Mathematical Modelling of Financial Time Series

via Black Sholes Model and Monte Carlo Methods

Siddharth Pathak

Roll no. - 211034

siddharthp21@iitk.ac.in

Statistics and Data Science

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

August 24, 2024

Abstract

This report delves into the mathematical modelling of financial time series with a particular focus on the Black-Scholes model and Monte Carlo methods. Starting with the fundamental definitions and properties of random variables and stochastic processes, the report progresses to the rigorous derivation of the Black-Scholes partial differential equation, a cornerstone in the field of financial mathematics. The Black-Scholes model is then applied to European options, providing analytical solutions for both call and put options. Additionally, the report explores Geometric Brownian Motion (GBM), a stochastic process commonly used to model stock prices. The characteristics and behaviour of GBM are examined, with a particular focus on its application in option pricing. Stochastic calculus, including Itô's Lemma, plays a pivotal role in deriving key results, and in the implementation part, I also touch upon Monte Carlo simulations as a numerical method to approximate the solutions of complex financial models. Overall, this report provides a comprehensive theoretical framework for understanding the modelling and pricing of financial derivatives, with practical implications for real-world financial markets.

Contents

Chapter 1

Random Variables and Their Properties

1.1 Algebra

Let S be a set. A collection Σ_o of subsets of S is called an algebra on S (or algebra of subsets of S) if

- $S \in \Sigma_o$
- $F \in \Sigma_o \Rightarrow F^c := S \setminus F \in \Sigma_o$
- $F, G \in \Sigma_o \Rightarrow F \cup G \in \Sigma_o$

1.2 σ Algebra

A collection Σ of subsets of S is called $\sigma - algebra$ on S (or $\sigma - algebra$ on subsets of S) if Σ is an algebra on S such that whenever $F_n \in \Sigma$, $(n \in \mathbb{N})$, then

$$\cup_n F_n \in \Sigma$$

1.3 Measure Space

Let (S, Σ) is a measurable space, so that Σ is a $\sigma - algebra$ on S. A map

$$\mu : \Sigma \mapsto [0, \infty]$$

is called a measure on (S, Σ) if μ is countably additive. The triple (S, Σ, μ) is then called a measure space.

1.4 Probability measure and Probability Space

A measure μ is called a probability measure if

$$\mu(S) = 1,$$

and (S, Σ, μ) is then called a Probability space.

1.5 Model for Experiment $(\Omega, \mathcal{F}, \mathbb{P})$

A model for an experiment involving randomness takes the form of a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ as we defined above.
Sample space: Ω is a set called sample space.
Sample point: A point ω of Ω is called a sample point.
Event: The σ Algebra $\mathcal{F}$ on Ω is called the family of events so that an event is an element of $\mathcal{F}$, that is, an $\mathcal{F} - measurable$ subset of Ω.

1.6 Almost surely

A statement S about outcomes is said to be true almost surely (a.s.), or with probability 1 (B.p.1), if

$$F := \{\omega : S(\omega) \ is\ true\} \in \mathcal{F} \ and \ \mathbb{P}(F) = 1$$

1.7 Σ - measurable functions

Suppose that $h : S \mapsto \mathbb{R}$. For $A \subseteq \mathbb{R}$, define

$$h^{-1}(A) := \{s \in S : h(s) \in A\}$$

Then h is called $\Sigma - measurable$ if $h^{-1} : \mathcal{B} \mapsto \Sigma$, that is, $h^{-1}(A) \in \Sigma$, $\forall A \in \mathcal{B}$. where $\mathcal{B}$ is the borel $\sigma - algebra$ on $\mathbb{R}$. We write $m\Sigma$ for the class of Σ measurable functions on S, and $(m\Sigma)^+$ for the class of non-negative elements in $m\Sigma$. We denote the class of bounded $\Sigma - measurable$ functions on S by $b\Sigma$.

1.8 Random Variables

Let $(\Omega, \mathcal{F})$ be our (sample space, family of events). A random variable is an element of $m\mathcal{F}$. Thus,

$$X : \Omega \mapsto \mathbb{R}, \quad X^{-1} : \mathcal{B} \mapsto \mathcal{F}$$

1.9 Law, Probability Distribution Function

Suppose that X is a random variable carried by some probability triple $(\Omega, \mathcal{F}, \mathbb{P})$. We have,

$$\Omega \xrightarrow{X} \mathbb{R}$$

$$[0,1] \xleftarrow{\mathbb{P}} \mathcal{F} \xleftarrow{X^{-1}} \mathcal{B}$$

or indeed

$$[0,1] \xleftarrow{\mathbb{P}} \sigma(X) \xleftarrow{X^{-1}} \mathcal{B}$$

Define the law $\mathcal{L}_X$ of X by

$$\mathcal{L}_X = \mathbb{P} \circ X^{-1}, \quad \mathcal{L}_X : \mathcal{B} \mapsto [0,1]$$

Note that $\mathcal{L}_X$ is a probability measure on $(\mathbb{R}, \mathcal{B})$. $F_X : \mathbb{R} \mapsto [0,1]$ defined as follows:

$$F_X(c) = \mathcal{L}_X(-\infty, c) = \mathbb{P}(X \leq c) = \mathbb{P}\{\omega : X(\omega) \leq c\}$$

Remark 1.9.1. *A Distribution function F is* ***Right continuous, Non-decreasing*** *with* $\mathbf{lim_{x\to\infty}F(x) = 1}$ *and* $\mathbf{lim_{x\to-\infty}F(x) = 0}$.

1.10 Probability Mass Function (PMF)

PMF stands for Probability Mass Function. It is a mathematical function that describes the probability distribution of a discrete random variable. The PMF of a discrete random variable assigns a probability to each possible value of the random variable. The probabilities assigned by the PMF must satisfy two conditions:

- The probability assigned to each value must be non-negative (i.e., greater than or equal to zero).
- The sum of the probabilities assigned to all possible values must equal 1.

A probability mass function (PMF) is a function that gives the probability that a discrete random variable is exactly equal to a certain value.

1.11 Probability Density Function (PDF)

PDF stands for Probability Density Function. It is a mathematical function that describes the probability distribution of a continuous random variable. Probability density is used for continuous random variables because in such cases, the probability of any specific value is infinitesimally small. This is because the number of possible values a continuous random variable can take is infinite, making it impossible to assign a non-zero probability to any individual value. Instead, we use probability density to describe the distribution of continuous random variables. For differentiable law we have,

$$f_X(x) = F_X^{'}(x)$$

1.12 Moments of a Random Variable

Moments provide significant insights into the characteristics of a random variable. The nth moment of a random variable X about the origin is defined as:

$$\mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\, dx,$$

where $\mathbb{E}$ denotes the expectation operator. Special cases include:

- The **mean** (first moment): $\mu = \mathbb{E}[X]$.
- The **variance** (second central moment): $\sigma^2 = \mathbb{E}[(X - \mu)^2]$.

Chapter 2

Stochastic Processes: Brownian Motion

2.1 Stochastic Process

A stochastic process is a family of random variables, $\{X(t) : t \in T\}$, where t usually denotes time. That is, at every time t in the set T, a random number $X(t)$ is observed. Depending on T, there are two types of stochastic processes:

- $\{X(t) : t \in T\}$ is a discrete-time process if the set T is finite or countable. In practice, this generally means $T = \{0, 1, 2, 3, \dots\}$. Thus a discrete-time process is $\{X(0), X(1), X(2), X(3), \dots\}$: a random number associated with every time $0, 1, 2, 3, \dots$.

- $\{X(t) : t \in T\}$ is a continuous-time process if T is not finite or countable. In practice, this generally means $T = [0, \infty)$, or $T = [0, K]$ for some K. Thus a continuous-time process $\{X(t) : t \in T\}$ has a random number $X(t)$ associated with every instant in time. (Note that $X(t)$ need not change at every instant in time, but it is allowed to change at any time; i.e. not just at $t = 0, 1, 2, \dots$, like a discrete-time process.)

The state space, S, is the set of real values that $X(t)$ can take.

Brownian motion or the Wiener process, is a fundamental stochastic process widely used in financial modelling. This chapter delves into the formal definition of Brownian motion and its applications in finance.

2.2 Definition of Brownian Motion

A stochastic process $\{B(t), t \geq 0\}$ is called a **Brownian motion** or Wiener process if it satisfies the following properties:

1. $B(0) = 0$ almost surely.

2. $B(t) - B(s) \sim \mathcal{N}(0, \sigma^2(t - s))$ for all $0 \leq s < t$ (independent increments).

3. $B(t)$ has continuous paths almost surely.

4. $B(t)$ has stationary increments.

The quantity σ^2 is called the variance of the Brownian motion. If $\sigma = 1$, the process is called a standard Brownian motion. The meaning of (4) is that if $0 \leq t_1 < t_2 < \cdots < t_n$, then $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent, and the distribution of $B_{t_i} - B_{t_{i-1}}$ depends only on $t_i - t_{i-1}$. According to (2), this distribution is normal with mean 0 and variance $\sigma^2(t_i - t_{i-1})$. We say simply that B is continuous to indicate that B has continuous paths as in (3). Also, Property (4) implies the increments are stationary, so a Brownian motion has stationary, independent increments.

2.3 Donsker's Theorem

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean $\mathbb{E}[X_i] = 0$ and variance $\mathrm{Var}(X_i) = 1$. Define the partial sum process $S_n(t)$ for $t \in [0, 1]$ by:

$$S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} X_i + \frac{(nt - \lfloor nt \rfloor)}{\sqrt{n}} X_{\lfloor nt \rfloor + 1}.$$

Then, as $n \to \infty$, the sequence of processes $\{S_n(t)\}_{t \in [0,1]}$ converges in distribution to a standard Brownian motion $\{W(t)\}_{t \in [0,1]}$ on $[0, 1]$. That is,

$$S_n(\cdot) \xrightarrow{d} W(\cdot) \quad \text{in } C[0, 1],$$

where $C[0, 1]$ denotes the space of continuous functions on the interval $[0, 1]$, and convergence in distribution means that for any finite collection of times $0 \leq t_1 < t_2 < \cdots < t_k \leq 1$,

$$(S_n(t_1), S_n(t_2), \ldots, S_n(t_k)) \xrightarrow{d} (B(t_1), W(t_2), \ldots, W(t_k))$$

where $\xrightarrow{d}$ denotes convergence in distribution.

Note 2.3.1. *Proof of the above theorem is beyond the scope of this project.*

2.3.1 Brownian motion as a limit of random walks

One way to construct a Brownian motion is as a limit of random walks. Let $X_1, X_2, \ldots$ be i.i.d. random variables with mean 0 and variance 1. For the sake of illustration, let's suppose that $X_i = \pm 1$ with equal probability; the argument below will hold for more general step distributions. Consider the sum

$$S_n = \sum_{j=1}^{n} X_j \quad \text{with } S_0 = 0.$$

This is a simple symmetric random walk on the integers. It is a discrete-time process, but we can make a continuous-time process by linearly interpolating between values of S_n. Consider the properties of $(S_t)_{t \in \mathbb{N}}$:

- $\mathbb{E}[S_t] = 0$
- $\mathrm{Var}(S_t) = t$
- $(S_t)_{t\in\mathbb{N}}$ has stationary increments. To see why, note that

$$S_t - S_s = X_{s+1} + \cdots + X_t, \quad S_{t-s} = X_1 + \cdots + X_{t-s}.$$

  Each of $S_t - S_s, S_{t-s}$ is a sum of $t-s$ i.i.d. random variables, so $S_t - S_s \sim S_{t-s}$.

- $(S_t)_{t\in\mathbb{N}}$ has independent increments. To see why, let $0 < q < r < s < t$, and write

$$S_t - S_s = X_{s+1} + \cdots + X_t, \quad S_r - S_q = X_{q+1} + \cdots + X_r.$$

  Each of $S_t - S_s, S_r - S_q$ is a sum of distinct, independent random variables, so they are independent.

- For t large, $S_t \approx \mathcal{N}(0,t)$. This follows from the Central Limit Theorem.

Therefore, $(S_t)_{t\in\mathbb{N}}$ has many of the properties of a Brownian motion. We might wonder if there is a way to scale it so it approaches a Brownian motion in some limit. We will construct such a limit by scaling space and time in a particular way. We scale spatial steps by Δx, and time steps by Δt. When the time steps are scaled by Δt, it means that instead of progressing by 1 unit of time for each step, time progresses by Δt. So, each step corresponds to a smaller time increment. When we say that the steps are scaled by Δx, it means that instead of moving by 1 unit, the particle moves by Δx units. So, if originally a step could be +1 or -1, after scaling, each step becomes $+\Delta x$ or $-\Delta x$. The rescaled process is

$$S_t^{\Delta t,\Delta x} = \Delta x \cdot S_{t/\Delta t} = \Delta x \left(X_1 + \cdots + X_{t/\Delta t} \right).$$

We want to consider the limit of the process $(S_t^{\Delta t,\Delta x})_{t\in\mathbb{N}}$ as $\Delta t, \Delta x \to 0$. How should these parameters be related? If the limit is to approach something finite, then the variance should be finite too. Since

$$\mathrm{Var}(S_t^{\Delta t,\Delta x}) = \frac{(\Delta x)^2}{\Delta t} t \quad \Rightarrow \quad \text{we should choose } \frac{(\Delta x)^2}{\Delta t} = \text{constant}.$$

This is an important point: for a diffusion process, space scales as the square root of time. We will call this *diffusive scaling*. It will come up again and again throughout the course.

Let's suppose the constant equals 1 so the limiting process has the same variance at a point as a Brownian motion. We write $\Delta t = 1/n, \Delta x = 1/\sqrt{n}$, and define a sequence of processes in terms of the parameter n. Since our original process was a discrete-time process, it is convenient to make a continuous-time process by linearly interpolating between the discrete values of t. The interpolated, rescaled process is

$$S_t^{(n)} = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} + \frac{(nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1}}{\sqrt{n}},$$

where $\lfloor nt \rfloor$ means the largest integer less than or equal to nt. Then Donsker's Theorem or Donsker's Invariance Principle says that $S^{(n)} \equiv (S_t^{(n)})_{t\in[0,1]}$ converges in distribution to a Brownian motion B on $[0,1]$.

2.4 Properties of Brownian Motion

Brownian motion has several useful and sometimes surprising properties, surveyed in this section.

2.4.1 Scaling Properties

(i) $(B_{ct})_{t\geq 0}$ is a Brownian motion (symmetry).

(ii) $(B_{t+s} - B_s)_{t\geq 0}$ for fixed s is a Brownian motion (translation property).

(iii) $\frac{1}{\sqrt{c}} B_{ct}$ with $c > 0$ is a fixed constant, is a Brownian motion (scaling).

(iv) $(-B_t)_{t\geq 0}$ is a Brownian motion (time-inversion).

Property (iii) shows that Brownian motion is like a fractal: it looks statistically "the same" at all scales, no matter how much you zoom in, provided that space and time are scaled in the right way (again, we see the diffusive scaling space $\sqrt{\text{time}}$). This property follows naturally from the construction of Brownian motion as a limit of random walks.

Proof. 1. (i), (ii), (iii) follow straightforwardly from the Definition, by checking the required conditions are satisfied. For example, for (iii): let $X_t = c^{-1/2} B_{ct}$. Then

(a) $X_0 = c^{-1/2} B_0 = 0$.

(b) X_t has independent increments—this is straightforward to check.

(c) X_t is normally distributed: for $t \geq s$, $X_t - X_s = c^{-1/2}(B_{ct} - B_{cs}) \sim c^{-1/2}\mathcal{N}(0, c(t-s)) \sim \mathcal{N}(0, t-s)$.

(d) Continuity—this follows from continuity of B_t.

To check (iv), we use Definition #2 of BM as a Gaussian process. We have that B_t is Gaussian, with mean 0. It has covariance function $\mathbb{E}[B_s B_t] = s \wedge t$. It is continuous for $t > 0$. It remains to check that it is continuous at 0. But $\lim_{t\to 0^+} t^{-1} B_t = \lim_{t\to 0^+} \frac{B_t}{\sqrt{t}} \cdot \frac{\sqrt{t}}{t}$ a.s., by a result in the next section. □

2.4.2 Behavior as $t \to \infty$

There are several ways to characterize Brownian Motion (BM) in the limit as $t \to \infty$:

Proposition 2.4.1. *(i)* $\lim_{t\to\infty} \frac{B_t}{t} = 0$ *a.s.*

(ii) $\limsup_{t\to\infty} \frac{B_t}{\sqrt{t}} = \infty, \quad \liminf_{t\to\infty} \frac{B_t}{\sqrt{t}} = -\infty$ *(both a.s.)*

(iii) ***(Law of the Iterated Logarithm)***

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = 1 \quad a.s., \qquad \limsup_{t\to+0} \frac{B_t}{\sqrt{2t\log\log\frac{1}{t}}} = 1 \quad a.s.$$

If $\limsup$ *is replaced by* $\liminf$ *in either of the above, the limits are* -1.

Proof. (i) This follows from the Strong Law of Large Numbers. For $n \in \mathbb{N}$ we can write $B_n = (B_1 - B_0) + (B_2 - B_1) + \cdots + (B_n - B_{n-1})$, which is a sum of i.i.d. random variables. By the SLLN, $B_n/n \to 0$ a.s. To obtain behavior at non-integer t, let

$$Z_k = \max_{0 \le s \le 1} |B(k+s) - B(k)|.$$

For $t \in [k, k+1]$,

$$\frac{B_t}{t} - \frac{B_k}{k} \le \frac{1}{k(k+1)}|B_k| + \frac{1}{k+1}Z_k.$$

The first term on the RHS $\to 0$ a.s., and Z_k has the same distribution as $\max_{0 \le s \le 1} |B_s|$. It can be shown that $\mathbb{E}[Z_k] < \infty$ (see this week's homework!), and that this implies $Z_k/k \to 0$ a.s.

(ii) This follows from the Law of the Iterated Logarithm.

(iii) One only needs to show one of these limits, since they are related to each other by the time inversion property (iv) of BM. The proof is long, I am omitting it.

□

2.4.3 Quadratic Variation

Recall that the concept of total variation from analysis:

Definition 2.4.2. *The total variation of a function $f(t)$ on an interval $[a, b]$ is defined by*

$$V_{[a,b]}(f) = \sup_\sigma \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|,$$

where the supremum is over all partitions $\sigma = \{t_0, t_1, \ldots, t_n\}$ of $[a, b]$ with $a = t_0 < t_1 < \cdots < t_n = b$. If $V_{[a,b]}(f) < \infty$, then f is said to be of bounded variation, and if $V_{[a,b]}(f) = \infty$ then f is said to be of infinite variation.

If a function is of bounded variation on $[a, b]$, then it has a derivative almost everywhere on $[a, b]$ (i.e., except for a set of measure zero). Conversely, if a function is nowhere differentiable, then it must have infinite variation on any interval.

Since Brownian motion is nowhere differentiable, it has infinite variation on any interval. However, it has finite *quadratic variation* (in a mean-square sense). This will turn out to be an important property when we construct the stochastic integral.

Definition 2.4.3. *The quadratic variation of a function f on $[0, t]$ with respect to a partition σ is*

$$Q_t^\sigma(f) = \sum_{i=0}^{n-1} \left[f(t_{i+1}) - f(t_i)\right]^2.$$

We would like to define the quadratic variation of a Brownian motion W on $[0, t]$ as $\sup_\sigma Q_t^\sigma(B)$. However, for a stochastic process, we need to be careful with how the supremum over partitions is calculated since not all types of stochastic convergence will give a finite result. We will use the following notion of convergence:

Definition 2.4.4. *A sequence of random variables* $X_1, X_2, \ldots$ *converges in mean-square to another random variable* X*, written* $X_n \xrightarrow{m.s.} X$ *or* $\lim_{n\to\infty} \mathbb{E}[|X_n - X|^2] \to 0$ *as* $n \to \infty$*.*

Remark 2.4.5. *Convergence in mean-square implies convergence in probability, which in turn implies convergence in distribution. Convergence in mean-square does not imply almost sure convergence nor vice versa. (Almost sure convergence does imply convergence in probability and hence convergence in distribution.)*

Lemma 2.4.6. *Let* $|\sigma| = \max_{0\le i\le n-1} |t_{i+1} - t_i|$*. The quadratic variation of Brownian motion* $Q_t^\sigma(B)$ *converges in mean-square to* t *as* $|\sigma| \to 0$*:*

$$\sum_{i=0}^{n-1} \left[B_{t_{i+1}} - B_{t_i}\right]^2 \xrightarrow{m.s.} t.$$

We can write this result formally as $(\Delta B)^2 = \Delta t$*. We see the diffusive scaling yet again.*

Proof. Write $\Delta B_i = B_{t_{i+1}} - B_{t_i}$, and $\Delta t_i = t_{i+1} - t_i$. Then

$$\begin{aligned}\mathbb{E}\left(Q_t^\sigma(B) - t\right)^2 &= \mathbb{E}\left(\sum_{i=0}^{n-1} \Delta B_i^2 - \Delta t_i\right)^2 \\ &= \sum_{i=0}^{n-1} \mathbb{E}\left(\Delta B_i^2 - \Delta t_i\right)^2 + 2\sum_{i,j=0,i\neq j}^{n-1} \mathbb{E}\left(\Delta B_i^2 - \Delta t_i\right)\mathbb{E}\left(\Delta B_j^2 - \Delta t_j\right).\end{aligned}$$

Now, we have

$$\mathbb{E}\left(\Delta B_i^2 - \Delta t_i\right)\mathbb{E}\left(\Delta B_j^2 - \Delta t_j\right) = \mathbb{E}\left(\Delta B_i^2 - \Delta t_i\right)\mathbb{E}\left(\Delta B_j^2 - \Delta t_j\right),$$

since $\Delta B_i, \Delta B_j$ are independent for $i \neq j$. Therefore

$$\mathbb{E}\left(Q_t^\sigma(B) - t\right)^2 = \sum_{i=0}^{n-1} \mathbb{E}\left(\Delta B_i^2 - \Delta t_i\right)^2 \le \sum_{i=0}^{n-1} \mathbb{E}\left(\Delta B_i^4\right)\Delta t_i^2,$$

since $(a-b)^2 \le a^2 + b^2$ for $a, b \ge 0$. We know that $\mathbb{E}\left(B_t - B_s\right)^4 = 3|t-s|^2$, so

$$\mathbb{E}\left(Q_t^\sigma(B) - t\right)^2 \le 4\sum_{i=0}^{n-1} \mathbb{E}\Delta t_i^2 = 4r|\sigma| \to 0 \text{ as } |\sigma| \to 0.$$

□

2.4.4 Martingale Property

To prove that Brownian motion is a martingale, we need to show that the process B_t satisfies the martingale property. Specifically, for a process B_t to be a martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\ge0}$, it must satisfy the following conditions:

1. B_t is adapted to the filtration $\{\mathcal{F}_t\}_{t\ge0}$.

2. B_t has finite expectation: $\mathbb{E}[|B_t|] < \infty$ for all $t \geq 0$.

3. $\mathbb{E}[B_t \mid \mathcal{F}_s] = B_s$ for all $0 \leq s \leq t$.

- **Adaptedness:** By definition, Brownian motion B_t is adapted to the natural filtration $\{\mathcal{F}_t\}$, where $\mathcal{F}_t$ is the sigma-algebra generated by the process up to time t. This means that for each t, B_t is measurable with respect to $\mathcal{F}_t$, so the first condition is satisfied.

- **Finite Expectation**

  The expectation of Brownian motion at any time t is zero, i.e., $\mathbb{E}[B_t] = 0$. Since B_t is a continuous process and normally distributed with mean zero and variance t, it has a finite expectation:
  $$\mathbb{E}[|B_t|] < \infty \quad \text{for all } t \geq 0.$$
  Thus, the second condition is also satisfied.

- **Martingale Property** To prove the martingale property, we need to show that for $0 \leq s \leq t$,
  $$\mathbb{E}[B_t \mid \mathcal{F}_s] = B_s.$$
  Using the properties of Brownian motion, we know that for $0 \leq s \leq t$,
  $$B_t = B_s + (B_t - B_s),$$
  where $B_t - B_s$ is independent of $\mathcal{F}_s$ and is normally distributed with mean 0 and variance $t - s$.

  Taking the conditional expectation given $\mathcal{F}_s$, we have:
  $$\mathbb{E}[B_t \mid \mathcal{F}_s] = \mathbb{E}\left[B_s + (B_t - B_s) \mid \mathcal{F}_s\right].$$
  Since B_s is $\mathcal{F}_s$-measurable, $\mathbb{E}[B_s \mid \mathcal{F}_s] = B_s$. Additionally, $B_t - B_s$ is independent of $\mathcal{F}_s$, so
  $$\mathbb{E}[B_t \mid \mathcal{F}_s] = B_s + \mathbb{E}[B_t - B_s \mid \mathcal{F}_s] = B_s + \mathbb{E}[B_t - B_s] = B_s + 0 = B_s.$$
  Thus, the third condition is satisfied.

Since Brownian motion B_t satisfies all three conditions, it is a martingale with respect to the natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

2.5 Ito's Lemma

It is easy to see that Brownian motion is not differentiable at all. We can see it intuitively by considering
$$Z = fracB(t + \Delta) - B(t)\Delta$$
$Var(Z) = 1/\Delta$, which becomes larger when $\Delta \to 0$. So the limitation does not exist.

While studying Brownian motions, or when using Brownian motion as a model, the situation of estimating the difference of a function of the type $f(B_t)$ over an infinitesimal time difference occurs quite frequently (suppose that f is a smooth function). To be more precise, we are considering a function $f(t, B_t)$ which depends only on the second variable One possible way to work around this problem is to try to describe the difference df in terms of the difference dB_t. In this case, the equation above becomes

$$df = f'(B_t)dB_t.$$

Our new formula at least makes sense, since there is no need to refer to the differentiation $\frac{dB_t}{dt}$ which does not exist. The only problem is that it does not quite work. Consider the Taylor expansion of f to obtain

$$f(x + \Delta x) - f(x) = (\Delta x)f'(x) + \frac{(\Delta x)^2}{2}f''(x) + \frac{(\Delta x)^3}{6}f'''(x) + \cdots .$$

To deduce Equation (1.1) from this formula, we must be able to say that the significant term is the first term $(\Delta x)f'(x)$ and all other terms are of smaller order of magnitude. Is this true for $x = B_t$? For $x = B_t$, we have

$$\Delta f = (\Delta B_t)f'(B_t) + \frac{(\Delta B_t)^2}{2}f''(B_t) + \frac{(\Delta B_t)^3}{6}f'''(B_t) + \cdots .$$

Now consider the term $(\Delta B_t)^2$. Since B_t is a Brownian motion, we know that $\mathbb{E}[(\Delta B_t)^2] = \Delta t$. Since a difference in B_t is necessarily accompanied by a difference in t, we see that the second term is no longer negligible. The theory of Itô calculus essentially tells us that we can make the substitution

$$(\Delta B_t)^2 = \Delta t,$$

and the remaining terms are negligible. Hence the equation above becomes

$$\Delta f = (\Delta B_t)f'(B_t) + \frac{\Delta t}{2}f''(B_t) + \cdots ,$$

which in terms of infinitesimals becomes

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt.$$

This equation, known as the Itô's lemma, is the main equation of Itô's calculus.

More generally, consider a smooth function $f(t, x)$ which depends on two variables, and suppose that we are interested in the differential of $f(t, B_t)$. In classical calculus, we will get

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx,$$

but in Itô calculus, we will have

$$df(t, B_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dB_t)^2.$$

Theorem 1.1. (Itô's lemma) Let $f(t,x)$ be a smooth function of two variables, and let X_t be a stochastic process satisfying $dX_t = \mu dt + \sigma dB_t$ for a Brownian motion B_t. Then

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

Proof. We have

$$df(t,X_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2.$$

Substituting $dX_t = \mu dt + \sigma dB_t$, we obtain

$$df(t,X_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}(\mu dt + \sigma dB_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\sigma dB_t)^2.$$

Since $(dB_t)^2 = dt$ and ignoring higher-order terms, we have

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

□

Chapter 3

Geometric Brownian Motion

3.1 Definition

Geometric Brownian motion (GBM) is a stochastic process used to model the dynamics of financial instruments such as stock prices. GBM is commonly used to model stock prices. It is a continuous-time stochastic process where the logarithm of the stochastic process follows a Brownian motion. Here is the formula for the Geometric Brownian Equation. It is defined by the stochastic differential equation (SDE):

$$dS(t) = \mu S(t)\, dt + \sigma S(t)\, dB(t),$$

where:

- $S(t)$ is the value of the process at time t,
- μ is the drift coefficient, it is analogous to the expected return of the asset. In a way, it aims to capture the general trend of the asset price over time.
- σ Known as the volatility coefficient, it quantifies the extent of the asset's price fluctuations.
- $B(t)$ is a standard Brownian motion.

This equation elegantly encapsulates both the deterministic (drift) and stochastic (diffusion) aspects of asset price dynamics, forming a bedrock for various financial modelling and derivative pricing endeavours.

The term $\sigma B(t)$ in the GBM formula introduces randomness into the stock price evolution, and it's driven by the Wiener process $B(t)$.

- **Asset Price Dynamics ($dS(t)$) :**The left-hand side $dS(t)$ represents the incremental change in the asset price S at time t .
- **Drift Term** ($\mu S(t)dt$) **:** The first term on the right-hand side, ($\mu S(t)dt$), is known as the drift term. It represents the deterministic trend in the asset price, often reflecting the expected return over a small time interval dt.

- **Diffusion Term ($\sigma S(t)dB(t)$) :** The second term on the right-hand side, $\sigma S(t)dB(t)$, is known as the diffusion term. It captures the stochastic (random) movements in the asset price. $dB(t)$ represents a Wiener Process increment over the small time interval dt, which embodies the random shock to the asset price.

This stochastic differential equation models asset price dynamics under Geometric Brownian Motion (GBM), which assumes that returns are normally distributed and the logarithm of asset prices forms a Brownian motion with drift. It's a crucial assumption underlying many financial models, including the Black-Scholes Model for options pricing.

3.1.1 Deriving the Geometric Brownian Motion (GBM) Formula from SDE

Here's a concise walkthrough of how the formula for GBM comes about from the Stochastic Differential Equation (SDE). The SDE for the continuous-time model for the dynamics of an asset price $S(t)$ is given by:

$$dS(t) = \mu S(t)\, dt + \sigma S(t)\, dB(t)$$

To get the formula for $S(t)$, we need to solve this SDE. The method of solving involves Itô's Lemma, a fundamental result in stochastic calculus. Applying Itô's Lemma, we can find a differential equation for $\ln(S(t))$. The procedure leads to the differential equation:

$$d(\ln(S(t))) = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dB(t)$$

Integrating this result over the interval $[0, t]$, we obtain:

$$\ln(S(t)) - \ln(S(0)) = \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B(t)$$

Taking the exponential of both sides to eliminate the natural logarithm, we get:

$$S(t) = S(0) \times \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B(t)\right)$$

This formula for $S(t)$ represents the asset price at time t in terms of the initial asset price $S(0)$, the drift rate μ, the asset volatility σ, and a standard Brownian motion $B(t)$ up to time t.

It provides a closed-form solution for the asset price under the assumptions of Geometric Brownian Motion, which is pivotal for various applications in finance, including options pricing via the Black-Scholes Model.

3.2 Properties of Geometric Brownian Motion

- **Log-Normal Distribution**: $S(t)$ is log-normally distributed, meaning $\log(S(t))$ follows a normal distribution. Since $B(t)$ is a standard Brownian motion, it follows that $B(t) \sim \mathcal{N}(0,t)$. Therefore, $\ln(S(t))$ is normally distributed with mean $\ln(S(0)) + \left(\mu - \frac{\sigma^2}{2}\right)t$ and variance $\sigma^2 t$:

$$\ln(S(t)) \sim \mathcal{N}\left(\ln(S(0)) + \left(\mu - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right).$$

  Since $\ln(S(t))$ is normally distributed, it follows that $S(t)$ is log-normally distributed. Specifically, $S(t)$ can be expressed as:

$$S(t) = S(0)\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right).$$

  Therefore, $S(t)$ follows a log-normal distribution, with the distribution given by:

$$S(t) \sim \text{LogNormal}\left(\ln(S(0)) + \left(\mu - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right).$$

- **Expectation of $S(t)$:** We know that $B(t) \sim \mathcal{N}(0,t)$, so $\sigma B(t) \sim \mathcal{N}(0, \sigma^2 t)$. The expectation of the exponential of a normal random variable is given by:

$$\mathbb{E}[\exp(X)] = \exp\left(\mathbb{E}[X] + \frac{1}{2}\text{Var}(X)\right).$$

  Applying this to $S(t)$:

$$\mathbb{E}[S(t)] = \mathbb{E}\left[S(0)\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right)\right].$$

  Since $S(0)$ is a constant, we can factor it out:

$$\mathbb{E}[S(t)] = S(0)\mathbb{E}\left[\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right)\right].$$

  Now, using the property of expectation for the normal distribution:

$$\mathbb{E}[S(t)] = S(0)\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \frac{1}{2}\sigma^2 t\right).$$

  Simplifying, we get:

$$\mathbb{E}[S(t)] = S(0)\exp(\mu t).$$

- **Variance of $S(t)$:** To calculate the variance of $S(t)$, we first need to compute $\mathbb{E}[S(t)^2]$:

$$\mathbb{E}[S(t)^2] = \mathbb{E}\left[S(0)^2\exp\left(2\left(\mu - \frac{\sigma^2}{2}\right)t + 2\sigma B(t)\right)\right].$$

Again, using the property of the expectation of the exponential of a normal random variable:

$$\mathbb{E}[S(t)^2] = S(0)^2 \exp\left(2\left(\mu - \frac{\sigma^2}{2}\right)t + 2\sigma^2 t\right).$$

Simplifying, we get:

$$\mathbb{E}[S(t)^2] = S(0)^2 \exp\left(2\mu t + \sigma^2 t\right).$$

Thus, the variance of $S(t)$ is given by:

$$\mathrm{Var}(S(t)) = \mathbb{E}[S(t)^2] - (\mathbb{E}[S(t)])^2 = S(0)^2 \exp(2\mu t + \sigma^2 t) - (S(0)\exp(\mu t))^2.$$

Simplifying further:

$$\mathrm{Var}(S(t)) = S(0)^2 \exp(2\mu t)\left[\exp(\sigma^2 t) - 1\right].$$

- **Long-Time Behavior of Expectation and Variance of $S(t)$:** The expectation of $S(t)$ is given by:

  $$\mathbb{E}[S(t)] = S(0)\exp(\mu t).$$

  As t becomes large, the expectation $\mathbb{E}[S(t)]$ grows exponentially with rate μ. Specifically:

  $$\lim_{t\to\infty} \mathbb{E}[S(t)] = \lim_{t\to\infty} S(0)\exp(\mu t).$$

  - If $\mu > 0$, then $\mathbb{E}[S(t)] \to \infty$ as $t \to \infty$.
  - If $\mu = 0$, then $\mathbb{E}[S(t)] = S(0)$, which remains constant over time.
  - If $\mu < 0$, then $\mathbb{E}[S(t)] \to 0$ as $t \to \infty$.

  Thus, the long-time behaviour of the expectation is heavily dependent on the sign of the drift coefficient μ.

  The variance of $S(t)$ is given by:

  $$\mathrm{Var}(S(t)) = S(0)^2 \exp(2\mu t)\left[\exp(\sigma^2 t) - 1\right].$$

  As t becomes large, the term $\exp(\sigma^2 t)$ dominates -1, so:

  $$\mathrm{Var}(S(t)) \approx S(0)^2 \exp(2\mu t)\exp(\sigma^2 t) = S(0)^2 \exp\left((2\mu + \sigma^2)t\right).$$

  Hence, the variance grows exponentially with rate $2\mu + \sigma^2$ as $t \to \infty$:

  $$\lim_{t\to\infty} \mathrm{Var}(S(t)) = \lim_{t\to\infty} S(0)^2 \exp\left((2\mu + \sigma^2)t\right).$$

  - If $2\mu + \sigma^2 > 0$, then $\mathrm{Var}(S(t)) \to \infty$ as $t \to \infty$.
  - If $2\mu + \sigma^2 = 0$, then $\mathrm{Var}(S(t))$ remains constant ($= S^2(0)$) over time.
  - If $2\mu + \sigma^2 < 0$, which is highly unusual in practical situations, then $\mathrm{Var}(S(t)) \to 0$ as $t \to \infty$.

Thus, the long-term behaviour of $S(t)$ depends critically on the values of μ and σ, with the possibility of significant growth or decay based on the sign and magnitude of these parameters.

- **Martingale Property of $S(t)$:** To check whether $S(t)$ is a martingale with respect to the natural filtration $\{\mathcal{F}_t\}$, we need to verify if:

  $$\mathbb{E}[S(t) \mid \mathcal{F}_s] = S(s) \quad \text{for all } 0 \leq s \leq t.$$

  Starting from the expectation:

  $$\mathbb{E}[S(t) \mid \mathcal{F}_s] = \mathbb{E}\left[S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B(t)\right) \mid \mathcal{F}_s\right].$$

  This can be rewritten as:

  $$\mathbb{E}[S(t) \mid \mathcal{F}_s] = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B(s)\right) \mathbb{E}\left[\exp\left(\sigma(B(t) - B(s))\right) \mid \mathcal{F}_s\right].$$

  Since $B(t) - B(s) \sim \mathcal{N}(0, t - s)$ and is independent of $\mathcal{F}_s$, we have:

  $$\mathbb{E}\left[\exp\left(\sigma(B(t) - B(s))\right) \mid \mathcal{F}_s\right] = \exp\left(\frac{1}{2}\sigma^2(t - s)\right).$$

  Substituting this into the expectation:

  $$\mathbb{E}[S(t) \mid \mathcal{F}_s] = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B(s) + \frac{1}{2}\sigma^2(t - s)\right).$$

  Simplifying, we get:

  $$\mathbb{E}[S(t) \mid \mathcal{F}_s] = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) s + \sigma B(s)\right) = S(s).$$

  Therefore, $S(t)$ is indeed a martingale with respect to $\{\mathcal{F}_t\}$.

Chapter 4

Black Scholes Model

4.1 Options

An option is a financial derivative that gives the buyer the right, but not the obligation, to buy or sell an underlying asset (such as a stock, bond, commodity, or other financial instruments) at a predetermined price, known as the strike price, within a specified period. The seller of the option is obligated to fulfil the contract if the buyer chooses to exercise the option. There are two main types of options:

- **Call Option:** Gives the holder the right to buy the underlying asset at the strike price before or on the expiration date.
- **Put Option:** Gives the holder the right to sell the underlying asset at the strike price before or on the expiration date.

Options are used for various purposes, including hedging against potential losses, generating income, or speculating on the future direction of an asset's price. To buy (Call) or sell (Put) an underlying asset or instrument at a specified strike price K on or before a specified expiry date T, depending on the style of the option.

European Option:

Let S_t represent the price of the underlying asset at time t, K represent the strike price of the option, and T represent the expiration time of the option. A European option provides the holder with the right to:

- **Call Option:** Buy the underlying asset at the strike price K at time T.
- **Put Option:** Sell the underlying asset at the strike price K at time T.

The payoff functions for a European option at maturity T are given by:

- **Call Option Payoff:**

$$\text{Payoff} = \max(S_T - K, 0)$$

- **Put Option Payoff:**
$$\text{Payoff} = \max(K - S_T, 0)$$

Where:

- S_T is the price of the underlying asset at maturity T.
- K is the strike price.

American Option

Let S_t represent the price of the underlying asset at time t, K represent the strike price of the option, and T represent the expiration time of the option. An American option provides the holder with the right to:

- **Call Option:** Buy the underlying asset at the strike price K at any time up to and including the expiration time T.
- **Put Option:** Sell the underlying asset at the strike price K at any time up to and including the expiration time T.

The payoff functions for an American option at any time t are given by:

- **Call Option Payoff:**
$$\text{Payoff} = \max(S_t - K, 0)$$
- **Put Option Payoff:**
$$\text{Payoff} = \max(K - S_t, 0)$$

Where:

- S_t is the price of the underlying asset at time t.
- K is the strike price.

4.2 Assumptions of the Black-Scholes Model

The Black-Scholes model for option pricing is built on the following assumptions:

1. The short-term interest rate, r, is known and constant through time. All securities share this short-term interest rate.
2. The stock price follows a Brownian Motion, and the variance rate of return on the stock is constant.
3. The parameters μ (drift) and σ (volatility) are functions of the stock price S.
4. The stock does not pay dividends or other distributions.

5. Short selling is allowed.
6. There are no arbitrage opportunities, and all security trading is continuous.
7. The option is a European option, meaning it has a maturity at time t and can only be exercised at expiration.

Note 4.2.1. *These assumptions form the foundation of the Black-Scholes model, making the complex problem of option pricing mathematically tractable. However, in real markets, some of these assumptions may not hold, leading to potential deviations between the model's predictions and actual market prices.*

4.3 Derivation of the Black-Scholes Equation

Let $V(S_t, t)$ be the value of the option at time t when the underlying asset price is S_t. The value of the option is a function of S_t and t, so we can apply Itô's lemma to $V(S_t, t)$:

$$dV(S_t, t) = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_t} dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^2}(dS_t)^2.$$

Substituting the stochastic differential equation for dS_t:

$$dS_t = \mu S_t\, dt + \sigma S_t\, dB_t,$$

we get:

$$dV(S_t, t) = \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2}\right) dt + \sigma S_t \frac{\partial V}{\partial S_t} dB_t.$$

Consider a portfolio, called the delta-hedge portfolio, consisting of being short one option and long $\frac{\partial V}{\partial S}$ shares at time t. The value of these holdings is:

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

Over the time period $[t, t + \Delta t]$, the total profit or loss from changes in the values of the holdings is:

$$\Delta\Pi = -\Delta V + \frac{\partial V}{\partial S}\Delta S$$

Now discretize the equations for dS/S and dV by replacing differentials with deltas:

$$\Delta S = \mu S \Delta t + \sigma S \Delta W$$

$$\Delta V = \left(\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) \Delta t + \sigma S \frac{\partial V}{\partial S}\Delta W$$

Substituting them into the expression for $\Delta\Pi$:

$$\Delta\Pi = \left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) \Delta t$$

Notice that the ΔW term has vanished. Thus uncertainty has been eliminated and the portfolio is effectively riskless. The rate of return on this portfolio must be equal to the rate of return on any other riskless instrument; otherwise, there would be opportunities for arbitrage. Now assuming the risk-free rate of return is r, we must have over the time period $[t, t + \Delta t]$:

$$\Delta\Pi = r\Pi\Delta t$$

If we now substitute our formulas for $\Delta\Pi$ and $\Pi = -V + \frac{\partial V}{\partial S}S$ into the above equation, we obtain:

$$\left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\Delta t = r\left(-V + \frac{\partial V}{\partial S}S\right)\Delta t$$

Simplifying, we arrive at the celebrated Black-Scholes partial differential equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

Solution of the Black-Scholes PDE

The Black-Scholes partial differential equation is given by:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

The solution for a European call option is:

$$V(S,t) = S_0\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

Here, S_0 is the current stock price, K is the strike price, T is the time to maturity, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. For a European put option, the solution is:

$$V(S,t) = Ke^{-r(T-t)}\Phi(-d_2) - S_0\Phi(-d_1)$$

Note 4.3.1. *Solving stochastic partial differential equations is beyond the scope of this project.*