$= \frac{e^{x+y}}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x+2y-ln(K)}{\sqrt{2y}}} e^{-\frac{p^2}{2}} dp = e^{x+y} N\left(\frac{x+2y-ln(K)}{\sqrt{2y}}\right).$$

By completing the squares in a similar way, we see that the second term reduces as follows

$$-\frac{1}{\sqrt{4\pi y}} \int_{ln(K)}^{\infty} K e^{-\frac{1}{4y}(x-z)^2} dz = -KN\left(\frac{x-ln(K)}{\sqrt{2y}}\right).$$

Finally we can solve the diffusion equation for G and thus for V

$$G(x,y) = e^{x+y} N\left(\frac{x+2y-ln(K)}{\sqrt{2y}}\right) - KN\left(\frac{x-ln(K)}{\sqrt{2y}}\right)$$

$$\Longrightarrow V(s,t) = e^{-r\tau + ln(s) + (r-\frac{\sigma^2}{2})\tau - \frac{\sigma^2}{2}\tau} N\left(\frac{ln(s) + (r-\frac{\sigma^2}{2})\tau + \sigma^2\tau - ln(K)}{\sigma\sqrt{\tau}}\right)$$

$$-Ke^{-r\tau} N\left(\frac{ln(s) + (r-\frac{\sigma^2}{2})\tau - ln(K)}{\sigma\sqrt{\tau}}\right)$$

$$\boxed{V(s,t) = sN\left(\frac{ln(s/K) + (r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\right) - Ke^{-r\tau} N\left(\frac{ln(s/K) + (r-\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\right).} \tag{6.11}$$

The above equation is the famed Black-Scholes-Merton formula, and it is the pinnacle result achieved in this paper. As we saw earlier, most sources will use the variables

$$d_1 = \frac{ln(s/K) + (r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}} \text{ and } d_2 = d_1 - \sigma\sqrt{\tau}$$

so that the value of an ECO at time $t = 0$ can be neatly written as

$$V_0^{call} = S_0 N(d_1) - Ke^{-rT} N(d_2). \tag{6.12}$$

In practice, financial analysts use this formula as only one of many tools to price options. The next section gives a quick look at some methods used in the real world that relate to the BSM formula.

7. Applications

This section will focus on two main applications of the BSM formula: implied volatility surfaces and the Heston Model. These topics were motivated by insight gained from Stivers (2017)[11]. Much of the detail in this section was found through additional research, however.

7.1. Implied Volatility Surfaces.

To begin, we must introduce the concept of implied volatility.

Definition 7.1. Let us consider a stock with stock price S_0 and risk-free interest rate r. Also, consider a European call option with strike price K, time to expiration T, and current price V_0. The **implied volatility** of the stock is the volatility σ_A that is required for the option to be priced at V_0. It can be found by solving for σ_A in the BSM formula

$$V_0 = S_0 N\left(\frac{\ln(S_0/K) + (r + \frac{\sigma_A{}^2}{2})T}{\sigma_A\sqrt{T}}\right) - Ke^{-rT} N\left(\frac{\ln(S_0/K) + (r - \frac{\sigma_A{}^2}{2})T}{\sigma_A\sqrt{T}}\right). \tag{7.1}$$

Clearly this definition only makes sense if exactly one volatility will solve the above equaiton. To check this, we consider the next proposition.

Proposition 7.2. *Equation 7.1 has exactly one solution σ_A for a given option price V_0.*

Proof. First, we want to look at the derivative

$$\frac{\partial V_0}{\partial \sigma_A} = S_0 \frac{\partial N}{\partial d_1}\frac{\partial d_1}{\partial \sigma_A} - Ke^{-rt}\frac{\partial N}{\partial d_2}\frac{\partial d_2}{\partial \sigma_A} \tag{7.2}$$

where d_1 and d_2 are defined as in the previous chapter. Now, we observe,

$$\frac{\partial N}{\partial d_2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(d_1 - \sigma_A\sqrt{T}\right)^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}d_1{}^2 + \sigma_A\sqrt{T}d_1 - \frac{1}{2}\sigma_A{}^2 T} = \frac{\partial N}{\partial d_1}\frac{1}{\sqrt{2\pi}} e^{\sigma_A\sqrt{T}d_1 - \frac{1}{2}\sigma_A{}^2 T}$$
$$= \frac{\partial N}{\partial d_1}\frac{1}{\sqrt{2\pi}} e^{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma_A{}^2\right)T - \frac{1}{2}\sigma_A{}^2 T} = \frac{\partial N}{\partial d_1}\frac{S_0}{K}e^{rT}.$$

Thus, we can rewrite equation 7.2 as

$$\frac{\partial V_0}{\partial \sigma_A} = S_0 \frac{\partial N}{\partial d_1}\frac{\partial d_1}{\partial \sigma_A} - Ke^{-rt}\frac{dN}{dd_1}\frac{S_0}{K}e^{rt}\frac{\partial d_2}{\partial \sigma_A} = S_0\frac{\partial N}{\partial d_1}\left(\frac{\partial d_1}{\partial \sigma_A} - \frac{\partial d_2}{\partial \sigma_A}\right).$$

Next, we consider

$$\frac{\partial d_1}{\partial \sigma_A} = \frac{\partial}{\partial \sigma_A}\left(\frac{\ln(S_0/K) + r}{\sigma_A\sqrt{T}}\right) + \frac{\partial}{\partial \sigma_A}\frac{\sigma_A\sqrt{T}}{2}$$

and

$$\frac{\partial d_2}{\partial \sigma_A} = \frac{\partial}{\partial \sigma_A}\left(\frac{\ln(S_0/K) + r}{\sigma_A\sqrt{T}}\right) - \frac{\partial}{\partial \sigma_A}\frac{\sigma_A\sqrt{T}}{2}.$$

Finally, it is clear that

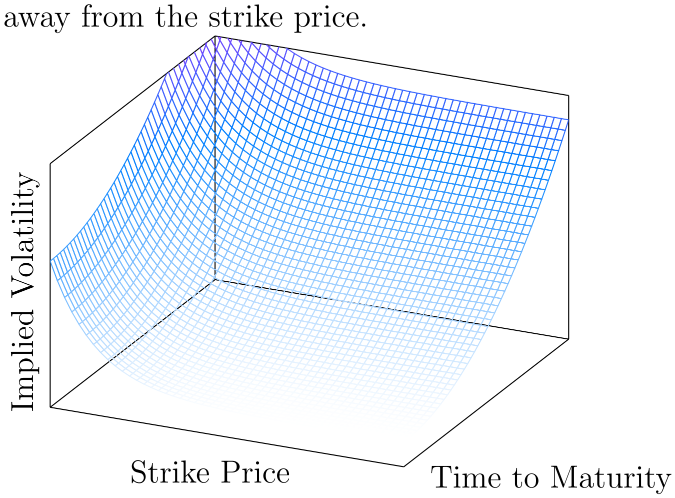
$$\frac{\partial V_0}{\partial \sigma_A} = 2S_0\frac{\partial N}{\partial d_1}\left(\frac{\partial}{\partial \sigma_A}\frac{\sigma_A\sqrt{T}}{2}\right) = S_0\sqrt{T}\frac{\partial N}{\partial d_1} > 0.$$

The last inequality follows from the facts that S_0 and T are always positive and N is an increase function. Thus, we have shown that V_0 is a strictly increasing function of the variable σ_A, and implied volatility is well defined above. This result can be found in Cosimano and Himonas (2017)[3]. □

It is important to note here that solving for the implied volatility in any realistic scenario is very difficult, and so it is done with special computing software. With this concept in mind, however, we can move on to implied volatility surfaces.

In actual markets, many different options with various strike prices and maturity dates will be available for purchase on any given stock. With this data, we can calculate the implied volatility of a stock given many option parameters. Upon doing so, we would see that implied volatility varies greatly depending on strike price and time to maturity. Thus we could plot the implied volatility of a stock on a three-dimensional graph with respect to strike price and time to maturity at many different points. The **implied volatility surface** is the surface that arises by continuously connecting each of these points. An example of such a surface is shown in figure 7 below, which is motivated by Kamal and Gatheral[6].

Figure 7. This is a very smooth example of an implied volatility surface. The lower left corner is the origin, and so this graph shows that implied volatility generally increases with increasing time to maturity and as the price of the option moves away from the strike price.



By using sophisticated graphing software and daily market information, traders are able to create implied volatility surfaces for any number of assets they wish. The goal of this process is to find a surface with a sharp local minimum point. In this way, implied volatility is used as a measure of relative cost. Lower implied volatility does not necessarily mean lower option price, but it means better expected return for your investment. For this reason, steep minimum points are rare and indicate an underpriced option in the market.

When investors see this, they will buy the undervalued option and sell options at close strike prices and maturity dates to hedge against loss. In this way, traders take advantage of a phenomenon known as **statistical arbitrage**, in which an investor can *almost* guarantee a profitable return from a portfolio requiring no initial capital without any risk of loss. These opportunities are not real arbitrage because there is always a very small chance of money loss, but it is so insignificant that investors treat these market inefficiencies the way they would normal arbitrage, and such opportunities quickly disappear. This is why traders are constantly looking for better software and updated market information.

7.2. **Heston's Model.** As mentioned earlier, the Black-Scholes-Merton equation is not often used to price options in practice. A much more common tool in industry for pricing individual options is known as Heston's Model. The math behind this model is much more complicated than that of the BSM model, and so we will only take a short look it here. The following discussion can be found in more detail in Heston (1993)[(5)].

The main improvement offered by the Heston Model is that it considers volatility of stocks to vary stochastically. In reality the volatility of stocks changes regularly, and thus Heston's model is able to track stocks on the market well. The following two stochastic equations are used to explain the behavior of stock prices and volatility, respectively,

$$dS_t = rS_0dt + \sqrt{v_t}S_0dB_1$$

and

$$d\sqrt{v_t} = -\beta\sqrt{v_t}dt + \delta dB_2$$

where r, β, and δ are fixed constants, $v_t = {\sigma_t}^2$ is the variance of the underlying asset as a funcion of time t. We also note that B_1 and B_2 are separate Brownian motion processes as usual. Notice that the first equation looks like the log-normal model, except volatility is a function of time. Also, by applying Itô's Lemma, we can arrive at the alternative equation

$$dv_t = \left[\delta^2 - 2\beta v_t\right] dt + 2\delta\sqrt{v_t}dB_2.$$

With these differential equations describing a stock, Heston derived the following PDE for a European call option with strike price K and time to maturity T

$$\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + 2\rho\delta\frac{\partial^2 V}{\partial S\partial v} + 2\delta^2 v\frac{\partial^2 V}{\partial v^2} + rS\frac{\partial V}{\partial S} + \left[\delta^2 - 2\beta v - \lambda\right]\frac{\partial V}{\partial v} - rV + \frac{\partial V}{\partial t} = 0 \tag{7.3}$$

where λ is the price of volatility risk and ρ is the correlation of B_1 with B_2. The price is also subject to the following initial conditions

$$V(S, v, T) = max(0, S - K),$$

$$V(0, v, t) = 0,$$

$$\frac{\partial V}{\partial S}(\infty, v, t) = 1,$$
$$rS\frac{\partial V}{\partial S}(S,0,t) + \delta^2\frac{\partial V}{\partial v}(S,0,t) - rV(S,0,t) + \frac{\partial V}{\partial t}(S,0,t) = 0,$$
$$V(S,\infty,t) = S.$$

Clearly this model is much more complicated than the BSM model, but by using the correct software, we can apply it in discrete time to determine the price of an option on the market under the Heston model. Because we use a discrete estimator, any calculation of price will be accompanied by an error term. This term tells us approximately how close an estimated price is to the price that would be given by the actual Heston model, which often cannot be calculated. To demonstrate the degree to which this model disagrees with the BSM model, we will create a table of prices given by each model under certain conditions. Because the Heston model has many more parameters (relating to the volatility of variance v), we will assume the rate of mean reversion of $\sqrt{v}$ is 0.1, the annualized volatility of $\sqrt{v}$ is 0.1, the correlation factor ρ is -0.3, and the long term mean of $\sqrt{v}$ is the same as the initial volatility σ. We will also list an error term associated with the Heston model, since the returned value is estimated using discrete methods.

BSM v. Heston

Stock Price	100	80	120	100	100	100
Strike Price	100	100	100	100	100	100
Years to Maturity	1	1	1	3	1	1
Risk-Free Rate	5%	5%	5%	5%	5%	5%
Initial volatility	20%	20%	20%	20%	10%	30%
BSM value	10.45	1.86	26.17	20.92	6.81	14.23
Heston value	18.42 ± 0.03	8.05 ± 0.01	32.22 ± 0.04	30.50 ± 0.08	13.88 ± 0.02	21.60 ± 0.03

All of the above Heston calculations were done using the calculator found at Yu-Ting (2016)[(2)]. These calculations show that the models in fact vary quite significantly.

8. References

(1) Black, Fischer, and Myron Scholes. "The Pricing of Options and Corporate Liabilities." *The Journal of Political Economy*, Volume 81, Issue 3, (May-Jun. 1973), 637-654.

(2) Chen, Yu-Ting. *QuantCalc: The calculator for all quants.* "Variance Reduction for Monte Carlo Simulation." 13 June, 2016, http://www.quantcalc.net/MCAV_vanilla_Heston.html. Accessed 19 Mar. 2017.

(3) Cosimano, Tom, and Alex Himonas. *Mathematical Methods in Financial Economics.* University of Notre Dame, Jan. 2017.

(4) Folland, Gerald B. *Real Analysis: Modern Techniques and Their Applications.* 2nd ed., John Wiley & Sons, Inc., 1999.

(5) Heston, Steven L. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *The Review of Financial Sutdies*, Volume 6, Issue 2 (1993), 327-343.

(6) Kamal, Michael, and Jim Gatheral. "Implied Volatility Surface." http://faculty.baruch.cuny.edu/jgatheral/impliedvolatilitysurface.pdf.

(7) Merton, Robert C. *Continuous-Time Finance.* Blackwell Publishing, 1990. http://faculty.baruch.cuny.edu/jgatheral/impliedvolatilitysurface.pdf.

(8) Shifrin, Theodore. *Multivariable Mathematics: Linear Algebra, Multivariable Calculus, and Manifolds.* Wiley, 2005.

(9) Shreve, Steven E. *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model.* Carnegie Mellon University. Springer, 2004.

(10) Shreve, Steven E. *Stochastic Calculus for Finance II: Continuous-Time Models.* Carnegie Mellon University. Springer, 2004.

(11) Stivers, Ed. Personal Interview. 21 February 2017.

(12) Strauss, Walter A. *Partial Differential Equations: An Introduction.* Brown University. 2nd ed., John Wiley & Sons, Inc, 2008.

(13) Wackerly, Dennis D, William Mendenhall III, and Richard L. Scheaffer. *Mathematical Statistics with Applications.* University of Florida. 7th ed., Brooks/Cole, 2008.