Topic 4 Martingales

1. Measure-theoretic view of probability theory
   - Sample space
   - σ-algebra
   - Probability triple
   - Filtration
   - Adaption
2. Motivation for introducing martingale
3. Definition of a martingale

Motivation

- The definition of a martingale involves conditional expectation, which is fundamentally defined relative to a sigma-algebra.
- A filtration is an increasing family of sigma-algebras that encodes the information available up to each time.
- In the martingale definition we condition on filtration.

Sample space

- Each possible outcome of a random experiment is called a sample point, usually denoted by ω; the set of all possible outcomes is called the sample space, usually denoted by Ω.
- E.g., tossing a fair coin twice, the sample space is
$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$$
- E.g., for experiments with continuous outcomes, the sample space is usually a continuum of numbers, often an interval of the real line.

Events

- We define events as subsets of a sample space. An event is said to occur if and only if one of the sample points it contains occurs.
- We regard Ω and $\emptyset$ as events.

- Let Ω be a nonempty set, and let $\mathcal{F}$ be a **collection of subsets** of Ω.

$$\Omega = \{1,2,3\}, \qquad \mathcal{F} = \{\emptyset, \{1\}, \{2,3\}, \{1,2,3\}\}$$

σ-algebra

Definition. We say that $\mathcal{F}$ is a σ-algebra (or σ-field) of Ω, provided that:

1. The empty set $\emptyset$ belongs to $\mathcal{F}$,
2. Whenever a set A belongs to $\mathcal{F}$, its complement A^c also belongs to $\mathcal{F}$, and
3. Whenever a sequence of sets A_1, A_2 ... belong to $\mathcal{F}$, their countable union $\bigcup_{n=1}^{\infty} A_n$ also belongs to $\mathcal{F}$.

$$\Omega = \{1,2,3\}, \qquad \mathcal{F} = \{\emptyset, \{1\}, \{2,3\}, \{1,2,3\}\}$$

Probability can be assigned to sets in the sigma-algebra later.

σ-algebra

E.g. Let $\Omega = \{1,2,3,4\}$. Determine whether or not each of the following is a σ −algebra on Ω.

$$\mathcal{F}_1 = \{\emptyset, \{1,2\}, \{3,4\}, \{1,2,3,4\}\}$$

$$\mathcal{F}_2 = \{\emptyset, \{3\}, \{4\}, \{1,2\}, \{3,4\}, \{1,2,3\}, \{1,2,4\}, \{1,2,3,4\}\}$$

$$\mathcal{F}_3 = \{\emptyset, \{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}, \{1,2,3,4\}\}$$

σ-algebra

- E.g., tossing a fair coin once $\{tails\} = \{\omega: X(\omega) = 0\}$ and $\{heads\} = \{\omega: X(\omega) = 1\}$ must belong to $\mathcal{F}$. Similarly, Ω and its complement $\emptyset$ should also be in $\mathcal{F}$.

- Consider a **continuous valued** event, not only the events $\{\omega: a < X(\omega) \leq b\}$, but also $\{\omega: b < X(\omega)\}$, $\{\omega: X(\omega) \leq a\}$ should also be contained within $\mathcal{F}$.

σ-algebra

E.g. Let $\mathcal{F}$ be a σ-algebra. If sets $A_i \in \mathcal{F},\ i = 1,2,\ldots$ Show that the finite intersection $\bigcap_{i=1}^{n} A_i \in \mathcal{F}, n \geq 1$.

- Proof:

$A_i \in \mathcal{F}$, so $\overline{A_i} \in \mathcal{F}$

$\bigcap_{i=1}^{n} A_i = \overline{\bigcup_{i=1}^{n} \overline{A_i}} \in \mathcal{F}$, so $\bigcap_{i=1}^{n} A_i \in \mathcal{F}$.

De Mogan's law : $A_1 \cap A_2 = \overline{\bar{A_1} \cup \bar{A_2}}$

The Borel σ-algebra

When the sample space contains real numbers, we cannot list all elements in $\mathcal{F}$. However, we could construct the σ-algebra by the following steps:

1. We start with a basic class of sets, such as all half-lines of the form $(-\infty, x]$, where x is any real number.
2. Using the defining properties of a σ-algebra (closure under countable unions and complements), we gradually extend this class of sets. This process incorporates various types of intervals (open, closed, singleton sets etc.).
3. Through repeated ~~application~~ of **countable unions, intersections, and complement** operations, we eventually obtain a **σ-algebra** that includes all open sets.

This σ-algebra is called the Borel σ-algebra, and its elements are called Borel sets.

it doesn't cotain all possible subsets of real line.

Probability measure and probability space

Let Ω be a nonempty set, and let $\mathcal{F}$ be a σ-algebra of subsets of Ω. A **probability measure** $\mathbb{P}$ is a function that, to **every set** $\boldsymbol{A} \in \boldsymbol{\mathcal{F}}$, assigns a number in [0,1], called the probability of A and written $\mathbb{P}(A)$. We require:

1. $\mathbb{P}(\Omega) = 1$, and

2. Whenever $A_1, A_2 \ldots$ is a sequence of **disjoint** sets in $\mathcal{F}$, then $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ (countable additivity)

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

Probability measure and probability space

- E.g. Consider the sample space of rolling a fair die $\Omega = \{1,2,3,4,5,6\}$. Consider this σ-algebra $\mathcal{F} = \{\emptyset, \{1,2,3\}, \{4,5,6\}, \{1,2,3,4,5,6\}\}$.
- A probability measure $\mathbb{P}$ is defined on $\mathcal{F}$ such that
  - $\mathbb{P}(\Omega) = 1$
  - $\mathbb{P}(\emptyset) = 0$
  - $\mathbb{P}(\{1,2,3\}) = \mathbb{P}(\{4,5,6\}) = 0.5$
- Now consider the subset $\{1,2\}$ which is not in the σ-algebra $\mathcal{F}$, it does not have a probability measure.
- Any set not in $\mathcal{F}$ is not measurable and thus does not have a probability.

Probability measure and probability space

- We shall assume that there is an underlying probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to which all further probability objects are defined. This assumption shall be universal that we will often not even mention it.

Random variable

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random variable is a function $\boldsymbol{X}$ from Ω to the real numbers $\mathbb{R}$ such that

$$\text{For every } x \in \mathbb{R}, \{\omega \in \Omega; \boldsymbol{X}(\omega) \leq x\} \in \mathcal{F}, \text{ or equivalently}$$

$$\{\omega \in \Omega; X(\omega) \in \text{ B}\} \in \mathcal{F}, \text{ for every Borel set B.}$$

$P\{a < X < b\}$

- A random variable $X = X(\omega)$ is a real-valued function defined on sample space Ω and relative to a σ-algebra $\mathcal{F}$ so that we can assign a probability measure to the random variable.

Random variable

- E.g., $\Omega \ = \ \{a, b, c, d\}$. $\mathcal{F}_1 = \{\emptyset, \{a, b\}, \{c, d\}, \Omega\}$. $\mathcal{F}_2 = \{\emptyset, \{a\}, \{b, c, d\}, \Omega\}$

Define a random variable $X: \Omega \to \mathbb{R}$ such that

$$X(a) = 1, X(b) = 1, X(c) = 0, X(d) = 0.$$

Show that X is a random variable with respect to $\mathcal{F}_1$ but not $\mathcal{F}_2$.

For $\{1\} \in$ B, the collection of events is $\{a, b\} \in \mathcal{F}_1$ but $\{a, b\} \notin \mathcal{F}_2$.

For $\{0\} \in$ B, the collection of events is $\{c, d\} \in \mathcal{F}_1$ but $\{c, d\} \notin \mathcal{F}_2$.

σ-algebra generated by random variable

- Let X be a random variable defined on a nonempty sample space Ω. The σ-algebra generated by X, denoted by $\sigma(X)$, is the collection of all subsets of Ω of the form $\{X \in B\}$, where B ranges over the Borel subsets of $\mathbb{R}$.
- E.g., $\Omega = \{a, b, c, d\}$. $\mathcal{F} = \{\emptyset, \{a, b\}, \{c, d\}, \{a, c\}, \{b, d\}, \Omega\}$.

Define a random variable $X: \Omega \to \mathbb{R}$ such that

$$X(a) = 1, X(b) = 1, X(c) = 0, X(d) = 0.$$

- $\sigma(X) = \{\emptyset, \{a, b\}, \{c, d\}, \Omega\}$
- $\sigma(X)$ only **contains the events** whose values can be described by X.

σ-algebra generated by random variable

- $\sigma(X)$ is the smallest σ-algebra for which the random variable X is measurable.
- It is "smallest" in the sense that it includes no more sets than necessary to ensure that all the events in $\sigma(X)$ can be used to describe the behaviour of the random variable X.

σ-algebra generated a by stochastic process

- For a stochastic process, the sample space refers to the set of all possible "paths". Each path represents a possible realization of the stochastic process at every point in time.

- For a stochastic process $\{X_t, t \geq 0\}$, $\sigma(X_t)$ is the smallest σ-algebra containing all sets of the form

$$\{\omega\text{: the sample path } X_t(\omega) \text{ belongs to B}\},$$

where B ranges over the Borel subsets of $\mathbb{R}$.

Conditional expectation revisit

- Any events that are relevant to a random variable X can be found in $\sigma(X)$.
- Let Y be a random variable or a stochastic process on Ω and $\sigma(Y)$ the σ-algebra generated by Y.
- The conditional expectation of a random variable X given Y is defined by
$$E(X|Y) = E(X|\sigma(Y))$$

Filtration

power set.

- Consider tossing a fair coin 3 times. The sample space Ω is $\{(H,H,H),(H,T,H),(H,H,T),(H,T,T),(T,H,H),(T,H,T),(T,T,H),(T,T,T)\}$
- Denote the events as $\omega = (\omega_1, \omega_2, \omega_3)$. $\mathcal{F}$ is the set of all subsets of Ω.

Before tossing the coins, what are the information we know?

- Before the first coin is tossed, we only know that for a sample point ω, $\omega \in \Omega$, $\omega \notin \emptyset$. At time point 0, $\mathcal{F}_0 = \{\Omega, \emptyset\}$.
- After the first coin is tossed, we obtain partial information of ω. For example, if it turns out that $\omega_1 = \mathrm{H}$, we know $\omega \in A_1 = \{$the first toss is H$\}$, $\omega \in \Omega$, and $\omega \notin A_0 = \{$the first toss is T$\}$, $\omega \notin \emptyset$.

Filtration

- On the other hand, if $\omega_1 = \mathrm{T}$, we know $\omega \in A_0 = \{$the first toss is T$\}$, $\omega \in \Omega$, and $\omega \notin A_1 = \{$the first toss is H$\}$, $\omega \notin \emptyset$.
- After the first coin is tossed, we have $\mathcal{F}_1 = \{\Omega, \emptyset, A_0, A_1\}$ and we can tell if $\omega \in A_0$ or $\omega \in A_1$.
- Also, $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}$.

F_1 contains the information of 1st toss.

- Similarly, after the second toss, suppose $\omega_1 = \mathrm{H}$, $\omega_2 = \mathrm{T}$, we know $\omega \in A_{10}$, $A_{10} = \{HTT, HTH\}$, $\omega \in \Omega$, $\omega \notin A_{01}$, $\omega \notin A_{00}$, $\omega \notin A_{11}$, $\omega \notin \emptyset$.
- At time point 2, $\mathcal{F}_2 = \{\Omega, \emptyset, A_0, A_1, A_{10}, A_{01}, A_{00}, A_{11}, {A_{10}}^C, {A_{01}}^C, {A_{00}}^C, {A_{11}}^C, A_{11} \cup A_{01}, A_{11} \cup A_{00}, A_{10} \cup A_{01}, A_{10} \cup A_{00}\}$.
- $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}$

ω

A, A^c

Filtration

- The collection of σ-algebras $\mathcal{F}_0$, $\mathcal{F}_1$, $\mathcal{F}_2$ is an example of **filtration**.
- Let Ω be a nonempty set. Let T be a fixed positive number, and assume that for each $t \in [0, T]$ there is a σ-algebra $\mathcal{F}(t)$. Assume further that if $s \leq t$, then every set in $\mathcal{F}(s)$ is also in $\mathcal{F}(t)$. Then we call the collection of σ-algebras $\mathcal{F}(t)$, $0 \leq t \leq T$, a filtration.
- When we get to time t, we will know whether the true sample point ω lies in specific sets of $\mathcal{F}(t)$.
- A filtration reflects an increasing stream of information.

Adaption

- The stochastic process $\{X_t, t \geq 0\}$ is said to be adapted to the filtration $\{\mathcal{F}_t, t \geq 0\}$ if

$$\sigma(X_t) \subset \mathcal{F}_t \text{ , } t \geq 0.$$

- The adoptedness of a stochastic process $\{X_t, t \geq 0\}$ means that X_ts do not carry more information than $\mathcal{F}_t$.
- The stochastic process $\{X_t, t \geq 0\}$ is always adapted to the natural filtration generated by $\{X_t, t \geq 0\}$.

$$\mathcal{F}_t = \sigma(X_s), s \leq t.$$

Adaption

We could work with different filtrations for the same stochastic process.

- Consider a stock price process $\{X_t, t \geq 0\}$.
- Define $\mathcal{F}_t = \sigma(X_s), s \leq t$, which includes only the historical stock price up to time t.
- Another filtration $\mathcal{G}_t, s \leq t$, which may include additional information such as relevant economic indicators, market news, and perhaps insider information.
- Our prediction of future value of $\{X_t\}$ may change if we consider different filtrations.

Adaption: motivation

- When discussing a stochastic process (e.g., assets price, portfolio processes (i.e., positions), wealth processes), it is generally important to specify that the process is adapted to a certain filtration.

- In Itô calculus, the integrand must be adapted to the filtration to ensure that the integral is well-defined and that the stochastic process behaves consistently with respect to the available information.

Motivation for introducing martingale

In the later chapters, you will see that

- Brownian motion is a martingale.

- Using the risk neutral probability, the discounted stock price and discounted European call option prices are martingales.

- Indefinite Itô integrals are martingales.

4.1 Martingales

- A stochastic process $\{Z_n | n \geq 1\}$ is said to be a *martingale* process if

$$① \; E[|Z_n|] < \infty \qquad \text{for all } n \qquad \text{(integrability)}$$

and

$\sigma(Z_1, Z_2, \cdots Z_n)$

$$② \; E[Z_{n+1} | Z_1, Z_2, \ldots, Z_n] = Z_n. \tag{4.1}$$

- A martingale is a generalized version of a fair game.
- If we interpret Z_n as a gambler's fortune after the nth game, then Eq. (4.1) states that his expected fortune after the $(n+1)$st gamble is equal to his fortune after the nth gamble no matter what may have previously occurred, i.e. no gain or loss.
- Key properties: 1. Knowledge of previous events does not improve our position. 2. No gain or loss is expected.

A more rigorous definition of martingale

The stochastic process $\{X_n, n = 0,1,\ldots\}$ is called a discrete-time martingale with respect to the filtration $\mathcal{F}_n$, $\quad n = 0,1,\ldots$ if

- $E|X_n| < \infty$ for all $n = 0,1,\ldots$
- X_n is adapted to $\mathcal{F}_n$
- $E[X_{n+1} | \mathcal{F}_n] = X_n$, for all $n = 0,1,\ldots$

In other words, X_n is the best prediction of X_{n+1} given $\mathcal{F}_n$.

Take-home messages

- σ-algebra is a collection of subsets of a sample space Ω that satisfies the definitive properties.
- The definition of a martingale involves conditional expectations. Filtration helps in defining these conditional expectations by specifying what information is available at each time step.
- Martingales are defined with respect to specific filtrations. The filtration describes how information is revealed over time.