A Theoretical Derivation of the Black-Scholes Formula

University of Texas at Austin - 2023 Summer Directed Reading Program

Julian Canales

Abstract

The Black-Scholes formula was the first widely accepted standard for calculating the theoretical value of an options contract. While the formula has been shown to have its limitations, it is still critical to understand for those who wish to enrich their knowledge in financial theory. This paper will derive the Black-Scholes formula for pricing options from scratch. In order to accomplish this, one must have adequate understanding of probability theory, measure theory, stochastic processes, Brownian motion, and stochastic calculus. This paper will streamline the derivation and only include concepts that are necessary in the derivation of the Black-Scholes formula. The intention is that by the end of this paper, one will have sufficient understanding of these concepts for their applications and to engage in discussion.

Contents

1 Probability Theory and Measure Theory

We begin our exploration with elementary definitions regarding probability theory.

Definition 1.1. An experiment is a process or procedure that can be repeated and has well-defined results.

Definition 1.2. An outcome, denoted as ω, is a result of an experiment.

Definition 1.3. An event, denoted as E, is a set of outcomes from an experiment(s).

Definition 1.4. A probability is the likelihood of an event or outcome. Intuitively, probabilities are typically defined by the number of times we observe an outcome/event over a large amount of realized outcomes.

Definition 1.5. The sample space, Ω, of an experiment is the set of all possible outcomes of an experiment. It is worth noting that the sample space Ω can be either discrete (finite or countably infinite) or continuous (uncountable). A discrete sample space could be the outcomes of flipping a coin, $\Omega_{\text{coin}} = \{$ Heads, Tails $\}$. A continuous sample space could be the changes of a stock price, $\Omega_{\text{change in stock price}} = \mathbb{R}$.

Definition 1.6. A σ-$field$ on a set Ω is a collection $\mathcal{F}$ of subsets of Ω which obeys the following properties:

(i) $\Omega \in \mathcal{F}$

(ii) if $A \in \mathcal{F}$, then $A^C \in \mathcal{F}$

(iii) If $(A_n)_{n=1}^{\infty}$ is a sequence in $\mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

To provide an intuitive explanation, a σ-$field$, $\mathcal{F}$, on a sample space, Ω, represents the events that are observed and that can be recorded when an experiment is performed. This means that if $\mathcal{F}_1$ and $\mathcal{F}_2$ are two σ-$fields$ on Ω, then $\mathcal{F}_1 \subset \mathcal{F}_2$ if and only if $\mathcal{F}_2$ contains more information than $\mathcal{F}_1$. The pair $(\Omega, \mathcal{F})$ is referred to as a measurable space.

Definition 1.7. A mapping $\mu : \mathcal{F} \to [0, +\infty]$, where $(\Omega, \mathcal{F})$ is a measurable space, is a measure if $\mu(\emptyset) = 0$ and for any pairwise disjoint sequence of sets in $\mathcal{F}$, $(A_n)_{n=1}^{\infty}$, the following property, known as countable additivity, holds:

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n). \tag{1.1}$$

The triple $(\Omega, \mathcal{F}, \mu)$ is called a measure space.

A measure can be intuitively thought of as a method to assign size to subsets of a set. If a set has measure zero, then it can be thought of to have negligible size. We can now define a probability space, independence, and conditional probability.

Definition 1.8. A probability space is a triple $(\Omega, \mathcal{F}, P)$ where Ω is the sample space, $\mathcal{F}$ is the event space (or σ-$field$) on Ω, and P, the probability measure, is a mapping from $\mathcal{F}$ into $[0, 1]$ such that

$$P(\Omega) = 1 \tag{1.2}$$

and if $(A_n)_{n=1}^{\infty}$ is any sequence of pairwise disjoint events in $\mathcal{F}$, then

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n). \tag{1.3}$$

As one might have guessed, probability is assumed a measure, where the entire space has measure 1.

We now explore independence.

Definition 1.9. If $(\Omega, \mathcal{F}, P)$ is a probability space, then $A \in \mathcal{F}$ and $B \in \mathcal{F}$ are independent events if $P(A \cap B) = P(A) \cdot P(B)$. However, we can also interpret independence by means of conditional probability. We introduce Bayes' Theorem, which characterizes conditional probability; the probability of event B happening conditioned on event A is given by

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}. \tag{1.4}$$

Definition 1.10. Let $(\Omega, \mathcal{F}, P)$ denote a probability space and let $\mathcal{F}_1$ and $\mathcal{F}_2$ denote σ-*fields* on Ω with $\mathcal{F}_1 \subset \mathcal{F}$ and $\mathcal{F}_2 \subset \mathcal{F}$. We say that $\mathcal{F}_1$ and $\mathcal{F}_2$ are independent σ-*fields* if every $A \in \mathcal{F}_1$ is independent of every $B \in \mathcal{F}_2$. If $\mathcal{F}_1$ and $\mathcal{F}_2$ are independent σ-*fields*, we write $\mathcal{F}_1 \perp \mathcal{F}_2$.

We are now interested in exploring what happens when you apply a function to a set that has a defined measure. This will allow us to confidently analyze experiments with defined probabilities and event spaces later on that will serve to be critical for the construction of our model. First, we can see that if you have a collection of subsets on your sample space, there exists a smallest σ-field containing those subsets.

Proposition 1.11. *If $\mathcal{A}$ is a collection of subsets of Ω, then there exists a σ-field on Ω containing $\mathcal{A}$ and contained in every σ-field that contains $\mathcal{A}$. This σ-field is unique and is the smallest σ-field on Ω containing $\mathcal{A}$. We call this σ-field the σ-field generated by $\mathcal{A}$ and denote it by $\mathcal{F}(\mathcal{A})$.*

Proof. Refer to page 45 of [1]. □

Proposition 1.12. *If $X : \Omega \to \Theta$ and $\mathcal{A}$ is a collection of subsets of Θ, then*

(i) If $\mathcal{A}$ is a σ-field on Θ then $X^{-1}(\mathcal{A})$ is a σ-field on Ω.

(ii) $X^{-1}(\mathcal{F}(\mathcal{A})) = \mathcal{F}(X^{-1}(\mathcal{A}))$.

Proof. Refer to pages $60 - 61$ of [1]. □

This proposition gives us an important relationship between a σ-*field* over the co-domain, Θ, and a σ-*field* over the domain, Ω. Intuitively, if we have a knowledge of outcomes/events over Θ, then we can utilize the functions pre-image to obtain knowledge of Ω. Let $\mathcal{F}$ be the σ-*field* over Θ. We denote the σ-*field* constructed over Ω by the pre-image of the function X by $\mathcal{F}_X$ and call it the σ-*field* generated by X.

Definition 1.13. The Borel field on $\mathbb{R}$, $\mathcal{B}(\mathbb{R})$, is the σ-*field* generated by the open intervals in $\mathbb{R}$. Subsets of $\mathbb{R}$ that belong to $\mathcal{B}(\mathbb{R})$ are called Borel sets.

Definition 1.14. A mapping $X : \Omega \to \mathbb{R}$, where $(\Omega, \mathcal{F})$ is a measurable space, is called $\mathcal{F}$ measurable if $X^{-1}(B) \in \mathcal{F}$ for every Borel subset $B \subset \mathbb{R}$.

In conjunction Proposition 1.9 and the notation following, we can re-write Definition 1.10 more practically. We do so in the following proposition.

Proposition 1.15. *A mapping $X : \Omega \to \mathbb{R}$ is $\mathcal{F}$ measurable if and only if $\mathcal{F}_X \subset \mathcal{F}$.*

Finally, we arrive at a method to determine whether a mapping is measurable with the following proposition.

Proposition 1.16. *If the collection $\mathcal{A}$ of Borel subsets of $\mathbb{R}$ generates the Borel field, then $X : \Omega \to \mathbb{R}$ is $\mathcal{F}$ measurable if and only if $X^{-1}(\mathcal{A}) \subset \mathcal{F}$.*

Proof. Suppose X is $\mathcal{F}$ measurable. Then

$$X^{-1}(\mathcal{A}) \subset X^{-1}(\mathcal{F}(\mathcal{A})) = X^{-1}(B(\mathbb{R})) = \mathcal{F}_X \subset \mathcal{F}.$$

Suppose $X^{-1}(\mathcal{A}) \subset \mathcal{F}$. As $\mathcal{F}$ is a $\sigma\text{-}field$, $\mathcal{F}(X^{-1}(\mathcal{A})) \subset \mathcal{F}$. By Proposition 1.9(i),

$$X^{-1}(B(\mathbb{R}) = X^{-1}(\mathcal{F}(\mathcal{A})) = \mathcal{F}(X^{-1}(\mathcal{A})) \subset \mathcal{F}.$$

So X is $\mathcal{F}$ measurable.

□

It is worth noting that measurable functions preserve many important properties/operations. For instance, measurable functions are closed under addition and multiplication. Scalars can also be applied to measurable functions and will remain measurable functions. We can now define random variables.

Definition 1.17. If $(\Omega, \mathcal{F}, P)$ is a probability space and $X : \Omega \to \mathbb{R}$ is measurable, we call X a random variable[1] on $(\Omega, \mathcal{F}, P)$.

Definition 1.18. The random variables X and Y on the probability space $(\Omega, \mathcal{F}, P)$ are independent if the $\sigma\text{-}fields$ they generate, $\mathcal{F}_X$ and $\mathcal{F}_Y$, are independent. If $\mathcal{G}$ is a $\sigma\text{-}field$ on Ω, then X and $\mathcal{G}$ are independent if the $\sigma\text{-}fields$ $\mathcal{F}_X$ and $\mathcal{G}$ are independent.

We now introduce a very important function in probability: the indicator function.

Definition 1.19. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $A \in \mathcal{F}$ be some event. Then the indicator function of A, denoted as $\mathbb{1}_A$, is defined as follows:

$$\mathbb{1}_A(w) = \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{if } w \notin A. \end{cases}$$

Before we can advance our analysis of probability theory, we must first have a good understanding of the Lebesgue integral. In order to fully understand this concept, one would need to read supplementary material. The material in this paper alone will not suffice for deep understanding. But we do aim to provide the bare minimum to reach our target goal of deriving the Black-Scholes formula.

The Lebesgue measure is a tool used to provide a "work-around" to the problem of assigning measure to any arbitrary subset of the reals. Most subsets of the reals indeed cannot be assigned a measure[2]. However, these immeasurable sets prove quite difficult to construct, typically requiring the axiom of choice. Hence, these sets are rare in practice and it is therefore possible to construct a measure that works for almost any set one would need it for. This is exactly what the Lebesgue measure accomplishes. We avoid discussing it's precise definition and construction. However, we do claim its existence.

Proposition 1.20. *There exists a unique measure $\boldsymbol{m}$ on $(\mathbb{R}, \mathcal{B})$ such that $\boldsymbol{m}([a, b]) = b - a$ for all closed intervals $[a, b] \subset \mathbb{R}$. We call $\boldsymbol{m}$ the Lebesgue measure on $\mathbb{R}$.*

[1] A random variable is actually a function. Don't let the name cause confusion.
[2] The Banach-Tarski Paradox provides a particularly famous example.

We can now introduce the Lebesgue integral.

Definition 1.21. A measurable function $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue integrable if

$$\int_{\mathbb{R}} |f| \, d\mathbf{m} = \sum_{n=-\infty}^{\infty} \int_{[n,n+1)} |f| \, d\mu([n, n+1)) < \infty.$$

If $f : [a, b] \to \mathbb{R}$ is Lebesgue integrable, we let

$$\int_{[a,b]} f \, d\mathbf{m} = \int_{\mathbb{R}} \mathbb{1}_{[a,b]} f \, d\mathbf{m}.$$

The function f is said to be a member of the $\mathcal{L}^1$ class on Lebesgue integrable functions.

We now compare the more well-known Riemann integral with the Lebesgue integral and show that they coincide. In fact, the first difference arises in the manner in which each integral executes partitioning. Lebesgue partitions the domain *and* the range of the function which proves to be more robust in limiting operations. Riemann partitions only the domain.

Proposition 1.22. *If f is a continuous function on $[a, b]$ then*

$$\int_a^b f(x) \, dx = \int_{[a,b]} f \, d\boldsymbol{m}.$$

Proof. Let $\{x_1, x_2, \ldots, x_{n+1}\}$ partition $[a, b]$ into n adjacent intervals, each with length $1/n$. Let $f_n = \Sigma_{i=1}^n f(x_i) \mathbb{1}_{[x_i, x_{i+1}]}$. Then we can state

$$\sum_{i=1}^{n} f(x_i) \cdot (x_{i+1} - x_i) \, dx = \int_{[a,b)} f_n \, d\mathbf{m} = \int_{[a,b]} f_n \, d\mathbf{m}.$$

since the Lebesgue measure of a singleton is 0. As f is continuous, it is Riemann integrable over $[a, b]$. So we can state

$$\sum_{i=1}^{n} f(x_i) \cdot (x_{i+1} - x_i) \, dx \longrightarrow \int_a^b f(x) \, dx$$

as $n \to \infty$. Now note that $(f_n)_{n=1}^{\infty}$ is bounded and converges to f as $n \to \infty$. By the Dominated Convergence Theorem[3], this sequence, in the probability space $([a, b], \mathcal{B}, \mu)$, converges to $\int_{[a,b]} f \, d\mathbf{m}$. $\square$

We now shift our attention to developing the expected values of random variables. There is a natural development that begins by considering the simplest case (and therefore imposes the greatest amount of restrictions to the random variable) and gradually progresses to a general case that encompasses almost any random variable one might come across. In order to capture the intuition and final result, we will focus on the first and last components of the progression. We will be very brief over intermediary steps. We must first define simple random variables.

[3] This theorem is a fundamental convergence result from real analysis. One can find its statement and proof on pages 137-138 of ...

Definition 1.23. A simple random variable X has canonical representation

$$X = \sum_{i=1}^{n} c_i \cdot \mathbb{1}_{A_i} \tag{1.5}$$

where $c_i \neq c_j$ for $i \neq j$, $A_i = X^{-1}(\{c_i\}) \in \mathcal{F}$ for each i and $(A_i)_{i=1}^{n}$ is a finite partition of Ω. The range of X is $\{c_i\}_{i=1}^{n}$.

Definition 1.24. The expected value, $\mathbb{E}[X]$, of the simple random variable X on the probability space $(\Omega, \mathcal{F}, P)$ with canonical representation (1.5) is given by

$$\mathbb{E}[X] = \sum_{i=1}^{n} c_i \cdot P(A_i). \tag{1.6}$$

However, in place of the canonical representation, the expected value can also be expressed in a more practical manner. Keeping all else equivalent, let $(A_i)_{i=1}^{n}$ be a pairwise disjoint finite collection of $\mathcal{F}$ measurable subsets of Ω (i.e. $(A_i)_{i=1}^{n}$ is a partition of the range of X). Then we can write

$$\mathbb{E}[X] = \sum_{i=1}^{n} c_i \cdot P(A_i).$$

Since we typically work with the range and domain of random variables, it is more natural to write the above equation in terms of our random variable. If X has range $(x_i)_{i=1}^{n}$ and $w_i \in X^{-1}(\{x_i\})$ for all i, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} X(w_i) \cdot P_X[\{X(x_i)\}]. \tag{1.7}$$

We can now elaborate on the properties that the expected value of simple random variables hold.

Proposition 1.25. *Let $(\Omega, \mathcal{F}, P)$ denote a probability space. Let X, Y be simple random variables on $(\Omega, \mathcal{F}, P)$. Let $c \in \mathbb{R}$ and $A \in \mathcal{F}$. Then*

(a) $\mathbb{E}[X \pm Y] = \mathbb{E}[X] \pm \mathbb{E}[Y]$*;*

(b) if X and Y are independent, then $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$*;*

(c) $\mathbb{E}[cX] = c\mathbb{E}[X]$*;*

(d) if $X \geq 0$, then $\mathbb{E}[X] \geq 0$*;*

(e) if $X \geq Y$, then $\mathbb{E}[X] \geq \mathbb{E}[Y]$*;*

(f) if $|X| \leq M$ on $A \in \mathcal{F}$, then $|\mathbb{E}[X \cdot \mathbb{1}_A]| \leq M \cdot P(A)$.

Proof. Refer to pages $112 - 114$ of [1]. □

We introduce a new notation of expected value:

$$\int_{\Omega} X\, dP := \mathbb{E}[X]. \tag{1.8}$$

This new notation and the $\mathbb{E}[\cdot]$ notation are both used, but typically for different purposes. The $\mathbb{E}[\cdot]$ notation is more used for conditional expectation and martingales, while the integral notation is more used when there are limiting procedures being applied (typically for finding fundamentally important results).

We now briefly give an overview of the intermediary progression when developing expectation for general functions. We first extend these properties to positive bounded random variables. A positive bounded random variable X on the probability space $(\Omega, \mathcal{F}, P)$ is one that satisfies $0 \leq X \leq m$, where $m \in \mathbb{R}$ is fixed[4]. While very technical, it can be shown that X has a canonical representation $(X_n)_{n=1}^{\infty}$ where each X_n is a simple positive random variable. Hence, it would logically follow that $\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[X_n]$. This representation allows us to extend our properties to positive bounded random variables. Using similar intuition, we can use positive bounded random variables as the approximating sequence for a positive random variable X. Finally, we can make a statement that tells us if a random variable is integrable. But since integrability is synonymous with expected value, this would also tell us if a random variable has a computable expected value.

Definition 1.26. A random variable X on a probability space $(\Omega, \mathcal{F}, P)$ is integrable if its positive and negative parts, X^+ and X^-, are both integrable, that is if both $\mathbb{E}[X^+]$ and $\mathbb{E}[X^-]$ are finite. If X is integrable we let

$$\mathbb{E}[X] := \mathbb{E}[X^+] - \mathbb{E}[X^-] = \int_{\Omega} X^+\, dP - \int_{\Omega} X^-\, dP = \int_{\Omega} X\, dP.$$

As we have previously seen, we call the Lebesgue integral of X over Ω with respect to P and let $\mathcal{L}^1(\Omega, \mathcal{F}, P)$ denote the set of all integrable random variables on $(\Omega, \mathcal{F}, P)$.

We now explore two important components of random variables: distribution and density functions. As one would imagine, the distribution function characterizes how the probabilities of a random variable are distributed on $\mathbb{R}$. Density functions provide a method to define Borel probability measures on $\mathbb{R}$. In other words, a density function defines a random variable's probability within a distinct interval of $\mathbb{R}$. We will soon see how these two functions are closely related.

Definition 1.27. The distribution function F_X of a random variable X on $(\Omega, \mathcal{F}, P)$ is defined as

$$F_X := P(\{\omega \in \Omega : X(\omega) \leq x\})$$

for all $x \in \mathbb{R}$. This function satisfies the following properties:

(i) F_X is increasing

(ii) $\lim_{x\to-\infty} F_X(x) = 0$

(iii) $\lim_{x\to+\infty} F_X(x) = 1$

(iv) $P(\{\omega \in \Omega : x - h < X(\omega) \leq x + h\}) = F_X(x + h) - F_X(x - h)$

[4]This notation means that for all $\omega \in \Omega$, we have that $0 \leq X(\omega) \leq m$.

for all $h > 0$.

If we have a sequence of random variables $(X_n)_{n=1}^{\infty}$ that converges almost surely to a random variable X as $n \to \infty$, then we say $(X_n)_{n=1}^{\infty}$ converges to X in distribution as $n \to \infty$ and write $X_n \xrightarrow{D} X$ as $n \to \infty$.

Definition 1.28. A positive-valued Borel measurable function f with domain $\mathbb{R}$ is called a density function if $\int_{\mathbb{R}} f\, d\mathbf{m} = 1$. Furthermore, a random variable X defined on the probability space $(\Omega, \mathcal{F}, P)$ has a density if there exists a density function f such that for all closed intervals $[a, b]$,

$$P[\omega \in \Omega : a \leq X(\omega) \leq b] = P([a, b]) = \int_{[a,b]} f\, d\mathbf{m}. \tag{1.9}$$

We denote the density of f by f_X.

We now introduce a well-known theorem that characterizes the relationship between distribution functions and density functions.

Theorem 1.29. *(Radon-Nikodym Theorem)Let P and Q be probability measures on the measurable space $(\Omega, \mathcal{F})$ Let $A \in \mathcal{F}$ be arbitrary. If $Q(A) = 0$ and $P(A) = 0$, then there exists a positive measurable function Y on Ω such that*

$$Q(A) = \int_A Y\, dP. \tag{1.10}$$

Moreover, any $\mathcal{F}$ measurable function on Ω satisfying (1.10) for all $A \in \mathcal{F}$ is equal to Y almost everywhere.

Let X have a continuously differentiable distribution function. Let $A = [a, b]$ in the Radon-Nikodym Theorem, then we see that (1.10) reduces to an extension of the Fundamental Theorem of Calculus:

$$F_X(b) - F_X(a) = \int_{[a,b]} F_X'(x)\, d\mathbf{m} = \int_{[a,b]} f_X(x)\, d\mathbf{m}.$$

Finally, we claim a very important property of expected values.

Proposition 1.30. *If the random variable X on $(\Omega, \mathcal{F}, P)$ has density function f_X and g is a Borel measurable function such that $g(X)f_X$ is a Riemann integrable, then $g(X)$ is an integrable random variable and*

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) f_X(x)\, d\boldsymbol{m}. \tag{1.11}$$

Proof. Refer to page 158 of [1]. □

A commonly used metric for random variables is known as the variance. It characterizes a random variables spread.

Definition 1.31. If X is a random variable on the probability space $(\Omega, \mathcal{F}, P)$ and X^2 is integrable, let

$$Var(X) = \sigma_X^2 = \mathbb{E}[(X - \mu_X)^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

We call $Var(X)$ the variance of X and let $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ denote the set of all random variables X on $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}[X^2] < \infty$.

The probability distribution relevant to the Black-Scholes model is the normal distribution.

Definition 1.32. Let X be normally distributed with expected value μ_X and variance σ_X^2. We say that X has distribution $N(\mu_X, \sigma_X^2)$. The density function is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\frac{-1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}}$$

where $x \in \mathbb{R}$. If a random variable X is $N(0,1)$ distributed, then it is the standard normal distribution. We denote the distribution function of the standard normal distribution by

$$N(x) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2}}\, dx. \tag{1.12}$$

Proposition 1.33. *If X is an $N(0, \sigma^2)$ random variable, then*

$$\mathbb{E}[e^X] = e^{\frac{1}{2}\sigma^2}. \tag{1.13}$$

Proof. As X has density $f_X(x) = (\sqrt{2\pi}\sigma)^{-1} \exp(-x^2/2\sigma^2)$, Proposition 1.29 implies

$$\begin{aligned}
\mathbb{E}[e^X] &= \int_{-\infty}^{+\infty} e^x f_X(x)\, dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^x e^{-x^2/2\sigma^2}\, dx \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{1}{2\sigma^2}(x^2 - 2x\sigma^2 + \sigma^4) + \sigma^2/2}\, dx \\
&= \frac{e^{\sigma^2/2}}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{1}{2\sigma^2}(x-\sigma^2)^2}\, dx \\
&= e^{\sigma^2/2}.
\end{aligned}$$

□

We conclude this section with the Central Limit Theorem.

Theorem 1.34. *(The Central Limit Theorem) Let $(X_n)_{n=1}^{\infty}$ be a sequence of independent identically distributed (i.i.d) random variables in $\mathcal{L}^2(\Omega, \mathcal{F}, P)$. Let $Y_n = \frac{1}{n}\Sigma_{i=1}^n X_i$ for all n. If $\mathbb{E}[X_i] = \mu$ and $Var(X_i) = \sigma^2$ for all i, then*

$$\lim_{n\to\infty} P\left[\frac{Y_n - \mu}{\sigma/\sqrt{n}} \le x\right] = \lim_{n\to\infty} P\left[\frac{\Sigma_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \le x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$

for all $x \in \mathbb{R}$.

2 Stochastic Processes and Martingales

We now explore the notion of a sequence of random variables, or in other words, a stochastic process.

Definition 2.1. A stochastic process X is a collection of random variables $(X_t)_{t\in T}$ on a probability space $(\Omega, \mathcal{F}, P)$, indexed by $T \subset \mathbb{R}$. It is useful to consider this indexing set T as time.

There is a distinction to be made between two types of stochastic processes. If the indexing set T is countable, it is a discrete stochastic process (or discrete-time process). If the indexing set T is uncountable, it is a continuous stochastic process (or continuous-time process). There are two variables that any stochastic process takes on: $t \in T$ and $\omega \in \Omega$. For each fixed $t \in T$, the mapping

$$\omega \in \Omega \mapsto X_t(\omega) \in \mathbb{R}$$

is a random variable. For each fixed $\omega \in \Omega$, the mapping

$$t \in T \mapsto X_t(w) \in \mathbb{R}$$

is referred to as a sample path and is not a random variable.

Definition 2.2. Let $(\Omega, \mathcal{F})$ be a measurable space.

(a) The discrete-time filtration on $(\Omega, \mathcal{F})$ is an increasing sequence of σ-$fields$ $(\mathcal{F})_{n=1}^{\infty}$ such that

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_i \subset \ldots \subset \mathcal{F}.$$

(b) A continuous-time filtration on $(\Omega, \mathcal{F})$ is a set of σ-$fields$ $(\mathcal{F}_t)_{t \in I}$, where I is an interval in $\mathbb{R}$, such that for all $t, s \in I$, $t < s$, we have

$$\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}.$$

We call $\mathcal{F}_n$ the history up to time n.

We now establish an important relationship between stochastic processes and filtrations.

Definition 2.3. If $X = (X_t)_{t \in T}$ is a stochastic process on $(\Omega, \mathcal{F}, P)$ and $(\mathcal{F}_t)_{t \in T}$ is a filtration on $(\Omega, \mathcal{F}, P)$, then X is adapted to the filtration if X_t is $\mathcal{F}_t$ measurable for all $t \in T$.

The preceding allows us to have a formal way of describing the history of a stochastic process using filtrations. The filtration that aligns with the history of the stochastic process is known as the natural filtration.

We now delve into the building block for martingales: conditional expectation.

A crucial aspect of conditional expectation is that it is itself a random variable. To drive home this fact, we consider a simple example. Let X denote the number of heads that appear when a fair coin is tossed three times in succession. Then $\mathbb{E}[X] = 1(1/2) + 1(1/2) + 1(1/2) = 1.5$. However, suppose we know information on the first toss. If the first toss is a head, then $\mathbb{E}[X|\text{First toss is head}] = 1(1) + 1(1/2) + 1(1/2) = 2$. If the first toss is a tail, then then $\mathbb{E}[X|\text{First toss is tail}] = 1(0) + 1(1/2) + 1(1/2) = 1$. The conditional expectation changes depending on the information that becomes available. We first define conditional expectation over a finite sample space, then extend the result with respect to a σ-$field$ generated by a countable partition, and finally characterize conditional expectation arbitrarily.

Definition 2.4. Let X be a random variable on $(\Omega, \mathcal{F}, P)$ with Ω being finite, $\mathcal{F} = 2^{\Omega}$, $P(\{\omega\}) > 0$ for all $\omega \in \Omega$. Let $A \in \mathcal{F}$, $0 < P(A) < 1$, then

$$\mathbb{E}[X|A] = \sum_{\omega \in \Omega} X(\omega) P([\{\omega\}|A)$$

denotes the expectation of X given A has occurred. Applying Bayes' Theorem, we can re-write this to the following:

$$\mathbb{E}[X|A] = \sum_{\omega\in\Omega} X(\omega)\frac{P(\{\omega\})}{P(A)} = \frac{1}{P(A)}\int_A X\,dP.$$

In order to advance our development, we must consider taking the expected value of a random variable conditioned on not just a singular event, but all the information of the random variable (i.e. the σ-*field* generated by the random variable). This will allow us to perform a more powerful analysis of conditional expectation.

Definition 2.5. Let $(\Omega, \mathcal{F}, P)$ denote a probability space and let $\mathcal{G}$ denote a σ-*field* generated by a countable partition $(G_n)_{n=1}^{\infty}$ of Ω. Suppose $\mathcal{G} \subset \mathcal{F}$ and $P(G_n) > 0$ for all n. If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$, let

$$\mathbb{E}[X|\mathcal{G}](\omega) = \frac{1}{P(G_n)}\int_{G_n} X\,dP \tag{2.1}$$

for all n and all $\omega \in G_n$. We call $\mathbb{E}[X|\mathcal{G}]$ the conditional expectation of X given $\mathcal{G}$. We also write $\mathbb{E}[X|Y]$ if $\mathcal{G}$ is generated by a random variable Y.

We can now begin to characterize conditional expectation more arbitrarily. We begin with the following proposition.

Proposition 2.6. *Let $(\Omega, \mathcal{F}, P)$ denote a probability space and let $\mathcal{G}$ denote a σ-field on Ω generated by a countable partition $(G_n)_{n=1}^{\infty}$ of Ω. We suppose $\mathcal{G} \subset \mathcal{F}$ and $P(G_n) > 0$ for all n. If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$, then $\mathbb{E}[X|\mathcal{G}]$ is the unique $\mathcal{G}$ measurable integrable random variable on $(\Omega, \mathcal{F}, P)$ satisfying*

$$\int_A \mathbb{E}[X|\mathcal{G}]\,dP = \int_A X\,dP \tag{2.2}$$

for all $A \in \mathcal{G}$.

Proof. Refer to page 173 of [1]. □

Finally, we use the preceding result to extend conditional expectation to arbitrary cases.

Definition 2.7. If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$ and $\mathcal{G}$ is a σ-*field* on Ω such that $\mathcal{G} \subset \mathcal{F}$, then there exists a $\mathcal{G}$ measurable integrable random variable on $(\Omega, \mathcal{F}, P)$, denoted by $\mathbb{E}[X|\mathcal{G}]$, such that

$$\int_A \mathbb{E}[X|\mathcal{G}]\,dP = \int_A X\,dP$$

for all $A \in \mathcal{G}$. Furthermore, if Y is any $\mathcal{G}$ measurable integrable random variable satisfying

$$\int_A Y\,dP = \int_A X\,dP$$

for all $A \in \mathcal{G}$, then $Y = \mathbb{E}[X|\mathcal{G}]$ almost surely in $(\Omega, \mathcal{F}, P)$. If Y is a random variable , we let $\mathbb{E}[X|Y] = \mathbb{E}[X|\mathcal{F}_Y]$. We call $\mathbb{E}[X|\mathcal{G}]$ and $\mathbb{E}[X|Y]$ the conditional expectation of X given $\mathcal{G}$ and Y, respectively.

Certain conditional expectation properties are critical for the study of martingales.

Proposition 2.8. *Let X and Y denote integrable random variables on the probability space $(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}$ and $\mathcal{H}$ denote σ-fields on Ω where $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$. The following hold:*

(a) ***Taking out what is known.*** *If $X \cdot Y$ is integrable and X is $\mathcal{G}$ measurable, then*

$$\mathbb{E}[X \cdot Y|\mathcal{G}] = X \cdot \mathbb{E}[Y|\mathcal{G}].$$

(b) ***Independence drops out.*** *If X and $\mathcal{G}$ are independent, that is if $\mathcal{F}_X \perp \mathcal{G}$, then*

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X].$$

(c) ***Tower Law.***

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}|\mathcal{H}]] = \mathbb{E}[X|\mathcal{H}].$$

Proof. Refer to pages $177 - 178$ of [1]. □

We now have the building blocks necessary to define martingales and conclude this section.

Definition 2.9. Let $(\mathcal{F}_n)_{n=1}^{\infty}$ be a filtration on the probability space $(\Omega, \mathcal{F}, P)$. A discrete-time martingale on $(\Omega, \mathcal{F}, P, (\mathcal{F}_n)_{n=1}^{\infty})$[5] is a sequence $(X_n)_{n=1}^{\infty}$ of integrable random variables on $(\Omega, \mathcal{F}, P)$, adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$ (i.e. X_n is $\mathcal{F}_n$ measurable for all n) such that

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n \tag{2.3}$$

for all $n \geq 1$.

We define a continuous-time martingale synonymous to how we define the discrete-time martingale. Let $(\mathcal{F}_t)_{t\in I}$ be a filtration on the probability space $(\Omega, \mathcal{F}, P)$, indexed by an interval I of real numbers. A continuous-time martingale on $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in I})$ is a sequence $(X_t)_{t\in I}$ of integrable random variables on $(\Omega, \mathcal{F}, P)$, adapted to the filtration $(\mathcal{F}_t)_{t\in I}$ (i.e. X_t is $\mathcal{F}_t$ measurable for all t) such that

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \tag{2.4}$$

for all $s, t \in I, s \leq t$. Equation (2.3) and (2.4) are called the martingale property.

3 Brownian Motion

In order to define Brownian motion, it would be insightful to first understand symmetric random walks.

Definition 3.1. Let $(X_n)_{n=1}^{\infty}$ denote a sequence of independent random variables on $(\Omega, \mathcal{F}, P)$ where $\Omega = \{-1, +1\}$, $\mathcal{F} = 2^{\Omega}, P(\{1\}) = P(\{-1\} = 1/2$ and X_n takes on the values ± 1 with probability $1/2$. It is clear that $\mathbb{E}[X_n] = 0$ and $Var(X_n) = 1$. A symmetric random walk is the sequence $Y_n = \Sigma_{i=1}^{n} X_i$ for all n.

Proposition 3.2. *A symmetric random walk is a martingale.*

[5]This quadruple is known as the filtered probability space.

Proof. Let $(X_n)_{n=1}^{\infty}$ denote a symmetric random walk on $(\Omega, \mathcal{F}, P)$. Let $Y_n = \Sigma_{i=1}^{n} X_i$ for all n. Since $|Y_n| \leq \Sigma_{i=1}^{n} |X_i|$ and each X_i is integrable, we deduce that Y_n is integrable. Note that Y_n is $\mathcal{F}_n$ measurable since Y_n is a sum of $\mathcal{F}_n$ measurable functions. If $i \leq n$, the random variable X_{n+1} is independent of X_i, and hence, also from $\mathcal{F}_n$. As independence drops out, we can write $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_{n+1}] = 0$. By taking out what is known, we may write

$$\begin{aligned}\mathbb{E}[Y_{n+1}|\mathcal{F}_n] &= \mathbb{E}[X_{n+1} + Y_n|\mathcal{F}_n] \\ &= \mathbb{E}[X_{n+1}|\mathcal{F}_n] + \mathbb{E}[Y_n|\mathcal{F}_n] \\ &= Y_n.\end{aligned}$$

Hence, $(Y_n)_{n=1}^{\infty}$ is a martingale. □

We now investigate Brownian motion, also known as the Wiener Process. To provide some intuition, Brownian motion can be thought of as a normalized symmetric random walk with very small time increments. However, given enough time, a random walk can cover the entire real line. To avoid this, certain properties must be put in place.

Definition 3.3. Brownian motion, $(W_t)_{t\geq 0}$, is a continuous stochastic process that has the following properties:

(a) $W_0 = 0$ almost surely;

(b) W_t is $N(0, t)$ distributed for all $t \geq 0$ (Gaussian increments);

(c) for any n and any $\{0 = t_0 < t_1 < \ldots < t_{n+1}\}$, $(W_{t_i} - W_{t_{i-1}})_{i=1}^{n}$ is a set of independent random variables (independent increments);

(d) the probability distribution of $W_t - W_s$ depends only on $t - s$ for $0 \leq s \leq t$ (stationary increments).

The process $(\mu t + \sigma W_t)_{t\geq 0}$ is Brownian motion with drift μ and volatility σ. The process $(\exp(\mu t + \sigma W_t))_{t\geq 0}$ is geometric Brownian motion. Also, an important note on Brownian motion is that it is typically used to model a phenomenon that is influenced by a large number of independent forces that individually are negligible. Phenomenon that this aligns with is the diffusion of gas, dispersion of smoke, and of course, stock prices.

We now establish important properties of Brownian motion that serve to be critical for the proof of the Black-Scholes Formula.

Proposition 3.4. *Let $(W_t)_{t\geq 0}$ be a Wiener process on $(\Omega, \mathcal{F}, P)$, and for $t \geq 0$ let $\mathcal{F}_t$ denote the σ-field generated by $(W_s)_{0\leq s\leq t}$. The following hold:*

(a) $(W_t)_{t\geq 0}$ and $(W_t^2 - t)_{t\geq 0}$ are martingales;

(b) If μ and σ are real numbers, then $(e^{\mu t + \sigma W_t})_{t\geq 0}$ is a martingale if and only if $\mu = -\sigma^2/2$;

(c) If $\gamma > 0$, $X_t := e^{-\frac{1}{2}\gamma^2 t \gamma W_t}$ for $t \geq 0$ and $P_\gamma(A) := \int_A X_\gamma \, dP$ when $A \in \mathcal{F}$, then $(\Omega, \mathcal{F}, P_\gamma$ is a probability space. If $\mathcal{G}$ is a σ-field on Ω with $\mathcal{G} \subset \mathcal{F}$ and Y is an integrable random variable on $(\Omega, \mathcal{F}, P_\gamma)$, then $\mathbb{E}_{P_\gamma}[Y|\mathcal{G}] = \mathbb{E}_P[X_\gamma \cdot Y|\mathcal{G}]$. If $0 \leq s \leq t \leq \gamma$ and Y is $\mathcal{F}_t$ measurable, then

$$\mathbb{E}_{P_\gamma}[Y|\mathcal{F}_s] = \mathbb{E}_P[X_t \cdot Y|\mathcal{F}_s]. \tag{3.1}$$

Proof. Refer to pages 222 – 224 of [1]. □

We now have all the components to prove the Black-Scholes formula. It will require some set-up making use of many of the ideas we have already discussed. In addition, necessary finance terms will be defined. The following section accomplishes all of this.

4 Black-Scholes Formula

Definition 4.1. A call option is an option to buy a certain asset on or before a certain date, the maturity date, for a certain price. If the call option is for a fixed quantity of shares, then the price per share at maturity, if the option is exercised, is called the strike price. If the option can only be exercised at the maturity date, it is called a European option. If the option can be exercised at any time prior to the maturity date, it is called an American option. The most common use-case for options is to transfer risk (i.e. either to hedge and reduce exposure to risk or to speculate and accept exposure to risk). The Black-Scholes formula only considers European options.

Definition 4.2. Arbitrage is used to describe any situation, opportunity, or price which allows for a guaranteed profit without risk. The market typically capitalizes on this and is forced back to a state of equilibrium. Therefore, there is minimal utility in finding arbitrage option prices. Hence, we are interested in finding an arbitrage-free price for options and is an assumption for the Black-Scholes formula.

Definition 4.3. A fair game does not have a precise definition. It typically entails considering each participants investment in the game being proportionally related to their probability of winning or losing. If we think about a stock in this manner, we can develop the notion of the fair game principle. Note that risk neutral probabilities are probabilities of future outcomes that take into account risk. The fair game principle mandates that the risk neutral probabilities be found by assuming that the expected return on investing directly in shares would, when discounted back to present, equal the initial investment.

We now begin the setup of the Black-Scholes formula proof. Consider a company with stock price of X_t at time t where $t = 0$ is the present and $t > 0$ is the future. Then we can express the price of the stock as $X_t = X_0 \cdot e^{A_t}$ where A_t represents the forces that drive price changes of the stock. A natural question to ask is what forces cause the price to change. The Black-Scholes formula assumes that there are two forces that govern the stock price: the drift of a company (i.e. the internal growth rate of the company) and the force many of independent investors. We let μ represent the drift of a company. We can also let Brownian motion, σW_t, represent the force of the independent investors with some volatility σ. Combining these components, we yield the following proposition.

Proposition 4.4. *Let $(X_t)_{t\geq 0}$ be a collection of random variables giving the share price of a stock at different times, where X_t denotes the random variable at time t. By the assumptions above, there is a constant μ known as the drift and constant σ known as the volatility of the stock such that*

$$X_t = X_0 \cdot \exp(\mu t + \sigma W_t)$$

where $\exp(\mu t + \sigma W_t)$ is geometric Brownian motion. If interest is a constant rate r, then the discounted share price is

$$e^{-rT} X_t = X_0 \exp((\mu - r)t + \sigma W_t). \tag{4.1}$$

We now have a sufficient model for the share price. But we must now create an environment where we are able to perform analysis and operations. In other words, we must have a filtered probability space. We define this now.

Proposition 4.5. *If a stock has drift μ and volatility σ, then there exists a measurable space $(\Omega, \mathcal{F})$, a filtration $(\mathcal{F}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F})$, and two probability measures on $(\Omega, \mathcal{F})$: W, the Wiener process measure, and, P_N, the risk neutral probability measure, so that the share price $(X_t)_{t\geq 0}$ is a stochastic process on $(\Omega, \mathcal{F})$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and also has the following properties:*

(a) under W, $X_t = X_0 \exp(\mu t + \sigma W_t)$ where $(W_t)_{t\geq 0}$ is a Wiener process;

(b) under P_N, $e^{-rt}X_t = X_0 \exp(-\frac{\sigma^2}{2}t + \sigma \tilde{W}_t)$ where $(\tilde{W}_t)_{t\geq 0}$ is a Wiener process.

We now show that the discounted share price is indeed a martingale

Proposition 4.6. *The discounted share price $(e^{-rt}X_t)_{t\geq 0}$ is a martingale on $(\Omega, \mathcal{F}, P_N, (\mathcal{F}_t)_{t\geq 0})$.*

Proof. Suffices to combine Proposition 3.4(b) and Proposition 4.5(b). □

The Black-Scholes formula makes more assumptions that have yet to be acknowledged. We provide the remaining assumptions here:

(i) If the seller or buyer need capital, any amount of the risk-free asset may be borrowed or lent;

(ii) Any amount of the stock may be bought or sold (including fractional shares);

(iii) There are no transactional costs

(iv) Transactions occur immediately;

(v) The risk-free rate is non-random;

(vi) The underlying stock does not pay dividends between now and the expiration date.

We can finally present and prove the Black-Scholes formula.

Theorem 4.7. ***Black-Scholes Formula.*** *Suppose the share price of a stock with volatility σ is X_0 today. For the buyer*

$$X_0 N\left(\frac{\log(\frac{X_0}{k}) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) - ke^{-rT} N\left(\frac{\log(\frac{X_0}{k}) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \tag{4.2}$$

is a fair price for a call option with maturity date T and strike price k given that r is the risk-free interest rate.

Proof. Note that $X_T - k$ is the difference between the stock price at maturity date and the strike price. If this becomes negative, then the option is worthless. So we are only interested when $X_T - k$ is positive. If we discount this difference for one sample path, it will give us the current value of the option under the singular sample path. But we need to find its value in general. We accomplish this by finding the expected value of the discounted difference. Since the Black-Scholes formula assumes the call option is a European option, the contract can only be executed at the maturity date. Hence, the only time (besides maturity date) that one must consider is at time $t = 0$. Let

V_0 denote the buyers fair price for the option. Putting together the commentary up to this point leads us to the following:

$$V_0 = \mathbb{E}_{P_N}[e^{-rT}(X_T - k)^+|\mathcal{F}_0] = \mathbb{E}_{P_N}[e^{-rT}(X_T - k)^+]$$

It suffices to show that V_0 reduces to equationz (4.2). By Proposition 4.5 and how $\mu - r = -\sigma^2/2$ in order for the discounted share price to be a martingale, we can write

$$e^{-rT}(X_T - k)^+ = e^{-rT}(X_0 \cdot e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Y} - k)^+$$

where Y is an $N(0,1)$ distributed random variable. By Proposition 1.30 and 1.33, we obtain

$$V_0 = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-rT}(X_0 \cdot e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Y} - k)^+ e^{-\frac{1}{2}x^2}\, dx.$$

We now seek to calculate $(X_0 \cdot e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Y} - k)^+$. We write

$$X_0 \cdot e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Y} - k \geq 0 \quad \longleftrightarrow \quad e^{\sigma\sqrt{T}x} \geq \left(\frac{k}{X_0} e^{-(r-\frac{1}{2}\sigma^2)T}\right)$$

Going further,

$$e^{\sigma\sqrt{T}x} \geq \left(\frac{k}{X_0} e^{-(r-\frac{1}{2}\sigma^2)T}\right) \quad \longleftrightarrow \quad x \geq \frac{1}{\sigma\sqrt{T}}\{\log(\frac{k}{X_0}) - (r - \frac{1}{2}\sigma^2)T\} =: T_1.$$

By substituting $y = x - \sigma\sqrt{T}$, we yield

$$\begin{aligned}
V_0 &= \frac{X_0 e^{-\frac{1}{2}\sigma^2 T}}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{\sigma\sqrt{T}x - \frac{1}{2}x^2}\, dx - \frac{ke^{-rT}}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{\frac{1}{2}x^2}\, dx \\
&= \frac{X_0}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{-\frac{1}{2}(x-\sigma\sqrt{T})^2}\, dx - ke^{-rT}(1 - N(T_1)) \\
&= \frac{X_0}{\sqrt{2\pi}} \int_{T_1-\sigma\sqrt{T}}^{\infty} e^{-\frac{1}{2}y^2}\, dy - ke^{-rT}N(-T_1) \\
&= X_0(1 - N(T_1 - \sigma\sqrt{T})) - ke^{-rT}N(-T_1) \\
&= X_0 N(\sigma\sqrt{T} - T_1) - ke^{-rT}N(-T_1).
\end{aligned}$$

By plugging in our value for T_1, we yield

$$\sigma\sqrt{T} - T_1 = \frac{1}{\sigma\sqrt{T}}(\log(\frac{X_0}{k}) + (r + \frac{1}{2}\sigma^2)T).$$

Hence,

$$N(\sigma\sqrt{T} - T_1) = N\left(\frac{\log(\frac{X_0}{k}) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right)$$

and

$$N(-T_1) = N\left(\frac{\log(\frac{X_0}{k}) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right).$$

Substituting these two formulae with the integral representation for V_0, we obtain the desired.

$\square$

5 Conclusion

While powerful, the Black-Scholes model makes many assumptions that do not align with financial markets in practice. However, it is still a useful metric that has been used since its founding. Extensions to this model include using stochastic integration and *Itô's* formula to hedge the call option

6 Acknowledgements

I would like to thank Erin Bevilacqua for supporting me through this summer in learning this content. I would also like to thank Lewis Bowen, John Luecke, Will Stagner and Amanda Wilkens for organizing the Directed Reading Program at UT Austin for Summer 2023. I would also like to thank Sean Dineen for his writing in *Probability Theory in Finance: A Mathematical Guide to the Black-Scholes Formula.* His work is splendid and made this paper possible.

References

[1] Dineen, S. (2013). *Probability Theory in Finance: A Mathematical Guide to the Black-Scholes Formula.* American Mathematical Society.