# Martingale Theory for Finance

Tusheng Zhang

September 25, 2015

## 1 Introduction

In this course, we will introduce the basic theory of martingales, which is a branch of modern probability and is a part of the mathematical foundations for the modern theory of finance. We will see later that martingales are mathematical models for fair games. The course consists of three parts.

1. Probability and Integration.
2. Martingales.
3. Applications in finance.

## 2 Probability spaces and $\sigma$-fields

A probability space is a mathematical model for an experiment involving random variations. For example, toss a coin. We never know the exact outcome in advance. More precisely, we have the following definition:

**Definition 2.1** *A probability space is a triple $(\Omega, \mathcal{F}, P)$, where*

*(a) the set $\Omega$ is called the sample space consisting of all possible outcomes of the random experiment,*

*(b) $\mathcal{F}$ is a collection of events ( subsets of $\Omega$ ), called a $\sigma$-field, which satisfies the following properties: (1). $\Omega \in \mathcal{F}$, (2). if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, and (3) if $A_1, A_2, ..., A_n, ...$ is a sequence of events belonging to $\mathcal{F}$ then*

$$\cup_{n=1}^{\infty} A_n = \{\omega : \omega \in A_n \quad \text{for some} \quad n \geq 1\}$$

*and*

$$\cap_{n=1}^{\infty} A_n = \{\omega : \omega \in A_n \quad \text{for all} \quad n \geq 1\}$$

*also belong to $\mathcal{F}$.*

*(c) A probability measure $P$ which assigns to each event $A$ a number $P(A)$ (the probability of $A$) in such a way that*

*(i) $0 \leq P(A) \leq 1$ for any event $A$,*

*(ii) $P(\Omega) = 1$.*

*(iii) If $A_1, A_2, ..., A_n, ...$ is a sequence of disjoint (mutually exclusive ) events, then*

$$P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n) = P(A_1) + P(A_2) + ...$$

**Example 2.2** *Toss a coin two times.*

In this case,

$$\Omega = \{HH, HT, TT, TH\}$$

$\mathcal{F}$ is the collection of all possible subsets of $\Omega$. $P(A)$ is the probability of $A$. If $A = \{HT, TH\}$, then $P(A) = \frac{1}{2}$.

**Example 2.3** *Choose a number randomly from the interval* $[0, 1]$.

In this case, $\Omega = [0, 1]$. The basic events are intervals $[a, b]$, $0 \leq a \leq b \leq 1$. Choose $\mathcal{F}$ so that it contains intervals. For the probability $P$ , we set

$$P([a, b]) = \text{length of} \quad [a, b] = b - a$$

This means that the chance that we select a number lying between $a$ and $b$ is proportional to the length of the interval. The following are some of the properties of a probability measure.

**Proposition 2.4** *If* $B_1, B_2, ..., B_n, ...$ *is a sequence of events, then*

$$P(\cup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} P(B_n)$$

**Proof**. Let $A_1 = B_1$, $A_2 = B_2 \setminus B_1$, $A_3 = B_3 \setminus (B_1 \cup B_2)$, ..., $A_n = B_n \setminus (\cup_{i=1}^{n-1} B_i)$,... Then, $A_n \subset B_n$, so $P(A_n) \leq P(B_n)$. Moreover, $\cup_{i=1}^{\infty} B_i = \cup_{i=1}^{\infty} A_i$ and $A_i$ are disjoint. Thus,

$$P(\cup_{n=1}^{\infty} B_n) = P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n) \leq \sum_{n=1}^{\infty} P(B_n).$$

We write $B_n \uparrow B$ if $B_1 \subset B_2 \subset B_3 ...$ and $B = \cup_{n=1}^{\infty} B_n$. Likewise, write $B_n \downarrow B$ if $B_1 \supset B_2 \supset B_3 ...$ and $B = \cap_{n=1}^{\infty} B_n$.

**Proposition 2.5** *If* $B_n \uparrow B (B_n \downarrow B)$*, then* $P(B_n) \uparrow P(B) (P(B_n) \downarrow P(B))$.

**Proof**. Assume $B_n \uparrow B$. Introduce $A_1 = B_1$, $A_2 = B_2 \setminus B_1$, ..., $A_n = B_n \setminus B_{n-1}$,...Then, $B_n = \cup_{k=1}^{n} A_k$ and

$$\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n = B$$

Since $A_i' s$ are disjoint, we have

$$P(B) = P(\cup_{n=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$$

Hence,

$$\sum_{k=1}^{n} P(A_k) \to P(B)$$

But,

$$\sum_{k=1}^{n} P(A_k) = \sum_{k=1}^{n} [P(B_k) - P(B_{k-1})] = P(B_n)$$

The result follows.

**Definition 2.6** *A statement is said to be true almost surely if it holds with probability one (full probability ).*

**$\sigma$-fields**.

**Definition 2.7**

A collection $\mathcal{F}$ of events ( subsets of $\Omega$ )(information ) is called a $\sigma$-field if it satisfies the following properties:

(1). $\Omega \in \mathcal{F}$,

(2). if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$,

(3). if $A_1, A_2, ..., A_n, ...$ is a sequence of events belonging to $\mathcal{F}$, then $\cup_{n=1}^{\infty} A_n$ also belong to $\mathcal{F}$.

**Examples**.

(i). $\mathcal{F} = \{\Omega, \emptyset\}$ is the smallest $\sigma$-field.

(ii). $\mathcal{F} = \{\Omega, A, A^c, \emptyset\}$ is a $\sigma$-field for any $A \subset \Omega$.

(iii). The collection of all subsets of $\Omega$ is the biggest $\sigma$-field.

**Definition 2.8** *For any collection $\mathcal{D}$ of subsets of $\Omega$, the smallest $\sigma$-field $\mathcal{G}$ that contains $\mathcal{D}$ is called the $\sigma$-field generated by $\mathcal{D}$.*

We write $\mathcal{G} = \sigma(\mathcal{D})$. In fact, $\mathcal{G}$ is the intersection of all $\sigma$-fields that contain $\mathcal{D}$.

For example, if $\mathcal{G} = \{A\}$, then $\mathcal{G} = \sigma(\mathcal{D}) = \{\Omega, A, A^c, \emptyset\}$.

# 3 Integration with respect to a probability measure.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Roughly speaking, random variables are functions of outcomes.

**Definition 3.1** *A function $X(\omega) : \Omega \to R$ is said to be $\mathcal{F}$- measurable (random variable, or $\mathcal{F}$-determined) if the events of the form $\{\omega; a < X(\omega) < b\}$ belong to $\mathcal{F}$ for all choices of $a, b$.*

In this section we will define the integral of a random variable $X$ against a probability measure, denoted by $\int_{\Omega} X(\omega) dP$. This will be carried out in three steps.

Step1. Integral of a non-negative discrete random variable.

Let $X$ be a non-negative discrete r.v. with values $x_1, x_2, ..., x_n, ....$ Let $A_i = \{\omega; X(\omega) = x_i\}$. Then

$$X(\omega) = \sum_{i=1}^{\infty} x_i I_{A_i}(\omega)$$

Define

$$\int_{\Omega} X(\omega)dP = E[X] = \sum_{i=1}^{\infty} x_i P(A_i) = \sum_{i=1}^{\infty} x_i P(X = x_i) \tag{3.1}$$

$\int_{\Omega} X(\omega)dP$ is called the integral of $X$ with respect to $P$.

Step 2. Integral of a non-negative random variable. Let $X$ be a non-negative r.v. We will introduce a sequence of non-negative discrete random variables $X_n$ that increases to $X$. For $n \geq 1$, set

$$A_{n,i} = \{\omega; \frac{i}{2^n} \leq X(\omega) < \frac{i+1}{2^n}\}, i = 0, 1, 2, ...$$

Then $A_{n,i}$ forms a partition of the sample space $\Omega$, i.e.,

$$\Omega = \cup_{i=0}^{\infty} A_{n,i}$$

Define the $n$-th approximation $X_n$ of $X$ by

$$X_n(\omega) = \frac{i}{2^n} \qquad \text{if} \quad \omega \in A_{n,i}, i = 0, 1, ...$$

so that

$$X(\omega) - \frac{1}{2^n} \leq X_n^*(\omega) \leq X(\omega)$$

for all $n \geq 1$. Thus, for every $\omega$, $X_n^*(\omega) \to X(\omega)$ as $n \to \infty$. Since $A_{n,i} = A_{n+1,2i} \cup A_{n+1,2i+1}$, it holds that for $\omega \in A_{n,i}$,

$$X_{n+1}^*(\omega) = \begin{cases} \frac{2i}{2^{n+1}} = X_n^*(\omega), & \text{if } \omega \in A_{n+1,2i} \\ \frac{2i+1}{2^{n+1}} > X_n^*(\omega), & \text{if } \omega \in A_{n+1,2i+1} \end{cases}$$

Hence, $X_{n+1}^*(\omega) \geq X_n^*(\omega)$. Since $X_n^*(\omega)$ is discrete, $E(X_n^*) = \int_{\Omega} X_n^*(\omega)dP$ is already defined as above. Now we define

$$\int_{\Omega} X(\omega)dP = \lim_{n \to} \int_{\Omega} X_n^*(\omega)dP \tag{3.2}$$

Note that $\int_{\Omega} X(\omega)dP = \infty$ is allowed in the definition.

Step 3. The general case.

In this step, we extend the definition to general random variables which could take both positive and negative values. The idea is to write a random variable as a difference of two non-negative random variables. For a random variable $X(\omega)$, define the positive part of $X$ as

$$X^+(\omega) = \begin{cases} X(\omega), & \text{if } X(\omega) > 0 \\ 0, & \text{if } X(\omega) \le 0 \end{cases}$$

and the negative part of $X$ similarly as

$$X^-(\omega) = \begin{cases} -X(\omega), & \text{if } X(\omega) < 0 \\ 0, & \text{if } X(\omega) \ge 0 \end{cases}$$

Then it is easy to see that

$$X(\omega) = X^+(\omega) - X^-(\omega), \quad |X(\omega)| = X^+(\omega) + X^-(\omega)$$

If at least one of $\int_\Omega X^+(\omega)dP$ and $\int_\Omega X^-(\omega)dP$ is finite, then we define

$$\int_\Omega X(\omega)dP = \int_\Omega X^+(\omega)dP - \int_\Omega X^-(\omega)dP \tag{3.3}$$

**Definition 3.2** *$X$ is said to be integrable with respect to $P$ if both $\int_\Omega X^+(\omega)dP$ and $\int_\Omega X^-(\omega)dP$ are finite.*

It follows immediately that $X$ is integrable if and only if $\int_\Omega |X|(\omega)dP < \infty$. For $A \in \mathcal{F}$, the indicator is defined as

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \in A^c \end{cases}$$

From the definition, we have

$$\int_\Omega I_A(\omega)dP = E(I_A) = 0 \times P(I_A = 0) + 1 \times P(I_A = 1) = P(A)$$

In general, the integral coincides with the expectation.

In the sequel, we will use $\int_A X(\omega)dP$ to denote the integral of $X$ on the event $A$, that is,

$$\int_A X(\omega)dP = \int_\Omega X(\omega) \cdot I_A(\omega)dP \tag{3.4}$$

In the following theorem, we will list some elementary properties of the integral:

**Theorem 3.3** *1. If $X, Y$ are integrable, then $aX + bY$ is integrable and*

$$\int_\Omega \{aX(\omega) + bY(\omega)\}dP = a\int_\Omega X(\omega)dP + b\int_\Omega Y(\omega)dP,$$

*where $a, b$ are constants.*

2. $|\int_\Omega X(\omega)dP| \le \int_\Omega |X|(\omega)dP$.
3. *If* $X \ge Y$, *then* $\int_\Omega X(\omega)dP \ge \int_\Omega Y(\omega)dP$.
4. *If* $X \ge 0$, *then* $\int_\Omega X(\omega)dP \ge 0$ *and if moreover* $\int_\Omega X(\omega)dP = 0$, *then* $X(\omega) = 0$ *almost surely.*
5. *If* $|X| \le Y$ *and* $Y$ *is integrable, then* $X$ *is integrable.*

These properties follows directly from the definition.

Question: Suppose $X_n(\omega) \to X(\omega)$ for each $\omega$. Is it true that $E[X_n] = \int_\Omega X_n(\omega)dP \to E[X] = \int_\Omega X(\omega)dP$ ?

The answer is no in general. Here is an example.

**Example 3.4** *Let* $\Omega = [0,1]$ *and* $P$ *be the generalized length. Define*

$$X_n(t) = \begin{cases} n, & \frac{1}{n} \le t \le \frac{2}{n}, \\ 0 & otherwise. \end{cases}$$

*Then* $X_n(t) \to X(t) = 0$ *for every* $t \in [0,1]$. *However,*

$$\int_\Omega X_n(\omega)dP = 1 \not\to \int_\Omega X(\omega)dP = 0$$

But there are many cases where the question above has a positive answer. Here are some theorems.

**Theorem 3.5** *(Monotone Convergence Theorem) Suppose* $X_n$ *are r.v.'s with* $0 \le X_1(\omega) \le X_2(\omega) \le X_3(\omega)...$ *and* $X(\omega) = \lim_{n\to\infty} X_n(\omega)$. *Then*

$$\lim_{n\to\infty} \int_\Omega X_n(\omega)dP = \int_\Omega X(\omega)dP$$

**Example 3.6** *Let* $\Omega = [0,1]$ *and* $P$ *be the generalized length. Define*

$$X_n(t) = t - \frac{cos(t)}{1+3n^2}$$

*Then* $X_n(t) \uparrow X(t) = t$ *for every* $t \in [0,1]$. *By the Monotone Convergence Theorem we conclude that*

$$\int_0^1 X_n(t)dt \to \int_0^1 X(t)dt = \frac{1}{2}$$

**Theorem 3.7** *(Dominated Convergence Theorem) Let* $X_n$ *are r.v.'s such that* $X(\omega) = \lim_{n\to\infty} X_n(\omega)$. *Suppose there is a fixed integrable r.v.* $Y$ *such that* $|X_n(\omega)| \le Y(\omega)$ *for all* $n \ge 1$. *Then*

$$\lim_{n\to\infty} \int_\Omega X_n(\omega)dP = \int_\Omega X(\omega)dP$$

**Example 3.8** *Let $\Omega = [0, 1]$ and $P$ be the generalized length. Define*

$$X_n(t) = 1 - \frac{sin(nt)}{n}$$

*Then $X_n(t) \to X(t) = 1$ for every $t \in [0, 1]$. Furthermore, $|X_n(\omega)| \leq 2$ for all $n \geq 1$ and $t \in \Omega$. Applying the Dominated Convergence Theorem we conclude that*

$$\int_0^1 X_n(t)dt \to \int_0^1 X(t)dt = 1$$

# 4 Conditional expectation.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\mathcal{G}$ be a $\sigma$-field( collection of events, information). Given two events $A$, $B$, we know how to define the conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Now given an integrable random variable $X(\omega)$. Next we will define the conditional expectation of $X$ given $\mathcal{G}$, denoted by $E[X|\mathcal{G}]$, which is another random variable regarded as the best estimate of $X$ based on the information provided by $\mathcal{G}$.

**Definition 4.1** *A random variable $Y(\omega)$ ia called the conditional expectation of $X$ given $\mathcal{G}$, written as $Y = E[X|\mathcal{G}]$, if*
*(i) $Y$ is $\mathcal{G}$-measurable (determined by $\mathcal{G}$),*
*(ii) for any event $A \in \mathcal{G}$,*

$$E[YI_A] = \int_A Y(\omega)dP = \int_A X(\omega)dP = E[XI_A] \tag{4.1}$$

*This means that the two random variables $X$, $Y$, have the same average over any event in $\mathcal{G}$.*

The following proposition collects some of the important properties of the conditional expectation.

**Proposition 4.2** *(1). $E[X_1 + X_2|\mathcal{G}] = E[X_1|\mathcal{G}] + E[X_2|\mathcal{G}]$.*
*(2).$E[cX|\mathcal{G}] = cE[X|\mathcal{G}]$.*
*(3).$E[E[X|\mathcal{G}]] = E[X]$.*
*(4). If $Z$ is a random variable determined by $\mathcal{G}$ ($\mathcal{G}$-measurable ), then*

$$E[ZX|\mathcal{G}] = ZE[X|\mathcal{G}].$$

*In particular,*

$$E[Z|\mathcal{G}] = Z.$$

*(5) If $X$ is independent of $\mathcal{G}$, then*

$$E[X|\mathcal{G}] = E[X]$$

*(6). If $\mathcal{G}_1$ is another $\sigma$-field such that $\mathcal{G}_1 \subset \mathcal{G}$, then*

$$E[E[X|\mathcal{G}]|\mathcal{G}_1] = E[X|\mathcal{G}_1]$$

**Proof**. Let us prove (3), (5) and (6).

(3). Let $Y = E[X|\mathcal{G}]$. Since $\Omega \in \mathcal{G}$, we have

$$E[Y] = \int_\Omega Y(\omega)dP = \int_\Omega X(\omega)dP = E[X]$$

(5). To show that $E[X|\mathcal{G}] = E[X]$, we need to check the two conditions in the definition.

(i) $Y = E[X]$ is a constant, which is certainly $\mathcal{G}$-measurable (determined by $\mathcal{G}$).

(ii) For any $A \in \mathcal{G}$, we need to show that $\int_A Y dP = \int_A X dP$. Since $I_A$ is independent of $X$, it follows that

$$\int_A Y dP = E[X]\int_A dP = E[X]E[I_A] = E[XI_A] = \int_A X(\omega)dP$$

(6). Let $Z = E[X|\mathcal{G}]$, $Y = E[X|\mathcal{G}_1]$. We need to show $Y = E[Z|\mathcal{G}_1]$. We check the two conditions. By the definition of the conditional expectation, $Y$ is $\mathcal{G}_1$ measurable. For any $A \in \mathcal{G}_1 \subset \mathcal{G}$, by the definition, $\int_A Z dP = \int_A X dP$, $\int_A Y dP = \int_A X dP$. Hence, $\int_A Z dP = \int_A Y dP$. The condition (2) is satisfied.