6.2 Continuous Time Models

Let us now turn our attention to continuous time models. One way to obtain models at a continuous time level is to start with a discrete time model defined on a partition $\{t_0, t_1, \ldots, t_n\}$, with $\Delta t = t_{k+1} - t_k = T/n$, of the interval $[0, T]$ and then take $n \to \infty$ (equivalently, $\Delta t \to 0$). In this section, we start with the logarithmic binomial model derived in the above subsection and show that in the limiting case, it tends to a continuous time model, which is called the *lognormal model* that follows a geometric Brownian motion.

6.2.1 Lognormal Model

Using (6.5) and (6.6), we get

$$\mathbb{P}\left(\frac{\ln\left(\frac{S(t_n)}{S_0}\right) - \mu}{\sigma} \leq r\right) = \mathbb{P}\left(\frac{\ln\left(\frac{S(t_n)}{S_0}\right) - n\tilde{\mu}}{\tilde{\sigma}\sqrt{n}} \leq r\right),$$

for any given real number $r > 0$. Since $\ln(S(t_n)/S_0)$ is the sum of n iid random variables R_k, $k = 1, \ldots, n$, each with mean $\tilde{\mu}$ and variance $\tilde{\sigma}^2$, we can use the central limit theorem to see that as the discrete time tends to the continuous time level (*i.e.* as $n \to \infty$), we get

$$\mathbb{P}\left(\frac{\ln\left(\frac{S(1)}{S_0}\right) - \mu}{\sigma} \leq r\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{\ln\left(\frac{S(1)}{S_0}\right) - n\tilde{\mu})}{\sigma\sqrt{n}} \leq r\right) = \Phi(r),$$

where Φ is the distribution function of the standard normal random variable. Hence,

$$\frac{\ln\left(\frac{S(1)}{S_0}\right) - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

This implies that

$$\ln\left(\frac{S(1)}{S_0}\right) \sim \mathcal{N}(\mu, \sigma^2),$$

which is equivalent to saying that the *total return* $\frac{S(1)}{S_0}$ is a lognormally distributed random variable, in a continuous-time framework, with parameters μ and σ^2. Since the *rate of return* $(S(1) - S_0)/S_0$ of the stock is just a translation of the total return, we can also conclude that the rate of return is also lognormally distributed in a continuous-time framework.

Using the central limit theorem, we can conclude that

$$\ln\left(\frac{S(T)}{S_0}\right) \sim \mathcal{N}(T\mu, T\sigma^2), \tag{6.12}$$

where μ and σ^2 are given in (6.5).

Problem 6.7.

Show that $E(S(T)) = S_0 e^{\left(\mu+\frac{\sigma^2}{2}\right)T}$ and $\mathrm{Var}(S(T)) = S_0^2 e^{(2\mu+\sigma^2)T}(e^{\sigma^2 T}-1)$.

Problem 6.8.

Let the total return per annum of a stock be governed by the lognormal distribution with parameters $\mu = 0.12$ and $\sigma = 0.2$.

1. Find the expectation of the stock price for 6 months.
2. Find the standard deviation of the stock price for 6 months.
3. If the current spot price of one share of the stock is ₹100 and the prevailing interest rate $r = \mu + \frac{\sigma^2}{2}$, then find per-unit price at $t = 0$ of the corresponding future contract with one month expiration.

Answer: (1) $1.0725\, S_0$, (2) $0.1523\, S_0$, (3) $f(0,T) = 107.25$.

Hints:
For the last part of the problem, since no other assumptions are provided, we should assume that the market is arbitrage-free and allows short selling.

6.2.2 Brownian Motion

For every time $t \in [0,T]$, the stock price $S(t)$ is a random variable on a probability space, denoted by $(\mathsf{S}, \mathscr{S}, \mathbb{P})$. Hence, the collection $\{S(t) \mid t \in [0,T]\}$ defines a *stochastic process* often denoted as $\{S_t\}$. We now present an important result that a stock price process $\{S_t\}$ follows *geometric Brownian motion*. First, we provide a general definition of a standard one-dimensional Brownian motion on $[0,T]$. To this end, we first need the concept of *filtration*.

A σ-algebra may be interpreted as a collection of information. This may be understood from the following familiar example:

Example 6.2.1.

Consider a game in which a player rolls a die. If an even number appears, the opponent pays the player ₹1; otherwise, the player pays ₹1 to the opponent.

The player's cash flow in every game can be defined by the random variable

$$X(s) = \begin{cases} 1, & \text{if } s \in \{2,4,6\}, \\ -1, & \text{if } s \in \{1,3,5\}. \end{cases}$$

Let S be the set of outcomes. The question now is, what is the suitable σ-field? Obviously, the power set $\mathscr{S}_*$ of S will define a σ-field. But the more relevant σ-field is $\sigma(X)$, the generated σ-field of X. It is easy to see (check it!) that

$$\sigma(X) = \Big\{\emptyset, \mathsf{S}, \{2,4,6\}, \{1,3,5\}\Big\}.$$

We can observe that this is the smallest σ-field (why?) that carries complete information about the win and loss scenarios of the player. Of course, $\mathscr{S}_*$ also contains the same information, but with many other irrelevant pieces of information. For instance, $\{1\} \in \mathscr{S}_*$ but there is no information in this set in the sense that $\{1\}$ does not define any rule in the game, rather it is only a part of a rule. Mathematically, there is no Borel set $B \subset \mathbb{R}$ such that $X^{-1}(B) = \{1\}$.

Since $\sigma(X) \subset \mathscr{S}_*$, X is $\mathscr{S}_*$-measurable. In this case, we say that X depends on the information contained in $\mathscr{S}_*$.

Definition 6.2.2 [Filtration].

A collection $\{\mathcal{F}_t \mid t \geq 0\}$ of σ-fields on S is called a ***filtration*** on a sample space $(\mathsf{S}, \mathscr{S})$ if it is an nondecreasing collection of sub-σ-algebras of $\mathscr{S}$.

A ***filtered probability space*** is a quadruple $(\mathsf{S}, \mathscr{S}, \mathbb{P}, \{\mathcal{F}_t\})$, where $(\mathsf{S}, \mathscr{S}, \mathbb{P})$ is a probability space, and $\{\mathcal{F}_t\}$ is a filtration on the sample space $(\mathsf{S}, \mathscr{S})$.

Remark 6.2.3 [Interpretation of filteration].

Each σ-field $\mathcal{F}_k$ is interpreted as the information available to the investor for the period $[t_k, t_{k+1}]$. The information structure provided by a filtration $\{\mathcal{F}_k\}$ is that the information does not decrease in time. Thus an information known at time $t_1 \in [0, T]$ is known (and not forgotten) for any future time $t_1 < t \leq T$.

Definition 6.2.4 [Adapted Process].

Let $\{\mathcal{F}_t\}$ be a filtration defined on a sample space $(\mathsf{S}, \mathscr{S})$. A stochastic process $\mathcal{X} = \{X_t \mid t \geq 0\}$ defined on $(\mathsf{S}, \mathscr{S})$ is said to be ***adapted*** to the filtration $\{\mathcal{F}_t\}$ if X_t is $\mathcal{F}_t$-measurable, for every $t \geq 0$.

Note

While the concepts of filtration and adapted process are defined on the time interval $[0, \infty)$, it is also meaningful to apply these terms within any bounded interval $[0, T]$. Furthermore, although we introduced them for a continuous time frame, they can be similarly defined for a discrete time frame.

Remark 6.2.5 [Generated Filtration].

Often in applications, we have a stochastic process $\mathcal{X}$ defined on a suitable sample space. It is important to work with a suitable filtration for which $\mathcal{X}$ is adapted. One natural choice of the filtration is the *generated filtration* $\{\mathcal{F}_t^{\mathcal{X}}\}$ of $\mathcal{X}$ defined as

$$\mathcal{F}_t^{\mathcal{X}} = \sigma(X_s, 0 \leq s \leq t), \text{ for all } t \geq 0,$$

the smallest σ-field with respect to which X_s is measurable for every $s \in [0, t]$.

Definition 6.2.6 [Standard Brownian motion].

Let $(\mathsf{S}, \mathscr{S}, \mathbb{P}, \{\mathcal{F}_t\})$ be a filtered probability space. An adapted stochastic process $\mathcal{W} = \{W_t \mid t \in [0, T]\}$ defined on this filtered probability space is called a *standard Brownian motion* (also known as a *Wiener process*) if it satisfies the following conditions:

1. $W_0 = 0$ ($\mathbb{P}$-a.s).
2. For each $s \in \mathsf{S}$, the map $t \longmapsto W_t(s)$ from $[0, T]$ to $\mathbb{R}$ is continuous. This is to say $\mathcal{W}$ is a continuous stochastic process ($\mathbb{P}$-a.s is sufficient).
3. *Independent increment property*: For any partition $t_0 = 0 < t_1 < t_2 < \ldots < t_n = T$, the random variables $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent.
4. For each $s, t \in [0, T]$ with $s < t$, $W_t - W_s \sim \mathcal{N}(0, t - s)$.

The following property follows directly from the definition.

Problem 6.9 [Stationary Increment Property].

Let $\{W_t\}$ be a Wiener process. Show that for every $t \in [0, T]$, $W_{t+\Delta t} - W_t \overset{\mathrm{d}}{=} W_{\Delta t}$.

Remark 6.2.7.

Define

$$W_t := \frac{\ln\left(\frac{S_t}{S_0}\right) - t\mu}{\sigma}, \quad t \in [0, T], \tag{6.13}$$

where μ and σ are the parameters as defined in the above discussion. We can see that for each $t \in [0, T]$, W_t is a random variable and hence the collection $\mathcal{W} = \{W_t \mid t \in [0, T]\}$ also forms a stochastic process.

Problem 6.10.

Show that the stochastic process $\mathcal{W}$ defined in Remark 6.2.7 is a Weiner process (*i.e.*, $\mathcal{W}$ is a Brownian motion).

Hints:
Proof involves the following steps:

1. Define a suitable sample space and a filtration on which $\mathcal{W}$ is a adapted process;
2. Check the conditions states in Definition 6.2.6 (first check Condition 3 and then go for Condition 2).

Remark 6.2.8.

Recall that we have already mentioned that the sequence $\left\{ \ln\left(\frac{S(t_k)}{S_0}\right) \;\middle|\; k = 0, 1, 2, \ldots, \right\}$ defines a *random walk* under EMH. Now, we have illustrated that under EMH and using the central limit theorem, the Brownian motion can be obtained as a limit of random walks.

Remark 6.2.9.

From (6.13), we see that the stock price at time t can be written as

$$S(t) = S_0 \exp(t\mu + \sigma W(t)).$$

Such a process is called a *geometric Brownian motion* with parameters μ and σ. Here, μ is called the *drift* and σ is called the *volatility*.

Problem 6.11.

Consider the stock price process $\{S_t \mid t \in [0, 1]\}$ governed by a geometric Brownian motion with drift $\mu = 0.1$, volatility $\sigma = 0.15$, and an initial price $S_0 = 100$.

1. Determine the probability that the stock price increases by at least 25% by one year.
2. Determine the probability that the stock price decreases by at least 15% by one year.

Answer: (1) 0.2061, (2) 0.0401

6.3 Price Dynamics

In the previous section, we derived the stock price process S_t, characterized as a geometric Brownian motion with drift μ and volatility σ. In this section, we will derive a stochastic differential equation for the stock price process, whose solution corresponds to the process obtained in Remark 6.2.9. Additionally, we will outline a computational procedure for simulating a stock price process.

For every $t \in [0, T]$, the total return of an initial investment of S_0 for the period $[0, t]$ under continuous compounding scheme with an annual interest rate μ is given by

$$S = S_0 e^{\mu t}.$$

The above expression can be equivalently written in the form of the initial value problem (IVP)

$$\frac{dS}{dt} = \mu S;\ \ S(0) = S_0.$$

If μ is known, then the above IVP represents a *deterministic model* for the total return of an investment. If we take $X(t) = \ln(S(t)/S_0)$ as the logarithmic return, the above model can be equivalently written as

$$dX = \mu dt. \tag{6.14}$$

Note

So far, we have used the notation r for interest rate, but now we changed it to μ intensionally.

6.3.1 Stochastic Models

The deterministic model (6.14) is insufficient for pricing risky assets. In such cases, we need to replace the deterministic variable μ with a random variable, denoted as R. More precisely, let us assume that the random variable is given by

$$R = \mu + \epsilon,$$

where μ is fixed and $\epsilon = \epsilon(t)$ varies randomly as time t varies. However, directly replacing μ with R in the model (6.14) is not straightforward due to the rapid changes in ϵ. Instead, we consider the model in the form

$$dX = \mu dt + \text{noise}, \tag{6.15}$$

where the term 'noise' needs to be determined.

The basic idea of the model is to assume the noise term to be *Gaussian white noise*.

Definition 6.3.1 [Gaussian White Noise].

A stochastic process $\{\epsilon_t \mid t \in [0,T]\}$ is called the ***Gaussian white noise*** if for every $t \in [0,T]$

1. $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$, for some constant $\sigma_\epsilon^2 > 0$;
2. $E(\epsilon_t) = 0$; and
3. $E(\epsilon_s \epsilon_t) = 0$, for every $s, t \in [0,T]$ with $s \neq t$.

White noise is physically unrealistic, as its mean-square value blows up to infinity. A convenient way to handle this situation is to use a Wiener process (also called Brownian motion), which can be formally represented as a *stochastic integral* of a white noise process. We omit the technical details of stochastic integrability of a random process and formally define

$$W(t) = \int_0^t \epsilon(s) ds. \tag{6.16}$$

Note

Recall, a random process $Z(t)$ is said to be *stochastically integrable* if the limit

$$\int_a^b Z(s)ds = \lim_{\Delta t_i \to 0} \sum_i Z(t_i)\Delta t_i$$

exists in the mean-square sense.

Let the interval $[0,T]$ be uniformly discretized into N sub-intervals with length Δt, where $t_{k+1} = t_k + \Delta t$, for $k = 0, 1, \ldots, N-1$, $t_0 = 0$, and $t_N = T$. The above definition motivates us to consider the approximation

$$W(t_k) \approx \sqrt{\Delta t} \sum_{j=0}^{k-1} \epsilon(t_j)$$

for (6.16). We have taken $\sqrt{\Delta t}$ in the above expression instead of Δt for a reason which will be clear shortly.

We use the notation W_k to denote the approximate value in the above expression (*i.e.* $W(t_k) \approx W_k$). We take $\epsilon(t_k) \sim \mathcal{N}(0,1)$ and write

$$W_{k+1} = W_k + \epsilon(t_k)\sqrt{\Delta t}, \tag{6.17}$$

for $k = 0, 1, \ldots, N-1$, with $W_0 = 0$. The process $\{W_k\}$ is called a *random walk*. The random walk and the corresponding discrete white noise (often referred to as *color noise*) are depicted in Figure 6.4 (a) and (b), respectively, for $T = 1$ and $N = 1000$.

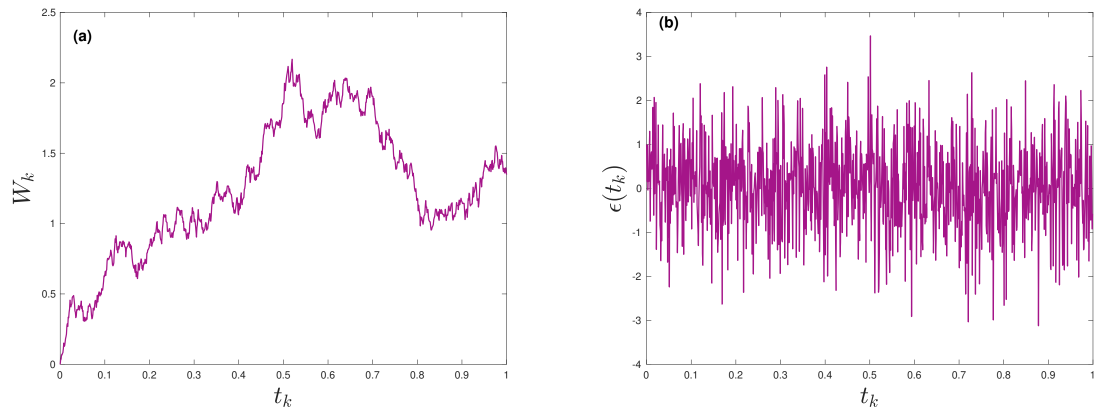

Figure 6.4: (a) An illustration of the random walk (6.17) in the interval $[0, 1]$ with $N = 1000$, (b) The corresponding discrete white noise.

For every $j, k = 0, 1, \ldots, N$ with $j < k$, we can write

$$W_k - W_j = \sum_{i=j}^{k-1} \epsilon(t_i)\sqrt{\Delta t},$$

which shows that $W_k - W_j$ is a normally distributed random variable and we have

$$E\big(W_k - W_j\big) = 0 \text{ and } \operatorname{Var}\big(W_k - W_j\big) = t_k - t_j.$$

Further, assume that the process $\{\epsilon_t\}$ consists of independent random variables. Then, for any $t_0 \le t_{k_1} < t_{k_2} < t_{k_3} \le t_N$, the random variable $W_{t_{k_2}} - W_{t_{k_1}}$ and $W_{t_{k_3}} - W_{t_{k_2}}$ are independent because these random variables are sum of two distinct set of ϵ's which are themselves independent.

From the above discussions, we can see that the process $\{W_k\}$ satisfies all the basic properties of a Wiener process except the continuity property. Continuity property can be achieved by taking $\Delta t \to 0$. We omit the details and accept that the random walk process defined above tends to a Wiener process $\{W_t \mid t \in [0, T]\}$ as $\Delta t \to 0$.

For every $t \in [0, T]$, we can write (6.17) as

$$W_{t+\Delta t} - W_t = \epsilon_t\sqrt{\Delta t},$$

for a small $\Delta t > 0$. In the limiting case of $\Delta t \to 0$, we write the above expression symbolically as

$$dW_t = \epsilon_t\sqrt{dt},$$

which governs the Wiener process $\{W_t\}$ that we are interested in (6.16).

Note
The above derivation of the Wiener process model is not rigorous because we have no assurance that the limiting operations are defined, but it provides a good intuitive description.

In view of the above discussions, we propose to take the noise term to be proportional to dW_t. Hence, we write the model (6.15) as

$$dX_t = \mu dt + \sigma dW_t, \tag{6.18}$$

where μ and $\sigma > 0$ are constants, and $\{W_t\}$ is a Wiener process. This is a *stochastic model*, which is a *stochastic differential equation*. The process $\{X_t\}$ governed by the above equation is called a *generalized Wiener process* with *drift* μ and *volatility* σ.

Recall that we have taken $X_t = \ln\big(S_t/S_0\big)$. Thus, the above equation can be written as

$$d(\ln S_t) = \mu\, dt + \sigma\, dW_t \tag{6.19}$$

We can also write the above equation in the form

$$dS_t = \tilde{\mu} S_t\, dt + \sigma S_t\, dW_t, \tag{6.20}$$

where $\tilde{\mu} = \mu + \sigma^2/2$, which follows from Itô Lemma 6.4.2 (also see Example 6.4.3).

A formal integration on both sides of (6.19) gives

$$\ln S_t = \mu t + \sigma W_t + \ln S_0, \quad t \in [0, T],$$

which can be rewritten as

$$S_t = S_0 \exp\big(\mu t + \sigma W_t\big).$$

This is precisely the geometric Brownian motion obtained from the lognormal model as the stock price process in Remark 6.2.9. A process $\{S(t)\}$ is called a *geometric Brownian motion* if it satisfies the stochastic differential equation (6.20) for some given $\tilde{\mu}$ and σ.

6.3.2 Monte Carlo Simulation

Monte Carlo simulation of a stochastic process is one of the most powerful tools to sample random outcomes of the process. In this section, we illustrate the method in simulating the stock price process given by the geometric Brownian motion (6.20).

For a positive integer N, let us consider the discretization of the interval $[0, T]$ by $t_{k+1} = t_k + \Delta t$, for $k = 0, 1, \cdots, N-1$, where $\Delta t = T/N$, $t_0 = 0$, and $t_N = T$. The

Day(t_k)	ϵ	S_k	Total Return	Rate of Return
0	-	1887.75	0	0
1	-1.32092	1825.89	0.97	-0.03
2	1.64723	1900.50	1.04	0.04
3	0.83594	1939.91	1.02	0.02
4	-0.82600	1900.16	0.98	-0.02
5	-0.20409	1890.54	0.99	-0.01

Table 6.1: Data simulated by the Monte Carlo method (6.21) for the parameters in Example 6.3.2.

Monte Carlo simulation equation for the process governed by the stochastic differential equation (6.20) is given by

$$S_{k+1} = \big(1 + \tilde{\mu}\Delta t + \sigma\epsilon(t_k)\sqrt{\Delta t}\big)S_k, \quad k = 0, 1, \dots N-1, \tag{6.21}$$

where S_k is an approximation to $S(t_k)$ obtained by the above formula for $k = 1, 2, \dots, N$, S_0 is a given initial value of the stock price at the present time $t = 0$, and $\epsilon(t_k) \sim \mathcal{N}(0, 1)$. Recall that $\tilde{\mu} = \mu + \sigma^2/2$, where μ is the drift (expected rate of return per annum) and σ is the volatility (standard deviation of the return) of the stock price, and are also given.

Note

In real-life, experts obtain the parameters μ and σ based many parameters and aspects. For instance, prevailing interest rate, fundamental analysis of a company, company's management commentaries on the earning perspectives, future events, general market conditions, and so on. However, one may also use some standard statistical estimators to obtain these parameters from the past price data of the stock. For generating values of ϵ at every time t_k one may use some standard normal random variable generators in-built in some standard softwares.

Example 6.3.2.

Consider a stock with drift $\tilde{\mu} = 0.1$ (expected rate of return 10%) and volatility $\sigma = 0.4$ (standard deviation of the return 40%) per annum. Given the stock price on 31$^{\text{st}}$ December 2021 as $S_0 = 1887.75$, we can generate a price process of the stock for a future period using the Monte Carlo simulation given by (6.21).

Let us first obtain the price process of the stock for one week on a daily basis. Considering 52 weeks per year, and 5 working days per week, we take

$$\Delta t = \frac{1}{260}.$$

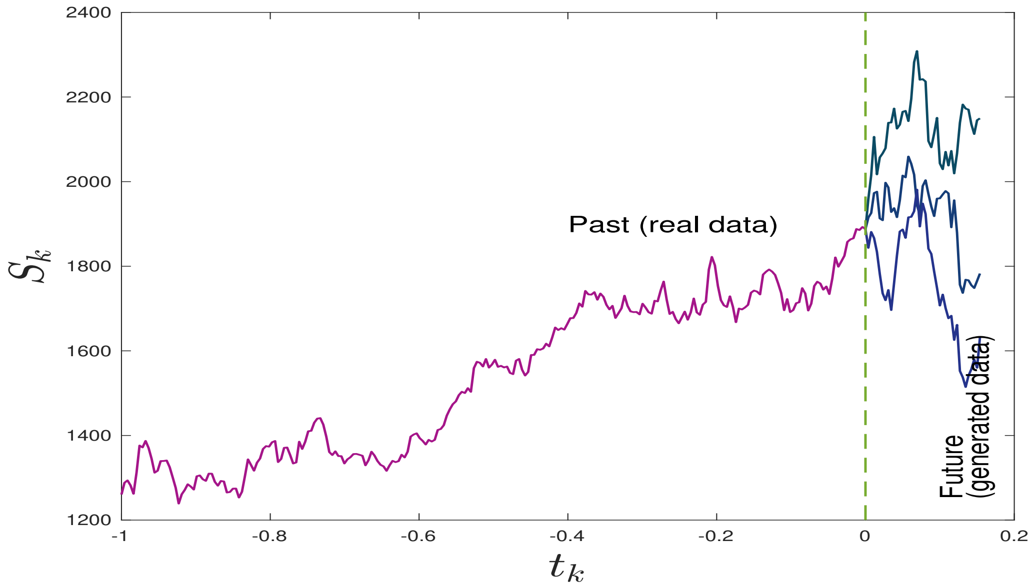

Figure 6.5: One year daily closing price of Infosys Technology (labeled 'Past') and the stock price process (labeled 'Future') simulated using the Monte Carlo method (6.21) with parameters given in Example 6.3.2.

A set of generated data is provided in Table 6.1.

Let us now obtain a price process of the stock for 8 weeks (from 2$^{\text{nd}}$ January to 27$^{\text{th}}$ February 2022) on a daily basis. The graphical representation of the simulated price process with three samples taken on a daily basis is depicted in Figure 6.5.

Problem 6.12.

Write the Monte Carlo simulation formula for the process governed by (6.19).

Project 6.1.

Write a python code to read a CSV file consisting of the past 200 days (at least) daily NIFTY50 data (can be downloaded by clicking `here`), and do the following:

1. Extract the column consisting of 'CLOSE' points of NIFTY50;
2. Find the 200 days simple moving average and take it as μ;
3. Find the standard deviation using 200 days simple moving average formula and take it as σ.
4. Predict the next day's NIFTY50 close price as per the Monte Carlo method.

6.4 Itô Process

In the previous section, we established a stochastic model for a stock price process, simplifying it into a generalized Wiener process with constant drift and volatility. However, real-world scenarios, such as stock prices and option prices, demand a more detailed approach, wherein these parameters can fluctuate with time and the underlying process itself. Thus, a more realistic stochastic model takes the form

$$dX = \mu(X,t)dt + \sigma(X,t)dW, \tag{6.22}$$

where W represents a Wiener process, and both the drift and volatility are functions of X and t. Such a stochastic process, denoted by $X = \{X_t\}$ governed by the above stochastic differential equation, is termed an *Itô process*. Unlike the straightforward solution found in the case of a generalized Wiener process, solving this equation is more involved.

Remark 6.4.1.

While stochastic differential equations offer the advantage of familiarity with basic calculus notation, they present challenges in interpretation due to the involvement of non-differentiable Brownian paths. Consequently, these equations are typically handled in integral form

$$X(t) = X_0 + \int_0^t \mu(X(s),s)ds + \int_0^t \sigma(X(s),s)dW,$$

where X_0 is a given constant. This form of equation is known as a *stochastic integral equation*. The integral in the third term on the right-hand side requires interpretation in the Itô sense, known as the *Itô integral*. To ensure meaningful interpretation of the second and third terms, we assume that $\mu \in L^1_{\text{loc}}$ and $\sigma \in L^2_{\text{loc}}$. We omit the details of the Itô integral and will not go further into the theory of stochastic integral equations.

6.4.1 Itô Lemma

The price of a derivative depends on both the underlying asset and time. Even in the stock price process, as we have seen, it is common to work with $\ln S$ rather than S. Thus, dealing with functions of an Itô process, such as $f(S,t)$, is inevitable. Caution is required in such situations as the differential $df(S,t)$ is not exactly analogous to what we encounter in ordinary calculus. In fact, $df(S,t)$ involves an additional term, and the precise formula is called the *Itô formula*. Here, we state the formula and provide only a brief outline of its proof.

Lemma 6.4.2 [Itô Lemma].

Let X be an Itô process in (6.22) and let $Y = F(X,t)$, where $F \in C^{2,1}(\mathbb{R}^2)$. Then, Y is an Itô process satisfying the equation

$$dY = \mu_F dt + \sigma_F dW,$$

with drift and volatility processes given by

$$\begin{aligned} \mu_F &= \mu\frac{\partial F}{\partial x} + \frac{\partial F}{\partial t} + \sigma^2\frac{1}{2}\frac{\partial^2 F}{\partial x^2}, \\ \sigma_F &= \sigma\frac{\partial F}{\partial x}. \end{aligned}$$

Note

In the ordinary calculus, the expression for μ_F does not include the term

$$\sigma^2\frac{1}{2}\frac{\partial^2 F}{\partial x^2}.$$

Proof.

We omit the rigorous proof of the lemma and present only a formal derivation of the formula.

Let us take $F = F(X,t)$, for $(X,t) \in \mathbb{R}^2$ and assume that F is smooth. Using Taylor expansion, we can write

$$\Delta F = \frac{\partial F}{\partial X}\Delta X + \frac{\partial F}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}\Delta X^2 + \frac{\partial^2 F}{\partial X \partial t}\Delta X \Delta t + \frac{1}{2}\frac{\partial^2 F}{\partial t^2}\Delta t^2 + \cdots \quad (6.23)$$

Let us consider discrete form of the Itô process

$$\Delta X = \mu(X,t)\Delta t + \sigma(X,t)\Delta W.$$

Since W satisfies the Wiener process, we have

$$\Delta W = \epsilon\sqrt{\Delta t}.$$

Therefore, we have (dropping the arguments for the sake of notational convenience)

$$\Delta X = \mu\Delta t + \sigma\epsilon\sqrt{\Delta t},$$

Squaring both sides, we get

$$\Delta X^2 = \sigma^2\epsilon^2\Delta t + o(\Delta t).$$

Substituting these expressions in (6.23), we get

$$\Delta F \quad = \quad \left(\mu\frac{\partial F}{\partial X} + \frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}\sigma^2\epsilon^2\right)\Delta t + \sigma\frac{\partial F}{\partial X}\Delta W + o(\Delta t).$$

Since $\epsilon \sim \mathcal{N}(0,1)$, we have

$$\mathrm{Var}(\epsilon^2 \Delta t) = \Delta t^2 E(\epsilon^4) - \Delta t^2.$$

Since $\epsilon \sim \mathcal{N}(0,1)$, we have from the basic probability theory that the fourth moment of the standard normal distribution is 3. That is, $E(\epsilon^4) = 3$. This implies that

$$\mathrm{Var}(\epsilon^2 \Delta t) = 2\Delta t^2.$$

Recall from the basic probability theory, that if a random variable has $\mathrm{Var}(X) = 0$, then $X = E(X)$ a.s. Since $\mathrm{Var}(\epsilon^2\Delta t) \to 0$ as $\Delta t \to 0$, we see that $\epsilon^2 \Delta t$ tends to dt a.s. as $\Delta t \to 0$.

Thus, neglecting $o(\Delta t)$ terms and then taking $\Delta t \to 0$, we get

$$dF = \left(\mu\frac{\partial F}{\partial X} + \frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial X^2}\right)dt + \sigma\frac{\partial F}{\partial X}dW.$$

Example 6.4.3.

Consider the stochastic differential equation (6.20)

$$dS = \tilde{\mu} S dt + \sigma S dW.$$

This equation can be viewed as Itô process with $\mu(S,t) = \tilde{\mu} S$ and $\sigma(S,t) = \sigma S$.

Now let us take $F(S,t) = \ln S$. Then, we have

$$\frac{\partial F}{\partial S} = \frac{1}{S}, \quad \frac{\partial F}{\partial t} = 0, \quad \frac{\partial^2 F}{\partial X^2} = -\frac{1}{S^2}.$$

Therefore, by Itô lemma, we have

$$d\ln S = \left(\tilde{\mu} S \times \frac{1}{S} + 0 + \frac{1}{2}\sigma^2 S^2\left(-\frac{1}{S^2}\right)\right)dt + \sigma S \times \frac{1}{S}dW.$$

This leads to the stochastic differential equation equivalent to the above equation as

$$d\ln S = \left(\tilde{\mu} - \frac{1}{2}\sigma^2\right)dt + \sigma dW,$$

which is the same as (6.19).