APPLICATIONS OF STOCHASTIC CALCULUS TO FINANCIAL MODELING

BEN RAPPORT

Abstract. Following a brief discussion of measure theory and probability, we introduce stochastic processes, specifically Brownian motion, as mathematical models of randomness. Building upon this foundation, we derive fundamental results of Itô Calculus and use these results to solve Stochastic Differential Equations (SDEs). Alongside no-arbitrage arguments, we employ these tools to reproduce the Black-Scholes-Merton (BSM) option pricing model and the Margrabe Formula. We conclude by discussing Monte Carlo methods as a means of simulating stochastic processes.

Contents

1. Introduction

Most real world processes contain elements of randomness. In specific contexts, the impact of this randomness can be disregarded because of its small relative magnitude. However, in other contexts, this randomness is essential to describing a process' evolution. These random, or stochastic, processes are difficult to model using traditional calculus techniques as most of these processes form non-differentiable paths. In an effort to translate traditional calculus techniques and theory into the stochastic setting, Kiyosi Itô developed Itô Calculus to provide these random processes with a rigorous formalization. Itô Calculus retained many properties from traditional calculus while adding key new constructions, chief among them the Itô Integral and Itô's Formula. These two results, which produce the stochastic reformulations of integration and the chain rule respectively, have major applications throughout applied mathematics.

Date: AUGUST 29, 2020.

In modern finance stochastic processes are used to model price movements of securities in the stock market. One of the earliest pricing models, the BSM model, produces a PDE which describes how the value of an option changes over time in an arbitrage-free market. For European options, the solution to this differential equation can be written as a closed-form expression in which the characteristics of the underlying asset can be explicitly observed and interpreted. Many years later, William Margrabe built upon this model and derived a more general version of this PDE that applies to a broader range of instruments and derivatives.

In this paper, we will begin with a brief review of measure theory and some basic probabilistic concepts. We will then give a formalized introduction to stochastic processes and discuss some pertinent properties of Brownian motion. Using this knowledge as our foundation, we will then begin our treatment of Itô Calculus in which we will derive the Itô Formula and explore stochastic differential equations as a means to model financial instruments that depend on random price movements. Finally, we will consider the application of these models and review some numerical methods for simulating the stochastic behavior of more complex derivatives.

2. Review of Measure Theory and Probability

To begin this paper, we introduce several measure-theoretic and probabilistic definitions and concepts for consideration – many of which can be found in [1]. We denote Ω to be the outcome space, the set of all possible outcomes, for a given procedure. For example, the outcome space for a singular die roll is given by $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Definitions 2.1. Let Ω be given. Then a *σ-algebra* $\mathcal{F}$ on Ω is a family of subsets of Ω such that

(1) $\emptyset \in \mathcal{F}$
(2) $F \in \mathcal{F} \implies F^{\mathcal{C}} \in \mathcal{F}$, where $F^{\mathcal{C}} = \Omega \setminus F$ is the compliment of F in Ω
(3) $A_1, A_2, ... \in \mathcal{F} \implies A := \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*. For any $\mathcal{U} = \bigcup\limits_i F_i \subset \Omega$, we define

$$\mathcal{H}_{\mathcal{U}} := \bigcap \{\mathcal{H}; \mathcal{H} \text{ is a } \sigma\text{-algebra of } \Omega \text{ such that } \mathcal{U} \subset \mathcal{H}\}$$

to be the σ-algebra generated by $\mathcal{U}$. As is shown in [1], $\mathcal{H}_{\mathcal{U}}$ is also the smallest σ-algebra which contains $\mathcal{U}$. The *Borel* σ-algebra $\mathcal{B}$ is the σ-algebra generated by the collection of all open subsets of $\mathbf{R^n}$. A *Borel set* is an element $B \in \mathcal{B}$.

Definitions 2.2. We define a *probability measure* on $(\Omega, \mathcal{F})$ to be the function $P : \mathcal{F} \to [0, 1]$ with the following properties:

(1) $P(\emptyset) = 1 - P(\Omega) = 0$
(2) if $A_1, A_2, ... \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$

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

Definitions 2.3. Given a probability space $(\Omega, \mathcal{F}, P)$, a set F is *$\mathcal{F}$-measurable* if $F \in \mathcal{F}$. A function $f : \Omega \to \mathbf{R^n}$ is said to be $\mathcal{F}$-measurable if for all open sets $U \in \mathbf{R^n}$ we have that

$$f^{-1}(U) := \{\omega \in \Omega; f(\omega) \in U\} \in \mathcal{F}.$$

Definitions 2.4. A *random variable* on $(\Omega, \mathcal{F}, P)$ is an $\mathcal{F}$-measurable function $X : \Omega \to \mathbf{R^n}$. The σ-algebra generated by the random variable X is given by $\mathcal{H}_X = \{X^{-1}(B); B \in \mathcal{B}\}$. For any random variable, X, we define its *distribution* to be $\mu_X(B) := P(X^{-1}(B)) = P(X \in B)$ where B is a *Borel set*.

Definitions 2.5. A collection of measurable sets $\{F_i\}$ is *independent* if

$$P(F_1 \cap F_2 \cap ... \cap F_n) = P(F_1) \cdot P(F_2) \cdots P(F_n).$$

Suppose $X_1, X_2, ..., X_n$ are random variables on $(\Omega, \mathcal{F}, P)$. The collection of random variables $\{X_1, X_2, ..., X_n\}$ is independent if the collection of their generated σ-algebras $\{\mathcal{H}_{X_1}, \mathcal{H}_{X_2}, ..., \mathcal{H}_{X_n}\}$ is independent.

Remark 2.6. Any two elements of an independent set are independent. Thus, if $\{X_1, X_2, ..., X_n\}$ is independent, then X_1 and X_2 are independent.

Definitions 2.7. Let X be a random variable which assumes values in $\mathbf{R^n}$. If

$$\int_\Omega |X(\omega)| dP(\omega) < \infty, \text{ then } E[X] := \int_\Omega X(\omega) dP(\omega) = \int_{\mathbf{R^n}} x d\mu_X(x)$$

where $E[X]$ is the *expectation* of X. If $f : \mathbf{R^n} \to \mathbf{R}$ is $\mathcal{B}$-measurable and

$$\int_\Omega |f(X(\omega))| dP(\omega) < \infty, \text{ then } E[f(X)] := \int_\Omega f(X(\omega)) dP(\omega) = \int_{\mathbf{R^n}} f(x) d\mu_X(x)\,.$$

If $\mathcal{H} \subset \mathcal{F}$ is a σ-algebra, then the *conditional expectation* of X given $\mathcal{H}$, written as $E[X|\mathcal{H}]$, is a function from Ω to $\mathbf{R^n}$ satisfying the following two properties:

(1) $E[X|\mathcal{H}]$ is $\mathcal{H}$ measurable
(2) $\int_H E[X|\mathcal{H}]\, dP = \int_H X dP$, for all $H \in \mathcal{H}$

The *variance* of X is defined as $Var[X] := E[X^2] - [E[X]]^2 = E[(X - E[X])^2]$.

Proposition 2.8. *For conciseness, we will now present several properties of expectation and variance. Their proofs can be found in* [2]. *Suppose X is given as above. Let $a, b \in \mathbf{R}$, $\mathcal{H} \in \mathcal{F}$ and $Y : \Omega \to \mathbf{R^n}$ be another random variable such that $E[|Y|] < \infty$. We then observe*

(1) $E[aX + bY|\mathcal{H}] = aE[X|\mathcal{H}] + bE[Y|\mathcal{H}]$
(2) $E[E[X|\mathcal{H}]] = E[X]$
(3) $E[X|\mathcal{H}] = X$ *if X is $\mathcal{H}$-measurable.*
(4) $E[X|\mathcal{H}] = E[X]$ *if X and $\mathcal{H}$ are independent.*
(5) $Var[aX + b] = a^2 Var[X]$

Definition 2.9. The *cumulative distribution function* (CDF) of a random variable X is the function $F_X : \mathbf{R^n} \to [0, 1]$ where $F_X(\mathbf{x}) = P(X \leq \mathbf{x}) = \mu_X(\leq \mathbf{x})$. A random variable is said to be continuous if its CDF is continuous.

Definition 2.10. Let X be a continuous random variable $X : \Omega \to \mathbf{R}$. If there exists a function $f : \mathbf{R} \to [0, 1]$ such that

$$F_X(x) = \int_{-\infty}^{x} f(t) dt\,,$$

then $f(x)$ is the *density function* of X.

Definitions 2.11. A *normally distributed* random variable $X : \Omega \to \mathbf{R}$ with E[X] $= \mu$ and Var[X] $= \sigma^2$ has cumulative distribution function

$$F_X(b) = \int_{-\infty}^{b} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx\,. \tag{2.12}$$

For such a random variable we write $X \sim N(\mu, \sigma^2)$. A random variable Z is said to have *lognormal* distribution if $log(Z) \sim N(\mu, \sigma^2)$. A random variable Y is said to have *standard normal* distribution if $Y \sim N(0, 1)$. The standard normal CDF and density function are given by

$$\Phi(b) = F_Y(b) = \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}\, dx = \int_{-\infty}^{b} f(x)\, dx\,.$$

3. Stochastic Processes

While it is possible to model randomness in higher dimensions, in this paper we will focus exclusively on 1-D models, in particular 1-D Brownian motion.

Definition 3.1. A *stochastic process* in $(\Omega, \mathcal{F}, P)$ over an interval T is a parameterized collection of random variables $\{X_t(\omega)\}_{t\in T}$ where $X : T \times \Omega \to \mathbf{R}$.

Definition 3.2. A *filtration* on $(\Omega, \mathcal{F})$ is an increasing sequence of σ-algebras $\mathcal{K} = \{\mathcal{F}_t\}_{t\geq 0}$ such that for all $s \in [0, t)$ we have that $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$.

Remark 3.3. Let $\mathcal{K} = \{\mathcal{F}_t\}_{t\geq 0}$ be a filtration and $f : T \times \Omega \to \mathbf{R}$ be a random variable on $(\Omega, \mathcal{F}, P)$. A process $\{f_t(\omega)\}_{t\geq 0}$ is $\mathcal{F}_t$-*adapted* if $f_t(\omega)$ is $\mathcal{F}_t$-measurable for every $t \geq 0$. We define the *natural filtration* of a process $\{X_t(\omega)\}_{t\geq 0}$ to be the collection of σ-algebras $\{\mathcal{F}_t\}$ generated by $\{X_t(\omega)\}_{t\geq 0}$. It follows that a stochastic process is adapted with respect to its natural filtration.

Definition 3.4. A *martingale* in $(\Omega, \mathcal{F}, P)$ is an $\mathcal{F}_t$-adapted stochastic process $\{M_t\}_{t\geq 0}$ with the following properties:

(1) $E[|M_t|] < \infty$ for all $t \geq 0$
(2) $E[M_s|\mathcal{F}_t] = M_t$ for all $s \geq t$, which implies $E[M_s - M_t|\mathcal{F}_t] = 0$

Remark 3.5. Consider a game whose value at time t is modeled by a martingale, M_t. Over any interval, the expected change in value, or *return*, of the game is 0. As such, the second martingale property is often refered to as the *fair game property*.

Definition 3.6. We define a *Brownian motion* about the origin with drift μ and variance σ^2 to be the stochastic process $\{B_t(\omega)\}_{t\geq 0}$ for which

(1) $B_0 = 0$
(2) For $s < t$, the random variable $Y(t, \omega) = B_t(\omega) - B_s(\omega)$ is normally distributed with $E[Y] = \mu(t - s)$ and $Var[Y] = \sigma^2(t - s)$
(3) If $s < t$, then the random variable $Y = B_t(\omega) - B_s(\omega)$ is independent of the natural filtration of $\{B_p(\omega)\}_{s\geq p\geq 0}$ (Independent Increments)
(4) $P[B : T \times \omega \to B_t(\omega)$ is continous with respect to $t] = 1$

A Brownian motion is standard if it has drift $\equiv 0$ and variance $\equiv 1$. From this definition, it can be shown that a standard Brownian motion B_t is a *Gaussian process* and thus

$$E[e^{iu(B_t - B_s)}] = e^{-\frac{u^2(t-s)}{2}} \text{ for } u \in \mathbf{R} \text{ and } i = \sqrt{-1}\,. \tag{3.7}$$

Given its continuous nature, Brownian motion can be thought of as a random path.

Remark 3.8. Any Brownian motion W_t about the origin with drift μ and variance σ^2 may be described by $W_t = \mu t + \sigma B_t$, where B_t is standard Brownian motion. Given this fact, B_t will be used to represent standard Brownian motion from this point onward.

Examples 3.9. Let $(\Omega, \mathcal{F}, P)$ be given. It follows from Proposition 2.8 that

$$E[(B_t - B_s)^2] = E[(B_t - B_s)^2 - E[(B_t - B_s)]^2] = Var[(B_t - B_s)] = t - s\,.$$

Often we will compute $E[(B_{t_j+1} - B_{t_j})^2]$ which, as shown above, equals $t_{j+1} - t_j$. Moreover, for $s \geq t > 0$, we observe

$$E[B_s|\mathcal{F}_t] = E[B_s - B_t|\mathcal{F}_t] + E[B_t|\mathcal{F}_t] = B_t.$$

Through manipulations of (2.12) it can be shown that $E[|B_t|] < \infty$ for all t, at which point we may conclude that standard Brownian motion is in fact a martingale. A standard Brownian motion may thus be thought of as a fair game whose variance over an interval T is exactly the length of that interval.

Proposition 3.10. $E[(B_t - B_s)^4] = 3(t-s)^2$. *A proof can be found in the Appendix.*

4. Itô Calculus

Before we begin our treatement of Itô Calculus, we must first establish a class of functions for which these concepts can be applied.

Definition 4.1. A function $f(t,\omega) : [0,\infty) \times \Omega \to \mathbf{R}$ is *nice* on $[S,T)$ if

(1) $(t,\omega) \to f(t,\omega)$ is $\mathcal{B} \times \Omega$-measurable ($\mathcal{B} :=$ Borel $\sigma - algebra$)
(2) $f(t,\omega)$ is $\mathcal{F}_t$-*adapted*
(3) $E[\int_S^T (f(t,\omega))^2\, dt] < \infty$
(4) $P[\int_S^T (f(t,\omega))^2\, dt < \infty] = 1$, for any probability measure P

The last condition is stricter than necesary, but saves space. Additonally, a nice function is said to be *elementary* if it has the form $\phi(t,\omega) = \sum_j e_j(\omega) \cdot \mathbb{1}_{\{t_j, t_{j+1}\}}(t)$. Similarly, $\int_S^T \phi(t,\omega)\, dB_t(\omega) := \sum_j e_j(\omega) \cdot \Delta B_j$. The notation ΔB_j is used extensively in this section and assumes the existence of a partition $\pi = \{t_o, .., t_n\}$. For any stochastic process we denote $\Delta X_j := X_{t_{j+1}} - X_{t_j}$, and $\Delta t_j := t_{j+1} - t_j$.

Definition 4.2. Suppose $u(t,\omega)$, $v(t,\omega)$ are nice on $[0,\infty)$. We define an *Itô process* to be the stochastic process $\{X_t\}_{t\geq 0}$ in $(\Omega, \mathcal{F}, P)$ which satisfies

$$X_0 \in \mathbf{R},\ dX_t = u(t,\omega)dt + v(t,\omega)dB_t \text{ for } t \geq 0\,. \tag{4.3}$$

Theorem 4.4 (The Itô Isometry for Elementary Functions)**.** *Let $\phi(t,\omega)$ be a bounded elementary function. Then*

$$E[(\int_S^T \phi(t,\omega)\, dB_t(\omega))^2] = E[(\int_S^T \phi(t,\omega)^2\, dt)]\,. \tag{4.5}$$

Proof. Observe, if $i \neq j$, then $E[e_i \cdot e_j \cdot \Delta B_j \cdot \Delta B_i] = E[e_i \cdot e_j \cdot \Delta B_i] \cdot E[\Delta B_j] = 0$ because Brownian motion has independent increments. Hence,

$$\begin{aligned} E[(\int_S^T \phi(t,\omega)\, dB_t(\omega))^2] &= E[(\sum_j e_j(\omega) \cdot \Delta B_j)(\sum_i e_i(\omega) \cdot \Delta B_i)] \\ &= \sum_{i,j} E[e_j \cdot e_i \cdot \Delta B_j \cdot \Delta B_i] \\ &= \sum_j E[e_j^2(\omega) \cdot (\Delta B_j)^2] \\ &= \sum_j E[e_j^2(\omega)] \cdot (\Delta t_j) \\ &= E[\int_S^T \phi(t,\omega)^2\, dt] \text{ as needed.} \end{aligned}$$

□

We will now present the Itô Integral and give an outline of its proof, noting the similarities in construction to the Lebesgue and Riemann Integrals. The following outline derives from the proof given in [1].

Theorem 4.6 (The Itô Integral)**.** *Let f be a nice function on $[S,T)$. Then the Itô Integral over that interval is*

$$\mathcal{I}[f](w) := \int_S^T f(t,\omega)\, dB_t(\omega) = \lim_{n\to\infty} \int_S^T \phi_n(t,\omega)\, dB_t(\omega) \tag{4.7}$$

where $\{\phi_n\}$ is a sequence of elementary functions for which

$$E[\int_S^T (f(t,\omega) - \phi_n(t,\omega))^2\, dt] \to 0 \text{ as } n \to \infty\,. \tag{4.8}$$

Proof. In the same way step functions are used to approximate bounded functions in the derivation of the Riemann Integral, the Itô Integral uses elementary nice functions to approximate nice functions. In doing so, we will generate a sequence of elementary functions which satisfy (4.8). The first step of the proof is to approximate nice functions using a sequence of bounded nice functions. Specifically, for any nice function f, we must produce a sequence $\{h_n\}$ of bounded nice functions such that

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

Next, we replicate this process to show that any bounded nice function can be approximated by a sequence of bounded and continuous nice functions. Finally, we would show that any bounded and continuous nice function can be approximated by a sequence of elementary nice functions. It then follows that this sequence of elementary nice functions $\{\phi_n\}$ yields the desired convergence in expectation. We thus define the Itô Integral of f on (S,T) as

$$\mathcal{I}[f](w) := \int_S^T f(t,\omega)\, dB_t(\omega) = \lim_{n\to\infty} \int_S^T \phi_n(t,\omega)\, dB_t(\omega)\,,$$

where the convergence of the rightmost expression follows from the Itô Isometry and the convergence of Cauchy sequences. □

A corollary to Theorem 4.4 found in [1] extends the Itô Isometry from elementary functions to all nice functions and will be used in later theorems. Moreover, given this construction, it follows that the Itô process given by (4.3) is also described by

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

Lemma 4.9. *Let $f(s,\omega)$ be a nice function on $[0,t)$ and define $f_j = f(s_j,\omega)$. Then as $\Delta s_j \to 0$, we observe $\sum_j f_j \cdot (\Delta B_j)^2 \to \sum_j f_j \cdot \Delta s_j = \int_0^t f_j\, ds$.*

Proof. Following from analysis presented in the Itô Isometry, we find that if $i \neq j$, then $((\Delta B_i)^2 - \Delta t_i)$ and $(f_i)(f_j)((\Delta B_j)^2 - \Delta t_j)$ are independent, and thus the expectation of their product is 0. Following from the construction of the Itô Integral, we compute

$$E[(\sum_j f_j(\Delta B_j)^2 - \sum_j f_j \Delta t_j)^2] = E[\sum_{j,i} f_j f_i((\Delta B_j)^2 - \Delta t_j)((\Delta B_i)^2 - \Delta t_i)]\,.$$

Next, we expand the square, recall (3.10), use the independence argument presented above, and observe

$$\begin{aligned}
E[\sum_j f_j^2 \cdot ((\Delta B_j)^2 - \Delta t_j)^2] &= E[\sum_j f_j^2 \cdot ((\Delta B_j)^4 - 2(\Delta B_j)^2(\Delta t_j) + (\Delta t_j)^2)] \\
&= \sum_j E[f_j^2] \cdot E[(\Delta B_j)^4 - 2(\Delta B_j)^2(\Delta t_j) + (\Delta t_j)^2] \\
&= \sum_j E[f_j^2] \cdot (3(\Delta t_j)^2 - 2(\Delta t_j)^2 + (\Delta t_j)^2) \\
&= \sum_j E[f_j^2] \cdot (2(\Delta t_j)^2) \to 0 \\
&\quad \text{as } \Delta t_j \to 0 \text{ given that } f \text{ is nice.}
\end{aligned}$$

□

Setting $f \equiv 1$, this proof illustrates a fundamental construction of Itô Calculus: $(dB_t)^2 = dt$. Another important computation is revealed in the final line; as $\Delta t_j \to 0$, the summation $\sum_j (\Delta t_j)^2 \to 0$. Similarly, as we will show in (4.15) and (4.16), as $\Delta t_j \to 0$, the summation $\sum_j (\Delta t_j)(\Delta B_j) \to 0$. This is expressed in differential notation through the equations $(dt)^2 = 0$ and $(dt)(dB_t) = 0$.

Theorem 4.10 (Itô's Formula)**.** *Let X_t be an Itô process. Suppose $g(t,x) \in C^2([0,\infty) \times \mathbf{R})$. Then $Y_t = g(t,X_t)$ is also an Itô process which satisfies the following Stochastic Differential Equation*

$$dY_t = \partial_t g(t,X_t)dt + \partial_x g(t,X_t)dX_t + \frac{1}{2}\partial_{xx} g(t,X_t)(dX_t)^2\,. \tag{4.11}$$

Proof. By Taylor's Theorem, we may rewrite $Y_t = g(t,X_t)$ as

$$\begin{aligned}
g(t,X_t) \;=\; g(0,X_0) \;+\; \sum_{j<n} \frac{\partial g}{\partial t}\,\Delta t_j \;+\; \sum_{j<n} \frac{\partial g}{\partial x}\,\Delta X_j \;+\; \sum_{j<n} \frac{1}{2}\frac{\partial^2 g}{\partial t^2}(\Delta t_j)^2 \\
+ \sum_{j<n} \frac{\partial^2 g}{\partial t \partial x}(\Delta X_j)(\Delta t_j) + \sum_{j<n} \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(\Delta X_t)^2 + \sum_{j<n} R_j
\end{aligned} \tag{4.12}$$

where R_j is the remainder term of each polynomial. We then demonstrate the desired result (4.11) by evaluating each summation separately. As $\Delta t_j \to 0$, we observe that

$$\sum_{j<n} \frac{\partial g}{\partial t}\,\Delta t_j \to \int_0^t \frac{\partial g(s,X_s)}{\partial s} ds \text{ and } \sum_{j<n} \frac{\partial g}{\partial x}\Delta X_j \to \int_0^t \frac{\partial g(s,X_s)}{\partial x} dX_s\,. \tag{4.13}$$

Using similar arguments to those found in the proof of Lemma 4.9, we conclude

$$\sum_{j<n} \frac{1}{2}\frac{\partial^2 g}{\partial t^2}(\Delta t_j)^2 \to 0 \text{ as } \Delta t_j \to 0 \text{ and } g(t,X_t) \in C^2. \tag{4.14}$$

Next, we reintroduce the technique of approximation via elementary functions as in Theorem 4.6 to calculate ΔX_j. Without loss of generality, we suppose $dX_t = u(t,X_t)dt + v(t,X_t)dB_t$. Approximating u and v with elementary nice functions, we deduce $\Delta X_j = u(t_j,\omega)\Delta t_j + v(t_j,\omega)\Delta B_j$. Set $u(t_j,\omega) = u_j$, $v(t_j,\omega) = v_j$. Thus we compute

$$\sum_{j<n} \frac{\partial^2 g}{\partial t \partial x}(\Delta X_j)(\Delta t_j) = \sum_{j<n} \frac{\partial^2 g}{\partial t \partial x} u(t_j,\omega)(\Delta t_j)^2 + \sum_{j<n} \frac{\partial^2 g}{\partial t \partial x} v(t_j,\omega)(\Delta B_j)(\Delta t_j)\,.$$

Using the fact that u is approximated by elementary nice functions and referencing Lemma 4.9 for technique, we then claim that the first term on the right hand side of the equality tends to 0 as $\Delta t_j \to 0$. Moreover,

$$E[(\sum_{j<n} \frac{\partial^2 g}{\partial t \partial x} v_j(\Delta B_j)(\Delta t_j) - 0)^2] = \sum_{j<n} E[(\frac{\partial^2 g}{\partial t \partial x} v_j)^2] E[(\Delta B_j)^2 (\Delta t_j)^2)] \tag{4.15}$$

$$= \sum_{j<n} E[(\frac{\partial^2 g}{\partial t \partial x} v_j)^2](\Delta t_j)^3 \to 0 \quad as\ \Delta t_j \to 0,\ g \in C^2 \text{ and v is nice.} \tag{4.16}$$

Hence, the second term tends to 0 as well. Next notice that $(\Delta X_j)^2 = (u_j \Delta t_j)^2 + 2u_j v_j \Delta t_j \Delta B_j + (v_j \Delta B_j)^2$; hence, the penultimate term of (4.12) becomes

$$\frac{1}{2}\sum_{j<n} \frac{\partial^2 g}{\partial x^2}(\Delta X_t)^2 = \frac{1}{2}\sum_{j<n} \frac{\partial^2 g}{\partial x^2}(u_j \Delta t_j)^2 + \sum_{j<n} \frac{\partial^2 g}{\partial x^2} u_j v_j \Delta t_j \Delta B_j + \frac{1}{2}\sum_{j<n} \frac{\partial^2 g}{\partial x^2}(v_j \Delta B_j)^2\,.$$

Following from the computations above, we find that the first two summations tend to 0 as $\Delta t_j \to 0$. Using Lemma 4.9, and setting $f = \frac{1}{2}\frac{\partial^2 g(s,\omega)}{\partial x^2}(v_j)^2$ we find

$$\frac{1}{2}\sum_{j<n} \frac{\partial^2 g}{\partial x^2}(v_j \Delta B_j)^2 \to \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(v_j)^2\, ds$$

$$\text{and as a result, } \frac{1}{2}\sum_{j<n} \frac{\partial^2 g}{\partial x^2}(\Delta X_t)^2 \to \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(v_j)^2\, ds\,.$$

Similar computations can be used to show that $\sum_{j<n} R_j \to 0$ as $\Delta t_j \to 0$. Using the result above and substituting (4.13), (4.14), and (4.16) into (4.12) we find

$$g(t,X_t) = g(0,X_0) + \int_0^t \frac{\partial g(s,X_s)}{\partial s} ds + \int_0^t \frac{\partial g(s,X_s)}{\partial x} dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(v_j)^2\, ds\,.$$

Recalling that $(dX_s)^2 = u^2(ds)^2 + 2uv(dB_s)(ds) + v^2(dB_s)^2 = v^2 ds$, we find

$$\int_0^t d(g) = \int_0^t \frac{\partial g}{\partial s} ds + \int_0^t \frac{\partial g}{\partial x} dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(dX_s)^2 .$$

Therefore we conclude

$$\begin{aligned} d(Y_t) &= d(g(t, X_t)) \\ &= \left(\frac{\partial g(t, X_t)}{\partial t} + \frac{v^2(t,\omega)}{2} \cdot \frac{\partial^2 g(t, X_t)}{\partial x^2}\right) dt + \frac{\partial g(t, X_t)}{\partial x} dX_t \\ &= \frac{\partial g(t, X_t)}{\partial t} dt + \frac{\partial g(t, X_t)}{\partial x} dX_t + \frac{1}{2}\frac{\partial^2 g(t, X_t)}{\partial x^2}(dX_t)^2 \text{ is an Itô process.} \quad \square \end{aligned}$$

For our proof of the Margrabe Formula in the Applications to Financial Modeling section, we use the 2-dimensional interpretation of the Itô Formula. The following definition and theorem can be found in [1].

Definition 4.17. Let μ_i, σ_i be nice functions ($i \in (1, 2)$). We define a 2-*dimensional Itô process* to be

$$d\mathbf{X_t} = \begin{cases} dX_t = \mu_1(t,\omega)dt + \sigma_1(t,\omega)dB_t \\ dY_t = \mu_2(t,\omega)dt + \sigma_2(t,\omega)dW_t \end{cases}$$

where B_t and W_t are both 1-dimensional standard Brownian motions.

Theorem 4.18. *Let $d\mathbf{X_t}$ be a 2-dimensional Itô process as defined above and let $g(t, \mathbf{X_t}) = g(t, X_t, Y_t)$ be a C^2 map $g : [0, \infty] \times R^2 \to R$. Then $Y_t = g(t, \mathbf{X_t})$ is also an Itô process which satisfies*

$$\begin{aligned} dY_t = \partial_t g(t, \mathbf{X_t})dt + \partial_x g(t, \mathbf{X_t})dX_t + \partial_y g(t, \mathbf{X_t})dY_t + \\ \frac{1}{2}[\partial_{xx} g(t, \mathbf{X_t})(dX_t)^2 + 2\partial_{xy} g(t, \mathbf{X_t})(dX_t)(dY_t) + \partial_{yy} g(t, \mathbf{X_t})(dY_t)^2] . \end{aligned} \tag{4.19}$$

5. Stochastic Differential Equations

While we will not discuss differential equations in depth, we present some examples of common SDEs in mathematical finance and their solutions. We will focus primarily on Itô processes.

Example 5.1. Let X_t be an Itô process satisfying

$$dX_t = \mu_t X_t dt + \sigma_t dB_t \text{ where } \mu_t = \mu(t, X_t),\ \sigma_t = \sigma(t, X_t) \in C^2([0, \infty) \times \mathbf{R})$$

To solve this SDE we multiply both sides of the equality by an integrating factor:

$$e^{-\mu_t t} dX_t = e^{-\mu_t t}(\mu_t X_t dt + \sigma_t dB_t) .$$

Thus, letting $g(t, X_t) = e^{-\mu_t t} X_t = Y_t$, we apply Itô's Formula to obtain

$$d(Y_t) = -\mu_t e^{-\mu_t t} X_t dt + e^{-\mu_t t} dX_t + 0 \cdot (dX_t)^2 .$$

Substituting dX_t into $d(Y_t)$, we find

$$d(Y_t) = -e^{-\mu_t t}\mu_t X_t dt + e^{-\mu_t t}\mu_t X_t dt + e^{-\mu_t t}\sigma_t dB_t .$$

Therefore $d(e^{-\mu_t t} X_t) = e^{-\mu_t t}\sigma_t dB_t$, and thus $\int_0^T d(e^{-\mu_t t} X_t) = \int_0^T e^{-\mu_t t}\sigma_t dB_t$. Solving this integral we find $e^{-\mu_T T} X_T - X_0 = \int_0^T e^{-\mu_t t}\sigma_t dB_t$ and equivalently

$$X_T = X_0 e^{\mu_T T} + \int_0^T e^{(\mu_T T - \mu_t t)}\sigma_t dB_t .$$

The above example is a variation of the famous Ornstein-Uhlenbleck equation. Altering this equation slightly, we get the differential equation for Geometric Brownian motion (GBM). Because GBM is a strictly positive stochastic processes ($X_t > 0$ for all t assuming $X_0 > 0$) and has lognormal distribution, it is used extensively to model financial market behaviors. We show the derivation of its solution below.

Example 5.2. Let X_t be a stochastic process satisfying the following SDE:

$$dX_t = X_t[\mu_t dt + \sigma_t dB_t] \quad \mu_t = \mu(t, X_t),\ \sigma_t = \sigma(t, X_t)\ \in C^2([0,\infty) \times \mathbf{R})$$

We solve this SDE via careful selection of g. Let $g(t, X_t) = ln(X_t) = Y_t$. Hence,

$$\begin{aligned} d(Y_t) &= \frac{1}{X_t} dX_t - \frac{1}{2(X_t)^2}(dX_t)^2 \\ &= \mu_t dt + \sigma_t dB_t - \frac{1}{2}[\mu_t dt + \sigma_t dB_t]^2 \\ &= \mu_t dt + \sigma_t dB_t - \frac{1}{2}[(\mu_t)^2(dt)^2 + 2\mu_t\sigma_t(dt)(dB_t) + {\sigma_t}^2(dB_t)^2] \\ &= (\mu_t - \frac{1}{2}{\sigma_t}^2)dt + \sigma_t dB_t\,. \end{aligned}$$

Integrating both sides of this expression we observe

$$\begin{aligned} ln(X_T/X_t) &= \int_t^T d(Y_s) \\ &= \int_t^T (\mu_s - \frac{1}{2}{\sigma_s}^2)ds + \int_t^T \sigma_s dB_s \end{aligned}$$

Consequently, we derive $X_T = X_t e^{\int_t^T (\mu_s - \frac{1}{2}{\sigma_s}^2)ds + \int_t^T \sigma_s dB_s}$. If μ and σ are constants in $\mathbf{R}$, then we define the stochastic process $\{X_T\}_{T \geq t}$ to be a Geometric Brownian motion with drift μ and variance σ^2. For such a Brownian motion we claim

$$\begin{aligned} X_T &= X_t e^{[(T-t)(\mu - \frac{1}{2}\sigma^2) + \sigma(B_T - B_t)]} \\ &= X_t e^{[(T-t)(\mu - \frac{1}{2}\sigma^2) + \sigma\sqrt{T-t}\cdot Y]} \end{aligned} \tag{5.3}$$

where $Y \sim N(0,1)$. As such, we compute the conditional expectation

$$\begin{aligned} E[X_T|\mathcal{F}_t] &= E[X_t e^{[(T-t)(\mu - \frac{1}{2}\sigma^2) + \sigma\sqrt{T-t}\cdot Y]}|\mathcal{F}_t] \\ &= X_t e^{(T-t)(\mu - \frac{1}{2}\sigma^2)} \cdot E[e^{\sigma\sqrt{T-t}\cdot Y}|\mathcal{F}_t] \\ &= X_t e^{(T-t)(\mu - \frac{1}{2}\sigma^2)} \cdot e^{\frac{1}{2}\sigma^2(T-t)} \\ &= X_t e^{\mu(T-t)}\,. \end{aligned} \tag{5.4}$$

Lastly, let $Z = (T-t)(\mu - \frac{1}{2}\sigma^2) + \sigma\sqrt{T-t}\cdot Y = log(\frac{X_T}{X_t})$. We then assert that $E[Z] = (\mu - \frac{\sigma^2}{2})(T-t)$ and $Var[Z] = \sigma^2(T-t)$. It follows that

$$Z = log(\frac{X_T}{X_t}) \sim N((\mu - \frac{\sigma^2}{2})(T-t), \sigma\sqrt{T-t})\,;$$

hence, we claim that X_T has lognormal distribution.

6. Applications to Financial Modeling

The stochastic modeling techniques developed in the previous sections are readily applied to finance. A *security* x is an object which has value $V(x) = X$ and can be bought and sold. An *asset* is a security whose change in value can be described using SDEs. A *derivative* of an asset is a security whose value is a function of time and the asset's value. A *market*, M, is the set of all securities. The *risk-free rate* is a security whose value R_t is a function of time and the other securities in the market $\bar{X}_t$. Formally, the risk free rate is modeled by $R_0 = 1$, $dR_t = r(t, \overline{X}_t)R_t dt$ and equivalently $R_T = R_t \cdot e^{\int_t^T r(s,\overline{X}_s)\,ds}$ for $T > t \geq 0$.

A portfolio is an allocation in a market given by $\theta = \{a^i\}$, where $a^i \in R$. In this way, a portfolio can be thought of as a quantity of each security in a market. A strategy is an algorithm which buys and sells securities in a market, thus generating a collection of portfolios $\{\theta_t\}$. The value of a portfolio at time t, is the sum of the values of its securities and is denoted $V(\theta_t) = \theta_t \cdot \mathbf{X_t} = \sum a_t^i X_t^i$ where $\mathbf{X_t} = (R_t, \bar{X}_t)$. The return of a portfolio over an interval $[t, T]$ is given by $V(\theta_T) - V(\theta_t)$.

A strategy in M is an arbitrage when $P[V(\theta_T) > V(\theta_t)(1+\frac{R_T}{R_t})] = 1$. Conversely, a market is arbitrage-free when no strategy can be an arbitrage. In this way, we can think of an arbitrage-free market as a market in which the risk-free rate offers the best guaranteed return. Finally, a portfolio is self-financing if $dV(\theta_t) = \sum_i a_t^i dX_t^i$.

Definition 6.1. A *trading space* is a pair (M, θ) for which

(1) The market M is arbitrage-free and includes the risk-free rate.
(2) Every portfolio θ is self-financing.
(3) Securities in the market can be bought and sold at any time (Continuity).
(4) Asset values in the market are modeled by GBM.

Remark. The value of an asset in a trading space can also be thought of as its *price*. For such an asset X_t, we describe the instantaneous change of its price as dX_t.

One common example of a financial derivative is an *option*. Given an asset x with value X_t, an x-option grants its owner the ability to buy 1 unit of x at a specific point in time, or over a specific time interval, for a fixed price K.[1] Notice that an option grants its owner the *right* to buy the underlying asset, but does not obligate its owner to buy the asset – the owner of the option may choose NOT to buy the asset. A European option is an option wherein the owner of the option has the right to buy the underlying asset only at a specific point in time T.[2]

Suppose we want to sell an option, but do not want to be exposed to risk. We must derive a strategy such that at time T, the value of our portfolio is guaranteed to be greater than or equal to the value of the option which we sell. This strategy is derived via the Black-Scholes-Merton equation and is shown below.

[1] The derivative refered to as an option throughout this paper is more specifically known as a call-option. Likewise, a put-option grants its owner the ability to *sell* 1 unit of a security at a specific point in time, or over a specific time interval, for a fixed price K.

[2] Options are often used in the financial services industry as insurance contracts. For example, suppose that one asset made up a large percentage of our portfolio. This being the case, we might want to purchase a put-option so that if the asset were to dramatically decrease in value, we would still retain a baseline value of K.

Theorem 6.2 (Black-Scholes-Merton Equation)**.** *Let $\mathcal{T} = (M, \theta)$ be a trading space with an asset y, whose price is given by $\{S_t\}$. Suppose $f(t, S_t) \in \mathcal{T}$ is the price of a y-option whose value at time T is given by $f(T, S_T) = F(S_T) = max(S_T - K, O)$, for $K \in \mathbf{R}$ fixed. Then for $x = x(t) = S_t$*

$$\partial_t f(t,x) = r(t,x)(f(t,x) - x\partial_x f(t,x)) - \frac{(\sigma(t,x))^2 x^2}{2}\partial_{xx} f(t,x)\,. \tag{6.3}$$

Proof. The BSM strategy $\{\theta_t\}$ first sells an option at time 0 for a price $f(0, S_0)$. Simultaneously, the strategy creates a portfolio, θ_0, allocated between the risk-free rate and the asset y, such that $V(\theta_0) = f(0, S_0) = a_0^1 R_0 + a_0^2 S_0$. For $t \in [0, T]$, the strategy generates $\{\theta_t\}$ such that the value of the portfolio at time T satisfies

$$P[V(\theta_T) \geq F(S_T)] = 1\,. \tag{6.4}$$

The above inequality asserts that at time T the BSM strategy produces a portfolio θ_T with value greater than or equal to the value of the sold option almost surely. Due to the arbitrage-free nature of a trading space, we refine (6.4) as follows. For the sake of contradiction, suppose that there exists a strategy $\{\theta_t^*\}$ given by $\theta_t^* = (b_t^1, b_t^2, 0)$ such that $P[V(\theta_T) > F(S_T)] = 1$. Now consider the strategy $\{\theta_t'\}$ given by $\theta_t' = (b_t^1, b_t^2, -1)$. We would then observe

$$P[V(\theta_T') = V(\theta_T) - F(S_T) > f(0, S_0) - V(\theta_0) = V(\theta_0') = 0] = 1$$

$$\text{implying that } P[V(\theta_T') - V(\theta_0') > V(\theta_0')(1 + R_T) = 0] = 1\,.$$

This result would indicate that $\{\theta_t'\}$ is an arbitrage, contradicting the arbitrage-free nature of a trading space. Hence, we assert that for $t \in [0, T]$ the strategy must generate $\{\theta_t\}$ such that

$$P[V(\theta_T) = F(S_T)] = 1 \text{ where } \theta_t = (a_t^1, a_t^2, 0) \text{ for } t \in [0, T]\,. \tag{6.5}$$

Next, we will use a similar arbitrage argument to extend this result, in turn proving that $V(\theta_t) = f(t, S_t)$ for all $t \in (0, T)$. For the sake of contradiction, suppose that $V(\theta_t) > f(t, S_t)$ for some fixed set of points $t \in E = \{t_0, t_1, ...\} \subseteq [0, T]$. In this case, the strategy defined by

$$\theta_t' = \begin{cases} (a_t^1, a_t^2, -1) & \text{if } t \leq t_0 \\ (-a_t^1 + [V(\theta_{t_0}) - f(t_0, S_{t_0})], -a_t^2, 1) & \text{if } t > t_0 \end{cases}$$

is an arbitrage as

$$P[V(\theta_T') = F(S_T) - V(\theta_T) + \delta = \delta > 0 = V(\theta_0) - f(0, S_0) = V(\theta_0') = 0] = 1$$

where $\delta = [V(\theta_{t_0}) - f(t_0, S_{t_0})] > 0$. Therefore, E must be empty as there cannot exist any points t for which $V(\theta_t) \not> f(t, S_t)$. A similar argument, employing the strategy defined by

$$\theta_t' = \begin{cases} (-a_t^1, -a_t^2, 1) & \text{if } t \leq t_0 \\ (a_t^1 + [f(t_0, S_{t_0}) - V(\theta_{t_0})], a_t^2, -1) & \text{if } t > t_0 \end{cases}$$

can be made to show $V(\theta_{t_0}) \not< f(t_0, S_{t_0})$. Therefore, by trichotomy, we find that $f(t, S_t) = V(\theta_t)$ for all $t \in [0, T]$, which yields the crucial equality

$$dV(\theta_t) = d(f(t, S_t)) \; for \; all \; t \in [0, T)\,. \tag{6.6}$$

Let S_t be a GBM represented by $dS_t = S_t[\mu_t dt + \sigma_t dB_t]$ where $\mu_t = \mu(t, X_t)$, $\sigma_t = \sigma(t, X_t) \in C^2([0, \infty) \times \mathbf{R})$. As every portfolio in a trading space is self-financing, we assert $dV(\theta_t) = a_1 dR_t + a_2 dS_t$. By (6.6) we claim

$$d(V(\theta_t)) = d(f(t, S_t)),$$

$$\text{and thus } a_1 dR_t + a_2 dS_t = \partial_t f(t, S_t)dt + \partial_x f(t, S_t)dS_t + \frac{1}{2}\partial_{xx} f(t, S_t)(dS_t)^2 .$$

Noticing, $(dS_t)^2 = [S_t(\mu_t dt + \sigma_t dB_t)]^2 = S_t^2 \sigma_t^2 dt$ we proceed to show

$$a_1 dR_t + [a_2 - \partial_x f(t, S_t)] \cdot S_t[\mu_t dt + \sigma_t dB_t] = \partial_t f(t, S_t)dt + \frac{1}{2}\partial_{xx} f(t, S_t) S_t^2 \sigma_t^2 dt .$$

Recall, that we wanted the BSM strategy to be riskless, i.e. have a guarenteed return. To do so, we must constantly update our portfolio such that the random term in the above differential equation is always eliminated. Setting $a_2 = \partial_x f(t, S_t)$, and consequentially $a_1 = \frac{f(t,S_t) - \partial_x f(t,S_t)}{R_t}$, we compute

$$[f(t, S_t) - \partial_x f(t, S_t)]r(t, S_t)dt = \partial_t f(t, S_t)dt + \frac{1}{2}\partial_{xx} f(t, X_t) S_t^2 \sigma^2 dt$$

$$\text{and } \partial_t f(t, S_t) = [f(t, S_t) - \partial_x f(t, S_t)]r(t, S_t) - \frac{\sigma_t^2 S_t^2}{2}\partial_{xx} f(t, X_t)$$

which is the Black-Scholes-Merton Equation. □

The PDE derived in this theorem yields the risk-free pricing of a y-option with strike price K and expiration T. We consider this pricing schema to be risk-free because we can construct a portfolio of the risk-free rate and the underlying stock, for which we can entirely eliminate randomnness. Even though this schema is risk-free, the y-option itself need not be risk-free – the y-option will almost surely not have a guaranteed value at time T. However, the y-option does have an expected value at time T. We can use this expected value to derive the closed form risk-neutral solution to the BSM model, wherein the expected return of the y-option equals the return of the risk-free rate.

Corollary 6.7. *Suppose $r(t, S_t)$ and $\sigma(t, S_t)$ are constant. Then the risk-neutral price of the y-option described above is*

$$f(t, S_t) = S_t \Phi(d_1) - e^{-r(T-t)} K \Phi(d_2) \tag{6.8}$$

where $K \in \mathbf{R}$ is from 6.2 and d_1, d_2 are given by

$$d_1 = \frac{ln(\frac{S_t}{K}) + (r + \frac{\sigma^2}{2})(T - t)}{\sigma\sqrt{T - t}} \text{ and } d_2 = \frac{ln(\frac{S_t}{K}) + (r - \frac{\sigma^2}{2})(T - t)}{\sigma\sqrt{T - t}} .$$

Proof. We begin with a risk-neutral argument. If we desire the price of our y-option to be risk-neutral, then the expected return of the y-option should be equal to the expected return of an initial portfolio of $\frac{f(t,S_t)}{R_t}$ units of the risk free rate. Mathematically, we assert that

$$\frac{f(t, S_t)}{R_t} \cdot R_T - f(t, S_t) = E[F(S_T) - f(t, S_t)|\mathcal{F}_t]. \tag{6.9}$$

As $r(t, S_t)$ is constant, $\frac{R_T}{R_t}$ is $\mathcal{F}_t$-measurable and thus we simplify (6.9) to be

$$f(t, S_t)e^{r(T-t)} = E[F(S_T)|\mathcal{F}_t] . \tag{6.10}$$

Recalling (5.4), another risk-neutrality argument can be made to show

$$Z = e^{(T-t)(\mu-\frac{1}{2}\sigma^2)+\sigma\sqrt{T-t}\cdot Y} = e^{(T-t)(r-\frac{1}{2}\sigma^2)+\sigma\sqrt{T-t}\cdot Y}. \tag{6.11}$$

Using this result, we compute

$$\begin{aligned}
E[F(S_T)|\mathcal{F}_t] &= \int_{-\infty}^{\infty} max(S_t e^z - K, 0) f(z) dz \\
&= \int_{-\infty}^{ln(\frac{K}{S_t})} max(S_t e^z - K, 0) f(z) dz + \int_{ln(\frac{K}{S_t})}^{\infty} max(S_t e^z - K, 0) f(z) dz \\
&= \int_{ln(\frac{K}{S_t})}^{\infty} S_t e^z - K f(z) dz \\
&= S_t \int_{ln(\frac{K}{S_t})}^{\infty} e^z f(z) dz - K \int_{ln(\frac{K}{S_t})}^{\infty} f(z) dz \\
&= S_t \int_{ln(\frac{K}{S_t})}^{\infty} e^{(T-t)(\mu-\frac{1}{2}\sigma^2)+\sigma\sqrt{T-t}\cdot y} e^{-y^2} dy - K \int_{ln(\frac{K}{S_t})}^{\infty} e^{-z^2} dz
\end{aligned}$$

which following several lines of computation simplifies to be

$$= S_t e^{r(T-t)} \Phi\left(\frac{ln(\frac{S_t}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right) - K\Phi\left(\frac{ln(\frac{S_t}{K}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right).$$

Finally, we isolate $f(t, S_t)$ to derive the closed form BSM pricing equation

$$f(t, S_t) = S_t \Phi(d_1) - e^{-r(T-t)} K \Phi(d_2). \qquad \square$$

Before continuing, it is worth discussing some assumptions and limitations of the Black-Scholes-Merton model, the majority of which derive from the construction of the trading space. Firstly, the BSM model assumes that securities are bought and sold continously. In reality, there are specific intervals in which markets are open and securities can be exchanged – after normal market hours, securities are still bought and sold, but there is a signifcant decrease in the volume of exchanged securities. This leads to a drastically different market environment. Additionally, arbitrage strategies exist in realistic markets, but are almost never present for more than a microsecond. This fact combined with our assumption of continous trading leads to a certain tension in our result: if we assume the continuous exchange of assets, then in a more realistic model there could be instantaneous arbitrage strategies.[3]

Lastly, our assumption that assets follow GBM with constant drift and variance is an over-generalization of market behaviors. It is very rare for an asset to have constant variance, and even rarer for its drift to remain constant. However, BSM is simply a model, and should not be viewed as an all encompassing formula. BSM is a framework.

Now we will consider a derivative, whose payoff does not depend on a singular asset, but instead on two assets. One such example is a spread-option. An x/y-spread-option grants its owner the ability to exchange 1 unit of an asset x for 1 unit of an asset y at a specific point in time. Like an plain y-option, a spread-option grants its owner the right to exchange the assets, but does not obligate the owner

[3] One could construct a heuristic measure-theoretic argument that these moments of inconsistency are so minute that we may treat them as a set of measure zero, but this lacks formalization.

to do so. Unlike a y-option, a spread-option does not have a strike price, its payoff depends solely on the price of the two underlying assets at time T. The formula generalizing BSM to this environment is known as the Margrabe Formula, and is shown below

Corollary 6.12 (The Margrabe Formula)**.** *Let $\mathcal{T} = (M, \theta)$ be a trading space with two assets x, y, whose prices satisfy*

$$dX_t = X_t[\mu_1 dt + \sigma_1 dW_t] \text{ and } dY_t = Y_t[\mu_2 dt + \sigma_2 dB_t] \text{ where } \mu_1, \sigma_1, \mu_2, \sigma_2 \in R$$

and W_t and B_t are independent standard Brownian motions. Suppose that $f(t, X_t, Y_t) \in \mathcal{T}$ is the price of a x/y-spread-option whose value at time T is given by $f(T, X_T, Y_T) = F(X_T, Y_T) = max(X_T - Y_T, O)$. Then the risk-neutral price of the x/y-spread-option described above is

$$f(t, X_t, Y_t) = X_t\Phi(q_1) - Y_t\Phi(q_2) \tag{6.13}$$

where q_1 and q_2 are given by

$$q_1 = \frac{ln(\frac{S_1}{S_2}) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}} \text{ and } q_2 = \frac{ln(\frac{S_1}{S_2}) - \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}, \text{ for } \sigma = \sqrt{\sigma_1^2 + \sigma_2^2}.$$

Proof. The proof of the Margrabe formula is a direct extension of the BSM equation. Consider a self-financing portfolio

$$V(\theta_t) = a_t^1 R_t + a_t^2 X_t + a_t^3 Y_t \text{ such that } V(\theta_T) = F(X_T, Y_T)\,. \tag{6.14}$$

Using the same type of arbitrage arguments as those found in BSM, we have that

$$V(\theta_t) = f(t, X_t, Y_t) \text{ for all } t \in [0, T]$$

which yields the crucial equality:

$$dV(\theta_t) = d(f(t, X_t, Y_t))\,.$$

Following the same procedure as in Theorem 6.2, this time using the 2-d Itô Formula, we find that

$$\begin{aligned} a_t^1 dR_t + a_t^2 dX_t + a_t^3 dY_t &= \partial_t f(t, \mathbf{X_t})dt + \partial_x f(t, \mathbf{X_t})dX_t + \partial_y f(t, \mathbf{X_t})dY_t \\ &+ \frac{1}{2}[\partial_{xx} f(t, \mathbf{X_t})(dX_t)^2 + 2\partial_{xy} f(t, \mathbf{X_t})(dX_t)(dY_t) + \partial_{yy} f(t, \mathbf{X_t})(dY_t)^2]\,. \end{aligned} \tag{6.15}$$

Using analogous riskless arguments from Theorem 6.2, we set

$$a_2 = \partial_x f(t, \mathbf{X_t}), a_3 = \partial_y f(t, \mathbf{X_t}) \text{ and } a_1 = \frac{f(t, \mathbf{X_t}) - \partial_x f(t, \mathbf{X_t}) - \partial_y f(t, \mathbf{X_t})}{R_t}.$$

Recalling that $dR_t = r(t, \mathbf{X_t})R_t dt$, we assert

$$a_t^1 dR_t = r(t, \mathbf{X_t})(f(t, \mathbf{X_t}) - \partial_x f(t, \mathbf{X_t}) - \partial_y f(t, \mathbf{X_t}))dt\,.$$

We then reduce (6.15) to be

$$\begin{aligned} &r(t, \mathbf{X_t})(f(t, \mathbf{X_t}) - \partial_x f(t, \mathbf{X_t}) - \partial_y f(t, \mathbf{X_t}))dt \\ &= \partial_t f(t, \mathbf{X_t})dt + \frac{1}{2}[\partial_{xx} f(t, \mathbf{X_t})(dX_t)^2 + 2\partial_{xy} f(t, \mathbf{X_t})(dX_t)(dY_t) + \partial_{yy} f(t, \mathbf{X_t})(dY_t)^2]\,. \end{aligned} \tag{6.16}$$

Next, we look to expand $(dX_t)^2, (dX_t)(dY_t)$, and $(dY_t)^2$. We observe that

$$\begin{aligned} (dX_t)^2 &= [X_t(\mu_1 dt + \sigma_1 dW_t)]^2 \\ &= X_t^2(\mu_1^2 (dt)^2 + 2\mu_1\sigma_1 dt dW_t + \sigma_1^2 (dW_t)^2) \\ &= X_t^2 \sigma_1^2 dt. \end{aligned}$$

Symmetrically, we also find that $(dY_t)^2 = Y_t^2\sigma_2^2 dt$. Recalling that W_t and B_t are independent, we compute

$$\begin{aligned}(dX_t)(dY_t) &= [X_t(\mu_1 dt + \sigma_1 dW_t)][Y_t(\mu_2 dt + \sigma_2 dB_t)] \\ &= X_t Y_t[\mu_1\mu_2(dt)^2 + \mu_1\sigma_2 dt dB_t + \mu_2\sigma_1 dW_t dt + \sigma_1\sigma_2(dW_t)(dB_t)] \\ &= 0\,.\end{aligned}$$

Therefore, (6.16) simplifies to be

$$\begin{aligned}&r(t,\mathbf{X_t})(f(t,\mathbf{X_t}) - \partial_x f(t,\mathbf{X_t}) - \partial_y f(t,\mathbf{X_t})) \\ &= \partial_t f(t,\mathbf{X_t}) + \frac{1}{2}[\partial_{xx} f(t,\mathbf{X_t})X_t^2\sigma_1^2 + \partial_{yy} f(t,\mathbf{X_t})Y_t^2\sigma_2^2]\,.\end{aligned} \tag{6.17}$$

We claim that $V(\theta_T) = F(X_T, Y_T) = max(X_T - Y_T, 0) = Y_T max(\frac{X_T}{Y_T} - 1, 0)$. Suppose that we standardized our market such that all asset values are measured in terms of Y_t – the new numeraire.[4] By asserting Y_t as the numeraire, the spread-option thus becomes an x-option with $K = 1$ and $r = 0$. Hence, we apply the BSM model to $Z_t = \frac{X_t}{Y_t}$, noting Z_t is a GBM with drift μ_1 and variance $\sigma_1^2 + \sigma_2^2$. As such, we obtain the equivalences

$$\frac{f(t,X_t,Y_t)}{Y_t} = f(t,Z_t) \text{ and } \frac{f(T,X_T,Y_T)}{Y_T} = F(Z_T) = max(Z_T - 1, 0)\,. \tag{6.18}$$

Applying the BSM schema to Z_t, (6.17) reduces to

$$0 = \partial_t f(t,Z_t) + \frac{1}{2}Z_t^2\ \partial_{zz} f(t,Z_t)(\sigma_1^2 + \sigma_2^2)\,. \tag{6.19}$$

Next we use Corollary 6.7 and observe

$$f(t,Z_t) = Z_t\Phi\left(\frac{ln(\frac{S_1}{S_2}) + \frac{\sigma_1^2+\sigma_2^2}{2}(T-t)}{\sqrt{(\sigma_1^2+\sigma_2^2)(T-t)}}\right) - 1e^{-0\cdot(T-t)}\Phi\left(\frac{ln(\frac{S_1}{S_2}) - \frac{\sigma_1^2+\sigma_2^2}{2}(T-t)}{\sqrt{(\sigma_1^2+\sigma_2^2)(T-t)}}\right).$$

Isolating $f(t,X_t,Y_t)$, we conclude

$$f(t,X_t,Y_t) = X_t\Phi(q_1) - Y_t\Phi(q_2)\,. \tag{6.20}$$

□

7. Stochastic Simulation

While it is sometimes possible to model stochastic process with closed form equations, more often than not, doing so is computationally intensive and inefficient. As a result, approximation schema known as Monte Carlo methods have become important tools for simulating Itô calculus in applied settings. In a Monte Carlo method, a sample value of an asset at time T is computed recursively from an initial condition via a step formula. The expected value of a process is hence computed as the average of a large number of sample values.

We will present two common Monte Carlo methods, the Euler method and the Milstein method, and use them both to simulate models of complex financial derivatives.

[4] Often this procedure of standardizing a market with respect to Y_t is formalized via applications of the Girsanov Theorem and the Radon-Nikodym Theorem, which provide a foundation for the change in measure induced by a change in the numeraire. While a more heuristic representation is shown here, formal proofs of both theorems and their application can be found in [2] and [7].

Construction 7.1 (Euler Method)**.** Let x be an asset whose price X_t is modeled by (4.3) and let Δt be the step size of the method. Given $X_0 = x_0 \in R_+$ we define the step formula of the Euler Method of X_t to be

$$\hat{X}_{t+1} = \hat{X}_t + u(t, \hat{S}_t)\Delta t + v(t, \hat{X}_t)\Delta B_t \text{ where } \Delta B_t \sim N(0, \Delta t)\,.$$

In the Euler Method, we see that the step formula follows almost directly from the asset value's SDE. This is one of the most fundamental Monte Carlo methods because of its simplistic formula and computation. As with any Monte Carlo method, as the number of sample values is increased, the approximation error decreases. Similarly, as we decrease the method's step size, our approximation error decreases. The Milstein Method, presented below, is a refinement of the Euler method and derives from the Itô Formula.

Construction 7.2 (Milstein Method)**.** Let x be an asset whose price X_t is also modeled by (4.3). Given $X_0 = x_0 \in R_+$ we define the step formula of the Milstein Method of X_t to be

$$\hat{X}_{t+1} = \hat{X}_t + u(t, \hat{X}_t)\Delta t + v(t, \hat{X}_t)\Delta B_t + \frac{1}{2}v(t, \hat{X}_t)\partial_x v(t, \hat{X}_t) \cdot ((\Delta B_t)^2 - \Delta t)$$

where $\Delta B_t \sim N(0, \Delta t)$ and Δt is the step size of the method.

We now present a number of simulations using these Monte Carlo methods. For each of the following examples our step size will be equal to $\frac{1}{100}$.

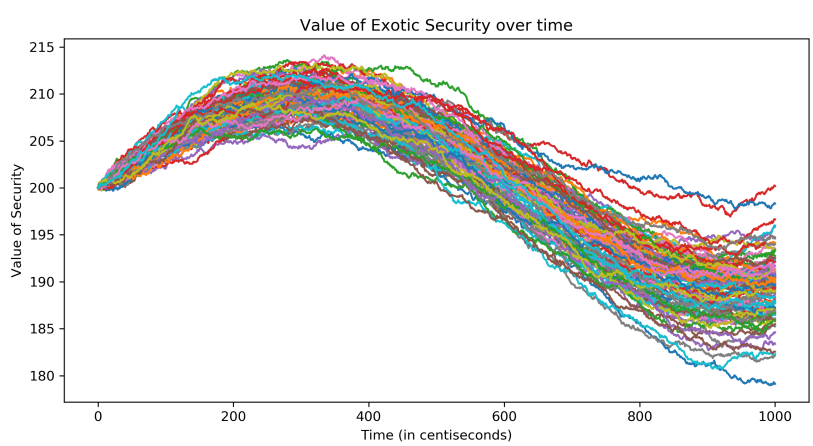
Example 7.3. Let x be an asset whose price X_t is modeled by

$$X_0 = 200,\ dX_t = \tfrac{1000}{\hat{X}_t} \cdot cos(\tfrac{t}{2000})dt + log(\hat{X}_t^2 + t)dB_t \text{ for } t \in [0, 10)\ \ .$$

The step formula of the Milstein Method of X_t is given by

$$\hat{X}_{t+1} = \hat{X}_t + \frac{10}{\hat{X}_t} \cdot cos(\frac{t}{2000}) + log(\hat{X}_t^2 + t)\Delta B_t + \frac{log(\hat{X}_t^2 + t) \cdot 2\hat{X}_t}{\hat{X}_t^2 + t} \cdot ((\Delta B_t)^2 - \frac{1}{100})$$

where $\Delta B_t \sim N(0, \frac{1}{100})$. A plot of 100 sample paths generated by the Milstein Method is shown below.



From the sample paths generated in this schema, we approximate the expected value of this security at time $T = 10$ to be 189.66 with a variance of 10.99.

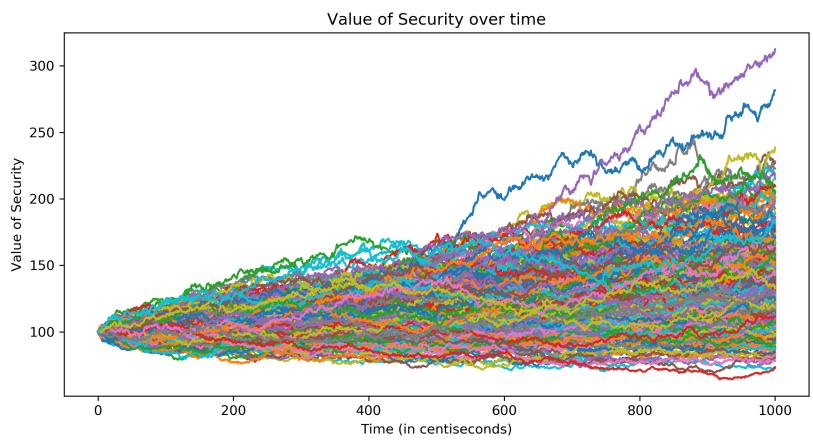
However, 100 sample paths is a relatively small sample size, which in turn produces an inexact estimate.

The purpose of the above model is not to represent any specific real life asset class, but rather to show the computational efficiency of Monte Carlo methods in handling complex stochastic structures. Next, we take a basic GBM and evaluate how close the Monte Carlo method approximation comes to the theoretic value derived in (5.5).

Example 7.4. Let x be an asset whose price X_t is modeled by

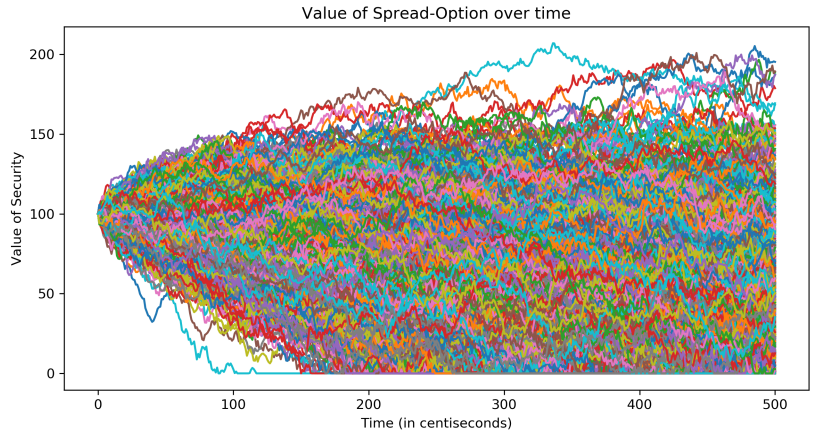
$$X_0 = 100,\ dX_t = .03X_t dt + 1.4X_t dB_t \text{ for } t \in [0, 10)\,.$$

A plot of 400 samples generated by the Milstein Method is shown below.



From the sample paths generated in this schema, we approximate the expected value of this asset at time $t = 10$ to be 134.17 with a variance of 1062.07. In (5.5), we derived the expected value of a GBM with mean μ and σ to be $X_0 e^{T\mu}$. As such, we compute $E[X_t] = 134.99$, which is incredibly close to our simulated value of 134.17. Next, we use Monte-Carlo methods to model a spread-option derivative similar to the one discussed in our presentation of the Margrabe formula.

Example 7.5. Let x and y be assets whose prices are modeled by $X_0 = 300$, $dX_t = .03X_t dt + .4X_t dB_t$ for $t \in [0, 5)$ and $Y_0 = 200$, $dY_t = .1Y_t dt + .8Y_t dB_t$ for $t \in [0, 5)$, respectively. Lastly, let z be a derivative whose value at time T, is given as $F(Z_T) = max(X_T - Y_T, 0)$. Using Euler methods, we can approximate the expected value of Z_T by simulating pairs of the assets X_t and Y_t. A plot of 1000 sample paths of Z_t is shown below.



From the sample paths generated in this schema, we approximate the expected value of this derivative at time $T = 10$ to be 39.27 with a variance of 1917.65. While the initial value of this derivative at time t is exactly 100, by time T we see that its expected value is significantly less than 100. Again, it is important to recall that a real life spread-option strategy using this model would run several thousand simulations so that the approximated price could converge to the theoretical expected value.

Acknowledgments

I would like to thank DeVon Ingram for his guidance throughout the program and for helping me understand the intricacies of measure theory. I would like to thank Sebastián Muñoz for first introducing me to Itô calculus during our spring DRP. I would like to thank Greg Lawler for his insightful lecture series on random walks and stochastic systems. I would like to thank Peter May for making this program not only possible, but exciting despite less than ideal conditions. Lastly, thank you to Mom, Dad, E and M – I love you guys.

References

[1] Øksendal, B. (1995). Stochastic Differential Equations, An Introduction with Applications. http://th.if.uj.edu.pl/ gudowska/dydaktyka/Oksendal.pdf
[2] Evans, L. (2013). An Introduction to Stochastic Differential Equations.
[3] Lawler, G. F. (2014). Stochastic Calculus: An Introduction with Applications. http://www.math.uchicago.edu/ lawler/finbook.pdf
[4] Bass, R. (2013). Real Analysis graduate students. http://bass.math.uconn.edu/rags010213.pdf
[5] Hull, J. C. (1989). Options, Futures, and Other Derivatives. *Pearson*.
[6] Suresh, E. (2019). Stochastic Calculus Applied to Arbitrage-Free Options Pricing. http://math.uchicago.edu/ may/REU2019/REUPapers/Suresh.pdf
[7] Margrabe, W. (1978). The Value of an Option to Exchange One Asset for Another. *Journal of Finance*. http://www.jstor.com/stable/2326358

Appendix

A. Proof for Proposition 3.10

Proof. Let $W_t = B_t - B_s \sim N(0, t-s)$. Using the Taylor expansion for e^x we find

$$\begin{aligned}
E[e^{iu(B_t-B_s)}] &= E[e^{iu(W_t)}] \\
&= E[\sum_{n=0}^{\infty} \frac{(iuW_t)^n}{n!}] \\
&= \sum_{n=0}^{\infty} E[\frac{(iuW_t)^n}{n!}] \\
&= E[1] + E[iuW_t] + E[\frac{-u^2W_t^2}{2}] + E[\frac{-iu^3W_t^3}{3!}] + \ E[\frac{u^4W_t^4}{4!}] \\
E[e^{iu(B_t-B_s)}] &= 1 + uE[iW_t] - \frac{u^2}{2}E[W_t^2] - \frac{u^3}{3!}E[iW_t^3] + \ \frac{u^4}{4!}E[W_t^4] + ...
\end{aligned}$$

Recall, by (3.6) we have $E[e^{iu(B_t-B_s)}] = e^{-\frac{u^2(t-s)}{2}}$. Expanding the right hand side of the equality we find

$$e^{-\frac{u^2(t-s)}{2}} = 1 - \frac{u^2(t-s)}{2} + \frac{u^4(t-s)^2}{2 \cdot 4} - \frac{u^6(t-s)^3}{3! \cdot 8} + \frac{u^8(t-s)^4}{4! \cdot 16} + ...$$

Setting the even powered u terms equal we find

$$(-1)^k u^{2k} \frac{E[W_t^{2k}]}{(2k)!} = (-1)^k \frac{u^{2k} \cdot (t-s)^k}{k! \cdot 2^k} \text{ for } k \in \mathbb{N}$$

$$\text{Thus, } E[W_t^{2k}] = \frac{(2k)!}{2^k \cdot k!} \cdot (t-s)^k \text{ for } k \in \mathbb{N}$$

$$\text{Therefore, } E[(B_t - B_s)^4] = E[W_t^4] = 3(t-s)^2. \qquad \square$$

B. Coding for Section 7 can be found on the following pages.

Stochastic Simulation Code

August 28, 2020

[1]:
```
import numpy as np
import math
import random
from matplotlib import pyplot as plt
from IPython.display import clear_output
PI = 3.1415926
e = 2.71828
```

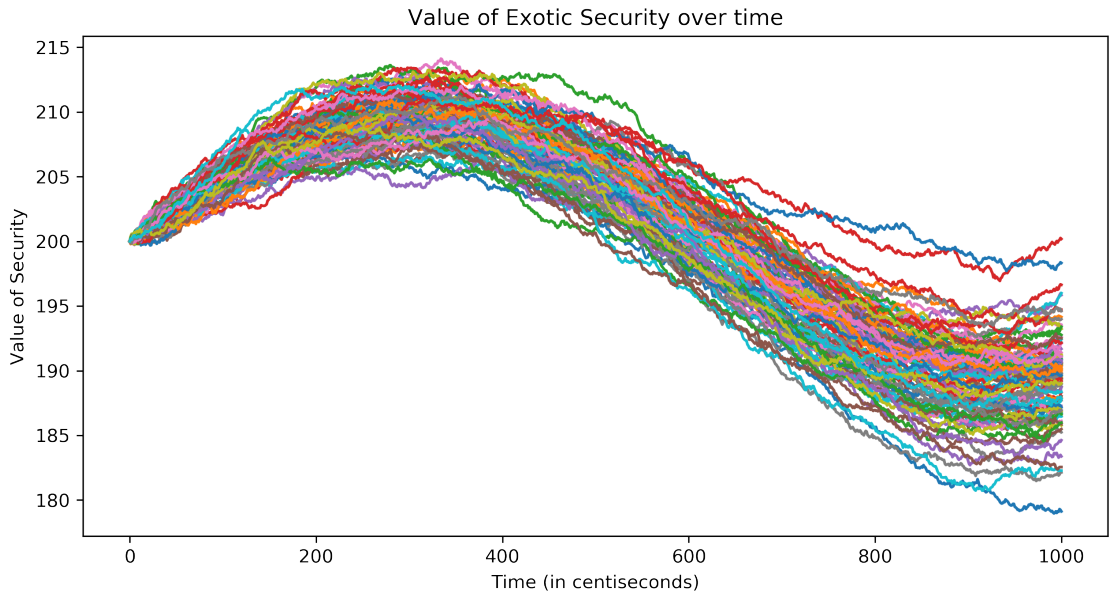
[2]:
```
def random_normal(sigma):
    mu = 0
    r = np.random.normal(mu, sigma)
    return r

print(random_normal(.01))
```

```
-0.0009863635516575134
```

[56]:
```
#EXAMPLE 1
#the u, v and v_prime functions are input manually...

def u_fcn(t,x,mu,sigma):
    value = 1000*np.cos(t/2000)/x
    return value

def v_fcn(t,x,mu,sigma):
    value = np.log(x*x+t)
    return value

def v_prime(t,x,mu,sigma):
    value = 2*x/(x*x+t) #= v_prime = (d/dx)(v_fcn)
    return value

testing=[]

def average(lst):
    return sum(lst)/len(lst)
```

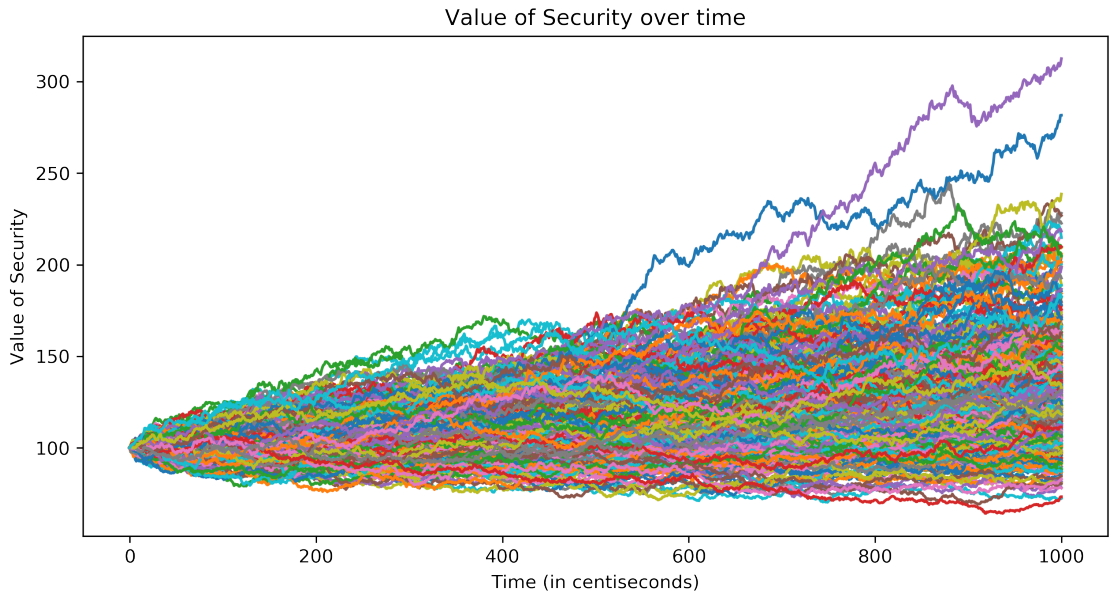
```
def milstein_method(mu, sigma, start, step, end):
    prices = [start]
    step_count = int(end/step)
    time = 0
    for i in range(step_count):
        current = prices[-1]
        time+=step_count/100
#notice here we divide by 100 so that the time value remains appropriately scaled
        rand = random_normal(step)
        new = current + u_fcn(time,current,mu,sigma)*step +␣
 ↪v_fcn(time,current,mu,sigma)*rand +\
        .
 ↪5*v_fcn(time,current,mu,sigma)*v_prime(time,current,mu,sigma)*(rand**2-step)
        prices.append(new)
    testing.append(current)
    plt.plot(prices)

plt.figure(figsize=(10,5),dpi=300)
plt.xlabel('Time (in centiseconds)')
plt.ylabel('Value of Security')
plt.title('Value of Exotic Security over time')

for i in range(100):
    milstein_method(mu = .1, sigma = 10, start = 200, step = .01, end = 10)
print('the simulated mean at t=10 is ' + str(average(testing)))
print('the simulated variance at t=10 is ' + str(np.var(testing)))
print('the simulated volatility at t=10 is ' + str(np.std(testing)))

plt.savefig("REUfig1.png")
```

```
the simulated mean at t=10 is 189.65625938675225
the simulated variance at t=10 is 10.989436532745883
the simulated volatility at t=10 is 3.315031905238
```



[48]:
```
#EXAMPLE 2

testing=[]

def euler_method(mu, sigma, start, step, end):
    prices = [start]
    step_count = int(end / step)
    for i in range(step_count):
        current = prices[-1]
        rand = random_normal(step)
        new = current + current*mu*step + current*sigma*rand
        prices.append(new)
    testing.append(current)
    plt.plot(prices)

plt.figure(figsize=(10,5),dpi=300)
plt.xlabel('Time (in centiseconds)')
plt.ylabel('Value of Security')
plt.title('Value of Security over time')

for i in range(400):
    euler_method(mu = .03, sigma = .7, start = 100, step = .01, end = 10)
print('the simulated mean at t=10 is ' + str(average(testing)))
print('the simulated variance at t=10 is ' + str(np.var(testing)))
print('the simulated volatility at t=10 is ' + str(np.std(testing)))

plt.savefig("REUfig2.png")
```

```
the simulated mean at t=10 is 134.1736349676156
the simulated variance at t=10 is 1062.0686338366047
the simulated volatility at t=10 is 32.58939449938591
```



[49]:
```
100*e**(10*.03)
```

[49]: 134.98585351801717

[51]:
```
#EXAMPLE 3
testing = []
def spreadoption_euler_method(mu1, sigma1, start1, mu2, sigma2, start2, step,␣
 ↪end):
    prices1 = [start1]
    prices2= [start2]
    prices3 = [start1-start2]

    step_count = int(end / step)
    for i in range(step_count):
        current1 = prices1[-1]
        rand1 = random_normal(step)
        new1 = current1 + current1*mu1*step + current1*sigma1*rand1
        prices1.append(new1)

        current2 = prices2[-1]
        rand2 = random_normal(step)
        new2 = current2 + current2*mu2*step + current2*sigma2*rand2
        prices2.append(new2)
```

```
        delta = new1 - new2
        new3 = max(delta,0)
        prices3.append(new3)

    plt.plot(prices3)
    testing.append(prices3[-1])

plt.figure(figsize=(10,5),dpi=300)
plt.xlabel('Time (in centiseconds)')
plt.ylabel('Value of Security')
plt.title('Value of Spread-Option over time')

for i in range(1000):
    spreadoption_euler_method(mu1 = 0.03, sigma1 = 0.4, start1 = 300, mu2 = .1,␣
 ↪sigma2 = 0.8, start2 = 200, step = .01, end = 5)

print('the simulated mean at t=5 is ' + str(average(testing)))
print('the simulated variance at t=5 is ' + str(np.var(testing)))
print('the simulated volatility at t=5 is ' + str(np.std(testing)))

plt.savefig("REUfig3.png")
```

```
the simulated mean at t=5 is 39.2667816646157
the simulated variance at t=5 is 1917.9965577027765
the simulated volatility at t=5 is 43.79493758076128
```