PRICING THE ASIAN CALL OPTION

VINH XUAN DANG,
SCOTT GLASGOW,
HARRISON POTTER,
STEPHEN TAYLOR

ABSTRACT. Background material on measure-theoretic probability theory and stochastic calculus is provided in order to clarify notation and inform the reader unfamiliar with these concepts. These fields are then employed in exploring two distinct but related approaches to fair option pricing: developing a partial differential equation whose solution, given specified boundary conditions, is the desired fair option price and evaluating a risk-neutral conditional expectation whose value is the fair option price. Both approaches are illustrated by example before being applied to the Asian call option.

Two results are obtained by applying the latter option pricing approach to the Asian call option. The price of an Asian call option is shown to be equal to an integral of an unknown joint distribution function. This exact formula is then made approximate by allowing one of the random variables to become a parameter of the system. This modified Asian call option is then priced explicitly, leading to a formula that is strikingly similar to the Black-Scholes-Merton formula, which prices the European call option. Finally, possible methods of generalizing the procedure to price the Asian call option both exactly and explicitly are speculated.

1. INTRODUCTION: FINANCIAL MOTIVATION

An understanding of the financial issues that give the mathematics problems presented here meaning is critical to following the overall reasoning. With this end in mind, a brief overview of relevant financial background material is provided.

There are essentially three forms that wealth can take in financial models: money in the money market, shares of an asset in the stock market, and stakes in an option. Any wealth in the money market will grow in accordance with a given interest rate. Wealth invested in the money market is not expected to have a particularly impressive growth rate, but it is considered to be a reliable investment as it will generally steadily increase in value. On the other hand, investing directly in an asset is an inherently risky endeavor as its value will fluctuate both up and down in a random manner; however, the potential for loss is counterbalanced by the potential for greater gain. This is what attracts casual investors with dreams of instantly becoming wealthy to bet on the stock market; however, this is not the way that most major financial institutions invest in assets.

From the perspective of a casual investor, purchasing an option from a financial institution is risky in much the same way that investing in an asset is risky: the final value of the option is dependent upon the asset's values over the duration of the contract, and there is thus again the potential for loss. It simply changes the way in which the risk is managed. But this, in fact, is the real value of an option: by enabling the reallocation of risk, options serve a purpose that is the financial world's analog of insurance. A business, for example, might buy an option that will increase in value if its competitors do well, but will decrease

Date: February 27, 2008.

in value if it does well. The risk is thereby transferred to the financial institution, as it must now pay a large return if the business's competitors do well, whereas the business will now be fine financially in either scenario. In order to be able to pay this return, the financial institution will need to charge an initial price for the option. This initial price will depend upon the nature of the option contract.

When a financial institution sells an option it must determine how it will invest the initial wealth gained from the sale in order to be able to settle its contract when the option is used. This is no trivial matter, and is precisely why the financial institution takes this task upon itself. Although it is theoretically possible for the business to replicate the option by investing in the stock and money markets, it is a practical impossibility because the business does not have the resources to invest so intelligently. The business, in essence, must get the financial institution to do this task on its behalf.

Ideally, the financial institution can perfectly hedge its position in the underlying asset by investing the wealth obtained from the sale of the option in both the underlying asset and the money market. A perfect hedge will result in the replication of the option: the financial institution will get precisely as much money from its investments in the underlying asset and the money market as the option-holder will get by invoking the option. Thus the financial institution can exactly pay off the option's value to the business, with no money left over. If for any given option it is possible for a financial institution to construct such a perfect hedge in an economy by determining a fair initial price for the option and then investing wisely, that economy is said to be complete. A fair initial price is one for which the financial institution does not profit by providing the option the business needs, if the option is optimally executed by the option-holder, but that also allows the financial institution to avoid losses, regardless of the specific changes in stock price over the duration of the contract. Although actual financial institutions will charge a small premium for providing the businesses with this service, the actual commercial price is based off of the option's theoretical fair price, and thus determining this fair price is still of great importance in actual practice.

The mathematical problem is to calculate the fair price and perfect hedging portfolio for any given option. This is done by using the techniques of measure-theoretic probability theory and stochastic calculus, as presented below.

2. Background Material

There are two main approaches to pricing options. For any given option, there exists a partial differential equation governing the option value. When certain boundary conditions are taken into consideration, this partial differential equation has a single solution that is the value of the option at any given time. An option can thus be priced by solving its partial differential equation. An alternative method for pricing a given option is to write the option value as a risk-neutral conditional expectation. If this expectation can be evaluated explicitly, then the option's value at any given time has been determined. The fair option price is simply the value of the option at the time of the sale.

In order to apply these two option pricing approaches to the asian call option, however, there are several mathematical techniques that must be presented. These will be developed in several steps. We begin by introducing σ-algebras and several closely related concepts, with those of measurability and independence being of particular importance, in order to provide a rigorous foundation for the necessary measure-theoretic probability theory. Lebesgue integration will then be used to extend this foundation to distributions, expectations, and conditional expectations. These tools will be utilized to define Brownian motion and state several of its important properties. A summary of Itó Calculus will then be given to justify later calculations involving Brownian motions. Finally, material on how to change the probability measure with respect to which calculations are being made will be presented so that the risk-neutral probability measure that is so critically important to the financial applications of stochastic calculus can be seen as a specific instance of a more general procedure. All of this material is presented below and is based on [3], where a more thorough development is available to the interested reader.

2.1. **Sigma Algebras, Measurability, and Independence.** Intuitively, a σ-algebra is just a way of writing all of the information known at a certain time as a set (more precisely, as a set of subsets of the power set of the sample space). Although this may sound like a peculiar and confusing concept from which to build up measure-theoretic probability theory, it is actually incredibly valuable because it enables the formalization of the intuitive concept of information. This is of particular importance when dealing with conditional expectations, which will be developed in a later subsection. The critical point here is that, on an intuitive level, the word "σ-algebra" can everywhere be replaced by the word "information". This rule of thumb lies at the heart of a deeper understanding of the mathematics of σ-algebras.

Definition 2.1. Let Ω be a nonempty set, and let $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is a *σ-algebra* provided that:

(i.) the empty set $\emptyset$ belongs to $\mathcal{F}$,
(ii.) whenever a set A belongs to $\mathcal{F}$, its complement in Ω , denoted by A^c, also belongs to $\mathcal{F}$, and
(iii.) whenever a sequence of sets $A_1, A_2, \ldots$ belongs to $\mathcal{F}$, their union $\bigcup_{n=1}^{\infty} A_n$ also belongs to $\mathcal{F}$.

Definition 2.2. Let Ω be a nonempty set, and let $\mathcal{F}$ be a σ-algebra on Ω. A *probability measure* $\mathbb{P}$ is a function that assigns to every set $A \in \mathcal{F}$ a number in $[0, 1]$, which is called the *probability of A under* $\mathbb{P}$ and is written $\mathbb{P}(A)$, and that satisfies the following properties:

(i.) $\mathbb{P}(\Omega) = 1$, and
(ii.) whenever $A_1, A_2, \ldots$ is a sequence of disjoint sets in $\mathcal{F}$, then *countable additivity* holds:

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space*.

Definition 2.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If a set $A \in \mathcal{F}$ satisfies $\mathbb{P}(A) = 1$, we say that A occurs *almost surely*.

Definition 2.4. Let S be a subset of the real numbers. The *Borel σ-algebra generated by* S, denoted $\mathfrak{B}(S)$, is the collection of all the closed intervals $[a, b]$ that are in S, along with all other subsets of S that must be included in $\mathfrak{B}(S)$ in order to make $\mathfrak{B}(S)$ a σ-algebra. The sets in $\mathfrak{B}(S)$ are called the *Borel subsets of S*.

Definition 2.5. Let $f(x)$ be a real-valued function defined on $\mathbb{R}$. If for every Borel subset B of $\mathbb{R}$ the set $\{x; f(x) \in B\}$ is also a Borel subset of $\mathbb{R}$, then f is called a *Borel-measurable function*.

Definition 2.6. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A *random variable* is a real-valued function X defined on Ω with the property that for every Borel subset B of $\mathbb{R}$, the subset of Ω given by

$$\{X \in B\} := \{\omega \in \Omega; X(\omega) \in B\}$$

is in the σ-algebra $\mathcal{F}$.

Definition 2.7. Let Ω be a nonempty set. Let T be a fixed positive number, and assume that for each $t \in [0, T]$ there is a σ-algebra $\mathcal{F}(t)$. Assume further that if $s \leq t$, then every set in $\mathcal{F}(s)$ is also in $\mathcal{F}(t)$. Then we call the collection of σ-algebras $\mathcal{F}(t)$, $0 \leq t \leq T$, a *filtration*.

Definition 2.8. Let X be a random variable defined on a nonempty sample space Ω. The *σ-algebra generated by X*, denoted $\sigma(X)$, is the collection of all subsets of the form $\{\omega \in \Omega; X(\omega) \in B\}$, where B ranges over all of the Borel subsets of $\mathbb{R}$.

Definition 2.9. Let X be a random variable defined on a nonempty sample space Ω. Let $\mathcal{G}$ be a σ-algebra of subsets of Ω. If every set in $\sigma(X)$ is also in $\mathcal{G}$, we say that X *is $\mathcal{G}$-measurable*.

Definition 2.10. Let Ω be a nonempty sample space equipped with a filtration $\mathcal{F}(t)$, $0 \leq t \leq T$. Let $X(t)$ be a collection of random variables indexed by $t \in [0, T]$. We say this collection of random variables is an *adapted stochastic process* if, for each t, the random variable $X(t)$ is $\mathcal{F}(t)$-measurable.

Definition 2.11. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$. We say that $\mathcal{G}$ and $\mathcal{H}$ are *independent σ-algebras* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \quad \text{for all } A \in \mathcal{G}, B \in \mathcal{H}.$$

Let X and Y be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. We say that the *random variables X and Y are independent* if the σ-algebras they generate, $\sigma(X)$ and $\sigma(Y)$, are independent. We say that a *random variable X is independent of the σ-algebra $\mathcal{G}$* if $\sigma(X)$ and $\mathcal{G}$ are independent.

Definition 2.12. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \ldots$ be a sequence of sub-σ-algebras of $\mathcal{F}$. For a fixed positive integer n, we say that the *n σ-algebras $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ are independent* if

$$\begin{aligned} \mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) &= \mathbb{P}(A_1)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n) \\ &\quad \text{for all } A_1 \in \mathcal{G}_1, A_2 \in \mathcal{G}_2, \ldots, A_n \in \mathcal{G}_n. \end{aligned}$$

We say that the *full sequence of σ-algebras $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \ldots$ is independent* if, for every positive integer n, the n σ-algebras $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ are independent.

Let $X_1, X_2, X_3, \ldots$ be a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. We say that the *$n$ random variables $X_1, X_2, \ldots, X_n$ are independent* if the σ-algebras $\sigma(X_1), \sigma(X_2), \ldots, \sigma(X_n)$ are independent. We say that the *full sequence of random variables $X_1, X_2, X_3, \ldots$ is independent* if, for every positive integer n, the n random variables $\sigma(X_1), \sigma(X_2), \ldots, \sigma(X_n)$ are independent.

Theorem 2.13. *Let X and Y be independent random variables, and let f and g be Borel-measurable functions on $\mathbb{R}$. Then $f(X)$ and $g(Y)$ are independent random variables.*

2.2. **Distributions, Expectations, and Beyond.** There are several quantities associated with any random variable that are of great interest, such as the mean, standard deviation, and distribution of the random variable. In this section we develop several ways of expressing and calculating information encoded in a random variable.

Definition 2.14. Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *distribution measure of X* is the probability measure μ_X that assigns to each Borel subset of $\mathbb{R}$ the mass $\mu_X(B) = \mathbb{P}\{X \in B\}$.

Definition 2.15. The *Lebesgue measure on* $\mathbb{R}$, denoted by $\mathcal{L}$, assigns to each $B \in \mathfrak{B}(\mathbb{R})$ a number in $[0, \infty)$ or the value ∞ such that the following conditions are satisfied:

(i.) $\mathcal{L}[a, b] = b - a$ whenever $a \leq b$, and

(ii.) if $B_1, B_2, \ldots$ is a sequence of disjoint sets in $\mathfrak{B}(\mathbb{R})$, then countable additivity holds:

$$\mathcal{L}\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mathcal{L}(B_n).$$

Definition 2.16. For any set A, the *indicator function of* A is

$$\mathbb{I}_A(\alpha) = \begin{cases} 1, & \alpha \in A \\ 0, & \alpha \notin A \end{cases}.$$

Definition 2.17. Let $\Pi = \{y_0, y_1, \ldots\}$ be a partition of the y-axis of the Cartesian plane where $0 = y_0 < y_1 < \ldots$, and let

$$\|\Pi\| = \max_{1 \leq k}(y_k - y_{k-1}).$$

Given any nonnegative real-valued function $g(x)$, let $A_k = \{x \in \mathbb{R}; y_k \leq g(x) < y_{k+1}\}$, and define the *lower Lebesgue sum* to be

$$LS_{\Pi}^{-}(g(x)) := \sum_{k=1}^{\infty} y_k \mathcal{L}(A_k).$$

The *Lebesgue integral of any nonnegative function* $g(x)$ *over* $\mathbb{R}$ is defined by

$$\int_{\mathbb{R}} g(x) d\mathcal{L}(x) := \lim_{\|\Pi\| \to 0} LS_{\Pi}^{-}(g(x)).$$

The *Lebesgue integral of any real-valued function* $f(x)$ *over* $\mathbb{R}$ is defined in terms of the Lebesgue integrals of the nonnegative functions $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$ as

$$\int_{\mathbb{R}} f(x) d\mathcal{L}(x) := \int_{\mathbb{R}} f^+(x) d\mathcal{L}(x) - \int_{\mathbb{R}} f^-(x) d\mathcal{L}(x),$$

assuming that both of the integrals on the right are finite. The *Lebesgue integral over any Borel subset* B *of* $\mathbb{R}$ *of any real-valued function* $f(x)$ is

$$\int_B f(x) d\mathcal{L}(x) := \int_{\mathbb{R}} \mathbb{I}_B(x) f(x) d\mathcal{L}(x)$$

The Lebesgue measure $\mathcal{L}$ provides a way of quantifying the size of subsets of the real numbers $\mathbb{R}$ just as a probability measure $\mathbb{P}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ provides a way of quantifying the size of subsets of the sample space Ω. This correspondence enables us to define an integral with respect to a probability space as a Lebesgue integral in which the Lebesgue measure has been replaced by the desired probability measure and the integration is over subsets of the sample space rather than over subsets of the real numbers.

Definition 2.18. A random variable X is *integrable* if

$$\int_{\Omega} X^+(\omega) d\mathbb{P}(\omega) < \infty \quad \text{and} \quad \int_{\Omega} X^-(\omega) d\mathbb{P}(\omega) < \infty,$$

where $X^+(\omega)$ and $X^-(\omega)$ are as defined above.

Theorem 2.19. *A random variable X is integrable if and only if*

$$\int_\Omega |X(\omega)| d\mathbb{P}(\omega) < \infty.$$

Lebesgue integrals play a central role in calculating quantities in measure-theoretic probability theory because calculating the probability that any *event* A in our *sample space* Ω will occur, where $A \subseteq \Omega$, involves calculating a Lebesgue integral over A with respect to the probability measure. Lebesgue integration is essential here because although in ordinary calculus the x-axis spans $\mathbb{R}$, in probability theory the x-axis spans Ω. A consequence of this is that we can not define a Riemann integral on a general probability space by partitioning the x-axis because there is simply no natural way to partition an arbitrary sample space Ω, which may or may not be composed of numeric quantities; however, as the random variables we will be considering take on numerical values along the y-axis, it is still reasonable to define a Lebesgue integral on a general probability space. This is the great value of Lebesgue integration in probability theory.

Definition 2.20. The distribution of a random variable X can be described in terms of its *cumulative distribution function* $F(x)$, defined as

$$F(x) := \mathbb{P}\{X \leq x\} = \mu_X(-\infty, x] \quad \text{for all } x \in \mathbb{R}.$$

Definition 2.21. In certain cases the distribution of a random variable X has a probability density function $f(x)$, which encodes the distribution of the random variable in more detail. A *probability density function* $f(x)$ is a nonnegative function that is defined for all $x \in \mathbb{R}$ and for which

$$\mu_X[a, b] = \mathbb{P}\{a \leq X \leq b\} = \int_a^b f(x)dx \quad \text{for all } -\infty < a \leq b < \infty.$$

Definition 2.22. Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *expectation, or expected value, of* X is

$$\mathbb{E}[X] := \int_\Omega X(\omega) d\mathbb{P}(\omega)$$

if X is integrable.

Note. A direct consequence of defining expectations as Lebesgue integrals is that expectations are linear. This fact proves useful for manipulating expectations algebraically.

Theorem 2.23. *Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $g(x)$ be a Borel-measurable function on $\mathbb{R}$. If X has probability density function $f(x)$, then*

$$\mathbb{E}[|g(x)|] = \int_{-\infty}^{\infty} |g(x)| f(x) dx.$$

If this quantity is finite, then

$$\mathbb{E}[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Definition 2.24. Let X and Y be random variables. The pair of random variables (X, Y) takes values in the plane $\mathbb{R}^2$, and the *joint distribution measure* of (X, Y) is given by

$$\mu_{X,Y}(B) = \mathbb{P}\{(X, Y) \in B\} \quad \text{for all Borel subsets } B \subseteq \mathbb{R}^2.$$

This is a probability measure. The *joint cumulative distribution function* of (X, Y) is

$$F_{X,Y}(a, b) = \mu_{X,Y}\big((-\infty, a] \times (-\infty, b]\big) = \mathbb{P}\{X \leq a, Y \leq b\}, \quad \text{for all } a, b \in \mathbb{R}.$$

We say that a nonnegative, Borel-measurable function $f_{X,Y}(x,y)$ is a *joint density* for the pair of random variables (X,Y) if

$$\mu_{X,Y}(B) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{I}_B(x,y) f_{X,Y}(x,y)dxdy \quad \text{for all Borel subsets } B \subseteq \mathbb{R}^2.$$

Theorem 2.25. *Let X and Y be random variables. If a joint density function $f_{X,Y}(x,y)$ exists, then the marginal densities exist and are given by*

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \quad \textit{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx.$$

Theorem 2.26. *Let X and Y be random variables. The following conditions are equivalent.*

(i) *X and Y are independent.*

(ii) *The joint distribution measure factors:*

$$\mu_{X,Y}(A \times B) = \mu_X(A)\mu_Y(B) \quad \textit{for all Borel subsets } A, B \subseteq \mathbb{R}.$$

(iii) *The joint cumulative distribution function factors:*

$$F_{X,Y}(a,b) = F_X(a)F_Y(b) \quad \textit{for all } a, b \in \mathbb{R}.$$

(iv) *The joint density factors, provided that there is a joint density to factor:*

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \textit{for almost every } x, y \in \mathbb{R}.$$

Additionally, any one of the prior equivalent conditions implies the following.

(v) *The expectation factors, provided $\mathbb{E}[|XY|] < \infty$:*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

Definition 2.27. Let X be a random variable whose expectation is defined. The *variance* of X, denoted Var(X), is

$$\text{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2,$$

where the second equality follows immediately from the linearity of expectations. The *standard deviation* of X is $\text{SD}(X) := \sqrt{\text{Var}(X)}$.

Let Y be another random variable whose expectation is defined. The *covariance* of X and Y, denoted Cov(X,Y), is

$$\text{Cov}(X,Y) := E[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y],$$

where the second equality follows immediately from the linearity of expectations. If Var(X) and Var(Y) are both positive and finite, then the *correlation coefficient* of X and Y, denoted $\rho(X,Y)$, is

$$\rho(X,Y) := \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

If $\rho(X,Y) = 0$, we say that X and Y are *uncorrelated*.

Definition 2.28. Let X be a random variable. X is a *normal random variable* with *mean* $\mu = E[X]$ and variance $\sigma^2 = \text{Var}(X)$ if it has probability density function

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

Definition 2.29. Let X be a random variable. X is a *standard normal random variable* if it has probability density function

$$f_X(x) = \varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

known as the *standard normal density*, and cumulative distribution function

$$F_X(x) = N(x) := \int_{-\infty}^{x} \varphi(\xi) d\xi,$$

known as the *cumulative normal distribution function*. A comparison with Definition 2.28 leads to the observation that a standard normal random variable has mean 0 and variance 1.

Definition 2.30. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and let X be a random variable that is either nonnegative or integrable. The conditional expectation of X given $\mathcal{G}$, denoted $\mathbb{E}[X|\mathcal{G}]$, is any random variable that satisfies the following two conditions.

(i) *Measurability*:
$\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable.

(ii) *Partial Averaging*:

$$\int_A \mathbb{E}[X|\mathcal{G}](\omega) d\mathbb{P}(\omega) = \int_A X(\omega) d\mathbb{P}(\omega) \quad \text{for all } A \in \mathcal{G}.$$

If $\mathcal{G}$ is the σ-algebra generated by some other random variable W, we generally write $\mathbb{E}[X|W]$ rather than $\mathbb{E}[X|\sigma(W)]$.

Note. Theorems regarding conditional expectations can be applied to full expectations as a full expectation is simply a conditional expectation that is conditioned on the trivial σ-algebra: $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Theorem 2.31. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then the following hold.*

(i) ***Linearity of conditional expectations:***
If X and Y are integrable random variables and c_1, c_2 are constants, then

$$\mathbb{E}[c_1 X + c_2 Y|\mathcal{G}] = c_1 \mathbb{E}[X|\mathcal{G}] + c_2 \mathbb{E}[Y|\mathcal{G}].$$

(ii) ***Taking out what is known:***
If X, Y, and XY are integrable random variables, and X is $\mathcal{G}$-measurable, then

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

(iii) ***Iterated conditioning:***
If $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{G}$ and X is an integrable random variable, then

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

(iv) ***Independence:***
If X is an integrable random variable that is independent of $\mathcal{G}$, then

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

(v) ***Conditional Jensen's Inequality:***
If $\psi(x)$ is a convex function of a dummy variable x and X is an integrable random variable, then

$$\mathbb{E}[\psi(X)|\mathcal{G}] \geq \psi(\mathbb{E}[X|\mathcal{G}]).$$

Lemma 2.32 (Independence)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Suppose the random variables $X_1, \dots, X_K$ are $\mathcal{G}$-measurable and the random variables $Y_1, \dots, Y_L$. are $\mathcal{G}$-independent. Let $f(x_1, \dots, x_K, y_1, \dots, y_L)$ be a function of the dummy variables $x_1, \dots, x_K$ and $y_1, \dots, y_L$, and define*

$$g(x_1, \dots, x_K) := \mathbb{E}[f(x_1, \dots, x_K, Y_1, \dots, Y_L)].$$

Then

$$\mathbb{E}[f(X_1, \dots, X_K, Y_1, \dots, Y_L)|\mathcal{G}] = g(X_1, \dots, X_K).$$

Definition 2.33. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\mathcal{F}(t)$, $0 \le t \le T$, be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an adapted stochastic process $M(t)$, $0 \le t \le T$.

(i) If

$$\mathbb{E}[M(t)|\mathcal{F}(s)] = M(s) \quad \text{for all } 0 \le s \le t \le T,$$

then $M(t)$ is a *martingale*. It has no tendency to rise or fall.

(ii) If

$$\mathbb{E}[M(t)|\mathcal{F}(s)] \ge M(s) \quad \text{for all } 0 \le s \le t \le T,$$

then $M(t)$ is a *submartingale*. It has no tendency to fall, but may have a tendency to rise.

(iii) If

$$\mathbb{E}[M(t)|\mathcal{F}(s)] \le M(s) \quad \text{for all } 0 \le s \le t \le T,$$

then $M(t)$ is a *supermartingale*. It has no tendency to rise, but may have a tendency to fall.

Definition 2.34. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\mathcal{F}(t)$, $0 \le t \le T$, be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an adapted stochastic process $X(t)$, $0 \le t \le T$. If for all $0 \le s \le t \le T$ and for every nonnegative, Borel-measurable function f, there is another Borel-measurable function g such that

$$\mathbb{E}[f(X(t))|\mathcal{F}(s)] = g(X(s)),$$

then X is a *Markov process*.

Note. A given process X is usually proven to be a Markov process by manipulating it so that the Independence Lemma can be invoked. Without the Independence Lemma it would be very difficult to prove that a given process was a Markov process.

2.3. **Brownian Motion.** In order to apply the powerful tools of mathematics to financial analysis, the financial processes being studied must be modeled by mathematical processes that are well understood. Brownian motion is such a mathematical process and lies at the heart of the models that we will consider.

Definition 2.35. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For each $\omega \in \Omega$, suppose that there is a continuous function $W(t)$ of $t \ge 0$ that satisfies $W(0) = 0$ and that depends on ω. Then $W(t)$, $t \ge 0$, is a *Brownian motion* if for all possible partitions $\Pi = \{t_0, t_1, \dots, t_m\}$ of $[0, t]$ such that $0 = t_0 < t_1 < \cdots < t_m = t$, the increments

$$W(t_1) = W(t_1) - W(t_0), W(t_2) - W(t_1), \dots, W(t_m) - W(t_{m-1})$$

are independent and each of these increments is normally distributed with

$$\begin{aligned} \mathbb{E}[W(t_{i+1}) - W(t_i)] &= 0 \\ \mathrm{Var}\big(W(t_{i+1}) - W(t_i)\big) &= t_{i+1} - t_i, \end{aligned}$$

where $i \in \{0, 1, \dots, m-1\}$.

Note. Brownian motion is denoted by a W rather than a B here because although Robert Brown first studied the physical three dimensional Brownian motions of pollen grains suspended in liquid, which were caused by the buffeting of the atoms of the liquid, as Albert Einstein later demonstrated, it was actually Norbert Wiener who first defined Brownian motion as a mathematical object and studied its properties rigorously. Thus the mathematical process defined above is actually called the Wiener process in honor of Norbert Wiener, although the related physical processes are known as Brownian motions in honor of Robert Brown. We refer to the processes here as Brownian motions, although they are indeed also Wiener processes, in order to unify terminology between both disciplines [1].

Definition 2.36. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which a Brownian motion $W(t)$, $t \geq 0$, is defined. A *filtration for the Brownian motion* $W(t)$ is a collection of σ-algebras $\mathcal{F}(t)$, $t \geq 0$, satisfying the following properties.

(i) **Information accumulates**:
For $0 \leq s < t$, every set in $\mathcal{F}(s)$ is also in $\mathcal{F}(t)$.

(ii) **Adaptivity**:
For each $t \geq 0$, the Brownian motion $W(t)$ at time t is $\mathcal{F}(t)$-measurable.

(iii) **Independence of future increments**:
For $0 \leq t < u$, the increment $W(u) - W(t)$ is independent of $\mathcal{F}(t)$.

Let $\Delta(t)$, $t \geq 0$, be a stochastic process. We say that $\Delta(t)$ is *adapted to the Brownian motion* $W(t)$, or that $\Delta(t)$ is *adapted to the filtration* $\mathcal{F}(t)$, if for each $t \geq 0$ the random variable $\Delta(t)$ is $\mathcal{F}(t)$-measurable.

Theorem 2.37. *Brownian motion is a martingale.*

Proof. Let $0 \leq s \leq t$ be given. Then by using Definition 2.36 and Theorem 2.31, we have

$$\begin{aligned}\mathbb{E}[W(t)|\mathcal{F}(s)] &= \mathbb{E}[(W(t) - W(s)) + W(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] + \mathbb{E}[W(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[W(t) - W(s)] + W(s) \\ &= 0 + W(s) = W(s),\end{aligned}$$

which, according to Definition 2.33, shows that any Brownian motion $W(t)$ is a martingale. □

Definition 2.38. Let $f(t)$ be any function defined on $[0, T]$, $\Pi = \{t_0, t_1, \ldots, t_n\}$ any partition of $[0, T]$ such that $0 = t_0 < t_1 < \ldots < t_n = T$, and

$$\|\Pi\| = \max_{0 \leq k \leq n-1} (t_{k+1} - t_k).$$

Then the *quadratic variation* of f up to time T is

$$[f, f](T) := \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \big(f(t_{j+1}) - f(t_j)\big)^2.$$

It is convenient and algebraically useful to record the quadratic variation of a function f using differential notation:

$$df(t)df(t) = d[f, f](t).$$

Definition 2.39. Let $f(t)$ and $g(t)$ be any functions defined on $[0, T]$, $\Pi = \{t_0, t_1, \ldots, t_n\}$ be any partition of $[0, T]$ such that $0 = t_0 < t_1 < \ldots < t_n = T$, and

$$\|\Pi\| = \max_{0 \leq k \leq n-1} (t_{k+1} - t_k).$$

Then the *cross variation* of f with g up to time T is

$$[f,g](T) := \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \big(f(t_{j+1}) - f(t_j)\big)\big(g(t_{j+1}) - g(t_j)\big).$$

It is convenient and algebraically useful to record the cross variation of a function f with g using differential notation:

$$df(t)dg(t) = dg(t)df(t) = d[f,g](t) = d[g,f](t).$$

Theorem 2.40. *The quadratic and cross variations of any Brownian motion $W(t)$ and time t are as follows.*

(i.) $dtdt = 0$
(ii.) $dW(t)dt = 0$
(iii.) $dW(t)dW(t) = dt$

2.4. **Itó Calculus.** Calculus involving functions of stochastic processes is not entirely the same as calculus involving only differentiable functions. As Brownian motion is a stochastic process, these differences need to be explored and understood so that the expressions involving Brownian motions obtained later can be properly manipulated. Several fundamental results of this stochastic calculus, known as Itó Calculus, are thus presented below.

Definition 2.41. Let $W(t)$ be a Brownian motion, $\Pi = \{t_0, t_1, \ldots, t_n\}$ any partition of $[0,T]$ such that $0 = t_0 < t_1 < \ldots < t_n = T$, and

$$\|\Pi\| = \max_{0\le k\le n-1} (t_{k+1} - t_k).$$

The *Itó integral* of $\Delta(t)$ over $[0,T]$ is defined to be

$$\int_0^T \Delta(t)dW(t) := \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big).$$

Theorem 2.42. *Let T be a positive constant and let $\Delta(t)$, $0 \le t \le T$, be an adapted stochastic process for which*

$$\mathbb{E}\left[\int_0^T \Delta^2(t)dt\right] < \infty.$$

Then any Itó integral $I(t) = \int_0^t \Delta(u)dW(u)$ has the following properties.

(i) ***Continuity:***
As a function of the upper limit of integration t, the paths of $I(t)$ are continuous.

(ii) ***Adaptivity:***
For each t, $I(t)$ is $\mathcal{F}(t)$-measurable.

(iii) ***Linearity:***
If $\Gamma(t)$, $0 \le t \le T$, is another adapted stochastic process and c_1, c_2 are constants, then

$$\int_0^t \big(c_1\Delta(u) + c_2\Gamma(u)\big)dW(u) = c_1 \int_0^t \Delta(u)dW(u) + c_2 \int_0^t \Gamma(u)dW(u).$$

(iv) ***Martingale:***
I(t) is a martingale.

(v) ***Itó Isometry:***

$$\mathbb{E}[I^2(t)] = \mathbb{E}\left[\int_0^t \Delta^2(u)du\right].$$

(vi) ***Quadratic Variation:***

$$[I, I](t) = \int_0^t \Delta^2(u)du.$$

Definition 2.43. Let $W(t)$, $t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$, be an associated filtration. An *Itó process* is a stochastic process of the form

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du,$$

where $X(0)$ is nonrandom and $\Delta(u)$ and $\Theta(u)$ are adapted stochastic processes.

Note. The use of differential notation makes the calculation of quadratic and cross variations an exercise in applying Theorem 2.40. The quadratic variation of an Itó integral, for example, is quickly seen to be

$$dI(t)dI(t) = \big(\Delta(t)dW(t)\big)^2 = \Delta^2(t)dt,$$

in verification of the conclusion of Theorem 2.42. The quadratic variation of an Itó process is calculated just as easily:

$$dX(t)dX(t) = \big(\Delta(t)dW(t) + \Theta(t)dt\big)^2 = \Delta^2(t)dt.$$

This method of manipulating differentials is incredibly useful and will be used extensively in later calculations.

Theorem 2.44 (General Itó-Doeblin Formula)**.** *Let $g(x_1, x_2, \ldots, x_n)$ be a smooth function of the n dummy variables $x_1, x_2, \ldots, x_n$, and let $X_1, X_2, \ldots, X_n$ be n potentially stochastic processes. The general Itó-Doeblin Formula states that*

$$\begin{aligned} dg(X_1, X_2, \ldots, X_n) &= \sum_{i=1}^n g_{x_i}(X_1, X_2, \ldots, X_n)dX_i \\ &\quad + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n g_{x_i x_j}(X_1, X_2, \ldots, X_n)dX_i dX_j \\ &:= \sum_{i=1}^n g_{x_i}(x_1, x_2, \ldots, x_n)dX_i \\ &\quad + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n g_{x_i x_j}(x_1, x_2, \ldots, x_n)dX_i dX_j\Big|_{(X_1, X_2, \ldots, X_n)}, \end{aligned}$$

where g_{x_i} denotes partial differentiation of the function g with respect to x_i.

Corollary 2.45. *Let t be a time variable, and $X(t)$,$Y(t)$ be stochastic processes. Then*

$$df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))dX(t) + \frac{1}{2}f_{xx}(t, X(t))dX(t)dX(t)$$

and

$$\begin{aligned} dg(t, X(t), Y(t)) &= g_t(t, X(t), Y(t))dt + g_x(t, X(t), Y(t))dX(t) \\ &\quad + g_y(t, X(t), Y(t))dY(t) \\ &\quad + \frac{1}{2} f_{xx}(t, X(t), Y(t))dX(t)dX(t) \\ &\quad + f_{xy}(t, X(t), Y(t))dX(t)dY(t) \\ &\quad + \frac{1}{2} f_{yy}(t, X(t), Y(t))dY(t)dY(t), \end{aligned}$$

where in both cases any double differential involving a dt *has been recognized as equal to* 0.

2.5. **Risk-Neutral Measure.** Option pricing benefits greatly from the ability to change the probability measure with respect to which calculations are made. This alternative probability measure is deemed the risk-neutral measure because under this new measure both the discounted stock price and the discounted wealth process are martingales. Essentially, the mean rate of return of the asset being considered is now simply the present interest rate in this new probability measure. These changes make the mathematics much easier and enable the development of several techniques for pricing an arbitrary option. Background material related to the general mathematical procedure of changing from one probability measure to another is provided in order to put this specific application of the technique in its proper context.

Theorem 2.46. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and let* Z *be an almost surely nonnegative random variable with* $\mathbb{E}[Z] = 1$. *For all* $A \in \mathcal{F}$, *define*

$$\widetilde{\mathbb{P}}(A) := \int_A Z(\omega) d\mathbb{P}(\omega).$$

Then $\widetilde{\mathbb{P}}$ *is a probability measure. Furthermore, if* X *is a nonnegative random variable, then*

$$\widetilde{\mathbb{E}}[X] := \int_\Omega X(\omega) d\widetilde{\mathbb{P}}(\omega) = \mathbb{E}[XZ].$$

Definition 2.47. Let Ω be a nonempty set and $\mathcal{F}$ a σ-algebra of subsets of Ω. Two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ are said to be *equivalent* if they agree which sets in $\mathcal{F}$ have probability zero.

Definition 2.48. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\widetilde{\mathbb{P}}$ be another probability measure on $(\Omega, \mathcal{F})$ that is equivalent to $\mathbb{P}$, and let Z be an almost surely positive random variable that relates $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ via Theorem 2.46. Then Z is called the *Radon-Nikodým derivative of* $\widetilde{\mathbb{P}}$ *with respect to* $\mathbb{P}$, and we write

$$Z = \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}.$$

Theorem 2.49 (Radon-Nikodým)**.** *Let* $\mathbb{P}$ *and* $\widetilde{\mathbb{P}}$ *be equivalent probability measures defined on* $(\Omega, \mathcal{F})$. *Then there exists an almost surely positive random variable* Z *such that* $\mathbb{E}[Z] = 1$ *and*

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega) \quad \text{for every } A \in \mathcal{F}.$$

Theorem 2.50 (Lévy Theorem)**.** *Let* $M(t)$, $t \geq 0$, *be a martingale relative to a filtration* $\mathcal{F}(t)$, $t \geq 0$. *If* $M(0) = 0$, $M(t)$ *has continuous paths, and* $[M, M](t) = t$ *for all* $t \geq 0$, *then* $M(t)$ *is a Brownian motion.*

Definition 2.51. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, T a fixed positive number, $\mathcal{F}(t)$, $0 \leq t \leq T$, be a filtration, and Z be the Radon-Nikodým derivative of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$. The *Radon-Nikodým derivative process* is

$$Z(t) := \mathbb{E}[Z|\mathcal{F}(t)] \quad \text{for all } t \in [0, T].$$

Lemma 2.52. *Let t satisfying $0 \leq t \leq T$ be given and Y be an $\mathcal{F}(t)$-measurable random variable. Then*

$$\widetilde{\mathbb{E}}[Y] = \mathbb{E}[YZ(t)].$$

Lemma 2.53. *Let s and t satisfying $0 \leq s \leq t \leq T$ be given and let Y be an $\mathcal{F}(t)$-measurable random variable. Then*

$$\widetilde{\mathbb{E}}[Y|\mathcal{F}(s)] = \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)].$$

Theorem 2.54 (Girsanov)**.** *Let $W(t)$, $0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t)$, $0 \leq t \leq T$, be a filtration for this Brownian motion. Let $\Theta(t)$, $0 \leq t \leq T$, be an adapted stochastic process. Define*

$$\begin{aligned} Z(t) &= \exp\left\{-\int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du\right\}, \\ \widetilde{W}(t) &= W(t) + \int_0^t \Theta(u)du, \end{aligned}$$

and assume that

$$\mathbb{E}\left[\int_0^T \Theta^2(u)Z^2(u)du\right] < \infty.$$

Set $Z = Z(T)$. Then $\mathbb{E}[Z] = 1$ and under the probability measure $\widetilde{\mathbb{P}}$ as defined in Theorem 2.46, the process $\widetilde{W}(t)$, $0 \leq t \leq T$, is a Brownian motion.

Note. Theorem 2.54 is incredibly important in option pricing theory because it enables us to consider the problem in the context of an alternative probability measure $\widetilde{\mathbb{P}}$. Many quantities of interest in option pricing theory are martingales under a new risk-neutral probability measure $\widetilde{\mathbb{P}}$ that are not martingales under the actual probability measure $\mathbb{P}$. This makes the risk-neutral picture, and thus Theorem 2.54, invaluable in option pricing.

Theorem 2.55 (Martingale Representation)**.** *Let $W(t)$, $0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t)$, $0 \leq t \leq T$, be the filtration generated by this Brownian motion. Let $M(t)$, $0 \leq t \leq T$, be a martingale with respect to this filtration. Then there is an adapted stochastic process $\Gamma(u)$, $0 \leq u \leq T$, such that*

$$M(t) = M(0) + \int_0^t \Gamma(u)dW(u).$$

3. Option Pricing

There are two different approaches to pricing a given option. The first method turns the problem of pricing an option into the problem of finding a solution to a given partial differential equation with a specific boundary condition. The second method expresses the option price as a risk-neutral conditional expectation. Using the techniques of stochastic calculus, an explicit expression for the price of the given option may then be found by simplifying and evaluating this expression. Although such complete simplification is not always possible, this latter method will be our main focus as it not only has the potential to provide an exact price for an option, but even in the event that this is not possible the option price can often be stated in a simpler form that may be more tractable from a calculational standpoint.

3.1. **Notation and Preliminary Material.** Both approaches to option pricing have a great deal of notation in common. In our treatment of these models, we will only be considering a single asset whose price $S(t)$ changes in time t over the period of the option contract, which has expiration time T, a positive constant. Thus $0 \leq t \leq T$ in all of our models. Furthermore, the asset price is modeled as a *geometric Brownian motion*, which means that our asset price satisfies

$$S(t) = S(0) \exp\left\{ \int_0^t \sigma(u) dW(u) + \int_0^t \big(\alpha(u) - \frac{1}{2}\sigma^2(u)\big) du \right\},$$

where the adapted stochastic processes $\sigma(t)$ and $\alpha(t)$ are the asset *volatility* and the asset's instantaneous *mean rate of return*, respectively. The changing interest rate $R(t)$ is an adapted stochastic process associated with the market in which our asset resides. For convenience, we further define a discount factor $D(t)$ as

$$D(t) := \exp\left\{ -\int_0^t R(u) du \right\},$$

yet another adapted stochastic process.

In order to invoke Theorem 2.54 to define the risk-neutral probability measure mentioned above, we define an adapted stochastic process called the *market price of risk* by

$$\Theta(t) := \frac{\alpha(t) - R(t)}{\sigma(t)},$$

where it is assumed that there is always a certain degree of volatility in the asset price, so that $\sigma(t) > 0$ for every value of $t \in [0, T]$. We can now rewrite the asset price as a geometric Brownian motion in terms of a Brownian motion under the risk-neutral probability measure defined by using the market price of risk as the adapted stochastic process in Theorem 2.54:

$$S(t) = S(0) \exp\left\{ \int_0^t \sigma(u) d\widetilde{W}(u) + \int_0^t \big(R(u) - \frac{1}{2}\sigma^2(u)\big) du \right\}.$$

We can now use Corollary 2.45 to calculate $dS(t)$ in terms of this risk-neutral Brownian motion.

Theorem 3.1.

$$dS(t) = S(t)\sigma(t) d\widetilde{W}(t) + S(t)R(t)dt$$

Proof. The desired result is obtained by invoking Corollary 2.45 with

$$f(t, x) = S(0) \exp\left\{ \int_0^t \sigma(u) dx + \int_0^t \big(R(u) - \frac{1}{2}\sigma^2(u)\big) du \right\}:$$

$$\begin{aligned} df(t, \widetilde{W}(t)) &= f_t dt + f_x d\widetilde{W}(t) + \frac{1}{2} f_{xx} d\widetilde{W}(t) d\widetilde{W}(t) \\ &= S(t)\big(R(t) - \frac{1}{2}\sigma^2(t)\big)dt + S(t)\sigma(t) d\widetilde{W}(t) + \frac{1}{2} S(t)\sigma^2(t) dt \\ &= S(t)\sigma(t) d\widetilde{W}(t) + S(t)R(t)dt. \end{aligned}$$

□

The total wealth of an individual portfolio's money market and asset market investments defines the adapted stochastic process $X(t)$, which is known as the *wealth process*. The differential change in this wealth process is given by summing the contribution from each of these component markets. The change in the value of the asset market investment is simply the change in the asset price multiplied by the amount of the asset that is held at time t. The change in the money market investment is simply the amount of wealth in the money market at time t multiplied by the interest rate. Define the adapted stochastic process $\Delta(t)$ to be the number of shares of the asset that are held at time $t \in [0, T]$. We then have the following theorem.

Theorem 3.2.

$$dX(t) = R(t)X(t)dt + S(t)\Delta(t)\sigma(t)d\widetilde{W}(t)$$

Proof. We use the reasoning articulated above to write the first equality. An application of Theorem 3.1 then leads to the desired result:

$$\begin{aligned} dX(t) &= \Delta(t)dS(t) + R(t)\big(X(t) - \Delta(t)S(t)\big)dt \\ &= \Delta(t)\big(S(t)\sigma(t)d\widetilde{W}(t) + S(t)R(t)dt\big) + R(t)\big(X(t) - \Delta(t)S(t)\big)dt \\ &= R(t)X(t)dt + S(t)\Delta(t)\sigma(t)d\widetilde{W}(t) \end{aligned}$$

□

Theorem 3.3. *A stochastic process $\Xi(t)$ whose filtration $\mathcal{F}(t)$, $0 \le t \le T$, is generated by a Brownian motion is a martingale if and only if*

$$d\Xi(t) = 0dt + \Phi(t)dW(t).$$

Proof. First consider a stochastic process $\Xi(t)$ for which

$$d\Xi(t) = 0dt + \Phi(t)dW(t).$$

Integrating this differential formula yields

$$\Xi(t) = \Xi(0) + \int_0^t \Phi(u)dW(u).$$

Theorem 2.42 states that all Itó integrals are martingales. Using this fact, along with the properties of conditional expectations given in Theorem 2.31, we have that for $0 \le s \le t \le T$,

$$\begin{aligned} E[\Xi(t)|\mathcal{F}(s)] &= E\left[\Xi(0) + \int_0^t \Phi(u)dW(u)\middle|\mathcal{F}(s)\right] \\ &= E[\Xi(0)|\mathcal{F}(s)] + E\left[\int_0^t \Phi(u)dW(u)\middle|\mathcal{F}(s)\right] \\ &= \Xi(0) + \int_0^s \Phi(u)dW(u) = \Xi(s). \end{aligned}$$

Thus, by Definition 2.33, $\Xi(t)$ is a martingale.

Now consider a stochastic process $\Xi(t)$ that is a known martingale, and whose filtration $\mathcal{F}(t)$, $0 \le t \le T$, is generated by a Brownian motion. According to Theorem 2.55, there is some adapted stochastic process $\Phi(u)$, $0 \le u \le T$, such that

$$\Xi(t) = \Xi(0) + \int_0^t \Phi(u)dW(u).$$

We take the differential of this equation to arrive at

$$d\Xi(t) = 0dt + \Phi(t)dW(t),$$

which completes the proof. □

Note. All of the adapted stochastic processes considered here are adapted to a filtration generated by a Brownian motion. This is a fundamental assumption inherent in modeling the asset price as a geometric Brownian motion: all of the random movements of the stock are a result of the associated movements of the underlying Brownian motion. Theorem 2.55 and all theorems that rely upon it will thus be universally applicable for the problems considered in this paper.

Lemma 3.4.

$$dD(t) = -D(t)R(t)dt$$

Proof. The desired result is obtained by invoking Corollary 2.45 with

$$f(t,x) = \exp\left\{-\int_0^t R(u)du\right\}:$$

$$\begin{aligned} df(t,\widetilde{W}(t)) &= f_t dt + f_x d\widetilde{W}(t) + \frac{1}{2} f_{xx} d\widetilde{W}(t) d\widetilde{W}(t) \\ &= -D(t)R(t)dt + 0d\widetilde{W}(t) + 0d\widetilde{W}(t)d\widetilde{W}(t) = -D(t)R(t)dt. \end{aligned}$$

□

Theorem 3.5. *The adapted stochastic process $D(t)X(t)$ is a martingale.*

Proof. We calculate the differential of $D(t)X(t)$ in the risk-neutral framework in order to show that the coefficient of the dt term is 0 when considered under the risk-neutral probability measure. We use

$$f(d,x) = xd$$

in Theorem 2.44 to calculate $D(t)X(t)$, invoking Theorem 3.2 and Lemma 3.4 where necessary:

$$\begin{aligned} d(D(t)X(t)) &= df(D(t),X(t)) \\ &= f_d dD(t) + f_x dX(t) \\ &\quad + \frac{1}{2} f_{dd} dD(t)dD(t) + f_{xd} dD(t)dX(t) + \frac{1}{2} f_{xx} dX(t)dX(t) \\ &= X(t)\big(-D(t)R(t)\big)dt \\ &\quad + D(t)\big(R(t)X(t)dt + S(t)\Delta(t)\sigma(t)d\widetilde{W}(t)\big) \\ &= D(t)S(t)\Delta(t)\sigma(t)d\widetilde{W}(t). \end{aligned}$$

$D(t)X(t)$ is thus a martingale under the risk-neutral probability measure by Theorem 3.3. □

The fact that the discounted wealth process is a martingale under the risk-neutral probability measure $\widetilde{\mathbb{P}}$ is important in option pricing theory. It underlies all of the option pricing theory contained in the rest of the paper.

3.2. **Partial Differential Equation Approach.** Consider the discounted wealth process $X(t)$, $0 \leq t \leq T$, that perfectly replicates the value of the option $V(t)$, $0 \leq t \leq T$, which is another adapted stochastic process. In order to perfectly replicate the option, the replicating wealth process $X(t)$ must at all times be exactly equal in value to the value of the option $V(t)$. If this were not the case, then there would be some time at which invoking the option would result in the financial institution either gaining or losing money. In either case the value of the option at that time would not be fair. This runs contrary to our assumptions, and thus it must be that $X(t) = V(t)$ for all $t \in [0, T]$.

The value of the option can reasonably be assumed to be a function of some number n of potentially stochastic variables, say $X_1, X_2, \ldots, X_n$. We may then write the option value at time t as $v(X_1, X_2, \ldots, X_n)$. But since $X(t)$ and $V(t)$ have been identified, Theorem 3.5 implies that the discounted option value is a martingale, and thus by Theorem 3.3 this differential has a coefficient of 0 in front of its dt term. This enables us to calculate the differential of the discounted option value explicitly in terms of partial derivatives of v using Theorem 2.44 and then set the dt term equal to 0 in order to obtain a partial differential equation governing the price of the option. It turns out that if we then also set the $d\widetilde{W}(t)$ term of this expression equal to the corresponding term in $d(D(t)X(t))$, we will obtain the perfect hedging portfolio that will lead to replication of the option value at all times. Thus not only does solving this partial differential equation allow us to determine the fair price of the option, but it also provides us with specific instructions as to how the money obtained by selling the option at this fair price can be invested in the asset and money markets in order to perfectly hedge our position to avoid losses.

The general procedure is best illustrated by some specific examples. Consider a simple option whose value is a function of only the present time t and the present price of the underlying asset $S(t)$, so that we have $v(t, S(t))$. This is a reasonable model for a variety of options, such as the *European call option*, which can only be executed at the expiration time T of the option contract, and has payoff

$$V(T) = (S(T) - K)^+,$$

where the *plus-function* is defined as

$$(x)^+ := \begin{cases} x, & \text{for all } x \geq 0 \\ 0, & \text{for all } x < 0 \end{cases},$$

and K is a constant called the *strike price*.

Theorem 3.6. *Consider a European call option whose value at time $t \in [0, T]$ is a function of only the present time t and the present price of the underlying asset $S(t)$: $v(t, S(t))$. Let the asset price be modeled by a geometric Brownian motion with constant volatility σ and constant mean rate of return α, and let the interest rate be a constant r.*

Consider the underlying function $v(t, x)$ that gives the option price at time t when evaluated at $x = S(t)$. This function satisfies the partial differential equation

$$v_t(t, x) + rxv_x(t, x) + \frac{1}{2}\sigma^2 x^2 v_{xx}(t, x) = rv(t, x)$$

subject to the terminal condition

$$v(T, x) = (x - K)^+,$$

and the perfect hedging portfolio for this option is given by

$$\Delta(t) = v_x(t, S(t)).$$

Proof. First we calculate the differential $d(e^{-rt}v(t,x))$:

$$\begin{aligned} d(e^{-rt}v(t,x)) &= e^{-rt}dv(t,x) - re^{-rt}v(t,x)dt \\ &= e^{-rt}\left(v_t dt + v_x dS(t) + \frac{1}{2}v_{xx}dS(t)dS(t)\right) - re^{-rt}vdt \\ &= e^{-rt}\Bigg(v_t dt + \sigma S(t)v_x d\widetilde{W}(t) + rS(t)v_x dt \\ &\quad + \frac{1}{2}\sigma^2 S^2(t)v_{xx}dt\Bigg) - re^{-rt}vdt \\ &= e^{-rt}\left(v_t + rxv_x + \frac{1}{2}\sigma^2 x^2 v_{xx} - rv\right)dt \\ &\quad + e^{-rt}\left(\sigma x v_x\right) d\widetilde{W}(t). \end{aligned}$$

We now equate this to $d(D(t)X(t))$ from Theorem 3.5, which it must equal because $X(t) = V(t) = v(t, S(t))$. We find that since $e^{-rt} \neq 0$ for any t, we must have

$$v_t + rxv_x + \frac{1}{2}\sigma^2 x^2 v_{xx} - rv = 0$$

and

$$\sigma S(t)v_x(t, S(t)) = \sigma S(t)\Delta(t),$$

which implies that

$$v_x(t, S(t)) = \Delta(t).$$

By considering the nature of our specific option, the European call option, we further impose the necessary boundary condition

$$v(T, x) = (x - K)^+$$

on our solution. This completes the proof. □

This example demonstrates the ease with which a partial differential equation characterizing a desired option can be determined. Our primary interest is with pricing the *Asian call option*, which, like the European call option, can only be executed at the expiration time T. Unlike the European call option, however, the Asian call option has payoff

$$\left(\frac{Y(T)}{T} - K\right)^+,$$

where we define

$$Y(t) := \int_0^t S(u)du, \quad \text{for all } t \in [0, T].$$

The value of the Asian call option at an arbitrary time $t \in [0, T]$ is generally a function of not only t and $S(t)$, but also $Y(t)$: $v(t, S(t), Y(t))$. It is natural to wonder what partial differential equation characterizes the Asian call option. We follow the same procedure as before in order to find out.

Note. It follows immediately from the definition of $Y(t)$ that $dY(t) = S(t)dt$. This is used in the proof of the following theorem.

Theorem 3.7. *Consider an Asian call option whose value at time $t \in [0, T]$ is a function of the present time t, the present price of the underlying asset $S(t)$, and $Y(t) := \int_0^t S(u)du$: $v(t, S(t), Y(t))$. Let the asset price be modeled by a geometric Brownian motion with* *constant volatility σ and constant mean rate of return α, and let the interest rate be a constant r.*

Consider the underlying function $v(t,x,y)$ that gives the option price at time t when evaluated at $x = S(t)$, and $y = Y(t)$. This function satisfies the partial differential equation

$$v_t(t,x,y) + rxv_x(t,x,y) + \frac{1}{2}\sigma^2 x^2 v_{xx}(t,x,y) + xv_y(t,x,y) = rv(t,x,y)$$

subject to the boundary condition

$$v(T,x,y) = (\frac{y}{T} - K)^+,$$

and the perfect hedging portfolio for this option is given by

$$\Delta(t) = v_x(t, S(t), Y(t)).$$

Proof. First we calculate the differential $d(e^{-rt}v(t,x,y))$:

$$\begin{aligned}
d(e^{-rt}v(t,x,y)) &= e^{-rt}dv(t,x,y) - re^{-rt}v(t,x,y)dt \\
&= e^{-rt}\left(v_t dt + v_x dS(t) + v_y dY(t) + \frac{1}{2}v_{xx}dS(t)dS(t)\right) \\
&\quad -re^{-rt}vdt \\
&= e^{-rt}\left(v_t dt + \sigma S(t)v_x d\widetilde{W}(t) + rS(t)v_x dt\right. \\
&\quad \left. + \frac{1}{2}\sigma^2 S^2(t)v_{xx}dt + S(t)v_y dt\right) - re^{-rt}vdt \\
&= e^{-rt}\left(v_t + rxv_x + \frac{1}{2}\sigma^2 x^2 v_{xx} + xv_y - rv\right)dt \\
&\quad + e^{-rt}\left(\sigma x v_x\right) d\widetilde{W}(t).
\end{aligned}$$

We now equate this to $d(D(t)X(t))$ from Theorem 3.5, which it must equal because $X(t) = V(t) = v(t, S(t), Y(t))$. We find that since $e^{-rt} \neq 0$ for any t, we must have

$$v_t + rxv_x + \frac{1}{2}\sigma^2 x^2 v_{xx} + xv_y - rv = 0$$

and

$$\sigma S(t) v_x(t, S(t), Y(t)) = \sigma S(t)\Delta(t),$$

which implies that

$$v_x(t, S(t), Y(t)) = \Delta(t).$$

By considering the nature of our specific option, the Asian call option, we further impose the necessary boundary condition

$$v(T,x,y) = \left(\frac{y}{T} - K\right)^+$$

on our solution. This completes the proof. □

We see that there was no difficulty at all in applying the method used to determine the partial differential equation associated with the European call option to determine the partial differential equation associated with the Asian call option. Indeed, the proof of Theorem 3.7 almost exactly follows the proof of Theorem 3.6. Such a natural generalization of this approach to option pricing is encouraging and leads us to hope that the other more explicit approach to option pricing, presented below, will also generalize naturally from the European call option to the Asian call option. If this were the case, then we would be able to price the Asian option exactly; unfortunately, this is not the case, as we shall see, and pricing the Asian call option is somewhat more difficult than pricing the European call option.

3.3. **Risk-Neutral Conditional Expectation Approach.** The fact that $D(t)X(t)$, and thus also $D(t)V(t)$, is a martingale under the risk-neutral probability measure can be used in a more direct fashion to price options. In particular, we have

$$D(t)V(t) = \widetilde{\mathbb{E}}[D(T)V(T)|\mathcal{F}(t)],$$

and thus

$$V(t) = \widetilde{\mathbb{E}}\left[e^{-\int_t^T R(u)du}V(T)|\mathcal{F}(t)\right],$$

since $D(t)$ is $\mathcal{F}(t)$-measurable.

The problem of pricing any given option has now been reduced to the problem of evaluating a risk-neutral conditional expectation. Evaluating the conditional expectation directly is not feasible; however, if we could find a way to break the argument into several different parts, with each part being either $\mathcal{F}(t)$-measurable or $\mathcal{F}(t)$-independent, we could invoke the Independence Lemma to turn the conditional expectation into a full expectation. If we could then further manipulate the argument to this expectation in such a way that the expectation was over independent random variables with known density functions, then we could express the price of the given option as an integral. This would be a significant simplification, and would provide a very useful explicit formula for the value of the given option at any given time.

This method of simplifying the conditional expectation is, in fact, the method that will be used here in order to price the Asian option as explicitly as possible. We first illustrate the method by using it to price the European call option, a case in which it yields a very nice result, known as the Black-Scholes-Merton formula, that is both exact and explicit. We then apply the method to pricing the Asian call option in the hope that the method generalizes without too much difficulty, just as the partial differential equations method presented above generalized easily. We postpone resorting to approximations for as long as possible in order to provide an exact, but not entirely explicit, result that is as simplified as our method allows. This is given in Theorem 3.9. We then extend our approach, sacrificing exact accuracy in order to obtain a more explicit formula that is akin to the Black-Scholes-Merton formula. This is given in Theorem 3.10.

We recall the definitions

$$\varphi(x) := \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

and

$$N(x) := \int_{-\infty}^{x} \varphi(\xi)d\xi$$

from Definition 2.29 as they are important in the following theorems.

Theorem 3.8 (Black-Scholes-Merton Formula). *Let the expiration time T be a positive constant so that the present time $t \in [0,T]$. Define the time until expiration to be $\tau := T-t$. Now consider a European call option with strike price K in a market with a constant interest rate r for which the underlying asset is modeled as a geometric Brownian motion with constant volatility σ. Further define*

$$d_-(\tau,x) := \frac{1}{\sigma\sqrt{\tau}}\left[\ln\left(\frac{x}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)\tau\right]$$

and

$$d_+(\tau, x) := d_-(\tau, x) + \sigma\sqrt{\tau} = \frac{1}{\sigma\sqrt{\tau}}\left[\ln\left(\frac{x}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)\tau\right].$$

The value of a European call option is then given by

$$v(\tau, S(t)) = S(t)N(d_+(\tau, S(t))) - Ke^{-r\tau}N(d_-(\tau, S(t))).$$

Proof. Since the discounted value of the option is a martingale under the risk-neutral probability measure, we have that

$$v(\tau, S(t)) = V(t) = \widetilde{\mathbb{E}}\left[e^{-r\tau}V(T)|\mathcal{F}(t)\right].$$

If we insert the final value of the European call option into this equation, we obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}(S(T) - K)^+|\mathcal{F}(t)\right].$$

By using the fact that the asset price is being modeled as a geometric Brownian motion, we can replace the final asset price to obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(S(t)\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau + \sigma\left(\widetilde{W}(T) - \widetilde{W}(t)\right)\right\} - K\right)^+\middle|\mathcal{F}(t)\right].$$

We now define the standard normal random variable

$$Z := -\frac{\widetilde{W}(T) - \widetilde{W}(t)}{\sqrt{T - t}} = -\frac{\widetilde{W}(T) - \widetilde{W}(t)}{\sqrt{\tau}}.$$

This enables us to obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(S(t)\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}Z\right\} - K\right)^+\middle|\mathcal{F}(t)\right]$$

through substitution. We then replace the random variable $S(t)$ with the dummy variable x to get

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(x\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}Z\right\} - K\right)^+\middle|\mathcal{F}(t)\right].$$

Every individual component of this expression is either $\mathcal{F}(t)$-measurable or $\mathcal{F}(t)$-independent. This makes it possible to invoke the Independence Lemma to obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(x\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}Z\right\} - K\right)^+\right].$$

This is now a full expectation and can thus be written as an integral in terms of the known standard normal density $\varphi(z)$ as

$$\int_{-\infty}^{\infty} e^{-r\tau}\left(x\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}z\right\} - K\right)^+\varphi(z)dz.$$

The function being integrated is zero for all z greater than $d_-(\tau, x)$. By changing the limits of integration to reflect this, we are able to drop the plus-function, leaving only

$$\int_{-\infty}^{d_-(\tau,x)} e^{-r\tau}\left(x\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}z\right\} - K\right)\varphi(z)dz.$$

The linearity of integration can now be used to separate this integral into the sum of several simpler integrals, namely

$$xe^{-\sigma^2\tau/2}\int_{-\infty}^{d_-(\tau,x)} e^{-\sigma\sqrt{\tau}z}\varphi(z)dz - Ke^{-r\tau}\int_{-\infty}^{d_-(\tau,x)}\varphi(z)dz.$$

A change of variables is used to simplify the first integral, which causes a change in its upper limit of integration. This simplification enables both integrals to be written in terms of the cumulative normal distribution function $N(x)$ as

$$xN(d_+(\tau,x)) - Ke^{-r\tau}N(d_-(\tau,x)).$$

We now simply replace the dummy variable x with the original random variable $S(t)$ to obtain the Black-Scholes-Merton formula:

$$v(\tau,S(t)) = S(t)N(d_+(\tau,S(t))) - Ke^{-r\tau}N(d_-(\tau,S(t))).$$

□

Theorem 3.9. *Let the expiration time T be a positive constant so that the present time $t\in[0,T]$. Define the time until expiration to be $\tau := T-t$. Define the process*

$$Y(t) := \int_0^t S(u)du$$

and use the random variable $\Gamma\in[t,T]$, defined implicitly by

$$\int_t^T S(u)du = S(\Gamma)(T-t)$$

using the mean value theorem, to define the random variable $\Lambda := \Gamma - t$. Now consider an Asian call option with strike price K in a market with a constant interest rate r for which the underlying asset is modeled as a geometric Brownian motion with constant volatility σ. Further define

$$d_-(\tau,\Lambda,x,y) := \frac{1}{\sigma\sqrt{\Lambda}}\left[\ln\left(\frac{\tau x}{KT-y}\right) + \left(r-\frac{1}{2}\sigma^2\right)\Lambda\right],$$

$$d_+(\tau,\Lambda,x,y) := d_-(\tau,\Lambda,x,y) + \sigma\sqrt{\Lambda} = \frac{1}{\sigma\sqrt{\Lambda}}\left[\ln\left(\frac{\tau x}{KT-y}\right) + \left(r+\frac{1}{2}\sigma^2\right)\Lambda\right],$$

and the standard normal random variable

$$Z := -\frac{\widetilde{W}(\Gamma)-\widetilde{W}(t)}{\sqrt{\Lambda}}.$$

The value of an Asian call option is then given by

$$v(\tau,S(t),Y(t)) =$$

$$\int_0^\tau\int_{-\infty}^{d_-(\tau,\lambda,S(t),Y(t))} e^{-r\tau}\left(\frac{Y(t)}{T}+\frac{\tau S(t)}{T}\exp\left\{\left(r-\frac{1}{2}\sigma^2\right)\lambda-\sigma\sqrt{\lambda}z\right\}-K\right)f_{\Lambda,Z}(\lambda,z)dzd\lambda,$$

where $f_{\Lambda,Z}(\lambda,z)$ is the joint density function for Λ and Z.

Proof. Since the discounted value of the option is a martingale under the risk-neutral probability measure, we have that

$$v(\tau,S(t),Y(t)) = V(t) = \widetilde{\mathbb{E}}\left[e^{-r\tau}V(T)|\mathcal{F}(t)\right].$$

If we insert the final value of the Asian call option into this equation, we obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(\frac{1}{T}Y(T)-K\right)^{+}\middle|\mathcal{F}(t)\right].$$

We need to simplify the $Y(T)$ term in order to continue with the calculation. We use the mean value theorem to achieve this simplification by writing

$$Y(T)=Y(t)+\int_t^T S(u)du=Y(t)+S(\Gamma)\tau,$$

where Γ is a random variable that is $\mathcal{F}(t)$-independent. We substitute this into the previous expression to obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(\frac{1}{T}\Big(Y(t)+S(\Gamma)\tau\Big)-K\right)^{+}\middle|\mathcal{F}(t)\right].$$

By using the fact that the asset price is being modeled as a geometric Brownian motion, we can replace the final asset price to obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(\frac{1}{T}\Big(Y(t)+\tau S(t)\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\Lambda+\sigma\big(\widetilde{W}(\Gamma)-\widetilde{W}(t)\big)\Big\}\Big)-K\right)^{+}\middle|\mathcal{F}(t)\right].$$

We now define the standard normal random variable

$$Z:=-\frac{\widetilde{W}(\Gamma)-\widetilde{W}(t)}{\sqrt{\Gamma-t}}=-\frac{\widetilde{W}(\Gamma)-\widetilde{W}(t)}{\sqrt{\Lambda}}.$$

This enables us to obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(\frac{1}{T}\Big(Y(t)+\tau S(t)\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\Lambda-\sigma\sqrt{\Lambda}Z\Big\}\Big)-K\right)^{+}\middle|\mathcal{F}(t)\right]$$

through substitution. We then substitute the dummy variables x and y for the random variables $S(t)$ and $Y(t)$, respectively, to get

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(\frac{1}{T}\Big(y+\tau x\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\Lambda-\sigma\sqrt{\Lambda}Z\Big\}\Big)-K\right)^{+}\middle|\mathcal{F}(t)\right].$$

Every individual component of this expression is either $\mathcal{F}(t)$-measurable or $\mathcal{F}(t)$-independent. This makes it possible to invoke the Independence Lemma to obtain

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(\frac{1}{T}\Big(y+\tau x\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\Lambda-\sigma\sqrt{\Lambda}Z\Big\}\Big)-K\right)^{+}\right].$$

This is now a full expectation and can thus be written as an integral in terms of the unknown joint density function $f_{\Lambda,Z}(\lambda,z)$ as

$$\int_0^\tau\int_{-\infty}^{\infty}e^{-r\tau}\left(\frac{1}{T}\Big(y+\tau x\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\lambda-\sigma\sqrt{\lambda}z\Big\}\Big)-K\right)^{+}f_{\Lambda,Z}(\lambda,z)dzd\lambda.$$

The function being integrated is zero for all z greater than $d_-(\tau,\lambda,x,y)$. By changing the limits of integration to reflect this, we are able to drop the plus-function, leaving only

$$\int_0^\tau\int_{-\infty}^{d_-(\tau,\lambda,x,y)}e^{-r\tau}\left(\frac{1}{T}\Big(y+\tau x\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\lambda-\sigma\sqrt{\lambda}z\Big\}\Big)-K\right)f_{\Lambda,Z}(\lambda,z)dzd\lambda.$$

We now simply replace the dummy variables x and y with the original stochastic variables $S(t)$ and $Y(t)$, respectively, to obtain the desired result:

$$\int_0^{\tau}\int_{-\infty}^{d_-(\tau,\lambda,S(t),Y(t))} e^{-r\tau}\left(\frac{1}{T}\Big(Y(t)+\tau S(t)\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\lambda-\sigma\sqrt{\lambda}z\Big\}\Big)-K\right)f_{\Lambda,Z}(\lambda,z)dzd\lambda.$$

□

Theorem 3.10. *Using the notation of Theorem 3.9, let the value of the random variable Λ be agreed upon prior to the sale of the contract so that the random variable Λ is now just a parameter λ. Then the exact value of this modified Asian call option is given by*

$$v(\tau,\lambda,S(t),Y(t)) =$$

$$\frac{\tau S(t)}{T}e^{-r(\tau-\lambda)}N(d_+(\tau,\lambda,S(t),Y(t))) + \frac{Y(t)-KT}{T}e^{-r\tau}N(d_-(\tau,\lambda,S(t),Y(t))).$$

This serves as a first approximation to the value of the true Asian call option, with equality holding for at least one choice of λ.

Proof. The proof is a continuation of the proof of Theorem 3.9. We begin with an option price equal to

$$\widetilde{\mathbb{E}}\left[e^{-r\tau}\left(\frac{1}{T}\Big(y+\tau x\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\lambda-\sigma\sqrt{\lambda}Z\Big\}\Big)-K\right)^+\right],$$

where the random variable Λ has been replaced with the parameter λ. As Z is the only random variable in this expression, this full expectation can be written as an integral in terms of the known standard normal density $\varphi(z)$ as

$$\int_{-\infty}^{\infty} e^{-r\tau}\left(\frac{1}{T}\Big(y+\tau x\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\lambda-\sigma\sqrt{\lambda}Z\Big\}\Big)-K\right)^+\varphi(z)dz.$$

The function being integrated is zero for all z greater than $d_-(\tau,\lambda,x,y)$. By changing the limits of integration to reflect this, we are able to drop the plus-function, leaving only

$$\int_{-\infty}^{d_-(\tau,\lambda,x,y)} e^{-r\tau}\left(\frac{1}{T}\Big(y+\tau x\exp\Big\{\Big(r-\frac{1}{2}\sigma^2\Big)\lambda-\sigma\sqrt{\lambda}Z\Big\}\Big)-K\right)\varphi(z)dz.$$

The linearity of integration can now be used to separate this integral into the sum of several simpler integrals, namely

$$\frac{y-KT}{T}e^{-r\tau}\int_{-\infty}^{d_-(\tau,\lambda,x,y)}\varphi(z)dz + \frac{\tau x}{T}e^{-r(\tau-\lambda)}e^{-\sigma^2\lambda/2}\int_{-\infty}^{d_-(\tau,\lambda,x,y)} e^{-\sigma\sqrt{\lambda}z}\varphi(z)dz.$$

A change of variables is used to simplify the second integral, which causes a change in its upper limit of integration. This simplification enables both integrals to be written in terms of the cumulative normal distribution function $N(x)$ as

$$\frac{y-KT}{T}e^{-r\tau}N(d_-(\tau,\lambda,x,y)) + \frac{\tau x}{T}e^{-r(\tau-\lambda)}N(d_+(\tau,\lambda,x,y)).$$

We now simply replace the dummy variables x and y with the original random variables $S(t)$ and $Y(t)$, respectively, to obtain the desired result:

$$\frac{Y(t)-KT}{T}e^{-r\tau}N(d_-(\tau,\lambda,S(t),Y(t))) + \frac{\tau S(t)}{T}e^{-r(\tau-\lambda)}N(d_+(\tau,\lambda,S(t),Y(t))).$$

□

4. CONCLUSION

Although the risk-neutral conditional expectation approach to option pricing does not generalize from the European call option to the Asian call option quite as easily as the partial differential equations approach to option pricing, this natural extension of the Black-Scholes-Merton method to the Asian call option shows some hope of utility on several different fronts.

Expressing the option price as a double integral over an unknown joint probability distribution opens the door to approximation schemes in which this unknown distribution is assumed to be of a certain standard type. The Asian option price can then be calculated for these various cases. This might provide a sufficiently accurate approximation to prove useful in actual practice for pricing the Asian call option. It is also possible that this joint distribution function could be determined explicitly, in which case Theorem 3.9 provides an exact price for the Asian option at any time.

A generalization of the method used to obtain Theorem 3.10 might also provide such an exact price for the Asian option at any time. The authors sought to generalize the method by partitioning the time interval $[t, T]$ into n subintervals. The average price of the asset over each subinterval was then assumed to be the value of the asset at the beginning of that subinterval. Numerous mutually independent standard normal random variables Z_i were then defined in analogy to the definition of Z given above and an expression involving manyfold integration over many standard normal densities $\varphi(z_i)$ was obtained. Unfortunately, this complicated expression was not worked out carefully enough to be included in this paper as the possibility of lingering miscalculations rendered the final equation unreliable; however, although unwieldy, if this expression were carefully derived, it is possible that by taking the limit as n approaches ∞ that an exact price for the Asian option could be obtained. Another possible approach would be to set the first derivative of the result of Theorem 3.10 with respect to λ equal to zero in order to calculate its extrema. These maximum and minimum values would then provide upper and lower bounds, respectively, on the value of the Asian option, which might serve to determine the fair price of the Asian option with sufficient accuracy for actual practice. Skillfully incorporating the use of a computer algebra system in either of these approaches would probably be essential to completing the necessary calculations accurately.

5. ACKNOWLEDGEMENTS

The authors were partially supported by the Brigham Young University REU and NSF grant DMS-0453421.

REFERENCES

[1] Robert Jarrow and Phillip Protter. A short history of stochastic integration and mathematical finance: the early years, 1880-1970. *IMS Lecture Notes*, 45:75–91, 2004.

[2] Steven E. Shreve. *Stochastic calculus for finance I: the binomial asset-pricing model*. Springer, New York, NY, 2004.

[3] Steven E. Shreve. *Stochastic calculus for finance II: continuous-time models*. Springer, New York, NY, 2004.