7
Stochastic Processes

7.1 Introduction

In our earlier chapters, we have been making frequent use of terms like *random variables* and *random vectors*. In fact, any first course on probability theory discusses these concepts in detail. The aim of this chapter is to go beyond the notion of a random vector, and introduce a collection of random variables parameterized by a parameter, say time. Such a collection of random variables is essentially a *stochastic process*.

Stochastic processes occur naturally in real life applications. For instance consider the (opening) exchange rate: Indian rupees (INR)/US dollar (USD) at every day between August 18, 2011 and February 2, 2012 except weekends. The actual exchange rate is shown in Fig. 7.1. As on any given day, the exchange rate is random, we can interpret this figure as a realization x_n of the random variable X_n, where n is the day $1, 2, \ldots$. In order to make a guess of the interest rate on a future date, it is reasonable to look at the whole evolution of X_n between August 18, 2011 and February 2, 2012. Therefore, there is a need into develop theory which provides almost continuous information about the process considered. A mathematical representation for describing such a phenomenon leads to the notion of stochastic process.

Stochastic processes play a vital role in financial mathematics, because the proper understanding of asset dynamics is crucial for making any meaningful financial decision. The derivation of Black-Scholes formula for option pricing illustrates this point very well. The derivative pricing and interest rate modeling can be studied in greater depth using the concept of stochastic process. In this chapter, we aim to present a brief introduction of stochastic processes and some of the related concepts keeping finance in view.

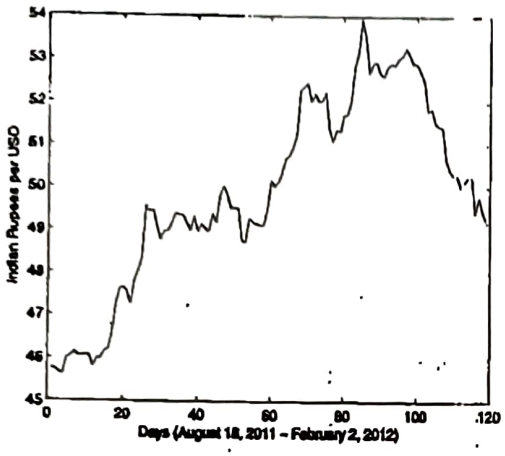

Fig. 7.1. Exchange Rate: Indian Rupees (INR) / US Dollar (USD) from August 18, 2011 to February 2, 2012

7.2 Definitions and Simple Stochastic Processes

In this section, we first review certain definitions for random variables and then proceed formally to define a stochastic process. We shall denote the random experiment by $\mathcal{E}$ and the corresponding sample space by Ω.

Definition 7.2.1 (σ-field) *A σ-field $\mathcal{F}$ (or σ-algebra) is a family of subsets of Ω which satisfies the following properties.*

(i) $\emptyset \in \mathcal{F}$,
(ii) if $A \in \mathcal{F}$ then its complement $A^c \in \mathcal{F}$,
(iii) if a countable sequence of sets $A_1, A_2, \ldots$ is in $\mathcal{F}$, then $\cup_i A_i \in \mathcal{F}$.

From the above it is simple to conclude that $\Omega \in \mathcal{F}$, and if a countable sequence of sets $A_1, A_2, \ldots$ is in $\mathcal{F}$ then $\cap_i A_i \in \mathcal{F}$.

Obviously, the smallest possible σ-field is $\{\emptyset, \Omega\}$. It is called the *trivial σ-field*. When the set Ω is finite, the collection of all subsets of Ω is a σ-field, which is same as the power set of Ω. It is called the *total σ-field* and is the largest σ-field on Ω. The set $\{\emptyset, A, A^c, \Omega\}$ where A is a proper subset of Ω ($\emptyset \subset A \subset \Omega$), is another σ-field on Ω; it is called the *σ-field generated by the set A*. The *σ-field $\mathcal{F}$ generated by two proper subsets A and B* is

$$\mathcal{F} = \{\emptyset, A, B, A^c, B^c, A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c, (A \triangle B)^c, (A \triangle B), A \cup B, A \cup B^c, A^c \cup B, (A \cap B)^c, \Omega\}.$$

Here $A \triangle B$ is defined as $(A \cap B^c) \cup (B \cap A^c)$ and is called the symmetric difference of sets A and B. Further it may be noted that given two σ-fields $\mathcal{F}_1$ and $\mathcal{F}_2$, $\mathcal{F}_1 \cap \mathcal{F}_2$ is always a σ-field but $\mathcal{F}_1 \cup \mathcal{F}_2$ may not be a σ-field.

Example 7.2.1 *Let $\Omega = \{1,2,3\}$ and $\mathcal{F} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{2,3\}, \{1,3\}, \Omega\}$. Verify that $\mathcal{F}$ is a σ-field on Ω.*

Solution We must show that properties (i), (ii) and (iii) of Definition 7.2.1 of σ-field are satisfied by given $\mathcal{F}$. We have $\emptyset \in \mathcal{F}$. Also $\emptyset^c = \Omega$, $\{1\}^c = \{2,3\}$, $\{1,2\}^c = \{3\}$ and so on. Hence (ii) is satisfied. Also, $\{1\} \cup \{2,3\} = \{1,2,3\}$, $\{1\} \cup \{3\} = \{1,3\}$. Further this property can be checked for other sets as well. Hence $\mathcal{F}$ satisfies (i), (ii) and (iii). Therefore, $\mathcal{F}$ is a σ-field on Ω. □

Example 7.2.2 *Let $\Omega = \{a,b,c,d\}$. Let $\mathcal{F}_1 = \{\emptyset, \{a\}, \{b,c,d\}, \Omega\}$ and $\mathcal{F}_2 = \{\emptyset, \{a,b\}, \{c,d\}, \Omega\}$. Show that $\mathcal{F}_1 \cup \mathcal{F}_2$ is not a σ-field. What is the largest σ-field in Ω?*

Solution By the definition of σ-field, $\mathcal{F}_1$ and $\mathcal{F}_2$ are σ-fields. Now, $\mathcal{F}_1 \cup \mathcal{F}_2 = \{\emptyset, \{a\}, \{a,b\}, \{c,d\}, \{b,c,d\}, \Omega\}$. Clearly, $\mathcal{F}_1 \cup \mathcal{F}_2$ is not a σ-field because $\{a\} \cup \{c,d\} = \{a,c,d\} \notin \mathcal{F}_1 \cup \mathcal{F}_2$. Further, $\mathcal{F} = \{\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a,b\}, \{c,d\}, \{a,c\}, \{b,d\}, \{a,d\}, \{b,c\}, \{a,b,c\}, \{b,c,d\}, \{a,c,d\}, \{a,b,d\}, \Omega\}$ is the largest σ-field. □

Let $\mathcal{F}$ be the given σ-field. Then $\mathcal{F}$ consists of those subsets of Ω which satisfy conditions (i), (ii) and (iii) of Definition 7.2.1. These subsets of Ω are called *$\mathcal{F}$-events* or *$\mathcal{F}$-measurable sets*. The pair $(\Omega, \mathcal{F})$ is called a *measurable space*. Measurable spaces are the building blocks for defining probability spaces.

Definition 7.2.2 (Probability Measure) *Let $\mathcal{F}$ be a σ-field over Ω. Let P be a real-valued function defined on $\mathcal{F}$ such that*

(i) $P(A) \geq 0$ for all $A \in \mathcal{F}$,
(ii) $P(\Omega) = 1$,
(iii) if $A_1, A_2, A_3, \ldots$ are mutually exclusive events, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Then P(.) is called a probability measure. The triplet $(\Omega, \mathcal{F}, P)$ is called a probability space.

In most of the random experiments, either it is not feasible or it is difficult to obtain the probability of an event using definition. However, it is possible to calculate the probability of an event and related probability distributions by mapping the set of all possible outcomes into real line. This leads to the definition of a random variable.

Definition 7.2.3 (Random Variable) *Let $(\Omega, \mathcal{F}, P)$ be given probability space. A function $X : \Omega \to \mathbb{R}$ is called a random variable if for any x in $\mathbb{R}$, $X^{-1}\{(-\infty, x]\}$ belongs to $\mathcal{F}$. Here*

$$X^{-1}\{(-\infty, x]\} = \{w \in \Omega : X(w) \in (-\infty, x]\}.$$

Remark 7.2.1 *For a given probability space $(\Omega, \mathcal{F}, P)$, when $\mathcal{F}$ is the largest σ-field on Ω, any real-valued function defined on Ω will be a random variable.*

Remark 7.2.2 *Let $\Omega = \mathbb{R}$ and $\mathcal{U}^{(1)} = \{(a, b] : -\infty < a < b < \infty\}$. Then the σ-field generated by $\mathcal{U}^{(1)}$ is called the Borel σ-field, and its elements are called Borel sets. Thus Borel σ-field is the σ-field generated by the family of semi-open intervals of type $(a, b]$, $-\infty < a < b < \infty$.*

Remark 7.2.3 *A function f from a measurable space $(\Omega, \mathcal{F})$ into $(\mathbb{R}, \mathcal{B})$ is a said to be measurable function if for each Borel set $B \in \mathcal{B}$, the set $\{w \in \Omega : f(w) \in B\} \in \mathcal{F}$. Here $\mathcal{B}$ is the Borel σ-field of subsets of $\mathbb{R}$. From the above definition, a random variable X is a measurable function from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B})$.*

Remark 7.2.4 *As $X : \Omega \to \mathbb{R}$, a statement of type $a \le X \le b$ is to be understood as $\{w \in \Omega : a \le X(w) \le b\}$. In that case $P(a \le X \le b) = P\{w \in \Omega : a \le X(w) \le b\}$. In general for a Borel set $\mathcal{B}$, $P(X \in \mathcal{B}) = P\{w \in \Omega : X(w) \in \mathcal{B}\}$. Here it is assumed that $(\Omega, \mathcal{F}, P)$ is the underlying probability space.*

The function which records the probabilities associated with the random variable is defined as below.

Definition 7.2.4 (Cumulative Distribution Function of a Random Variable) *The cumulative distribution function (CDF) of a random variable X is given as*

$$F_X(x) = P(X \le x) = P(w \in \Omega : X(w) \le x), \quad -\infty < x < \infty.$$

For $a < b$, we have

$$P(a < X \le b) = P(w \in \Omega : a < X(w) \le b) = F_X(b) - F_X(a).$$

Example 7.2.3 *Let $\Omega = \{a, b, c, d\}$ and $\mathcal{F} = \{\emptyset, \{a\}, \{b, c, d\}, \Omega\}$. Let $(\Omega, \mathcal{F}, P)$ be the given probability space where $P(\{a\}) = 1/4$ and $P(\{b, c, d\}) = 3/4$. Define*

$$X(w) = \begin{cases} 0, & w = a \\ 1, & w = b \\ 2, & w = c \\ 3, & w = d, \end{cases}$$

and

$$Y(w) = \begin{cases} 0, & w = a \\ 1, & w \in \{b, c, d\}. \end{cases}$$

Check whether X and Y are random variables? If so, find the corresponding CDFs.

Solution As $X^{-1}\{(-\infty, 1]\} = \{a, b\} \notin \mathcal{F}$, the function X is not a random variable. Hence CDF of X is not defined.

Next, for Y we have

$$Y^{-1}\{(-\infty, y]\} = \begin{cases} \emptyset, & -\infty < y < 0 \\ \{a\}, & 0 \le y < 1 \\ \Omega, & 1 \le y < \infty. \end{cases}$$

Since, for every $y \in \mathbb{R}$, $Y^{-1}\{(-\infty, y]\} \in \mathcal{F}$, Y is a random variable. The CDF of Y is given by

$$F_Y(y) = \begin{cases} 0, & -\infty < y < 0 \\ \frac{1}{4}, & 0 \le y < 1 \\ 1, & 1 \le y < \infty. \end{cases}$$

□

To understand uncertainty in a static system, we employ standard probability theory. But to handle uncertainty in a dynamic system, e.g. the price of a share of stock, we need to go beyond to capture evolution over time. This can be studied via a collection of random variables namely a stochastic process with time as index set. Now, we define the stochastic process formally.

Definition 7.2.5 (Stochastic Process) *Let $(\Omega, \mathcal{F}, P)$ be a given probability space. A collection of random variables $\{X(t), t \in T\}$ defined on the probability space $(\Omega, \mathcal{F}, P)$ is called a stochastic process.*

A stochastic process is also called a random process or a chance process.

Remark 7.2.5 *Given a probability space $(\Omega, \mathcal{F}, P)$, the stochastic process $\{X(t), t \in T\}$ can be identified as a real-valued function $X : T \times \Omega \to \mathbb{R}$ of two independent variables $t \in T, w \in \Omega$ such that $X^{-1}(t)\{(-\infty, x]\}$ belongs to $\mathcal{F}$ for every $x \in \mathbb{R}$ and for every $t \in T$. Here*

$$X^{-1}(t)\{(-\infty, x]\} = \{w \in \Omega : X(t, w) \in (-\infty, x]\} .$$

In this sense, the stochastic process $\{X(t), t \in T\}$ can be thought of as $\{X(t, w) : t \in T, w \in \Omega\}$.

Remark 7.2.6 *The mapping X gives rise to two mappings, $X(\cdot, w)$ and $X(t, \cdot)$ as follows. For fixed $w \in \Omega$, the mapping $X(., w) : T \to \mathbb{R}$ takes t to $X(t, w)$. The sequence of real numbers $\{X(t, w), t \in T\}$ for fixed $w \in \Omega$ is called the realization or trajectory or sample path of the stochastic process. This is not a random variable. Next, for each fixed $t \in T$, the mapping $X(t, .) : \Omega \to \mathbb{R}$ takes w to $X(t, w)$ which is a random variable.*

Definition 7.2.6 (Parameter Space and State Space) *Let $\{X(t), t \in T\}$ be a given stochastic process. The set $\{t \in T\}$ is called the parameter space or index set. The collection of all possible values of $X(t)$ for $t \in T$ is called the state space. The state space is denoted by S.*

A given stochastic process is classified based on the values of state space and parameter space. For example, if both the state space and parameter space are finite or countably infinite, then it is called discrete time discrete space stochastic process. We shall be denoting a *discrete time stochastic process* by $\{X_t, t = 0, \pm 1, \pm 2, \ldots\}$ while the *continuous time stochastic process* will be denoted by $\{X(t), t \in T\}$. Note that, a time series $\{X_t, t = 0, \pm 1, \pm 2, \ldots\}$ is a discrete time stochastic process with $T = \mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$.

Based on parameter space T and state space S, we now present some of the examples of different stochastic processes.

(i) Continuous time discrete space stochastic process:
$T = \mathbb{R}_+ = [0, \infty)$, $S = \mathbb{N}$, e.g. total number of shares $\{X(t), t \in [0, \infty)\}$ held by an investor at any time t,

(ii) Discrete time continuous space stochastic process:
$T = \mathbb{N}$, $S = \mathbb{R}_+ = [0, \infty)$, e.g. $\{X_n, n = 1, 2, 3, \ldots\}$ value of one US dollar (USD) in rupees at the end of n-th day.

(iii) Continuous time continuous space stochastic process:
$T = \mathbb{R}^+ = [0, \infty)$, $S = \mathbb{R}_+ = [0, \infty)$, e.g. $\{X(t), t \in [0, \infty)\}$ stock price of a particular item at any time t.

(iv) Discrete time discrete space stochastic process:
$T = \mathbb{N}$, $S = \mathbb{N}$, e.g. total number of shares $\{X_n, n = 1, 2, 3 \ldots\}$ held by a particular person at the end of n-th day.

We already know that a random variable X can be characterized by its CDF $F_X(.)$. The main question now is to characterize a stochastic process. For this we need to define *finite dimensional distributions of the given process.*

Definition 7.2.7 (Finite Dimensional Distributions of a Stochastic Process) *Let $\{X(t), t \in T\}$ be a given stochastic process. Then finite dimensional distributions of the given process are the distributions of the finite dimensional random vectors $(X(t_1), X(t_2), \ldots, X(t_n))$, $t_1, t_2, \ldots, t_n \in T$, for all possible choices of times $t_1, t_2, \ldots, t_n \in T$ and for all $n \geq 1$.*

It can be shown that the stochastic process $\{X(t), t \in T\}$ is known completely if all its finite dimensional distributions are known.

Now we introduce certain properties of stochastic process $\{X(t), t \in T\}$.

Definition 7.2.8 (Strict Sense Stationary Stochastic Process) *The stochastic process $\{X(t), t \geq 0\}$ is called strict sense stationary stochastic process if, for arbitrarily chosen n, and for $0 < t_1 < t_2 < \ldots < t_n$, the finite dimensional random vectors $(X(t_1), X(t_2), \ldots, X(t_n))$ and $(X(t_1 + h), X(t_2 + h), \ldots, X(t_n + h))$ have the same joint distributions for all $h > 0$ and all $t_1, t_2, \ldots, t_n$.*

Note that a strict sense stationary stochastic process is invariant with respect to time translation. It is also called a *strong stationary stochastic process.*

Definition 7.2.9 (Wide Sense Stationary Stochastic Process) *The stochastic process $\{X(t), t \geq 0\}$ is wide sense stationary if it satisfies the following*

(i) $E(X(t))$ is independent of t,
(ii) $Cov(X(t), X(s))$ depends only on the time difference $|t - s|$ for all t, s,
(iii) $E((X(t))^2) < \infty$ (finite second order moment).

A wide sense stationary stochastic process is also called a *covariance stationary* or *weak stationary* or *second order stationary* stochastic process.

In other words, a stochastic process with finite second order moments is said to be weak stationary if its mean (or expectation) function is independent of t and covariance function is invariant under time shift.

Example 7.2.4 *Consider the process $X(t) = A\cos(\theta t) + B\sin(\theta t)$ where A and B are uncorrelated random variables with mean 0 and variance 1 and θ is a positive constant. Is $\{X(t), t \geq 0\}$ covariance or wide sense stationary stochastic process?*

Solution Since $E(A) = 0$ and $E(B) = 0$, we have, for a fixed t,

$$E(X(t)) = \cos(\theta t)E(A) + \sin(\theta t)E(B) = 0.$$

Hence $E(X(t))$ is independent of t.

Now for $s < t$,

$$\begin{aligned} Cov(X(t), X(s)) &= E((X(t) - 0)(X(s) - 0)) \\ &= E(X(t)X(s)) \\ &= E((A\cos(\theta t) + B\sin(\theta t))(A\cos(st) + B\sin(st))). \end{aligned}$$

Since $E(A^2) = 1$ and $E(AB) = E(A)E(B) = 0$, we get

$$\begin{aligned} Cov(X(t), X(s)) &= \cos(\theta t)\cos(\theta s) + \sin(\theta t)\sin(\theta s) \\ &= \cos(t - s)\theta. \end{aligned}$$

Hence $Cov(X(t), X(s))$ depends only on $|t - s|$.

Next for fixed t,

$$\begin{aligned} E(X^2(t)) &= Var(X(t)) - (E(X(t)))^2 \\ &= \cos^2(\theta t)Var(A) + \sin^2(\theta t)Var(B) + 2\cos(\theta t)\sin(\theta t)Cov(A, B) \\ &= \cos^2(\theta t) + \sin^2(\theta t) \\ &= 1. \end{aligned}$$

Hence, by Definition 7.2.9, $\{X(t), t \geq 0\}$ is a wide sense stationary process. □

Definition 7.2.10 (Stationary Increments) *The process $\{X(t), t \in T\}$ is said to have stationary increments if $X(t+\tau) - X(t)$ is a strict sense stationary process for all $\tau > 0$, i.e. whenever $s < t$, $X(t) - X(s)$ has the same distribution of $X(t-s) - X(0)$.*

Definition 7.2.11 (Independent Increments) *If for all n and for $t_1 < t_2 < \ldots < t_n$, $X(t_2) - X(t_1), X(t_3) - X(t_2), \ldots, X(t_n) - X(t_{n-1})$ are independent random variables, then the process is said to have independent increments.*

Definition 7.2.12 (Independent Process) *If for all n and for $t_1 < t_2 < \ldots < t_n$ the CDF satisfies $F(X, T) = \prod_{i=1}^{n} F(X(t_i), t_i)$ where $T = (t_1, t_2, \ldots, t_n)$ and $X = (X(t_1), X(t_2), \ldots X(t_n))$, then the process is called an independent process. Here $F(X, T)$ is the joint distribution function of the random vector X.*

Note that, a stochastic process having independent increments is different from a stochastic process which is an independent process.

Definition 7.2.13 (Markov Property) *A given process $\{X(t), t \in T\}$ is said to have Markov property if for all n and for all $0 < t_1 < t_2 < \ldots < t_n < t$, the conditional CDF satisfies*

$$P(X(t) \leq x / X(0) = x_0, X(t_1) = x_1, \ldots, X(t_n) = x_n) = P(X(t) \leq x / X(t_n) = x_n).$$

In other words, if the future prediction depends only on the current state of the stochastic process and does not depend on the past information, then it has the Markov property.

Definition 7.2.14 (Markov Process) *A given stochastic process $\{X(t), t \in T\}$ is said to be a Markov process if it satisfies Markov property.*

A Markov process is a stochastic process with property that, given the value of $X(s)$, the values of $X(t)$, $t > s$, do not depend on the values of $X(u)$, $u < s$, i.e. the probability of any particular future behavior of the process, when it's present state is known exactly, is not altered by additional knowledge concerning it's past behavior. For instance, binomial process, Poisson process and Brownian motion are examples of Markov process whereas Gaussian process is an example of non Markov process. All these processes will be discussed later in this chapter.

Stochastic processes can be classified according to different criteria. One of them is the distribution of finite dimensional random variables. Based on this, we list few stochastic processes which are related to stochastic study of financial mathematics.

Definition 7.2.15 (Bernoulli Process) *A discrete time discrete space stochastic process $\{X_n, n = 1, 2, \ldots\}$ is called a Bernoulli process if for each n, X_n is a Bernoulli random variable with parameter p, $0 < p < 1$.*

Here we may note that the Bernoulli process is strict sense as well as wide sense stationary process.

Definition 7.2.16 (Binomial Process) *Let for $n = 1, 2, \ldots$, $S_n = X_1 + X_2 + \ldots + X_n$ where X_i's are mutually independent Bernoulli distributed random variables with parameter p, $0 < p < 1$. Then $\{S_n, n = 1, 2, \ldots\}$ is called a binomial process.*

Here we may note that $\{S_n, n = 1, 2, \ldots\}$ is a discrete time discrete space stochastic process. Also each S_n is a binomial distributed random variable.

Lemma 7.2.1 *The binomial process $\{S_n, n = 1,2,\ldots\}$ is a Markov process.*

Proof. For the binomial process $\{S_n, n = 1,2,\ldots\}$, we have, for $n = 2,3,\ldots$,

$$S_n = S_{n-1} + X_n .$$

Hence,

$$P(S_n = x/S_{n-1} = x_{n-1},\ldots,S_1 = x_1) = \frac{P(S_n = x, S_{n-1} = x_{n-1},\ldots,S_1 = x_1)}{P(S_{n-1} = x_{n-1},\ldots,S_1 = x_1)}$$

$$= \frac{P(S_n = x/S_{n-1} = x_{n-1})P(S_{n-1} = x_{n-1}/S_{n-2} = x_{n-2})\ldots P(S_2 = x_2/S_1 = x_1)P(S_1 = x_1)}{P(S_{n-1} = x_{n-1}/S_{n-2} = x_{n-2})\ldots P(S_2 = x_2/S_1 = x_1)P(S_1 = x_1)}$$

$$= \frac{P(X_n = x - x_{n-1})P(X_{n-1} = x_{n-1} - x_{n-2})\ldots P(X_2 = x_2 - x_1)P(S_1 = x_1)}{P(X_{n-1} = x_{n-1} - x_{n-2})\ldots P(X_2 = x_2 - x_1)P(S_1 = x_1)}$$

$$= P(X_n = x - x_{n-1})$$

$$= P(S_n = x/S_{n-1} = x_{n-1}) .$$

Thus $\{S_n, n = 1,2,\ldots\}$ satisfies the Markov property. Hence $\{S_n, n = 1,2,\ldots\}$ is a Markov process.

□

Definition 7.2.17 (Gaussian Process) *A given stochastic process $\{X(t), t \in T\}$ is said to be a Gaussian process if, for $t_1, t_2, \ldots, t_n$ and all n, $(X(t_1), X(t_2), \ldots, X(t_n))$ is a multivariate normal distributed random vector.*

We may observe the following for a Gaussian process.

(i) The joint pdf of $(X(t_1), X(t_2), \ldots, X(t_n))$ is given by

$$f_X(x) = \frac{1}{(2\pi)^{n/2}(det(\Sigma))^{1/2}} \exp\left[-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right] ,$$

where

$$\mu = E(X) = (E(X(t_1)), E(X(t_2)), \ldots, E(X(t_n))) ,$$
$$\Sigma = (Cov(X(t_i), X(t_j)); i, j = 1,2,\ldots,n),$$

and

$$Cov(X(t_i), X(t_j)) = E[(X(t_i) - E(X(t_i))(X(t_j) - E(X(t_j)))] .$$

(ii) A Gaussian process which is weak stationary is also strict sense stationary. Further, in any Gaussian process, its finite dimensional random variables are independent if and only if those random variables are uncorrelated, i.e. the corresponding covariance matrix Σ is a diagonal matrix.

Definition 7.2.18 (Symmetric Random Walk) *Consider a random experiment of tossing a fair coin infinitely many times. Let the successive outcomes be denoted as $w = (w_1, w_2, w_3, \ldots)$ e.g. $w = (w_1, w_2, w_3, \ldots) = (H, T, T, \ldots)$ or $(T, H, T, \ldots)$ etc. We now define, for $j = 1,2,\ldots$*

$$X_j = \begin{cases} 1, & \text{if } w_j = H \\ -1, & \text{if } w_j = T, \end{cases}$$

and

$$P(X_j = 1) = P(w_j = H) = 0.5 ; \quad P(X_j = -1) = P(w_j = T) = 0.5 .$$

Set $S_0 = 0$. Let

$$S_k = \sum_{j=1}^{k} X_j, \quad (k = 1,2,\ldots) .$$

Then, $\{S_k, k = 0,1,\ldots\}$ is known as a symmetric random walk.

Now, we present the sample path of a random walk $\{S(k), k = 0,1,\ldots\}$ for fixed $w \in \Omega$. Suppose $w = (w_1, w_2, w_3, \ldots) = (H,H,H,T,H,H,T,H,H,H,H,\ldots)$. This realization is shown in Fig. 7.2. This symmetric random walk $\{S_k, k = 0,1,\ldots\}$ is a discrete time, discrete space stochastic process. By considering continuous time and continuous space, we get Brownian motion which will be discussed in the next section.

Next, we present the following results for a symmetric random walk.

Theorem 7.2.1 *Let $\{S_k, k = 0,1,2,\ldots\}$ be a symmetric random walk. Then*

(i) for each k, $E(S_k) = 0$ and $Var(S_k) = k$,
(ii) it has independent increments,
(iii) it has stationary increments,
(iv) it is a Markov process.

Proof.

(i) For $j = 1,2,\ldots$, we have $E(X_j) = 0$ and $Var(X_j) = 1$. Therefore, for $k = 1,2,\ldots$,

$$E(S_k) = 0 \text{ and } Var(S_k) = k .$$

Here we are using the fact that X_j's are mutually independent.

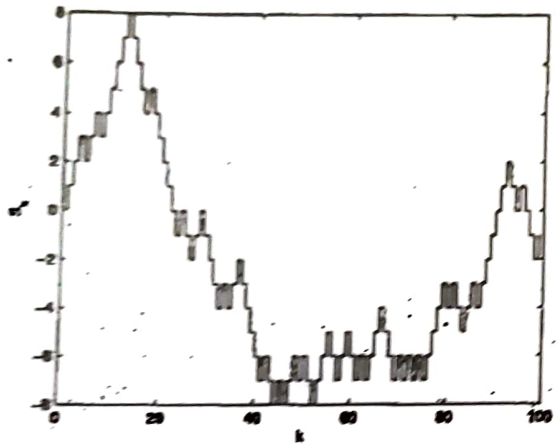

Fig. 7.2. Sample path of symmetric random walk

(ii) We choose an arbitrary positive integer n and then choose non-negative integers $0 = k_0 < k_1 < \ldots < k_n$. Then

$$S_{k_{i+1}} - S_{k_i} = \sum_{j=k_i+1}^{k_{i+1}} X_j .$$

Since X_j are independent and identically distributed (i.i.d) random variables having Bernoulli distribution, $S_{k_1} - S_{k_0}, S_{k_2} - S_{k_1}, \ldots, S_{k_n} - S_{k_{n-1}}$ are mutually independent variables. Hence the stochastic process $\{S_k, k = 0, 1, \ldots\}$ has the independent increment property.

(iii) Choose non-negative integers $k_1 < k_2$. Then

$$S_{k_2} - S_{k_1} = \sum_{j=k_1}^{k_2} X(j) .$$

Since X_j are i.i.d random variables having Bernoulli distribution, $S_{k_2} - S_{k_1}$ has the same distribution of $S_{k_2-k_1} - S_0$. Hence the stochastic process $\{S_k, k = 0, 1, \ldots\}$ has the stationary increment property.

(iv) We have for $k = 1, 2, \ldots$

$$S_k = S_{k-1} + X_k .$$

Now,

$$P(S_k \le x / S_{k-1} = x_{k-1}, \ldots, S_1 = x_1 = \frac{P(S_k \le x, S_{k-1} = x_{k-1}, \ldots, S_1 = x_1)}{P(S_{k-1} = x_{k-1}, \ldots, S_1 = x_1)}$$

$$= \frac{P(S_k \le x / S_{k-1} = x_{k-1}) \ldots P(S_2 = x_2 / S_1 = x_1) P(S_1 = x_1)}{P(S_{k-1} = x_{k-1} / S_{k-2} = x_{k-2}) \ldots P(S_2 = x_2 / S_1 = x_1) P(S_1 = x_1)}$$

$$= P(S_k \le x / S_{k-1} = x_{k-1}) .$$

From the above discussion we can easily conclude that $\{S_k, k = 1, 2, \ldots\}$ is a Markov process. □

Definition 7.2.19 (Poisson Process) *A stochastic process $\{N(t), t \ge 0\}$ is said to be a Poisson process with parameter $\lambda > 0$, if it satisfies the following properties*

(i) $N(0) = 0$,
(ii) For all n and for all $0 \le t_0 < t_1 < t_2 < \cdots < t_n$, increments $N(t_i) - N(t_{i-1})$, $i = 1, 2, \ldots, n$ are independent and stationary,
(iii) For $0 \le s < t$, $N(t) - N(s)$ is a Poisson distributed random variable with parameter $\lambda(t - s)$, i.e.,

$$P(N(t) - N(s) = n) = \frac{(\lambda(t-s))^n e^{-\lambda(t-s)}}{n!}, \quad (n = 0, 1, 2, \ldots) .$$

The sample path of a Poisson process with parameter $\lambda = 2$ is shown in Fig. 7.3. It is observed that in any finite interval, the sample path has finite jumps of size one and in the interval $[0, \infty)$ it has countably infinite jumps.

Let $\{N(t), t \ge 0\}$ be a Poisson process. Then, for each t, $E(N(t)) = \lambda t$ and $Var(N(t)) = \lambda t$.

Remark 7.2.7 *By property (ii) of Definition 7.2.19, for every $s < t$, the increment $N(t) - N(s)$ is independent from the history of the process up to time s and has the same law as $N(t - s)$. Therefore $\{N(t), t \ge 0\}$ satisfies Markov property. Hence, it is a Markov process.*

Remark 7.2.8 *Suppose $N(t)$ denotes the number of events that occur in the interval $[0, t]$. Let $T_i, i = 1, 2, \ldots$ denote the occurrence of events. The inter arrival of successive events $T_1, T_2 - T_1, \ldots, T_{i+1} - T_i$ are independent and are exponentially distributed with parameter λ.*

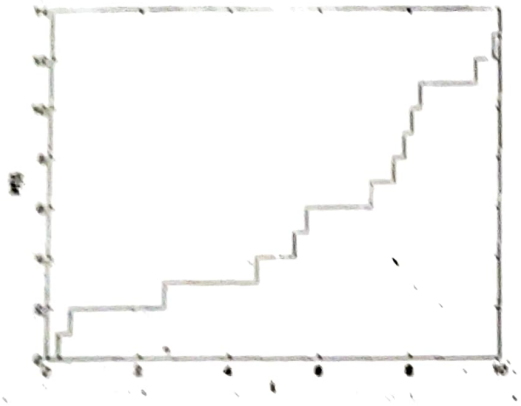

Fig. 7.5. Sample path of Poisson process with $\lambda = 2$

Example 7.2.5 *Consider a financial risk model that may occur in an insurance company. Let $N(t)$ denote the number of claims received by time t, $t \geq 0$. Thus, $N(t)$ counts the number of claims received in $(0, t]$. Suppose the process $\{N(t), t \geq 0\}$ is a Poisson process with claim arrival rate 9 per month of 30 days. In a randomly chosen month of 30 days,*

(i) what is the probability that there are exactly 4 claims were received in the first 15 days?

(ii) given that exactly 4 claims were received in the first 15 days, what is the probability that all the four claims were received in the last 7 days out of these 15 days?

Solution We have

$$P(N(t) = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}, \quad (k = 0, 1, \ldots),$$

where $\lambda = \frac{9}{30}$ per day.

(i) The required probability is

$$P(N(15) = 4) = \frac{e^{-\frac{3}{2}}9^4}{384}.$$

(ii) The required probability is

$$\frac{P(N(8) = 0)P(N(7) = 4)}{P(N(15) = 4)} = \frac{e^{-36/15}e^{-63/30}(63/30)^4}{e^{-9/2}(9/2)^4} = \left(\frac{7}{15}\right)^4.$$

□

7.3 Brownian Motion and its Properties

The long studied model known as Brownian motion, is named after the English botanist Robert Brown. In 1827, Brown described the unusual motion exhibited by a small particle totally immersed in a liquid or a gas. Early nineteenth century, Brownian motion is introduced to model the price movements of stocks and commodities. A formal mathematical description of Brownian motion and its properties was first given by the great mathematician Norbert Wiener beginning 1918.

Definition 7.3.1 (Brownian Motion) *A stochastic process $\{W(t), t \geq 0\}$ is said to be a Brownian motion (BM), if it satisfies the following properties*

(i) $W(0) = 0$,

(ii) for $t > 0$, the sample path of $W(t)$ is continuous,

(iii) for all n and for all $0 \leq t_0 < t_1 < t_2 < \cdots < t_n$, increments $W(t_i) - W(t_{i-1})$, $i = 1, 2, \ldots, n$ are independent and stationary,

(iv) for $0 \leq s < t < \infty$, $W(t) - W(s)$ is normally distributed random variable, with mean 0 and variance $(t - s)$.

A Brownian motion is also called a Wiener process. The sample path of Wiener process is shown in Fig. 7.4. Wiener process can also be visualized as a scaling limit of a symmetric random walk.

Remark 7.3.1 *We observe that, the sample paths are always continuous functions. In other words, the sample paths do not have any jumps. On the other hand, these paths are also essentially nowhere differentiable, that is, it is not possible to define a unique tangent line at any point on a curve. We can show this by using the convergence in the second order moment. For this we shall like to show that*

$$\lim_{\Delta t \to 0} Var\left(\frac{W(t_0 + \Delta t) - W(t_0)}{\Delta t}\right)$$

does not exist for any arbitrary point t_0.

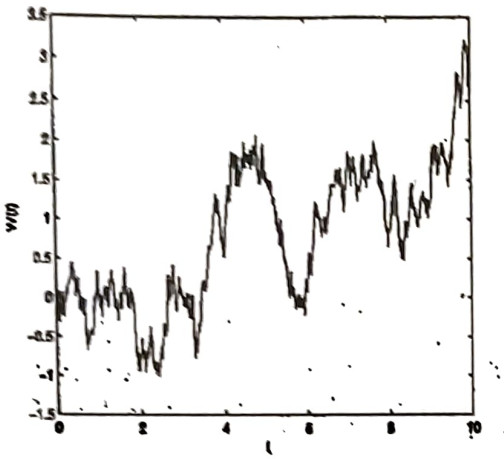

Fig. 7.4. Sample path of Wiener process

We know that, $W(t_0+\Delta t)-W(t_0)$ has normal distribution with mean zero and variance Δt, i.e. $N(0,\Delta t)$. Hence,

$$\lim_{\Delta t\to 0} Var\left(\frac{W(t_0+\Delta t)-W(t_0)}{\Delta t}\right)=\lim_{\Delta t\to 0}\frac{1}{(\Delta t)^2}\times\Delta t=\lim_{\Delta t\to 0}\frac{1}{\Delta t}.$$

Thus the above limit does not exist. Hence, we conclude that Brownian motion has non-differentiability at a fixed point. We observe that nowhere differentiability of Brownian motion requires a more careful argument than non-differentiability at a fixed point. The precise mathematical proof is not presented in this book but rather the above argument may convince to conclude nowhere differentiability of Brownian motion (refer Karatzas and Shreve [75] for proof).

Remark 7.3.2 *The Wiener process is not wide sense stationary. This is because for $s<t$, the covariance function $Cov(W(t),W(s))$ is not a function of $(t-s)$. In fact*

$$\begin{aligned}Cov(W(t),W(s)) &= E[(W(t)-E(W(t)))(W(s)-E(W(s)))]\\ &= E[W(t)W(s)]\\ &= E[(W(t)-W(s)+W(s))W(s)]\\ &= E[W(t)-W(s)]E[W(s)]+E[(W(s))^2]\\ &= (0\times 0)+s=s.\end{aligned}$$

Hence, $Cov(W(t),W(s))=Min(s,t)$.

Remark 7.3.3 *Given $W(t)$, the future $W(t+h)$ for any $h>0$ only depends on the increment $W(t+h)-W(t)$ and this is independent of the past. Hence by Definition 7.3.1, the Markov property is satisfied. Thus, $\{W(t),t\geq 0\}$ is a Markov process.*

Lemma 7.3.1 *The joint probability density function of $(W(t_1),W(t_2))$ is given by*

$$f(x_1,x_2)=\frac{1}{2\pi\sqrt{t_1(t_2-t_1)}}\exp\left[-\frac{1}{2}\left\{\frac{x_1^2}{t_1}+\frac{(x_2-x_1)^2}{(t_2-t_1)}\right\}\right].$$

Proof. From Definition 7.2.17, a Wiener process is also a Gaussian process. Since $\{W(t),t\geq 0\}$ is a Markov process as well as a Gaussian process, we have

$$\begin{aligned}P(W(t)\leq x/W(t_n)=x_n) &= P(W(t)-W(t_n)\leq x-x_n)\\ &= \int_{-\infty}^{x-x_n}\frac{1}{\sqrt{2\pi(t-t_n)}}\exp\left[-\frac{s^2}{2(t-t_n)}\right]ds.\end{aligned}$$

Let us now consider the joint distribution of $(W(t_1),W(t_2))$. We know that, $W(t_1)$ and $(W(t_2)-W(t_1))$ are independent. Also, $W(t_1)$ is $N(0,t_1)$ and $W(t_2)-W(t_1)$ is $N(0,t_2-t_1)$. Therefore, the joint probability density function of $(W(t_1),W(t_2))$ is

$$f(x_1,x_2)=p(x_1,t_1)p(x_2-x_1,t_2-t_1),$$

where

$$p(x,t)=\frac{1}{\sqrt{2\pi t}}\exp\left[-\frac{x^2}{2t}\right],\ t>0,\ -\infty<x<\infty.$$

Thus,

$$f(x_1,x_2)=\frac{1}{2\pi\sqrt{t_1(t_2-t_1)}}\exp\left[-\frac{1}{2}\left\{\frac{x_1^2}{t_1}+\frac{(x_2-x_1)^2}{(t_2-t_1)}\right\}\right].$$

□

Remark 7.3.4 *The above Lemma can be generalized in the sense that for $0<t_1<t_2<\cdots<t_n$, $X=(W(t_1),W(t_2),\ldots,W(t_n))$ is jointly normally distributed with the joint probability density function given by*

$$f_X(x)=\frac{1}{(2\pi)^{n/2}(det(\Sigma))^{1/2}}\exp\left[-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right],$$

where

$$\mu = E(X) = (E(W(t_1)), E(W(t_2)), \ldots, E(W(t_n))) \in \mathbb{R}^n$$

and

$$\Sigma = [Cov(W(t_i), W(t_j)); i, j = 1, 2, \ldots, n] \in \mathbb{R}^{n \times n}.$$

We now present some important properties of Brownian motion and refer to Shreve [122] for their proofs.

Result 7.3.1 *Let $\{W(t), t \geq 0\}$ be a Wiener process. Then,*

(i) $\{-W(t), t \geq 0\}$ is a Wiener process. (symmetric property)

(ii) $\{\frac{1}{\sqrt{c}}W(ct), t \geq 0\}$ is a Wiener process for each fixed $c > 0$. (scaling property)

(iii) Taking $\widetilde{W}(0) = 0, \widetilde{W}(t) = tW(\frac{1}{t}), t > 0$, the process $\{\widetilde{W}(t), t \geq 0\}$ is a Wiener process. (time inversion property)

Example 7.3.1 *Let $\{W(t), t \geq 0\}$ be a Wiener process. Find the conditional distribution of $W(s)$ for $0 < s < t$, given $W(t) = x$.*

Solution There are two possible approaches to solve this problem. In the first approach we find the joint distribution of $(W(t), W(s))$ and then find the conditional distribution. In the second approach we use the time inversion property in Result 7.3.1. The required conditional distribution is $sW(\frac{1}{s})$ given that $tW(\frac{1}{t}) = x$. This is same as the distribution of $sW(\frac{1}{s}) - sW(\frac{1}{t}) + sW(\frac{1}{t})$ given $W(\frac{1}{t}) = \frac{x}{t}$. Hence, we need to find the distribution of $s\left(W(\frac{1}{s}) - W(\frac{1}{t})\right) + \frac{sx}{t}$. This distribution is a normal distribution with mean $\frac{sx}{t}$ and variance $s^2\left(\frac{1}{s} - \frac{1}{t}\right)$.

□

7.4 Processes Derived from Brownian Motion

A general Brownian motion need not have $W(0) = 0$ and $\sigma^2 = 1$. Therefore we define a general Brownian motion with drift μ and variance σ^2 as follows.

Definition 7.4.1 **(Brownian motion with drift μ and volatility σ)** *A stochastic process $\{X(t), t \geq 0\}$ is said to be a Brownian motion with drift μ and volatility σ if $X(t) = \mu t + \sigma W(t)$ where (i) $W(t)$ is a standard Brownian motion, (ii) $-\infty < \mu < \infty$ is a constant and (iii) $\sigma > 0$ is a constant.*

This is a generalization of standard Brownian motion. In this process, the mean function $E(X(t)) = \mu t$ and covariance function $Cov(W(s), W(t)) = \sigma^2 \text{Min}(s, t)$, $s, t \geq 0$. Fig. 7.5 shows a sample path of Brownian motion with drift $\mu = 0.1$ and volatility $\sigma^2 = 1.5$.

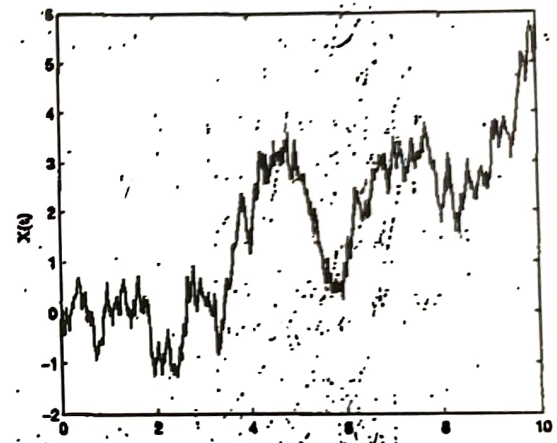

Fig. 7.5. Sample path of Brownian motion with $\mu = 0.1$ $\sigma^2 = 1.5$

Definition 7.4.2 **(Brownian Bridge)** *A standard Brownian bridge $\{X(t), 0 \leq t \leq 1\}$ is defined as $X(t) = W(t) - tW(1)$, where $W(t)$ is a standard Brownian motion. Clearly, $X(0) = 0 = X(1)$.*

Since for $0 < t < 1$, $W(t) \sim N(0, t)$, $X(t) \sim N(0, t(1-t))$. The covariance function is $s(1-t)$ for $0 \leq s \leq t \leq 1$. Therefore, the Brownian bridge is a Gaussian process but is not a Brownian motion. For fixed $T > 0$, the general Brownian bridge $\{X(t), 0 \leq t \leq T\}$ can be defined as

$$X(t) = W(t) - \frac{t}{T}W(T), \quad 0 \leq t \leq T.$$

The covariance function is given by

$$
\begin{aligned}
Cov(X(s), X(t)) &= E[X(s)X(t)] - E(X(s))E(X(t)) \\
&= E\left[W(t)W(s) + \frac{st}{T^2}W^2(T) - \frac{t}{T}W(s)W(T) - \frac{s}{T}W(t)W(T)\right] \\
&= \text{Min}\{s,t\} + \frac{st}{T} - \frac{t}{T}\text{Min}\{s,T\} - \frac{s}{T}\text{Min}\{t,T\} \\
&= \text{Min}\{s,t\} - \frac{st}{T}.
\end{aligned}
$$

Brownian motion model for stock prices is open to the objection that prices are by definition positive quantities while $W(t)$, being normally distributed, is negative with strictly positive probability. In 1965, Samuelson [117] introduced what has now become the standard model, namely geometric Brownian motion.

Definition 7.4.3 (Geometric Brownian Motion) *A stochastic process $\{X(t), t \geq 0\}$ is said to be a geometric Brownian motion (GBM) if $X(t) = X(0)\, e^{W(t)}$ where $W(t)$ is a standard Brownian motion.*

$X(0)\,e$

Result 7.4.1 *For any $h > 0$, we have*

$$
\begin{aligned}
X(t+h) &= X(0)\, e^{W(t+h)} \\
&= X(0)\, e^{W(t)+W(t+h)-W(t)} \\
&= X(t)\, e^{W(t+h)-W(t)}.
\end{aligned}
$$

We note that, BM has independent increments. Hence given $X(t)$, the future $X(t+h)$ only depends on the future increment of the BM. Thus future is independent of the past and therefore the Markov property is satisfied. Hence, $\{X(t), t \geq 0\}$ is a Markov process.

Because a geometric Brownian motion is nonnegative, it provides for a more realistic model of stock prices. Also, the GBM model considers the ratio of stock prices to have the same normal distribution. Therefore, the percentage change in price as opposed to the absolute change in price is modeled by a GBM. Fig. 7.6 shows a sample path of geometric Brownian motion.

How does geometric Brownian motion relate to stock prices? One possibility is to think of modeling the rate of return of the stock price as a Brownian motion. *Suppose that the stock price $S(t)$ at time t is given by*

$$S(t) = S(0)\, e^{H(t)},$$

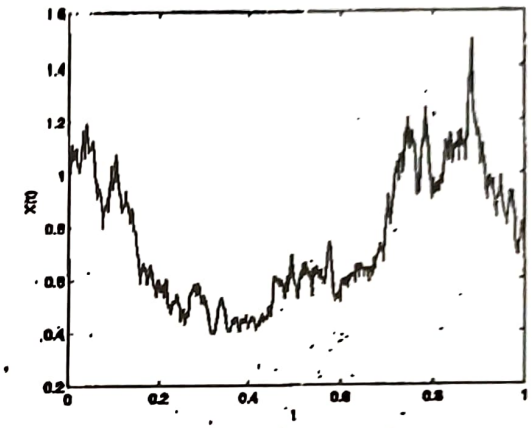

Fig. 7.6. Sample path of geometric Brownian motion

where $S(0)$ is the initial price and $H(t) = \mu t + \sigma W(t)$ is a Brownian motion with drift. In this case, $H(t)$ represents a continuously compounded rate of return of the stock price over the period of time $[0, t]$. Here, $H(t)$ refer to the logarithmic growth of the stock price, satisfies

$$H(t) = ln\left(\frac{S(t)}{S(0)}\right).$$

This gives

$$ln(S(t)) = ln(S(0)) + H(t).$$

Therefore, $ln(S(t))$ has a normal distribution with mean $\mu t + ln(S(0))$ and variance $\sigma^2 t$. As we have seen, if a random variable X has the property that lnX has normal distribution, then the random variable X is said to have a lognormal distribution. Accordingly, $S(t)/S(0)$ is lognormal distributed random variable.

Example 7.4.1 *Suppose that the stock price $S(t)$ at time t is given by $S(t) = S(0)\, e^{H(t)}$ where $S(0)$ is the initial price and $H(t) = \mu t + \sigma W(t)$ is a Brownian motion with drift μ and volatility σ. Prove that*

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

(ii) $Var(S(t)) = \left(S(0)\, \exp\left(\left(\mu + \frac{\sigma^2}{2}\right)t\right)\right)^2 (\exp(\sigma^2 t) - 1)$.

Solution We know that, if a random variable X is normally distributed with mean μ and variance σ^2, then the moment generating function of X is given by

$$M_X(\theta) = E\left(e^{\theta X}\right) = \exp\left(\mu\theta + \frac{1}{2}\sigma^2\theta^2\right). \tag{7.1}$$

Since for every t, $W(t)$ is normally distributed with mean zero and variance t, $H(t) = \mu t + \sigma W(t)$ is normally distributed with mean μt and variance $\sigma^2 t$. Hence,

$$M_{H(t)}(\theta) = E\left(e^{\theta H(t)}\right) = \exp\left(\mu t\theta + \frac{1}{2}\sigma^2 t\theta^2\right). \tag{7.2}$$

(i) Now

$$\begin{aligned} E(S(t)) &= E\left(S(0)e^{H(t)}\right) \\ &= S(0)E\left(e^{H(t)}\right) \end{aligned}$$

Using (7.2) and substituting $\theta = 1$, we obtain

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

(ii)

$$\begin{aligned} Var(S(t)) &= E(S^2(t)) - (E(S(t)))^2 \\ &= E\left(S^2(0)e^{2H(t)}\right) - S^2(0)\exp\left(2\left(\mu + \frac{\sigma^2}{2}\right)t\right) \end{aligned}$$

Using (7.2) and substituting $\theta = 2$, we obtain

$$\begin{aligned} Var(S(t)) &= S^2(0)\exp\left(2\left(\mu + \sigma^2\right)t\right) - S^2(0)\exp\left(2\left(\mu + \frac{\sigma^2}{2}\right)t\right) \\ &= \left(S(0)\exp\left(\left(\mu + \frac{\sigma^2}{2}\right)t\right)\right)^2\left(\exp\left(\sigma^2 t\right) - 1\right). \end{aligned}$$

□

Remark 7.4.1 *Letting $\bar{r} = \mu + \frac{1}{2}\sigma^2$, we get*

$$E(S(t)) = S(0)\, e^{\bar{r}t}$$

$$Var(S(t)) = \left(S(0)\, e^{(\mu+\frac{\sigma^2}{2})t}\right)^2\left(e^{\sigma^2 t} - 1\right).$$

Here we observe that, the expected stock price depends not only on the drift μ of $H(t)$ but also on the volatility σ. Further, it shows that, the expected price grows like a fixed-income security with continuously compounded interest rate $\bar{r}$. In real scenario, r is much lower than $\bar{r}$, the real fixed-income interest rate, that is why one invests in stocks. But the stock has variability due to the randomness of the underlying Brownian motion and hence a risk is involved here.

Example 7.4.2 *Suppose that stock price $\{S(t), t \geq 0\}$ follows geometric Brownian motion with drift $\mu = 0.12$ per year and volatility $\sigma = 0.24$ per annum. Assume that, the current price of the stock is $S(0) =$ Rs 40. What is the probability that a European call option having four years to exercise time and with a strike price $K =$ Rs 42, will be exercised?*

Solution We have,

$$\begin{aligned} P(S(4) > 42) &= P\left(\frac{S(4)}{40} > \frac{42}{40}\right) \\ &= P\left(ln\left(\frac{S(4)}{40}\right) > ln\left(\frac{42}{40}\right)\right). \end{aligned}$$

Since $ln\left(\frac{S(4)}{40}\right)$ follows normal distribution with mean 0.48 and variance $(0.48)^2$, we get

$$\begin{aligned} P\left(ln\left(\frac{S(4)}{40}\right) > ln\left(\frac{42}{40}\right)\right) &= P\left(\frac{ln\left(\frac{S(4)}{40}\right) - 0.48}{0.48} > \frac{ln\left(\frac{42}{40}\right) - 0.48}{0.48}\right) \\ &= 1 - \Phi\left(\frac{ln\left(\frac{42}{40}\right) - 0.48}{0.48}\right) \\ &= 1 - \Phi(-0.8983) \\ &= \Phi(0.8983) = 0.3133, \end{aligned}$$

where Φ is the cumulative standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy.$$

□

Definition 7.4.4 **(Ornstein-Uhlenbeck Process)** *Let $\{W(t), t \geq 0\}$ be a Wiener process. Define*

$$X(t) = X(0)e^{-at} + be^{-at}W(e^{2at} - 1),$$

where a and b are strictly positive real numbers and $X(0)$ is independent of $W(t)$. Then, we say $\{X(t), t \geq 0\}$ is an Ornstein-Uhlenbeck process.

We note that, $\{X(t), t \geq 0\}$ is a Markov process. It is also a Gaussian process if $X(0) = x(0)$ is fixed or $X(0)$ is Gaussian.

In the "real" world, we observe that asset price processes have jumps or spikes and risk-managers have to take them into account. We need processes that can describe the observed reality of financial markets in a more accurate way than models based on Brownian motion. Levy processes provide us with the appropriate framework to model both in the "real" and in the "risk-neutral" world.

Definition 7.4.5 **(Levy Process)** *A stochastic process $\{X(t), t \geq 0\}$ is said to be a Levy process if it satisfies the following properties*

(i) $X(0) = 0$,
(ii) for all n and for $0 \leq t_0 < t_1 < t_2 < \ldots < t_n$, increments $X(t_i) - X(t_{i-1})$, $i = 1, 2, \ldots, n$, are independent and stationary,
(iii) for $a > 0$, $P(|X(t) - X(s)| > a) \to 0$ when $t \to s$.

Remark 7.4.2 *Let b be a constant and $X(t) = bt$. Then $\{X(t), t \geq 0\}$ is a Levy process.*

Remark 7.4.3 *A Wiener process $\{W(t), t \geq 0\}$ defined in Definition 7.3.1, is a Levy process in $\mathbf{R}$ that has continuous paths and has the Gaussian distribution with mean zero and variance Δt for its increments $W(t + \Delta t) - W(t)$. The most general continuous Levy process in $\mathbf{R}$ has the form $X(t) = bt + cW(t)$, $t \geq 0$, where b and c are real constants.*

Remark 7.4.4 *A Poisson process $\{N(t), t \geq 0\}$ with parameter λ defined in Definition 7.2.19 is a Levy process that is a counting process having the Poisson distribution with mean $\lambda\Delta t$ for its increments $N(t + \Delta t) - N(t)$.*

The above three processes namely deterministic process $X(t) = bt$, Wiener process and Poisson process are Levy processes. It turns out that all Levy processes can be built up out of these building blocks. We may refer to Applebaum [4] for further reading in this regard.

7.5 Summary and Additional Notes

- As early as 1900, Louis Bachelier proposed Brownian motion (also, named as Wiener process) as a model of the fluctuations of stock prices. Since the 1970s, the Wiener process has been widely applied in financial mathematics and economics to model the evolution in time of stock prices and bond interest rates.
- The earlist attempt to model Brownian motion mathematically can be traced to three different sources. The first was that of T.N. Thiele in 1880 who effectively created a model of Brownian motion while studying time series; the second was that of L. Bachelier (mentioned above) and the third was that of Einstien, who in 1905 proposed a model of the motion of small particles suspended in a liquid so as to convince other physicists of the molecular nature of matter. The readers may refer to Vassilion [144] for further historical details.
- Section 7.2 presents the basics of stochastic process while the Brownian motion is introduced in Section 7.3.
- In Section 7.4, processes derived from Brownian motion are presented. It includes geometric Brownian motion which has been of tremendous use in modern finance theory.

7.6 Exercises

Exercise 7.1 *Consider the binomial model for trading in stock, $t = 1, 2$, where at each time the stock can go up by the factor u or down by the factor d. The sample space $\Omega = \{(u,u), (u,d), (d,u), (d,d)\}$. Create one non-trivial σ-field and the largest σ-field on Ω.*

Exercise 7.2 *Construct an example of σ-fields $\mathcal{F}(\infty)$ and $\mathcal{F}(\epsilon)$ such that $\mathcal{F}(\infty) \cup \mathcal{F}(\epsilon)$ is not a σ-field.*

Exercise 7.3 *Prove that there does not exist a σ-field which contains exactly 6 elements.*

Exercise 7.4 *Consider $\Omega = \{1,2,3,4\}$. Let σ-algebra $\mathcal{F} = \{\emptyset, \{1\}, \{2,3,4\}, \Omega\}$. Construct a random variable on the measurable space $(\Omega, \mathcal{F})$.*

Exercise 7.5 *Consider a random experiment of tossing an unbiased coin three times. Let Ω denote the set of all possible outcomes. Let the random variable X be the number of heads observed. Find the cumulative distribution function $F(X, x)$.*

Exercise 7.6 *Let X_n, for n even take values $+1$ and -1 each with probability 0.5, and for n odd, take values $\sqrt{a}, \frac{-1}{\sqrt{a}}$ with probability $\frac{1}{a+1}, \frac{a}{a+1}$ respectively ($a > 0$, $a \neq 1$). Further, let X_n's be independent, show that the stochastic process $\{X_n, n \geq 1\}$ is wide sense stationary but not strict stationary.*

x

Exercise 7.7 *Let $\{X(t), t \geq 0\}$ be a stochastic process with independent increments and $X(0) = 0$. Show that $Cov(X(s), X(t)) = Var(X(Min(s,t)))$, for any s, t $(t, s \geq 0)$.*

Exercise 7.8 *Let X and Y be i.i.d. random variables each having uniform distribution on the interval $(-\pi, \pi)$. Let $Z(t) = \cos(tX + Y), t \geq 0$. Is $\{Z(t), t \geq 0\}$ wide sense stationary process?*

Exercise 7.9 *Consider an urn containing 100 red balls and 100 black balls. Balls are drawn one by one without replacement. Let X_n be the number of red balls remaining in the urn after the nth ball is drawn. Is $\{X_n, n = 0, 1, \ldots\}$ a Markov process?*

Exercise 7.10 *Let $Y_n = a_0 X_n + a_1 X_{n-1}, n = 1, 2, \ldots$ where a_0, a_1 are constants and X_n, $(n = 0, 1, \ldots)$ are i.i.d random variables with mean 0 and variance σ^2. Is $\{Y_n, n = 1, 2, \ldots\}$ a Markov process?*

Exercise 7.11 *Let $\{N(t), t \geq 0\}$ be a Poisson process with parameter λ. Suppose $N(t)$ denotes the number of events that occur in the interval $[0, t]$, prove that the inter arrival of successive events are independent and are exponentially distributed with parameter λ.*

Exercise 7.12 *Let $\{W(t), t \geq 0\}$ be the Brownian motion. Prove that, $E((W(t) - W(s))^4) = 3(t - s)^2$.*

Exercise 7.13 *Let $\{W(t), t \geq 0\}$ be the Brownian motion. Prove that $(W(t_1), W(t_2), \ldots, W(t_n))$ is jointly normal distributed with CDF for $0 < t_1 < t_2 < \cdots < t_n$ is given by*

$$P(W(t_1) \leq a_1, W(t_2) \leq a_2, \ldots, W(t_n) \leq a_n) = \frac{1}{\sqrt{(2\pi)^n t_1 (t_2 - t_1) \cdots (t_n - t_{n-1})}} \times$$

$$\int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} exp\left[-\frac{1}{2}\left(\frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \cdots + \frac{(x_n - x_{n-1})^2}{t_n - t_{n-1}}\right)\right] dx_1 \cdots dx_n .$$

Exercise 7.14 *Let X be a normally distributed random variable with mean μ and variance σ^2. Let u be a fixed number in $\mathbb{R}$ and define the convex function $\phi(x) = e^{ux}$ for all $x \in \mathbb{R}$. Prove that*

(i) $E(\phi(X)) = e^{u\mu + \frac{1}{2}u^2\sigma^2}$.
(ii) Verify the Jensen's inequality holds $E(\phi(X)) \geq \phi(E(X))$.

Exercise 7.15 *Let $\{W(t), t \geq 0\}$ be a Wiener process. Find the conditional distribution of $W(t)$ given that $W(s) = c$ (where c is a constant) when $s < t$.*

Exercise 7.16 *Let $\{W(t), t \geq 0\}$ be a Wiener process. Find the conditional distribution of $W(s/2)$ given that $W(s) = x$.*

Exercise 7.17 *Show that for any $T > 0$, $V(t) = W(t + T) - W(T)$ is a Wiener process if $W(t)$ is a Wiener process.*

Exercise 7.18 *Let $\{W(t), t \geq 0\}$ be a Brownian motion. Prove that $\{tW(1/t), t \geq 0\}$ where $tW(1/t)$ is taken to be zero when $t = 0$, is a Brownian motion.*

Exercise 7.19 *Let $\{W(t), t \geq 0\}$ be a Wiener process. Prove that distribution of $\{W(t), 0 \leq t \leq 1\}$ and $\{W(1) - W(1 - t), 0 \leq t \leq 1\}$ are the same.*

Exercise 7.20 *Consider the process*

$$Y(t) = \begin{cases} W(t), & t < 1 \\ W(t) + Z, & t \geq 1 \end{cases}$$

where $W(t)$ is a Wiener process and $Z \sim N(0, 1)$ and is independent of $W(t)$. Show that $\{Y(t), t \geq 0\}$ not a Levy process.

8
Filtration and Martingale

8.1 Introduction

Conditional expectation is an extremely important concept in probability theory. Traditionally in a typical probability course, the well known *gambler's ruin problem* is presented as a motivational example. But now the concept of conditional expectation has found much favor in financial mathematics because it provides the basis for two of the most important concepts namely, the *filtration* and the *martingale*.

In gambler's ruin problem, a gambler starts with an amount of Rs N. Then an unbiased coin is flipped, landing head with probability 0.5 and landing tail with probability 0.5. If the coin lands head, he/she gains *Rs* 1, otherwise he/she loses *Rs* 1. The game continues until he/she loses all his/her amount of Rs N. Let X_n be the fortune at the nth game, and S_n be his/her capital after n games. An interesting question here is to know his/her fortune, on an average, on the next game given his/her current fortune. To answer this question, we need the conditional expectation of the random variables $\{X_n, n = 0, 1, \ldots\}$ given the information up to $r, r < n$.

We next present another example which is more relevant to us. Let $X(t)$ denote the share price of a particular stock at time t. Assume that, we have the information about the share price up to time $s > 0$. We may like to know the expected or average share price of the stock at a future time, say, $s + 5$ given that $X(s) = 10$ (say). To make such statements precise and also to answer them, we need the conditional expectation of random variables $\{X(t), t \geq 0\}$ given the information up to time s, $s < t$.

The above two examples certainly make a case for the study of conditional expectation. But for a better understanding and deeper study of various other

aspects we need to develop the concepts of *filtration* and *martingale* property of the underlying stochastic process.

A filtration is an increasing sequence of σ-fields on a measurable space $(\Omega, \mathcal{F})$. Since a σ-field defines the set of events that can be measured, a filtration is often used to represent the change in the set of events that can be measured through gain or loss of information. In financial mathematics, a filtration represents the information available at each time t, and it becomes more and more precise as information from the present becomes available.

The concept of martingale in probability theory is referred to a class of betting strategies. It says that, the conditional expected value of the next observation, given all the past observations is equal to the last observation. It is important to note that the property of being a martingale involves both the filtration and the probability measure. It is possible that one stochastic process could be a martingale with respect to one probability measure but not with respect to another one.

In this chapter, we first discuss the conditional expectation and then proceed to study filtrations and martingales in detail.

8.2 Conditional Expectation $E(X/Y)$ and Calculation Rules

Let $(\Omega, \mathcal{F}, P)$ be the given probability space and X be the given random variable. Then its expectation $E(X)$ is defined as

$$E(X) = \int_\Omega X(w)dP(w) , \tag{8.1}$$

where the integration is to be understood in the Lebesgue-Stieltjes sense. In case X is a discrete random variable taking values $\alpha_1, \alpha_2, \ldots, \alpha_i, \ldots$ we have

$$E(X) = \sum_i \alpha_i P(X = \alpha_i) ,$$

where $P(X = \alpha_i) = P(w \in \Omega : X(w) = \alpha_i)$. If X is a continuous random variable having probability density function (p.d.f.) $f(x)$ then

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx . \tag{8.2}$$

We say that the expectation $E(X)$ exists provided $E(X) < \infty$. Obviously not all random variables have expectation. For example, if the discrete random variable X takes value 2^n, $n = 1, 2, \ldots$ with probability mass function (p.m.f.)

$$P(X = 2^n) = 2^{-n}, \ (n = 1, 2, \ldots) ,$$

then $E(X)$ does not exist. In a similar manner, if the continuous random variable X has the p.d.f.

$$f(x) = \frac{1}{\pi(1 + x^2)} , \quad -\infty < x < \infty ,$$

then $E(X)$ does not exist.

Now, we define the conditional expectation and present some of its properties.

Definition 8.2.1 (Conditional Expectation of a Random Variable) *For a discrete random variable X and a discrete random variable Y, the conditional expectation of X given the event $Y = y$ is defined as*

$$E(X/Y = y) = \sum_x x \, P(X = x/Y = y) .$$

Remark 8.2.1 *Looking at the expression of $E(X/Y = y)$, it is simple to note that it is a function of y. If we write this function of y as $f(y) = E(X/Y = y)$, then we write $E(X/Y)$ for $f(Y)$. This is a discrete random variable. It is called the conditional expectation of X given Y and is denoted by $E(X/Y)$.*

Remark 8.2.2 *Definition 8.2.1 of conditional expectation can be extended in an obvious manner even when X is continuous random variable. We note that Y is still a discrete random variable. Therefore if we define*

$$E(X/Y)(w) = E(X/A_i) \ \text{for } w \in A_i \ (i = 1, 2, \ldots) ,$$

and $A_i = \{w \in \Omega : Y = y_i\}$, then it takes care of both discrete as well as continuous cases of the random variable X. Here again $E(X/Y)$ is a discrete random variable.

Remark 8.2.3 *Definition 8.2.1 of conditional expectation can further be extended when the event $Y = y$ is replaced by a general event B for which $P(B) > 0$. Thus when $P(B) > 0$, the conditional expectation of the random variable X given B is defined as*

$$E(X/B) = \frac{E(X1_B)}{P(B)} ,$$

where

$$1_B(w) = \begin{cases} 1, & w \in B \\ 0, & w \notin B, \end{cases}$$

denotes the indicator function of the set B.

Example 8.2.1 *Let X denote the outcome of tossing of an unbiased die. Let Y be a random variable which takes the value +1 if the outcome is an even number and it takes the value −1 if the outcome is an odd number.*

(i) Find $E(X/Y=1)$ *and* $E(X/Y=-1)$.
(ii) Let $B = \{1,2,3\}$. *Find* $E(X/B)$.

Solution Here $\Omega = \{1,2,3,4,5,6\}$ and $P(X=i) = \frac{1}{6}$ for $i = 1,2,3,4,5,6$.

(i) We have $A_1 = \{w \in \Omega : Y(w) = 1\} = \{2,4,6\}$ and $A_2 = \{w \in \Omega : Y(w) = -1\} = \{1,3,5\}$. Hence

$$E(X/Y=1) = E(X/A_1) = \frac{E(X1_{A_1})}{P(A_1)}.$$

But

$$E(X1_{A_1}) = 2\times\frac{1}{6} + 4\times\frac{1}{6} + 6\times\frac{1}{6},$$

and

$$P(A_1) = P(\{2,4,6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6}.$$

Hence,

$$E(X/Y=1) = \frac{E(X1_{A_1})}{P(A_1)} = 4.$$

Similarly,

$$E(X/Y=-1) = \frac{E(X1_{A_2})}{P(A_2)} = 3.$$

(ii) We have

$$E(X/B) = \frac{E(X1_B)}{P(B)} = 2.$$

□

Properties of Conditional Expectation

Using the above definition of conditional expectation we present the following properties for $E(X/Y)$. We do not prove these results here and shall refer to Mikosch [97] for the same.

(i) For random variables X_1, X_2 and reals a_1, a_2, we have

$$E((a_1X_1 + a_2X_2)/Y) = a_1E(X_1/Y) + a_2E(X_2/Y).$$

This property is called the *linear property* of $E(X/Y)$.
(ii) $E(E(X/Y)) = E(X)$. This property is called the *average property* of $E(X/Y)$.
(iii) For Borel measurable function g, we have $E(Xg(Y)/Y) = g(Y)E(X/Y)$. This property is called the *out property*.
(iv) $E(E(X/Y,Z)/Z) = E(X/Z)$. This property is called the *tower property*.
(v) If X and Y are independent random variables then $E(X/Y) = E(X)$.

Remark 8.2.4 *Conditional expectation tells us how to use the observation of the random variable Y to estimate another random variable X. This conditional expectation is the best guess about X given Y if we want to minimize the mean square error.*

Example 8.2.2 *Consider two i.i.d random variables X and Y each having uniform distribution between the intervals 0 and 1. Define* $Z = X + Y$. *Find*

(i) $E(Z/X)$.
(ii) $E(XZ/X)$.

Solution Since X and Y are uniformly distributed random variables, we have $E(X) = E(Y) = \frac{1}{2}$.

(i) We have

$$E(Z/X) = E[(X+Y)/X] = E(X/X) + E(Y/X) = X + E(Y) = X + \frac{1}{2}.$$

(ii) Using property (iii) of conditional expectation, we have

$$E(XZ/X) = XE(Z/X) = X\left(X + \frac{1}{2}\right).$$

□

Example 8.2.3 *Consider a 2-period binomial model. Let S(n) be the stock price at period n. Assume that,* $P(up) = p$ *and* $P(down) = 1-p$ *such that* $0 < p < 1$. *Find* $E(S(2)/S(1))$.

Solution The stock price $S(1)$ is a discrete random variable taking two values $uS(0)$ and $dS(0)$ with probabilities p and $1-p$ respectively. This can be formed as partition $\mathcal{B} = \{B_1, B_2\}$ with $B_1 = \{S(1) = uS(0)\}$ and $B_2 = \{S(1) = dS(0)\}$. Then the forecast of $S(2)$ given the information B_1 is given by

$$E(S(2)/S(1) = uS(0)) = pS(0)u^2 + (1-p)S(0)ud\ .$$

In a similar manner the forecast of $S(2)$ given the information B_2 is given by

$$E(S(2)/S(1) = dS(0)) = pS(0)ud + (1-p)S(0)d^2\ .$$

Note that, this forecast differs depending on whether B_1 or B_2 actually occurs in period $n = 1$. □

Example 8.2.4 *Consider Example 7.3.1. Find the conditional expectation of $W(s)$ given $W(t) = x$ for $0 < s < t$.*

Solution We know that the conditional distribution of $W(s)$ given $W(t) = x$ for $0 < s < t$ is a normal distribution with mean $\frac{sx}{t}$ and variance $s^2\left(\frac{1}{s} - \frac{1}{t}\right)$. Hence,

$$E(W(s)/W(t) = x) = \frac{sx}{t}\ .$$

□

8.3 More on σ-Fields

In the previous section we have defined the conditional expectation $E(X/Y)$ where Y is a discrete random variable. We have observed that $E(X/Y)$ is a function of Y which is again a discrete random variable. The definition required the construction of sets $A_i = \{w \in \Omega : Y(w) = y_i\}$ and then define

$$E(X/Y)(w) = E(X/A_i) \quad \text{for } w \in A_i \ (i = 1, 2, \ldots).$$

Thus $E(X/Y)$ can be understood as a discrete random variable constructed from a collection of subsets A_i of Ω.

In Chapter 7, we have studied σ-fields and also learnt how to generate a σ-field from a given collection of subsets of Ω. Let us recall the following definitions.

Definition 8.3.1 (σ-Field) *A σ-field on Ω is a collection $\mathcal{F}$ of subsets of Ω satisfying*

(i) $\emptyset \in \mathcal{F}$,
(ii) if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$,
(iii) if a countable collection of sets $A_1, A_2, \ldots$ is in $\mathcal{F}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition 8.3.2 (σ-Field Generated by a Collection of Subsets of Ω) *Let $\mathcal{U}$ be a collection of subsets of Ω. Then the smallest σ-field containing $\mathcal{U}$ is called the σ-field generated by the collection $\mathcal{U}$ of subsets of Ω and it is denoted by $\sigma(\mathcal{U})$.*

Based on Definition 8.3.2, in this section we wish to introduce *σ-field generated by a random variable, σ-field generated by a random vector* and *σ-field generated by a stochastic process*. This is important because we need the σ-fields generated by stochastic processes to define filtration and martingale.

Definition 8.3.3 (σ-Field Generated by a Random Variable) *Let Y be a discrete random variable taking values y_i, $(i = 1, 2, \ldots)$. Let $A_i = \{w \in \Omega : Y(w) = y_i\}$ $(i = 1, 2, \ldots)$, and $\mathcal{U} = \{A_1, A_2, \ldots\}$ be the corresponding family of subsets of Ω. Then the σ-field generated by the random variable Y is defined as the σ-field generated by the family $\mathcal{U}$. In that case we denote $\sigma(Y) = \sigma(\mathcal{U})$.*

Remark 8.3.1 *The σ-field $\sigma(Y)$ certainly consists of all sets of form $A = \cup_{i \in I} A_i$, where I is any subsets of $\mathbf{N} = \{1, 2, \ldots\}$ including $I = \emptyset$. For $I = \emptyset$, we get $A = \emptyset$ and for $I = \mathbf{N}$ we get $A = \Omega$. Also $\sigma(Y)$ consists of all subsets of the form*

$$A_{a,b} = \{Y \in (a, b]\} = \{w \in \Omega : a < Y(w) \le b\}, -\infty < a < b \le \infty\ ,$$

because

$$A_{a,b} = \cup_{i \in I}\{w : Y(w) = y_i\} \in \sigma(Y)\ ,$$

where $I = \{i : a < y_i \le b\}$.

Remark 8.3.2 *Recall, if $\Omega = \mathbf{R}$ and $\mathcal{U}^{(1)} = \{(a, b] : -\infty < a < b < \infty\}$, then the σ-field generated by $\mathcal{U}^{(1)}$ is called the Borel σ-field, and its elements are called Borel sets. Thus the Borel σ-field is the σ-field generated by the family of semi-open intervals of type $(a, b]$, $-\infty < a < b < \infty$.*

In case $\Omega = \mathbf{R}^n$, we can consider the family of rectangles

$$\mathcal{U}^{(n)} = \{((a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n]), -\infty < a_i < b_i < \infty, (i = 1, 2, \ldots, n)\}\ ,$$

and the σ-field generated by $\mathcal{U}^{(n)}$, i.e. $\sigma(\mathcal{U}^{(n)})$. This σ-field is called the n-dimensional Borel field. Thus n-dimensional Borel field is the σ-field generated by the collection of rectangles in $\mathbf{R}^n$.

Example 8.3.1 *Consider a binomial model with $t = 1, 2$. Let S_t be the stock price at time t. Let $\Omega = \{(u, u), (u, d), (d, u), (d, d)\}$. Define a discrete random variable X as*

$$X(w) = \begin{cases} 0.5, & w \in \{(u, u), (u, d)\} \\ 1.5, & w \in \{(d, u), (d, d)\}\ . \end{cases}$$

Find $\sigma(X)$.

Solution The random variable X is a discrete random variable on $\mathcal{F}_1$. We have

$$A_1 = \{w \in \Omega : X(w) = 0.5\} = \{(u,u),(u,d)\},$$

and

$$A_2 = \{w \in \Omega : X(w) = 1.5\} = \{(d,u),(d,d)\}.$$

Hence to determine the required σ-field we need to consider the family $\mathcal{U} = \{A_1, A_2\}$ of subsets of Ω. Thus

$$\mathcal{U} = \{\{(u,u),(u,d)\},\{(d,u),(d,d)\}\}.$$

Then the σ-field generated by the random variable X is defined as the σ-field generated by the family $\mathcal{U}$. In this case,

$$\begin{aligned}\sigma(X) &= \sigma(\mathcal{U})\\ &= \{\emptyset, \{(u,u),(u,d)\}, \{(d,u),(d,d)\}, \Omega\}.\end{aligned}$$

□

Definition 8.3.4 (σ-Field Generated by a Random Vector) *Let $Y = (Y_1, \ldots, Y_n)$ be an n-dimensional random vector. Then the σ-field generated by the random vector $Y = (Y_1, Y_2, \ldots, Y_n)$ is the smallest σ-field containing all n-dimensional Borel sets, i.e. sets of the form*

$$\{w \in \Omega : a_i < Y_i(w) \le b_i \ (i = 1,2,\ldots,n), -\infty < a_i < b_i < \infty \ (i = 1,2,\ldots,n)\}.$$

This σ-field is denoted by $\sigma(Y) = \sigma(Y_1, Y_2, \ldots, Y_n)$.

Example 8.3.2 *Consider a 2-period binomial model. Let S_n be the stock-price at period n. Let $\Omega = \{(u,u),(u,d),(d,u),(d,d)\}$. Define a discrete random vector $X = (X_1, X_2)$ where*

$$X_1(w) = \begin{cases} 0.5, & w \in \{(u,u),(u,d)\} \\ 1.5, & w \in \{(d,u),(d,d)\}, \end{cases}$$

and

$$X_2(w) = \begin{cases} 0.25, & w = (u,u) \\ 0.5, & w = (u,d) \\ 0.75, & w = (d,u) \\ 1.00, & w = (d,d). \end{cases}$$

Find the σ-field generated by the random vector $X = (X_1, X_2)$.

Solution We have

$$A_1 = \{w \in \Omega : X_1(w) = 0.5\} = \{(u,u),(u,d)\},$$

and

$$A_2 = \{w \in \Omega : X_1(w) = 1.5\} = \{(d,u),(d,d)\}.$$

Similarly, we have

$$\begin{aligned} B_1 &= \{w \in \Omega : X_2(w) = 0.25\} = \{(u,u)\} \\ B_2 &= \{w \in \Omega : X_2(w) = 0.5\} = \{(u,d)\} \\ B_3 &= \{w \in \Omega : X(w) = 0.75\} = \{(d,u)\}, \end{aligned}$$

and

$$B_4 = \{w \in \Omega : X_2(w) = 1.00\} = \{(d,d)\}.$$

Hence σ-field generated by $X = (X_1, X_2)$ is the σ-field generated by the family $\mathcal{U}$ where

$$\mathcal{U} = \{A_1, A_2, B_1, B_2, B_3, B_4\}.$$

Therefore the desired σ-field is

$$\begin{aligned}\sigma(X) = \sigma(X_1, X_2) = \{&\emptyset, A_1, A_2, B_1, B_2, B_3, B_4, B_1 \cup B_3, B_1 \cup B_4, B_2 \cup B_3, B_2 \cup B_4, \\ &B_1 \cup B_2 \cup B_3, B_1 \cup B_2 \cup B_4, B_2 \cup B_3 \cup B_4, B_1 \cup B_3 \cup B_4, \Omega\}.\end{aligned}$$

□

Definition 8.3.5 (σ-Field Generated by a Stochastic Process) *Let $\{Y(t), t \in T\}$ be the given stochastic process. Then the σ-field generated by the stochastic process $\{Y(t), t \in T\}$ is the smallest σ-field containing all sets of the form*

$$\{w : \text{the sample path } (Y(t,w), t \in T) \text{ belongs to } C\}$$

for all suitable sets C of functions on T.

Remark 8.3.3 *The set C in Definition 8.3.5 probably needs some more specifications. But at this stage and with our limited background, it is not possible to answer this question for a general stochastic process. However we shall present the σ-field generated by a Brownian motion, as this is the one which we shall be using most often.*

Definition 8.3.6 (σ-Field Generated by a Brownian Motion) *Let $W = \{W(s), 0 < s \le t\}$ be the given Brownian motion on $[0,t]$. Then the σ-field generated by all sets of the form*

$$A_{t_1,t_2,\ldots,t_n} = \{w \in \Omega : (W(t_1,w), W(t_2,w), \ldots, W(t_n,w)) \in C\},$$

for any n-dimensional Borel set C, and for any choice of $t_i \in [0,t]$, $i \ge 1$ is called the σ-field generated by the Brownian motion W.

Remark 8.3.4 *For a random variable, a random vector or a stochastic process Y on Ω, the σ-field $\sigma(Y)$ generated by Y contains all the essential information about the structure of Y as a function of $w \in \Omega$. It consists of all subsets $\{w : Y(w) \in C\}$ for all suitable sets C. In general, C has to be any n-dimensional Borel set ($n \ge 1$), n being equal to 1 for the case of a random variable. In this situation we agree to the terminology that Y contains information represented by $\sigma(Y)$ or Y carries the information $\sigma(Y)$.*

8.4 Filtration

Filtration is important in financial mathematics, because it allows to model the flow of information. Of course, the information increases as time goes by. Before we give the formal definition of a filtration, we consider below the given situation.

Let us consider the random experiment of tossing a coin three times. Here the set of all possible outcomes is

$$\Omega = \{(HHH), (HHT), (HTH), (HTT), (THH), (THT), (TTH), (TTT)\}.$$

We choose Ω as the universal or global set as it contains all information on the process. Let E_H be consists of all possible outcomes starting with head in 1st toss and E_T consists of all possible outcomes starting with tail in 1st toss. Now, we consider the family $\mathcal{F}_1 = \{\emptyset, \{E_H\}, \{E_T\}, \Omega\}$ of subsets of Ω. Then it can be verified that $\mathcal{F}_1$ is a σ-field. Note that, $\mathcal{F}_1$ is the set of events that have been decided by the end of 1st toss. Similarly, let E_{HH} consists of all possible outcomes starting with head in 1st toss and head in 2nd toss. We define E_{HT}, E_{TH} and E_{TT} analogously. Then, we consider the following family $\mathcal{F}_2 = \{\emptyset, \{E_H\}, \{E_T\}, \{E_{HH}\}, \{E_{HT}\}, \{E_{TH}\}, \{E_{TT}\}, \{E^c_{HH}\}, \{E^c_{HT}\}, \{E^c_{TH}\}, \{E^c_{TT}\}, \{E_{HH} \cup E_{TH}\}, \{E_{HH} \cup E_{TT}\}, \{E_{HT} \cup E_{TH}\}, \{E_{HT} \cup E_{TT}\}, \Omega\}$ of subsets of Ω. We can again verify that $\mathcal{F}_2$ is a σ-field. Observe that, this σ-field $\mathcal{F}_2$ contains the information learned by observing the first two tosses of a coin.

Now, after the coin is tossed three times, every subset of Ω is resolved. Hence the family $\mathcal{F}_3$ of 256 subsets of Ω which is the set of all subsets of Ω. That is, $\mathcal{F}_3$ is the total σ-field. Since $\mathcal{F}_i$ is the set of events that have been decided by the end of ith toss, we have more information at the end of $(i+1)^{th}$ toss than at the end of ith toss.

Taking motivation from the above example we give the definition of filtration in discrete time.

Definition 8.4.1 (Filtration in a Discrete Time) *Let Ω be the set of all possible outcomes of a random experiment and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Then a filtration in discrete time is an increasing sequence $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots$ of σ-fields, one per time instant.*

The σ-field $\mathcal{F}_n$ may be thought of as the events of which the occurrence is determined at or before time n, the "known events" at time n. The interpretation of $\mathcal{F}_0 = \{\emptyset, \Omega\}$ is that, at the beginning, one has no information. Further the interpretation of $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots$ is that the state of information increases over time.

Example 8.4.1 *Consider $\Omega = \{a,b,c,d\}$. Construct σ-fields $\mathcal{F}_i$, $(i = 0,1,2)$, such that $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2$.*

Solution. Obviously $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Let $\mathcal{F}_1 = \{\emptyset, \{a,b\}, \{c,d\}, \Omega\}$ and $\mathcal{F}_2 = \{\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a,b\}, \{c,d\}, \{a,d\}, \{a,c\}, \{b,c\}, \{b,d\}, \{a,b,c\}, \{c,d,a\}, \{d,a,b\}, \{b,c,d\}, \Omega\}$. Then, $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2$.

□

Definition 8.4.2 (Filtration in a Continuous Time) *Let Ω be the set of all possible outcomes of a random experiment. Let T be a fixed positive number and assume that for each $t \in [0,T]$, there is a σ-field $\mathcal{F}_t$. Assume further that, if $s \le t$, then every set in $\mathcal{F}_s$ is also in $\mathcal{F}_t$. Then, the collection of σ-fields $\{\mathcal{F}_t, 0 \le t \le T\}$ is called a filtration in continuous time.*

Thus a collection of σ-fields $\{\mathcal{F}_t, t \ge 0\}$ is called a filtration in continuous time if $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \le s \le t$.

Remark 8.4.1 *Filtration is used to model the flow of information over time. As an example, we think of X_t as the price of some asset at time t and $\mathcal{F}_t$ as the information obtained by watching all the prices in the market up to time t.*

Definition 8.4.3 (Natural Filtration) *The natural filtration of a discrete time stochastic process $\{X_0, X_1, \ldots\}$ is defined by the filtration $\{\mathcal{F}_n, n = 0,1,\ldots\}$ where $\mathcal{F}_n$ is the σ-field generated by the random vector $(X_0, X_1, \ldots, X_n)$.*

Thus $\mathcal{F}_n$ contains all events that depend on the first $(n+1)$ elements of the stochastic process. It gives the "history" of the process up till time n. A convenient notation to describe a σ-field corresponding to observing a random vector X is $\sigma(X)$. Thus $\sigma(X)$, the σ-field generated by X, consists of all events that can be expressed in X: events of the type $\{X \in C\}$, C being $(n+1)$ dimensional Borel sets. In this notation, the natural filtration of a discrete time stochastic process $\{X_0, X_1, \ldots\}$ can be written as $\mathcal{F}_n = \{(X_0, X_1, \ldots, X_n)\}$.

The natural filtration of a continuous time stochastic process $\{X(t), t \geq 0\}$ is defined by the filtration $\{\mathcal{F}_t, t \geq 0\}$ where $\mathcal{F}_t = \sigma(X(s), s \leq t)$.

Definition 8.4.4 (Adapted Process) *We say that a discrete time stochastic process $\{X_0, X_1, \ldots\}$ is adapted to a given filtration $\{\mathcal{F}_n, n = 0, 1, \ldots\}$ if the σ-field generated by X_n is a subset of $\mathcal{F}_n$ means $\sigma(X_n) \subset \mathcal{F}_n$, for every n. In a similar manner, a continuous time stochastic process $\{X(t), t \geq 0\}$ is said to be adapted to a given filtration $\mathcal{F}_t, t \geq 0\}$ if $\sigma(X(t)) \subset \mathcal{F}_t$ for all $t \geq 0$.*

Thus the events connected to an adapted process up to time n are known at time n. For instance, suppose S_n is the price of a stock at the end of nth day then the price process $\{S_n, n = 0, 1, 2, \ldots\}$ is adapted to natural filtration $\{\mathcal{F}_n, n = 0, 1, 2, \ldots\}$ where $\mathcal{F}_n$ is the history up to the end of nth day, and $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\Omega = [0, \infty)$. In a similar manner for continuous time stochastic process $\{S(t), t \geq 0\}$ is adapted to $\{\mathcal{F}_t, t \geq 0\}$ where $\mathcal{F}_t$ is the history up to time t.

Remark 8.4.2 *The natural filtration corresponding to a process is the smallest filtration to which it is adapted. If the process $\{Y_0, Y_1, \ldots\}$ is adapted to the natural filtration of a stochastic process $\{X_0, X_1, \ldots\}$ then for each n the variable Y_n is a function $\sigma(X_0, X_1, \ldots, X_n)$ of the sample path of the process X up till time n.*

Example 8.4.2 *Consider Example 8.4.1. Define*

$$X_0(w) = \{1,\ w \in \{a, b, c, d\}\},$$

$$X_1(w) = \begin{cases} 1, & w \in \{a, b\} \\ -1, & w \in \{c, d\}, \end{cases}$$

and

$$X_2(w) = \begin{cases} 1, & w = a \\ 2, & w = b \\ 3, & w = c \\ 4, & w = d. \end{cases}$$

Verify that $\{X_i, i = 0, 1, 2\}$ is an adapted process to $\mathcal{F}_i$, $(i = 0, 1, 2)$.

Solution From Example 8.4.1, we have $\mathcal{F}_i$, $(i = 0, 1, 2)$ are the natural filtration of the stochastic process $\{X_i, i = 0, 1, 2\}$. The σ-field generated by the random variable X_0 is defined as the σ-field generated by the family $\{a, b, c, d\}$.

$$\sigma(X_0) = \{\emptyset, \Omega\} \subseteq \mathcal{F}_0.$$

Similarly, the σ-field generated by the random variable X_1 is defined as the σ-field generated by the family $\{\{a, b\}, \{c, d\}\}$.

$$\sigma(X_1) = \{\emptyset, \{a, b\}, \{c, d\}, \Omega\} \subseteq \mathcal{F}_1.$$

Finally, the σ-field generated by the random variable X_2 is defined as the σ-field generated by the family $\{\{a\}, \{b\}, \{c\}, \{d\}\}$.

$$\begin{aligned} \sigma(X_2) = \{&\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{c, d\}, \{a, d\}, \{a, c\}, \{b, c\}, \{b, d\}, \\ &\{a, b, c\}, \{c, d, a\}, \{d, a, b\}, \{b, c, d\}, \Omega\} \\ \subseteq\ &\mathcal{F}_2. \end{aligned}$$

Hence, $\{X_i, (i = 0, 1, 2)\}$ is an adapted process to $\mathcal{F}_i$ $(i = 0, 1, 2)$. □

Remark 8.4.3 *If a stochastic process $\{X(t), t \geq 0\}$ is adapted to the natural filtration $\{\mathcal{F}_t, t \geq 0\}$ of the Brownian motion $\{W(t), t \geq 0\}$, then it means that $X(t)$ is a function of $W(s)$, $s \leq t$. For example, $X(t) = W^2(t) - t$ and $Y(t) = \sup_{s \leq t} W(s)$ are adapted processes whereas $Z(t) = \sup_{s \leq t+1} W(s)$ and $U(t) = W(1) - W(t)$ are not adapted processes to the natural filtration $\{\mathcal{F}_t, t \geq 0\}$ of the Brownian motion $\{W(t), t \geq 0\}$.*

Remark 8.4.4 *Many stochastic calculus texts specify a probability space as $(\Omega, \{\mathcal{F}_t, t \geq 0\}, P)$ thereby, specifying the filtration explicitly from the very beginning.*

Remark 8.4.5 *The natural filtration is the name given to the filtration for which $\mathcal{F}_t$ consists of those sets that can be decided by observing trajectories of the specified process up to time t.*

8.5 Conditional Expectation $E(X/\mathcal{F})$ and Calculation Rules

In Section 8.2, we discussed the conditional expectation $E(X/Y)$ while in Section 8.3, we defined the σ-field generated by the random variable Y. In this section, we study the conditional expectation $E(X/\sigma(Y))$.

Definition 8.5.1 (Conditional Expectation of a Random Variable given σ-field) *Let X be a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{G}$ be a σ-field of subsets of Ω. A random variable Z is called the conditional expectation of X given the σ-field $\mathcal{G}$ if*

(i) Z does not contain more information than that contained in $\mathcal{G}$, i.e. $\sigma(Z) \subset \mathcal{G}$.
(ii) For every $A \in \mathcal{G}$, Z satisfies $E(X1_A) = E(Z1_A)$ where 1_A is the indicator function of the event A.

We denote Z by $E(X/\mathcal{G})$ and call the same as the conditional expectation of X given $\mathcal{G}$.

Remark 8.5.1 *Thus we may have Z and Z' such that $Z = E(X/\mathcal{G})$ and also $Z' = E(X/\mathcal{G})$, but then we have $Z = Z'$ almost surely. In that sense there is a unique $\mathcal{G}$-measurable random variable Z which satisfies $E(X1_A) = E(Z1_A)$ for every $A \in \mathcal{G}$.*

Remark 8.5.2 *The interpretation of (i) in the above definition is that, the information in $\mathcal{G}$ is sufficient to determine Z. Then, we say that Z is $\mathcal{G}$-measurable.*

Remark 8.5.3 *Obviously, every random variable is $\mathcal{F}$-measurable where $\mathcal{F}$ is the largest or total σ-field on Ω. Also, if X is $\mathcal{G}$-measurable, then $h(X)$ is also $\mathcal{G}$-measurable for any Borel-measurable function h. The other extreme is when a random variable is independent of a σ-field. In this case, the information contained in the σ-field gives no clue about the random variable.*

Remark 8.5.4 *When $\mathcal{G} = \sigma(Y)$, the conditional expectation $E(X/Y)$ is a special case of the above definition of $E(X/\mathcal{G})$. We know that, every element A of $\sigma(Y)$ is of the form*

$$A = \cup_{i \in I} A_i = \cup_{i \in I} \{w \in \Omega : Y(w) = y_i\},\ I \subset \mathbf{N}.$$

Hence,

$$E(X1_A) = E\left(X \sum_{i \in I} 1_{A_i}\right) = \sum_{i \in I} E(X1_{A_i}).$$

For the case of discrete random variable $Z1_A$, its expectation is given by

$$E(Z1_A) = \sum_{i \in I} E(X/A_i)P(A_i) = \sum_{i \in I} E(X1_{A_i}).$$

From Definition 8.5.1, we conclude that when $\mathcal{G}$ is the σ-field generated by Y, then we write $E(X/\mathcal{G})$ for the random variable $E(X/Y)$. Thus $E(X/\mathcal{G})$ is the expected value of X given the information $\mathcal{G}$. This result is valid for the continuous random variable also.

Example 8.5.1 *Consider a binomial model with $t = 1, 2$. Let S_t be the stock price at time t. Let $\Omega = \{(u,u),(u,d),(d,u),(d,d)\}$. Let σ-field $\mathcal{F}$ be the power set of Ω and $P(w) = \frac{1}{4}$ for all $w \in \Omega$. Let $\mathcal{G} = \{\emptyset, \{(u,u),(u,d)\}, \{(d,u),(d,d)\}, \Omega\}$. Define a discrete random variable X as*

$$X(w) = \begin{cases} 0.25, & w = (u,u) \\ 0.5, & w = (u,d) \\ 0.75, & w = (d,u) \\ 1.00, & w = (d,d). \end{cases}$$

Find $E(X/\mathcal{G})$.

Solution Define

$$Z(w) = \begin{cases} \frac{3}{8}, & w \in \{(u,u),(u,d)\} \\ \frac{7}{8}, & w \in \{(d,u),(d,d)\}. \end{cases}$$

Then we have

$$Z^{-1}\{(-\infty, z]\} = \begin{cases} \emptyset, & -\infty < z < \frac{3}{8} \\ \{(u,u),(u,d)\}, & \frac{3}{8} \le z < \frac{7}{8} \\ \Omega, & \frac{7}{8} \le z < \infty. \end{cases}$$

Hence, Z is $\mathcal{G}$-measurable.

For $A_1 = \{(u,u),(u,d)\}$, we have

$$E(X1_{A_1}) = 0.25 \times \frac{1}{4} + 0.50 \times \frac{1}{4} = \frac{3}{16},$$

and for $A_2 = \{(d,u),(d,d)\}$, we have

$$E(X1_{A_2}) = 0.75 \times \frac{1}{4} + 1.00 \times \frac{1}{4} = \frac{7}{16}.$$

Now, for $A_1 = \{(u,u),(u,d)\}$, we have

$$E(Z1_{A_1}) = \frac{3}{8} \times \left(\frac{1}{4} + \frac{1}{4}\right) = \frac{3}{16},$$

and for $A_2 = \{(d,u),(d,d)\}$, we have

$$E(Z1_{A_2}) = \frac{7}{8} \times \left(\frac{1}{4} + \frac{1}{4}\right) = \frac{7}{16}.$$

By Definition 8.5.1, the random variable Z is the conditional expectation of X given the σ-field $\mathcal{G}$.

□

Result 8.5.1 *Let X and Y be two integrable random variables and a and b be two real numbers, and $\mathcal{G}$ be the sub-σ-field of $\mathcal{F}$. Then,*

(i) $E((aX + bY)/\mathcal{G}) = aE(X/\mathcal{G}) + bE(Y/\mathcal{G})$,

(ii) If X and $\mathcal{G}$ are independent, then $E(X/\mathcal{G}) = E(X)$. *It means that we do not gain any information about X, if we know $\mathcal{G}$ and vice versa.*

(iii) If σ-field $\sigma(X) \subset \mathcal{G}$, *then* $E(X/\mathcal{G}) = X$. *This means that the information contained in $\mathcal{G}$ provides us with the whole information about the random variable X. Hence, X can be treated as a constant.*

(iv) If X is $\mathcal{G}$-measurable, then $E(XY/\mathcal{G}) = XE(Y/\mathcal{G})$. *Given $\mathcal{G}$, we can deal with X as if it is a constant, hence we can pull X(w) out of the updated expectation and write it in front of* $E(Y/\mathcal{G})$.

(v) If $\mathcal{G}_1$ and $\mathcal{G}_2$ are two sub σ-fields of $\mathcal{F}$ with $\mathcal{G}_1 \subseteq \mathcal{G}_2$, then $E(E(X/\mathcal{G}_1)/\mathcal{G}_2) = E(E(X/\mathcal{G}_2)/\mathcal{G}_1) = E(X/\mathcal{G}_1)$.

For the proofs of above results we shall refer to Mikosch [97].

Example 8.5.2 *Suppose $\mathcal{G} = \{\emptyset, \Omega\}$. Find $E(X/\mathcal{G})$.*

Solution The given σ-field $\{\emptyset, \Omega\}$ is the trivial σ-field containing no information. The only random variables which are measurable with respect to the trivial σ-field are constants. Hence, $E(X/\mathcal{G}) = E(X) = c$, where c is a constant.

□

Example 8.5.3 *Consider a discrete random variable X which takes value $-\sqrt{b}$ and $\sqrt{b}$ with probabilities $\frac{1}{3}$ and $\frac{2}{3}$ respectively. Find $E(X/\sigma(X^2))$. Here $\sigma(X^2)$ is the σ-field generated by the random variable X^2.*

Solution We have

$$P(X = \sqrt{b}/X^2 = b) = \frac{2}{3}; \text{ and } P(X = -\sqrt{b}/X^2 = b) = \frac{1}{3}.$$

Hence,

$$E(X/\sigma(X^2)) = \frac{2\sqrt{b}}{3} - \frac{\sqrt{b}}{3} = \frac{\sqrt{b}}{3}.$$

□

8.6 Martingales

The theory of martingale plays a very important and useful role in the study of financial mathematics. A formal definition is given below.

Definition 8.6.1 (Discrete Time Martingale) *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{X_n, n = 0, 1, \ldots\}$ be a stochastic process and $\{\mathcal{F}_n, n = 0, 1, \ldots\}$ be the filtration. The stochastic process $\{X_n, n = 0, 1, \ldots\}$ is said to be a martingale corresponding to the filtration $\{\mathcal{F}_n, n = 0, 1, \ldots\}$ if it satisfies the following conditions*

(i) For every n, $E(X_n)$ exists.

(ii) Each X_n is $\mathcal{F}_n$-measurable.

(iii) For every n, $E(X_{n+1}/\mathcal{F}_n) = X_n$.

Remark 8.6.1 *The definition of martingale depends on the collection of σ-fields $\mathcal{F}_n$. For clarity, one can say that $(X_n, \mathcal{F}_n)$ is a martingale.*

Remark 8.6.2 *From the definition of martingale and using the properties of conditional expectation, we observe that if $\{X_n\}$ is a martingale then $E(X_{n+1}) = E(X_n)$ for every n. This implies that $E(X_n) = c$, a constant. Therefore if, for some $n > 0$, $E(X_n) < \infty$ and the increments $X_{n+1} - X_n$ of the martingale $\{X_n\}$ are bounded, then $E(X_n) = E(X_0)$.*

Result 8.6.1 *We can generate martingale sequences by the following procedure. Given any increasing family of σ-fields $\{\mathcal{F}_n\}$, and any integrable random variable X on $(\Omega, \mathcal{F}, P)$, we take $X_n = E(X/\mathcal{F}_n)$ and it is easy to check that $\{(X_n, \mathcal{F}_n)\}$ is a martingale sequence. Of course, every finite martingale sequence is generated this way for we can always take X to be X_n, the last one.*

Result 8.6.2 *In equation (8.6.1), if '=' is replaced by '≥' then $X_n : n = 0, 1, 2, \ldots$ is called a submartingale while if it is replaced by '≤' then $X_n : n = 0, 1, 2, \ldots$ is called a supermartingale. Obviously, $\{X_n\}$ is a supermartingale if and only if $\{-X_n\}$ is a submartingale.*

Result 8.6.3 *Let $\{X_n, n = 0, 1, \ldots\}$, $\{Y_n, n = 0, 1, \ldots\}$ be discrete time stochastic processes. We say $\{X_n\}$ is a martingale with respect to $\{Y_n\}$ if $E(|X_n|) < \infty$ and $E(X_{n+1}/Y_0, Y_1, \ldots, Y_n) = X_n$. We may think of $Y_0, Y_1, \ldots, Y_n$ as the information or history upto stage (n+1). It may include more information than just $X_0, X_1, \ldots, X_n$.*

By calculating the conditional expectation value of X_{n+1} given the information about X_n (or Y_n) upto time n, we are making a forecast for the random variable. The martingale relation implies that the "best" forecast for the next value of the random variable is its current value.

Result 8.6.4 *Martingale theory provides a classification scheme for the time series. If a time series exhibits no discernible trend then it has a martingale like behavior. On the other hand, if the trend is an increasing (decreasing) one, then the time series behaves like submartingale (supermartingale).*

Example 8.6.1 *Consider the gambler's ruin problem. A gambler starts with Rs N. Then a coin is flipped, landing head with probability p and landing tail with probability $1-p$. If the coin lands heads, he/she gains one rupee, otherwise he /she loses a rupee. The game continues until he/she loses all his/her N rupees. Let X_n be fortune at the nth game, and S_n be his/her capital after n games. What will be his/her fortune, on an average, on the next game given that his/her current fortune?*

Solution For this we suppose that S_n is adapted to the filtration $\mathcal{F}_n$. Then we can interpret a martingale as a fair game. Thinking $(S_n - S_r)$ as the net winnings of the game per unit stake in time frame $(r,n]$, the best prediction of the net winnings given the information at the time $r < n$ has value

$$E((S_n - S_r)/\mathcal{F}_r) = E(S_n/\mathcal{F}_r) - S_r .$$

If S_n is a martingale then $E(S_n/\mathcal{F}_r) - S_r = 0$. This means that the best prediction of the future net winnings per unit stake in the interval $(r,n]$ is zero. This is exactly what we expect to be a fair game. It says that gambler expected capital after one more game played with the knowledge of the entire past and present is exactly equal to his/her current capital.

□

Example 8.6.2 *Let $X_1, X_2, \ldots$ be a sequence of i.i.d random variables each taking two values +1 and -1 with equal probabilities. Let us define $S_0 = 0$ and $S_n = \sum_{j=1}^{n} X_j$, $(n = 1,2,\ldots)$. This discrete time stochastic process $\{S_n, n = 0,1,\ldots\}$ is a symmetric random walk. Prove that, $\{S_n, n = 0,1,\ldots\}$ is a martingale with respect to $\{X_n, n = 1,2,\ldots\}$.*

Solution We have $E(|S_n|) \le E(|X_1|) + E(|X_2|) + \cdots + E(|X_n|) < \infty$. Also

$$\begin{aligned}
E(S_{n+1}/X_1, X_2, \ldots, X_n) &= E((S_n + X_{n+1})/X_1, X_2, \ldots, X_n) \\
&= E(S_n/X_1, X_2, \ldots, X_n) + E(X_{n+1}/X_1, X_2, \ldots, X_n) \\
&= S_n + E(X_{n+1}) \quad \text{(using independent of } X_1, X_2, \ldots, X_n, X_{n+1}) \\
&= S_n + 0 \\
&= S_n .
\end{aligned}$$

Hence $\{S_n, n = 0,1,\ldots\}$ is a martingale with respect to $\{X_n, n = 1,2,\ldots\}$. Suppose $\mathcal{F}_k$ is the σ-field of information corresponding to the first k random variables X_k, we have for non-negative integers $k < n$, $E(S_n/\mathcal{F}_k) = S_k$.

□

Example 8.6.3 *Consider a symmetric random walk $\{S_n, n = 0,1,\ldots\}$ which is a martingale with respect to the filtration $\{\mathcal{F}_n, n = 0,1,\ldots\}$ where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n), (n \ge 1)$, is the σ-field of information corresponding to the first n random variables X_n. Verify if $\{S_n^2, n = 0,1,\ldots\}$ is a martingale with respect to the filtration $\{\mathcal{F}_n, n = 0,1,\ldots\}$.*

Solution For each $n = 1,2,\ldots$, S_n^2 is $\mathcal{F}_n$-measurable. Also

$$E(S_n^2) = \sum_{i=1}^{n} E(X_i^2) < \infty .$$

Now,

$$\begin{aligned}
E(S_{n+1}^2/\mathcal{F}_n) &= E\left[(S_{n+1} - S_n + S_n)^2/\mathcal{F}_n\right] \\
&= E\left[(S_{n+1} - S_n)^2/\mathcal{F}_n\right] - 2E\left[S_n(S_{n+1} - S_n)/\mathcal{F}_n\right] + E(S_n^2/\mathcal{F}_n) \\
&= E(X_{n+1}^2/\mathcal{F}_n) - 2E(X_{n+1}S_n/\mathcal{F}_n) + E(S_n^2/\mathcal{F}_n).
\end{aligned}$$

Since X_{n+1} is independent of $\mathcal{F}_n$ and since S_n^2 is $\mathcal{F}_n$-measurable, we have

$$\begin{aligned}
E(S_{n+1}^2/\mathcal{F}_n) &= E(X_{n+1}^2) - 2S_n E(X_{n+1}) + S_n^2 \\
&= 1 - 0 + S_n^2 \\
&= 1 + S_n^2 .
\end{aligned}$$

Hence, $E(S_{n+1}^2/\mathcal{F}_n) \ne S_n^2$. Therefore, $\{S_n^2, n = 0,1,\ldots\}$ is not a martingale. Since $E(S_{n+1}^2/\mathcal{F}_n) \ge S_n^2$, $(n = 1,2,\ldots)$, hence $\{S_n^2, n = 0,1,\ldots\}$ is a submartingale.

□

Example 8.6.4 *Let $Y_0 = 0$ and let $Y_1, Y_2, \ldots$ be a sequence of i.i.d random variables with $E(|Y_k|) = 0$ and $E(Y_k^2) = \sigma^2$. Let $X_0 = 0$ and $X_n = (\sum_{k=1}^{n} Y_k)^2 - n\sigma^2$. Show that $\{X_n, n = 0,1,\ldots\}$ is a martingale with respect to $\{Y_n, n = 0,1,\ldots\}$.*

Solution We have

$$X_{n+1} = \left(\sum_{k=1}^{n+1} Y_k\right)^2 - (n+1)\sigma^2$$
$$= \left(\left(\sum_{k=1}^{n} Y_k\right) + Y_{n+1}\right)^2 - n\sigma^2 - \sigma^2$$
$$= X_n + 2Y_{n+1}\left(\sum_{k=1}^{n} Y_k\right) + Y_{n+1}^2 - \sigma^2 .$$

$$E[X_{n+1}/Y_0, Y_1, \ldots, Y_n] = E\left[\left(X_n + 2Y_{n+1}\left(\sum_{k=1}^{n} Y_k\right) + Y_{n+1}^2 - \sigma^2\right)/(Y_0, Y_1, \ldots, Y_n)\right]$$
$$= E[X_n/(Y_0, Y_1, \ldots, Y_n)]$$
$$+2E\left[Y_{n+1}\sum_{k=1}^{n} Y_k/(Y_0, Y_1, \ldots, Y_n)\right]$$
$$+E\left[Y_{n+1}^2/(Y_0, Y_1, \ldots, Y_n)\right] - \sigma^2$$
$$= X_n + 2\sum_{k=1}^{n} Y_k E(Y_{n+1}) + E(Y_{n+1}^2) - \sigma^2$$
$$= X_n .$$

Thus $\{X_n, n = 0, 1, \ldots\}$ is a martingale with respect to $\{Y_n, n = 0, 1, \ldots\}$.

□

Example 8.6.5 *Let a person start with Rs 1. A fair coin is tossed infinitely many times. For nth toss, if it turns up 'head', the person gets Rs 2, but if turn up 'tail', the person does not get any amount. Let Y_n be his/her fortune at the end of nth toss. Prove that Y_n is a martingale.*

Solution Let $X_1, X_2, \ldots$ be a sequence of i.i.d random variables each defined by

$$X_i = \begin{cases} 2, & \text{with probability } 0.5 \\ 0, & \text{with probability } 0.5 . \end{cases}$$

Since the game is double or nothing, his/her fortune at the end of nth toss is given by

$$Y_n = X_1 X_2 \cdots X_n \ (n = 1, 2, \ldots) .$$

Let $\mathcal{F}_n$ be the σ-field generated by $X_1, X_2, \ldots, X_n$. We note that $0 \le Y_n \le 2^n$ and $E[X_{n+1}] = 1$. Now,

$$E[Y_{n+1}/\mathcal{F}_n] = E[Y_n X_{n+1}/\mathcal{F}_n]$$
$$= Y_n E[X_{n+1}/\mathcal{F}_n]$$
$$= Y_n E[X_{n+1}]$$
$$= Y_n .$$

Hence, $\{Y_n, n = 1, 2, \ldots\}$ is a martingale.

□

Example 8.6.6 *Consider a binomial lattice model. Let S_n be the stock price at period n and*

$$S_{n+1} = \begin{cases} uS_n, & \text{with probability } p \\ dS_n, & \text{with probability } 1-p . \end{cases}$$

Define a related process R_n as

$$R_n = ln(S_n) - n\ [p\ ln(u) + (1-p)\ ln(d)] .$$

Prove that $\{ln(S_n), n = 1, 2, \ldots\}$ is not a martingale whereas $\{R_n, n = 1, 2, \ldots\}$ is a martingale with respect to $\{S_n, n = 1, 2, \ldots\}$. Also, prove that the discounted stock process $\{S_0, e^{-r}S_1, e^{-2r}S_2, \ldots\}$ is a martingale only if

$$p = \frac{e^r - d}{u - d} ,$$

where r is the nominal interest rate.

Solution In this binomial lattice model $\{S_0, S_1, \ldots\}$ with the natural filtration $\{\mathcal{F}_0, \mathcal{F}_1, \ldots\}$, we have

$$P(S_{n+1} = uS_n/\mathcal{F}_n) = 1 - P(S_{n+1} = dS_n/\mathcal{F}_n) = p .$$

Hence,

$$E(S_{n+1}/\mathcal{F}_n) = p\ uS_n + (1-p)\ dS_n = S_n[p\ u + (1-p)\ d] .$$

We consider the variable $ln(S_n)$ and observe that

$$E\left(ln\left(\frac{S_n}{S_{n-1}}\right)/S_{n-1}, S_{n-2}, \ldots, S_0\right) = p\ ln(u) + (1-p)\ ln(d) .$$

Therefore,

$$E\left(ln(S_n)/S_{n-1}, S_{n-2}, \ldots, S_0\right) = ln(S_{n-1}) + p\ ln(u) + (1-p)\ ln(d)\ . \qquad (8.3)$$

Here $\{ln(S_n), n = 1, 2, \ldots\}$ is not a martingale and depending upon the values of p, u and d it may be either a submartingale or a supermartingale. Next, we consider the process R_n.

$$E(R_n/R_{n-1}, R_{n-2}, \ldots, R_0) = E\left(ln(S_n) - n\left[p\ ln(u) + (1-p)\ ln(d)\right]/R_{n-1}, R_{n-2}, \ldots, R_0\right)$$

Using equation (8.3), and noting that the history of $S_{n-1}, S_{n-2}, \ldots, S_0$ yields the history of $R_{n-1}, R_{n-2}, \ldots, R_0$ and vice-versa, we get

$$\begin{aligned} E(R_n/R_{n-1}, R_{n-2}, \ldots, R_0) &= ln(S_{n-1}) - (n-1)\left[p\ ln(u) + (1-p)\ ln(d)\right] \\ &= R_{n-1}\ . \end{aligned}$$

Therefore $\{R_n, n = 1, 2, \ldots\}$ is martingale.

Now, consider the discounted process $\{S_0, e^{-r}S_1, e^{-2r}S_2, \ldots\}$ where r is the interest rate. We have

$$E(e^{-(n+1)r}S_{n+1}/\mathcal{F}_n) = p\ ue^{-(n+1)r}S_n + (1-p)\ de^{-(n+1)r}S_n\ .$$

The discounted process is a martingale only if the right hand side of the above equation is equal to $e^{-nr}S_n$. That is,

$$e^{-nr}S_n = p\ ue^{-(n+1)r}S_n + (1-p)\ de^{-(n+1)r}S_n$$

or

$$e^r = p\ u + (1-p)\ d\ .$$

Thus, the discounted process is a martingale only if

$$p = \frac{e^r - d}{u - d}.$$

□

Remark 8.6.3 *It may be noted that in Chapter 3 the RNPM for the binomial lattice model has been derived via replicating portfolio arguments. But the approach presented here in Example 8.6.6 is based on the fact that discounted stock process under RNPM form a martingale. This approach is more general and is applicable in a large variety of derivative pricing problems.*

Now, one of the important results of martingale theory is Doob's decomposition, which connects the martingale and submartingale. This result is stated below. We shall refer to Roman [113] for its proof.

Lemma 8.6.1 (Doob's Decomposition) *Let $\{X_n, n = 0, 1, \ldots\}$ be a submartingale with respect to the filtration $\{\mathcal{F}_n, n = 0, 1, \ldots\}$. Then there exists a martingale $M = \{M_n, n = 0, 1, \ldots\}$ with respect to the filtration $\{\mathcal{F}_n, n = 0, 1, \ldots\}$ and a stochastic process $A = \{A_n, n = 0, 1, \ldots\}$ such that:*

(i) $X_n = M_n + A_n$, $(n = 0, 1, \ldots)$.
(ii) A is an increasing stochastic process, i.e. $A_n \le A_{n+1}$, $(n = 0, 1, \ldots)$.
(iii) A_n is $\mathcal{F}_{n-1}$- measurable, for all n.

Remark 8.6.4 *In Example 8.6.6, $ln(S_n)$ is not a martingale whereas R_n is a martingale. This is simply an illustration (and not a consequence) of Lemma 8.6.1.*

Example 8.6.7 *Prove that, under risk neutral probability measure Q the discounted stock price process $\{(1+r)^{-k}S_k,\ k = 1, 2, \ldots\}$ is a martingale.*

Solution This is a simple interest version of the second part of Example 8.6.1. We are proving it separately even at the cost of repetition. Let Ω be all sequences of length n of heads H and tails T. Let S_0 be a fixed number. Define $S_k(w) = u^j d^{k-j} S_0$ if the first k elements of a given $w \in \Omega$ has j occurrences of H and $k-j$ occurrences of T. It means that if H occurs then the stock price goes up by a factor u; if T occurs then it goes down by a factor d. Let $\mathcal{F}_k$ be the σ-field generated by $S_0, S_1, \ldots, S_k$. Let $\overline{p} = \dfrac{(1+r)-d}{u-d}$, $\overline{q} = \dfrac{u-(1+r)}{u-d}$ and define $Q(w) = (\overline{p})^j(\overline{q})^{n-j}$ if w has j appearances of H and $n-j$ appearances of T. We note that under probability measure Q, the random variables $S_{k+1}/S_k = Y_k$ (say) are independent and identically distributed with probability mass function

$$P(Y = u) = \overline{p},\ \ P(Y = d) = \overline{q}\ .$$

Since Y_k i.i.d random variables, they satisfy

$$P(Y_1 = y_1, \ldots, Y_n = y_n) = P(Y_1 = y_1)\ldots P(Y_n = y_n) = (\overline{p})^j(\overline{q})^{n-j}\ .$$

Also we may note that Y_k is the factor by which the stock price goes up or down at time k. Since the random variable Y_k is independent of $\mathcal{F}_k$, we have

$$E_Q\left[(1+r)^{-(k+1)}S_{k+1}/\mathcal{F}_k\right] = (1+r)^{-k}S_k(1+r)^{-1}E_Q\left[(S_{k+1}/S_k)/\mathcal{F}_k\right]\ .$$

But

$$E_Q[(S_{k+1}/S_k)/\mathcal{F}_k] = E_Q[(S_{k+1}/S_k)] = \bar{p}\,u + \bar{q}\,d = (1+r)\,.$$

Therefore on substitution, we conclude

$$E_Q\left[(1+r)^{-(k+1)}S_{k+1}/\mathcal{F}_k\right] = (1+r)^{-k}S_k\,.$$

This proves that the process $\{(1+r)^{-k}S_k, k = 1,2,\ldots\}$ is a martingale.

□

Wealth Process

Let Δ_k be the number of shares of a stock held between time k and $k+1$. We assume that Δ_k is $\mathcal{F}_k$-measurable and X_0 is the amount of money we have started with time $t = 0$. If we have Δ_k shares between time k and $k+1$, then at time $k+1$ those shares will be worth $\Delta_k S_{k+1}$, where S_{k+1} is the share price at time $k+1$. The amount of cash we hold between time k and $k+1$ is X_k minus the amount held in stock, that is $X_k - \Delta_k S_k$. Hence, the worth of this amount at time $k+1$ is $(1+r)[X_k - \Delta_k S_k]$. Therefore, the amount of money we have at time $k+1$ is

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)[X_k - \Delta_k S_k]\,.$$

When $r = 0$, this reduces to

$$X_{k+1} - X_k = \Delta_k(S_{k+1} - S_k)\,.$$

Thus,

$$X_{k+1} = X_0 + \sum_{i=0}^{k} \Delta_i(S_{i+1} - S_i)\,.$$

The stochastic process $\{X_k, k = 0,1,\ldots\}$ is called the *wealth process*.

We shall now show that under risk neutral probability measure Q the discounted wealth process is a martingale.

$$\begin{aligned} E_Q[X_{k+1} - X_k/\mathcal{F}_k] &= E_Q[\Delta_k(S_{k+1} - S_k)/\mathcal{F}_k] \\ &= \Delta_k E_Q[(S_{k+1} - S_k)/\mathcal{F}_k] \quad (\Delta_k \text{ is } \mathcal{F}_k - \text{ measurable}) \\ &= 0 \quad (S_k \text{ is a martingale})\,. \end{aligned}$$

Now writing X_{k+1} as $X_k + \Delta_k(S_{k+1} - S_k)$ and noting that $r > 0$, we have

$$\begin{aligned} E_Q\left[(1+r)^{-(k+1)}X_{k+1} - (1+r)^{-k}X_k)/\mathcal{F}_k\right] &= E_Q\left[\Delta_k\left[(1+r)^{-(k+1)}S_{k+1} - (1+r)^{-k}S_k\right]/\mathcal{F}_k\right] \\ &= \Delta_k E_Q\left[(1+r)^{-(k+1)}S_{k+1} - (1+r)^{-k}S_k/\mathcal{F}_k\right] \\ &= 0 \quad (S_k \text{ is a martingale under } Q)\,. \end{aligned}$$

Hence, discounted wealth process $\{(1+r)^{-k}X_k, k = 1,2,\ldots\}$ is a martingale.

Definition 8.6.2 (Continuous Time Martingale) *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{X(t), t \geq 0\}$ be a stochastic process and $\{\mathcal{F}_t, t \geq 0\}$ be a filtration. The stochastic process $\{X(t), t \geq 0\}$ is said to be a martingale corresponding to the filtration $\{\mathcal{F}_t, t \geq 0\}$ if it satisfies the following conditions*

(i) For every t, $E(X(t))$ exists.
(ii) Each $X(t)$ is $\mathcal{F}_t$-measurable.
(iii) For every $0 < s < t$,

$$E(X(t)/\mathcal{F}_s) = X(s)\,. \tag{8.4}$$

Example 8.6.8 *Prove that $\{W(t), t \geq 0\}$ is a martingale.*

Solution For $0 < s < t$,

$$\begin{aligned} E(W(t)/\mathcal{F}_s) &= E(W(t) - W(s) + W(s)/\mathcal{F}_s) \\ &= E(W(t) - W(s)/\mathcal{F}_s) + E(W(s)/\mathcal{F}_s) \\ &= 0 + W(s) \quad \text{(from the property of Brownian motion)}\,. \end{aligned}$$

Therefore, $\{W(t), t \geq 0\}$ is a martingale.

□

Example 8.6.9 *Prove that the Poisson process $\{N(t), t \geq 0\}$ is not a martingale.*

Solution First $E(|N(t)|) \leq \lambda t < \infty$. For $0 < s < t$,

$$\begin{aligned} E(N(t)/\mathcal{F}_s) &= E((N(t) - N(s) + N(s))/\mathcal{F}_s) \\ &= E((N(t) - N(s))/\mathcal{F}_s) + E(N(s)/\mathcal{F}_s) \\ &= \lambda(t-s) + N(s) \quad \text{(from the property of Poisson process)}\,. \end{aligned}$$

Hence, $E(N(t)/\mathcal{F}_s) \neq N(s)$, $\{N(t), t \geq 0\}$ is not a martingale. Note that, $\{N(t), t \geq 0\}$ is a submartingale since $E(N(t)/\mathcal{F}_s) \geq N(s)$, for $0 < s < t$.

□

Example 8.6.10 *Show that $\exp\left(W(t) - \frac{t}{2}\right)$ is a martingale.*

Solution Let $0 \leq s < t$. Since $W(t) - W(s)$ is independent of $\mathcal{F}_s$ and $W(s)$ is $\mathcal{F}_s$-measurable, we have

$$\begin{aligned} E\left(e^{W(t)}/\mathcal{F}_s\right) &= E\left(e^{W(t)-W(s)}e^{W(s)}/\mathcal{F}_s\right) \\ &= e^{W(s)}E\left(e^{W(t)-W(s)}/\mathcal{F}_s\right) \\ &= e^{W(s)}E\left(e^{W(t)-W(s)}\right)\,. \end{aligned}$$

Since $W(t) - W(s)$ has normal distribution with mean zero and variance $(t-s)$, we have

$$E\left(e^{W(t)-W(s)}\right) = e^{\frac{t-s}{2}}.$$

Hence,

$$E\left(e^{W(t)}/\mathcal{F}_s\right) = e^{W(s)}e^{\frac{t-s}{2}}.$$

This gives, for $0 \le s < t$,

$$E\left(e^{W(t)-\frac{t}{2}}/\mathcal{F}_s\right) = e^{\frac{-t}{2}}E\left(e^{W(t)}/\mathcal{F}_s\right) = e^{W(s)-\frac{s}{2}}.$$

It follows that $\exp\left(W(t) - \frac{t}{2}\right)$ is a martingale.

□

The cornerstone of martingale theory is Doob's optional sampling theorem. This states, roughly, that "stopping" a martingale at a random time τ does not alter the expected "payoff", provided the decision about when to stop is based solely on information available up to τ. Such random times are called stopping times. *Stopping times* are also called as *Markov times* or *optional times*.

If τ is optional and $c > 0$ is a positive constant, then $\tau + c$ is a stopping time. Before stating Doob's optional sampling theorem we give below the definitions.

Definition 8.6.3 (Stopping Time) *A stopping time relative to a filtration $\{\mathcal{F}_n, n \ge 0\}$ is a non-negative integer-valued random variable τ such that for each n the event $\tau = n \in \mathcal{F}_n$. In continuous case, a random variable $\gamma : \Omega \to [a,b]$ is called a stopping time with respect to a filtration $\{\mathcal{F}_t, a \le t \le b\}$ if $\{w, \gamma(w) \le t\} \in \mathcal{F}_t$ for all $t \in [a,b]$.*

From the definition, we can think of γ as the time to stop playing a game. The decision to stop playing the game before or at time t should be determined by the information provided by $\mathcal{F}_t$ which will give the condition for γ to a stopping time.

Definition 8.6.4 (Hitting Time) *Let $\{X(t), t \ge 0\}$ be a stochastic process. Let A be Borel set in $\mathbb{R}^n$. Define*

$$\tau = inf\{t > 0 : X(t) \in A\}.$$

Then τ is called a hitting time of A for the stochastic process $\{X(t),\ t \ge 0\}$.

Example 8.6.11 *Consider a symmetric random walk $\{S_n, n = 0,1,\ldots\}$. Let T be the first time that $S_n = 2$. Prove that T is a stopping time.*

Solution Since T is the first time that $S_n = 2$, we have

$$T = \begin{cases} \text{Min}\{n \ge 0 : S_n = 2\}, & \text{if } S_n = 2 \text{ for some } n \in N \\ \infty, & \text{otherwise} \end{cases}$$

In other words,

$$T = inf\{n \ge 0 : S_n = 2\}.$$

Further

$$\{T \le n\} = \bigcup_{i=0}^{n}\{S_i = 2\} \in \mathcal{F}_n$$

where $\{S_i = 2\} \in \mathcal{F}_i \subseteq \mathcal{F}_n$ for $i \le n$ because $\{\mathcal{F}_n, n = 0,1,\ldots\}$ is a filtration. Hence it follows that T is a stopping time.

□

Result 8.6.5 *A stopping time T is said to be bounded if there exists a constant $K < \infty$ such that $P(T \le K) = 1$, and is said to be finite a.s. if $P(T < \infty) = 1$. If T is a bounded stopping time and $\{X(t), t \ge 0\}$ is a martingale, then $E(X(T)) = E(X(0))$. For instance, in the above example, T is a finite a.s. but not a bounded stopping time. Hence $E(S(T)) = 2 \ne 0 = E(S(0))$.*

Example 8.6.12 *Let τ be a stopping time. Prove that $W(t+\tau) - W(\tau)$ is a Brownian motion.*

Solution: By observing the Brownian motion, we can determine whether or not $\tau \le t$. For a fixed a, the hitting time of a is defined by

$$\tau = inf\{t > 0 : W(t) = a\}.$$

Since the sample paths of Brownian motion are continuous, it is easy to see that τ is a stopping time. That is,

$$\{\tau \le t\} = \{W(s) = a, \text{for some } s,\ 0 \le s \le t\},$$

which depends only on $\{W(s), 0 \le s \le t\}$. Also, note that, if $\tau < \infty$, then $W(\tau) = a$ by continuity of sample paths. Since τ is a stopping time, $W(t+\tau) - W(\tau)$ is a Brownian motion.

□

The following theorem is an important result in martingale theory. We shall refer to Roman [113] for the proof of the above theorem.

Theorem 8.6.1 (Doob's Optional Sampling Theorem) *Let $\{X_n, n = 0, 1, \ldots\}$ be a martingale with respect to $\{Y_n, n = 0, 1, \ldots\}$ and let τ be a stopping time. Then for any $n = 1, 2, \ldots$*

$$E(X_{\tau \wedge n}/Y_0, Y_1, \ldots Y_{n-1}) = X_{\tau \wedge n-1}$$

where $\tau \wedge n = Min(\tau, n)$. i.e. the stopped process is also a martingale.

We shall refer to Roman [113] for the proof of the above Theorem. This theorem is also known as optional stopping theorem.

Remark 8.6.5 *In case $\{X_n, n = 0, 1, \ldots\}$ is a submartingale with respect to $\{Y_n, n = 0, 1, \ldots\}$ then*

$$E(X_{\tau \wedge n}/Y_0, Y_1, \ldots Y_{n-1}) \geq X_{\tau \wedge n-1}.$$

Remark 8.6.6 *In case $\{X_n, n = 0, 1, \ldots\}$ is a supermartingale with respect to $\{Y_n, n = 0, 1, \ldots\}$ then*

$$E(X_{\tau \wedge n}/Y_0, Y_1, \ldots Y_{n-1}) \leq X_{\tau \wedge n-1}.$$

Example 8.6.13 *Suppose that two players A and B start respectively with Rs a and Rs b, and they bet against each other one rupee at a time by tossing a fair coin. What is the probability that player B will be broke and A ends up with all the money?*

Solution Let X_n be player A fortune at the nth game, and S_n be his/her capital after n games. We have

$$S_n = a + \sum_{i=1}^{n} X_n,$$

where X_i are i.i.d random variables with

$$P(X_n = \pm 1) = \frac{1}{2}.$$

Let

$$\tau_0 = inf\{n : S_n = 0\}$$
$$\tau_{a+b} = inf\{n : S_n = a + b\},$$

and $T = \tau_0 \wedge \tau_{a+b}$. Therefore, T is the first time that player A either makes b extra rupees or goes ruin i.e.

$$T = inf\{n : S_n = 0 \text{ or } S_n = b + a\}.$$

Hence, T is the stopping time with respect to the martingale $\{S_n, n = 0, 1, \ldots\}$. By using optional stopping theorem, we have

$$E(S_T) = S_0 = a.$$

But

$$E(S_T) = 0 \times P(T = \tau_0) + (a + b) \times P(T = \tau_{a+b}).$$

Hence,

$$P(T = \tau_{a+b}) = \frac{a}{a + b}.$$

This is exactly the probability that player B goes broke and player A wins all. □

8.7 Summary and Additional Notes

- Conditional expectation plays a key role in derivative pricing models. In Section 8.2, the definition of conditional expectation and its properties are presented.
- In Section 8.3, σ-fields generated by a random variable, a random vector and a stochastic process are defined. Finally, σ-field generated by Brownian motion is presented.
- Filtration is presented in Section 8.4 with simple examples. Both discrete and continuous time filtration are explained with some examples.
- We provide a detailed evaluation procedure of conditional expectation with respect to the filtration. Further, their properties are presented along with some examples.
- We know that a conditional expectation is itself a random variable. The complicated conditional expectations such as conditional expectation of a random variable with respect to the filtration arise in martingale theory and the concept of martingale theory is used in a betting strategy. Kolmogorov (1932) invented the martingale theory whereas its first systematic treatment was provided by Doob (1942).

- Two important results in martingale theory namely Doob's decomposition and Doob's optional sampling theorem are presented. Optional sampling concept reveals that the expectation of a martingale remains constant. It has immediate implications concerning the pathwise behaviour of martingales, submartingales and supermartingales. We may refer to Roman [113] for some more details on martingale theory.

8.8 Exercises

Exercise 8.1 *Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables each having uniform distribution with values -2, -1, 0, 1, 2, 3. Let $\{N(t), t \geq 0\}$ be a Poisson process with parameter 2. Find the mean and variance of $\sum_{i=1}^{N(t)} X_i$.*

Exercise 8.2 *Consider two i.i.d random variables X and Y each having uniform distribution between the intervals 0 and 1. Define $Z = X + Y$. Prove that $E(X/Z) = \frac{Z}{2}$.*

Exercise 8.3 *Consider the successive rolling of an unbiased die. Let X and Y denote the number of rolls necessary to obtain a two and a three respectively. Obtain (i) $E(X/Y = 2)$, (ii) $E(X/Y = 4)$.*

Exercise 8.4 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let X be an integrable random variable and let $\mathcal{G} \subset \mathcal{F}$ be a σ-field. Prove that, the conditional expectation $E(X/\mathcal{G})$ exists. (Hint: Use Radon-Nikodym theorem)*

Exercise 8.5 *Consider $\Omega = \{a, b, c, d\}$. Construct 4 distinct σ-fields $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4$ such that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \mathcal{F}_4$.*

Exercise 8.6 *Construct an example of σ-fields $\mathcal{F}_1$ and $\mathcal{F}_2$ such that $\mathcal{F}_1 \not\subset \mathcal{F}_2$ and $\mathcal{F}_2 \subset \mathcal{F}_1$.*

Exercise 8.7 *Consider a binomial model with $t = 1, 2, 3$. Let S_t be the stock price at time t. Let $\Omega = \{(u,u,u), (u,u,d), (u,d,u), (u,d,d), (d,u,d), (d,u,u), (d,d,u), (d,d,d)\}$. Let σ-field $\mathcal{F}$ be the power set of Ω and $P(w) = \frac{1}{8}$ for all $w \in \Omega$. Let $\mathcal{G} = \{\emptyset, \{(u,u,u), (u,u,d), (u,d,u), (u,d,d)\}, \{(d,u,d), (d,u,u), (d,d,u), (d,d,d)\}, \Omega\}$. Define a discrete random variable X as*

$$X(w_i) = i, \quad i = 1, 2, \ldots, 8.$$

Find $E(X/\mathcal{G})$.

Exercise 8.8 *Let $\{X_n, n = 0, 1, \ldots\}$ and $\{Y_n, n = 0, 1, \ldots\}$ be stochastic processes. We say $\{X_n\}$ is a martingale with respect to $\{Y_n\}$ if*

(i) $E(|X_n|) < \infty$.
(ii) $E(X_{n+1}/Y_0, Y_1, \ldots, Y_n) = X_n$.

Prove that $\{X_n, n = 0, 1, \ldots\}$ is a martingale with respect to $\{Y_n, n = 0, 1, \ldots\}$ where $X_n = Y_1 + Y_2 + \ldots + Y_n$, $n \geq 1$, $Y_0 = 0$, $\{Y_i, i = 1, 2, \ldots\}$ are independent random variables with $E(Y_n) = 0$.

Exercise 8.9 *Let $\{S_n, n = 0, 1, \ldots\}$ be a symmetric random walk and $\mathcal{F}_n$ be a filtration. Show that $Y_n = (-1)^n \cos(\pi S_n)$ is a martingale with respect to $\mathcal{F}_n$.*

Exercise 8.10 *Consider the tossing of an unbiased coin three times. Suppose that the toss of a head wins Rs.1 while an outcome of a tail loses Rs 1. Let X_n denote the sum of the winnings at time n. Prove that $\{X_n, n = 1, 2, 3\}$ is an adapted process with respect to $\{\mathcal{F}_i, i = 1, 2, 3\}$ defined in Section 7.4. Further prove that $\{X_n, n = 1, 2, 3\}$ is a martingale with respect to $\{\mathcal{F}_i, i = 1, 2, 3\}$.*

Exercise 8.11 *$\{N(t), t \geq 0\}$ be a Poisson process with parameter λ. Prove that, $\{N(t) - \lambda t, t \geq 0\}$ is a martingale.*

Exercise 8.12 *Let $X(t) = \mu t + \sigma W(t)$, $-\infty < \mu < \infty$, $0 < \sigma < \infty$. Prove that $\{X(t), t \geq 0\}$ is a martingale for $\mu = 0$. Also prove that $\{X(t), t \geq 0\}$ is a submartingale for $\mu > 0$, and a supermartingale for $\mu < 0$.*

Exercise 8.13 *Show that the maximum of two submartingales (with respect to the same filtration) is a submartingale.*

Exercise 8.14 *Let $Y_1, Y_2, \ldots$ be a sequence of positive independent random variables with $E(Y_i) = 1$ for all i. Define $X_0 = 1$, and $X_n = \prod_{i=1}^{n} Y_i$ for $n \geq 1$. Show that $\{X_n, n = 0, 1, \ldots\}$ is a martingale with respect to its natural filtration.*

Exercise 8.15 *Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d random variables with $P(Y_i = 1) = p$ and $P(Y_i = 0) = 1 - p$. Define $X_n = \prod_{i=1}^{n} \frac{Y_i}{p}$. Prove that, $E(X_{n+1}/Y_1, Y_2, \ldots, Y_n) = X_n$.*

Exercise 8.16 *Let X be a random variable and let φ be a convex function on R. Suppose both X and φ(X) are integrable, and $\mathcal{G} \subset \mathcal{F}$ is a σ-field. Prove that, $\phi(E(X/\mathcal{G})) \leq E(\phi(X)/\mathcal{G})$.*

Exercise 8.17 *Let $X_1, X_2, \ldots$ be a sequence of square integrable random variables. Show that if $\{X_n, n = 1,2,\ldots\}$ is a martingale with respect to a filtration $\mathcal{F}_n$, then X_n^2 is a sub martingale with respect to the same filtration.*
(Hint: Use Jensen's inequality with convex function $\phi(x) = x^2$)

Exercise 8.18 *Prove that, if $\{X_n, n = 1,2,\ldots\}$ is a martingale, then $|X_n|$, X_n^2, e^{X_n} and e^{-X_n} are all submartingales.*

Exercise 8.19 *Prove that, if $X_n > 0$ and is a martingale, then $\sqrt{X_n}$ and $\ln(X_n)$ are supermartingales.*

Exercise 8.20 *Prove that, if X_n is a submartingale and K a constant, then $Max\{X_n, K\}$ is a submartingale, while if X_n is a supermartingale, so is $Min\{X_n, K\}$.*

Exercise 8.21 *Let $X_1, X_2, \ldots$ be i.i.d random variables each having normal distribution with zero mean and unit variance. Show that the sequence $Y_n = \exp\left(\left(\sum_{i=1}^n X_i\right) - \frac{1}{2}n\right)$ forms a martingale.*

Exercise 8.22 *Prove that $\{W^2(t) - t, t \geq 0\}$ is a martingale, where $\{W(t), t \geq 0\}$ is a Brownian motion.*

Exercise 8.23 *Let $\{W(t), t \geq 0\}$ be a Wiener process. Is $\exp(\sigma W(t) - \frac{\sigma^2}{2}t)$ a martingale where σ is a positive constant?*

Exercise 8.24 *Let $\{N(t),\ t \geq 0\}$ be a Poisson process with parameter 1. Which of the following are martingales.*

(i) $\{N(t) - t,\ t \geq 0\}$.
(ii) $\{N(t)^2 - t,\ t \geq 0\}$.
(iii) $\{(N(t) - t)^2 - t,\ t \geq 0\}$.

Exercise 8.25 *Let X be a integrable random variable and $\mathcal{G}_1$ and $\mathcal{G}_2$ be two sub-σ-fields of $\mathcal{F}$. Prove that, if $\mathcal{G}_1 \subseteq \mathcal{G}_2$, then $E(E(X/\mathcal{G}_1)/\mathcal{G}_2) = E(E(X/\mathcal{G}_2)/\mathcal{G}_1) = E(X/\mathcal{G}_1)$.*

Exercise 8.26 *Let $\{\mathcal{F}_n, n = 1,2,\ldots\}$ be a filtration. Let X be any random variable with $E(|X|) < \infty$. Define $X_n = E(X/\mathcal{F}_n), n = 1,2,\ldots$ Prove that, $\{X_n, (n = 1,2,\ldots)\}$ is a martingale with respect to the filtration $\{\mathcal{F}_n, n = 1,2,\ldots\}$.*

Exercise 8.27 *Let $\{Y_n, (n = 0,1,\ldots)\}$ be an arbitrary sequence of random variables. Suppose X is a random variable with $E(|X|) < \infty$. Define*

$$X_n = E(X/Y_0, Y_1, \ldots, Y_n),\ (n = 0,1,\ldots)\ .$$

Show that $\{X_n, n = 0,1,\ldots\}$ forms a martingale with respect to $\{Y_n, n = 0,1,\ldots\}$. We say that, $\{X_n, n = 0,1,\ldots\}$ is a Doob's martingale.

Exercise 8.28 *Let $\{W(t), t \geq 0\}$ be a Weiner process and let τ be a stopping time. Define*

$$\tilde{W}(t) = \begin{cases} W(t), & t \leq \tau \\ 2W(\tau) - W(t), & t > \tau\ . \end{cases}$$

Prove that $\{\tilde{W}(t), t \geq 0\}$ is a Weiner process.

Exercise 8.29 *Let τ is a stopping time for the Weiner process. Using the optimal sampling theorem, show that $E(W(\tau)) = 0$. Use this to deduce that $P(W(\tau) = a) = \frac{b}{a+b}$ and $P(W(\tau) = b) = \frac{a}{a+b}$.*

Exercise 8.30 *For $a \in \mathbb{R}$, let τ_a be the first time that $\{W(t), t \geq 0\}$ hits a. Suppose that $a > 0 > b$. By considering the stopped martingale $\{(W(t))^{\tau_a \wedge \tau_b}, t \geq 0\}$, prove that $P(\tau_a < \tau_b) = \frac{-b}{a-b}$.*