5 Stochastic Processes & Martingales

Basic Concepts of Random Processes

- In many real-world problems, we are not just facing a single random variable $X : \Omega \to \mathbb{R}$. We may have *a sequence of* random variables $X_1, X_2, X_3, \dots$ over the same sample space evolving over time, which is called a **random process (or stochastic process)**.
- Here, we informally define a random process as a (finite or infinite) sequence of random variables $\{X_t : t \in \mathcal{T}\}$. If $\mathcal{T} = \{1, 2, \dots\}$, then it is called a discrete-time random process; if $\mathcal{T} = [0, T] \subseteq \mathbb{R}$, then it is called continuous-time. We will only study discrete-time random processes in this class.
- Given a random process $\{X_t : t = 1, 2, \dots\}$, if we fix the time t, then we get a random variable. If we fix the sample point $\omega \in \Omega$, then we get a **sample path** (or a trajectory).

Motivating Example: Random Walk

- A sequence of i.i.d. random variables $X_1, X_2, \dots$ is random process but this is not very interesting. A simple random process with interesting properties is the **random walk**. Suppose $X_1, X_2, \dots$ are i.i.d. random variables such that $P(X_i = 1) = P(X_i = -1) = 0.5$. Then we can define a sequence of new random variables $S_1, S_2, \dots$ by

$$S_n = X_1 + X_2 + \cdots + X_n$$

- A key question in random process is as follows: what is the sample space of $\{S_i\}$? Suppose that we only consider random walk with finite horizon of length $T \in \mathbb{N}$. Then we can construct the so called canonical probability space by

$$\Omega = \underbrace{\{-1,1\} \times \{-1,1\} \times \cdots \times \{-1,1\}}_{T\text{-time}} = \{-1,1\}^T$$

  Since the individual "coin tosses" need to be fair and independent. This dictates that any two sequences be equally likely. There are 2^T possible sequences of length T, so we can define the probability measure by

$$\forall w \in \Omega : P(\{\omega\}) = \left(\frac{1}{2}\right)^T$$

  Then we can define each random variable $S_i : \Omega \to \mathbb{Z}$ by:

$$\forall w = (c_1, c_2, \dots, c_T) : S_n(\omega) = c_1 + c_2 + \cdots + c_n$$

  Then $\{S_n : n = 1, \dots, T\}$ is the desired random process for random walk.
- Based on the above construction, one may better understand why fixing $\omega \in \Omega$ yields a sample path. Because $\{S_n : n = 1, \dots, T\}$ may live in a big sample space containing all information. For example, fixing $w = (1, 1, 1, 1)$, we get sample path $1, 2, 3, 4$ and by fixing $w = (1, -1, -1, -1)$, we get sample path $1, 0, -1, -2$.

Filtrations

- Note that although $X_1, X_2, \dots$ are defined over the same sample space Ω, the events we can tell at each time instant are different. For example, for S_2 in the random walk, it makes sense to consider $\{(1,1,1,1)\}$ and $\{(-1,-1,1,1)\}$ as two different events because we already have the outcomes of the first two tosses at time $t = 2$. However, it does not make much sense to consider $\{(1,1,1,1)\}$ and $\{(1,1,-1,-1)\}$ as two different events for S_2 because we cannot predict the future!

- Therefore, we are only interested in identifying the smallest event set that is sufficient enough to describe the current information. Recall that our information can be encoded as a σ-field. Furthermore, as time goes by, we should be able to tell more and more precise events, i.e., the information set should be non-decreasing. This leads to the definition of **filtrations**.

Definition: Filtrations

A sequence of σ-fields $\mathcal{F}_1, \mathcal{F}_2, \dots$ on Ω is called a filtration if

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$$

- Recall that σ-field generated by $X_1, \dots, X_n$, denoted by $\sigma(X_1, \dots, X_n)$ is the smallest σ-field containing all events of form $\{X_i \in B\}$, where $B \in \mathcal{B}(\mathbb{R})$ is a Borel set. For a sequence $X_1, X_2, \dots$ of coin tosses, we take $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Then consider the following event

$$A = \{\text{the first 5 tosses produce at least 2 heads}\}$$

Then we should have $A \in \mathcal{F}_5$ but $A \notin \mathcal{F}_4$. This is because this information cannot be determined at instant $n = 4$ but can be determined at $n = 5$. If we consider event

$$B = \{\text{there exists at least 1 head in the sequence } X_1, X_2, \dots\}$$

Then we have $B \notin \mathcal{F}_n$ for any n. Also, if we consider event

$$C = \{\text{there are no more 2 heads and 2 tails among the first 5 tosses}\}$$

Then we have $B \in \mathcal{F}_1$ because $B = \emptyset$.

- Based on the above discussion, at time n, random variable X_n at least needs σ-field $\mathcal{F}_n$ to support its information. This leads to the definition of adapted filtration.

Definition: Adapted Stochastic Processes

We say a sequence of random variables $X_1, X_2, \dots$ is adapted to a filtration $\mathcal{F}_1, \mathcal{F}_2, \dots$ if for any $n = 1, 2, \dots$, X_n is $\mathcal{F}_n$-measurable. In some books, we also call the process $\{X_n\}$ an $\{F_n\}$-adapted process.

- We can check easily (as a homework) that, by considering $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$, random process $X_1, X_2, \dots$ is adapted to filtration $\mathcal{F}_1, \mathcal{F}_2, \dots$. In fact, such an $\{\mathcal{F}_n\}$ is the smallest filtration to which $\{X_n\}$ is adapted.

Random Walk

- Suppose we consider the above random walk for $T = 3$. Then

$$\begin{aligned}\mathcal{F}_1 =& 2^{\{\{HHH,HHT,HTH,HTT\},\{THH,THT,TTH,TTT\}\}}\\ \mathcal{F}_2 =& 2^{\{\{HHH,HHT\},\{HTH,HTT\},\{THH,THT\},\{TTH,TTT\}\}}\\ \mathcal{F}_3 =& 2^{\{\{HHH\},\{HHT\},\{HTH\},\{HTT\},\{THH\},\{THT\},\{TTH\},\{TTT\}\}}\end{aligned}$$

Gambling Example

- Let us consider the following game. You can play a sequence of independent fail coin tosses. For each play, you can set c as your stake. If the result is T, then you lose ¥c; if the result is H, then you win ¥c. There is a well-known strategy called the **martingale strategy** saying that you double your stake every time a loss is faced. For example, initially your stake is $c_1 = 1$. If you lose, then set $c_2 = 2$ and if you loss again and again, then set $c_3 = 4, c_4 = 8, c_5 = 16$ and so forth. For such a strategy, we will always win ¥1: if H appears in the nth toss, then

$$2^{n-1} - 2^{n-2} - \cdots - 4 - 2 - 1 = 1$$

  Then by keeping playing the martingale strategy, you will win arbitrary amount of money.

- The problem of the above "sure winning" strategy is that (i) you are assumed to have an infinite amount of capital; and (ii) you are allowed to play the game for an infinite time horizon. In practice, you need a mechanism to quit the game either when the time is out or you reach a desired capital or you are bankrupt.

- Suppose that you are allowed to play the martingale strategy for only at most n times. Then your expected payoff is

$$1 \times \frac{1}{2} + 1 \times \frac{1}{4} + \cdots + 1 \times \frac{1}{2^n} - (1 + 2 + \cdots + 2^{n-1}) \times \frac{1}{2^n} = 0$$

- Question: Suppose you have unlimited time but your initial capital is a. Furthermore, you set your stake as $c = 1$ constantly for each toss. What is your probability of bankrupt? If you will quit game when your capital reaches some pre-defined value $b > a$, what is your probability of bankrupt? (This is essentially a random walk with **absorbing barriers**)

What We Need to Describe a Game by Stochastic Processes

- First, we need to have a **game dynamic**. For this, we have already understood that it can be described as a process $\{X_n\}$ adapted to filtration $\{\mathcal{F}_n\}$.

- Second, we need to have a **staking strategy** that determines $\{c_n\}$. Each c_n can be a random variable lives in the information set up to time $n - 1$.

- Third, we need a **stopping strategy** that determines when we quit the game. How to formally describe this?

- Finally, our ultimate goal is to evaluate if a game is **inherently fair** meaning that *you cannot beat the system no matter what you play.*

Martingales

Definition: Martingales

A discrete time process $\{X_n\}$ is called a **martingale** with respect to filtration $\{\mathcal{F}_n\}$ if

- $\{X_n\}$ is $\{\mathcal{F}_n\}$-adapted;
- $E(|X_n|) < \infty$ for any $n \geq 1$;
- $E(X_n \mid \mathcal{F}_{n-1}) = X_{n-1}$ almost surely, for all $n \geq 2$.

- ▶ Recall that $X = Y$ *almost surely* if $P(X = Y) = P(\{\omega : X(\omega) = Y(\omega)\}) = 1$.
- ▶ In the last condition:
  - if "=" is replaced by "≤", then $\{X_n\}$ is called a **super-martingale**;
  - if "=" is replaced by "≥", then $\{X_n\}$ is called a **sub-martingale**.

Let us discuss some consequences of the above definition

- ▶ The last condition essentially says that: given the information available up to time $n-1$, i.e., $\mathcal{F}_{n-1}$, the expectation for the value of the process is unchanged. By taking what is known, we can see more clearly that

$$E(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) = E(X_n \mid \mathcal{F}_{n-1}) - E(X_{n-1} \mid \mathcal{F}_{n-1}) = E(X_n \mid \mathcal{F}_{n-1}) - X_{n-1} = 0$$

- ▶ Note that $\{\mathcal{F}_n\}$ is a filtration. If we use the tower property, then for any $m < n$, we have

$$E(X_n \mid \mathcal{F}_m) = E(E(X_n \mid \mathcal{F}_{n-1}) \mid \mathcal{F}_m) = E(X_{n-1} \mid \mathcal{F}_m) = \cdots = E(X_m \mid \mathcal{F}_m) = X_m$$

  By taking the expectation, we have

$$E(X_n) = E(X_0), \forall n \geq 0$$

- ▶ Therefore, if we consider $\{X_n\}$ as the capital of a gambler at time n, then this can be interpreted that **the game is fair** in the sense that the expected profit in each step is 0. In the super-martingale case, this value is ≤ 0 meaning that the game is unfavourable.

- ▶ **A Simple Example of Martingale**
  Let $X_1, X_2, \ldots$ be a sequence of independent integrable random variables such that $E(X_n) = 0, \forall n \geq 1$. We define

$$S_n = X_1 + \cdots + X_n \text{ and } \mathcal{F}_n = \sigma(X_1, \ldots, X_n)$$

  Then $\{S_n\}$ is a **martingale** with respect to filtration $\{\mathcal{F}_n\}$ because

$$E(S_{n+1} \mid \mathcal{F}_n) = E(X_{n+1} \mid \mathcal{F}_n) + E(S_n \mid \mathcal{F}_n) = E(X_{n+1}) + S_n = S_n$$

  Note that here we just require independent not i.i.d.. As a special case, random walk is also a martingale process.

Games of Chance

- Let $W_1, W_2, \ldots$ be a sequence of random variables, where W_n are your *winnings (or losses) per unit stake* in game n. If your stake in each game is one, then your total winnings after n games will be
$$X_n = W_1 + W_2 + \cdots W_n$$
We take the filtration $\mathcal{F}_n = \sigma(W_1, \ldots, W_n)$ with $X_0 = 0$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

- Suppose that you can vary the stake to be C_n in game n. In particular, C_n may be zero if you refrain from playing the nth game; it may even be negative if you own the casino and can accept other people's bets. When the time comes to decide your stake C_n, you will know the outcomes of the first $n-1$ games. Therefore it is reasonable to assume that C_n is $\mathcal{F}_{n-1}$-measurable, where $\mathcal{F}_{n-1}$ represents your knowledge accumulated up to and including game $n-1$. In particular, since nothing is known before the first game, we take $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Definition: Previsible

A sequence of random variables $C_1, C_2, \ldots$ is said to be **previsible** with respect to filtration $\mathcal{F}_1, \mathcal{F}_2, \cdots$ if for any $n \geq 1$, C_n is $\mathcal{F}_{n-1}$-measurable, where $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Such a sequence in gambling is also called a **gambling strategy**.

- If you follow strategy $\{C_n\}$, then your total winnings after n games will be
$$Y_n = C_1 W_1 + \cdots + C_n W_n = C_1(X_1 - X_0) + \cdots + C_n(X_n - X_{n-1}) =: (C \bullet X)_n$$
The process $\{(C \bullet X)_n\}$ is called the **martingale transform** of $\{X_n\}$ by $\{C_n\}$. Then the big question now is: Can you choose $\{C_n\}$ such that your expected total winnings are positive? The following theorem shows that this is impossible if the system itself is inherently fair!

Theorem: You Cannot Beat the Systems

Let $\{C_n\}$ be a bounded previsible process, i.e., there exists $C > 0$ such that $|C_n(\omega)| \leq C$ for any $n \geq 1$ and $\omega \in \Omega$. Then, if $\{X_n\}$ is a (super or sub) martingale, then so is $\{(C \bullet X)_n\}$.

Proof: Since C_n and $(C \bullet X)_{n-1}$ are $\mathcal{F}_{n-1}$-measurable, by "taking out what is known", we have
$$\begin{aligned} E((C \bullet X)_n \mid \mathcal{F}_{n-1}) =& E((C \bullet X)_{n-1} + C_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}) \\ =& (C \bullet X)_{n-1} + C_n(E(X_n \mid \mathcal{F}_{n-1}) - X_{n-1}) \end{aligned}$$
Since X_n is a martingale, we have
$$C_n(E(X_n \mid \mathcal{F}_{n-1}) - X_{n-1}) = 0$$
Therefore, $E((C \bullet X)_n \mid \mathcal{F}_{n-1}) = (C \bullet X)_{n-1}$.

More Examples of Martingales

- **Random Product**
  Let $X_1, X_2, \dots$ be a sequence of independent integrable random variables such that $E(X_n) = 1, \forall n \geq 1$. We define

$$M_n = X_1 \times \cdots \times X_n \text{ and } \mathcal{F}_n = \sigma(X_1, \dots, X_n)$$

  Then $\{M_n\}$ is martingale with respect to filtration $\{\mathcal{F}_n\}$ because

$$E(M_{n+1} \mid \mathcal{F}_n) = E(X_{n+1}M_n \mid \mathcal{F}_n) = M_n E(X_{n+1} \mid \mathcal{F}_n) = M_n E(X_{n+1}) = M_n$$

- **Accumulating Data about a Random Variable**
  Let $\{\mathcal{F}_n\}$ be an arbitrary filtration and X is an integrable random variable. We define the data accumulated about X at time n by

$$X_n = E(X \mid \mathcal{F}_n)$$

  Then X_n is a martingale because of the tower property of conditional expectation

$$E(X_{n+1} \mid \mathcal{F}_n) = E(E(X \mid \mathcal{F}_{n+1}) \mid \mathcal{F}_n) = E(X \mid \mathcal{F}_n) = X_n$$

- **Polya's Urn**
  At time 0, an urn contains one black ball and one white ball. At each time $n = 1, 2, \dots,$ a ball is chosen at random from the urn and we add a new ball of the same colour. Just after time n, we have $n + 2$ balls in the urn, of which $B_n + 1$ black balls, where B_n is the number of black balls chosen by time n. We consider the proportion of black balls in the urn after time n

$$M_n = \frac{B_n + 1}{n + 2}$$

  Then $\{M_n\}$ is a martingale with respect to filtration $\mathcal{F}_n = \sigma(B_1, \dots, B_n)$. To see this, we consider a sequence of random variables $\{X_n\}$, where X_n takes value 1 if a black ball is chosen at n and zero otherwise. Then $B_n = X_1 + \cdots + X_n$. We have

$$P(X_1 = 1) = \frac{1}{2}, P(X_n = 1 \mid B_{n-1} = k) = \frac{k + 1}{n + 1}$$

  In terms of conditional expectation, this means that

$$E(X_n \mid \mathcal{F}_{n-1}) = \frac{B_{n-1} + 1}{n + 1}$$

  Then we have

$$\begin{aligned} E(M_n \mid \mathcal{F}_{n-1}) =& E\left(\frac{B_{n-1} + X_n + 1}{n + 2} \mid \mathcal{F}_{n-1}\right) = \frac{B_{n-1} + 1}{n + 2} + E\left(\frac{X_n}{n + 2} \mid \mathcal{F}_{n-1}\right) \\ =& \frac{B_{n-1} + 1}{n + 2} + \frac{B_{n-1} + 1}{(n + 1)(n + 2)} = \frac{B_{n-1} + 1}{n + 1} \\ =& M_{n-1} \end{aligned}$$