STOCHASTIC CALCULUS

MATH60646 LECTURE NOTES

*SHIHAO (OWEN) TONG**

HEC Montreal

CONTENTS

For more visit shttong.com.

*tongshihaoowne@outlook.com

1 PROBABILITY SPACE AND MEASURE

1.1 DEFINITION. The sample space $\Omega = \{\omega_i : i \in \mathcal{I}\}$ is the set of all possible outcomes of a random experiment. Let $\mathcal{I}$ be a set of indices. For example, $\mathcal{I} = \{0, 1, 2, \ldots, T\}, \mathcal{I} = \{0, 1, 2, \ldots\}, \mathcal{I} = [0, \infty)$, etc.

A event is a subset of the sample space. Notice for example we write an event

$$A = \{\omega_1, \omega_2\}$$

it means ω_1 **or** ω_2 but not and.

1.2 DEFINITION. (Probability Measure) $\mathbb{P}$ is a probability measure on the space Ω if :

- $\mathbb{P}(\Omega) = 1$
- For any event A in Ω, $0 \leq \mathbb{P}(A) \leq 1$
- For any mutually disjoint events $A_1, A_2, \ldots$,

$$\mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) = \sum_{i\geq 1} \mathbb{P}(A_i)$$

  where two events A_i and A_j are disjoint if $A_i \cap A_j = \emptyset$.

Notice, disjoint is exactly the same as mutually exclusive (as one happens, the other doesn't happen for certain). The measure $\mathbb{P}$ is a set function (e.g take a set, which is a event as input) to $\mathbb{R}$. Actually in measure theory, the measure $0 \leq \mathbb{P}(A) \leq \infty$. Given a sample space, the probability measure is not unique.

1.3 THEOREM. *When* $\mathrm{Card}(\Omega) < \infty$, *let's say* $\Omega = \{\omega_1, \ldots, \omega_n\}$ *then the three conditions in definition (1.2) of the partial definition of a probability measure are equivalent to the following three conditions:*

- $\forall i \in \{1, \ldots, n\}, \mathbb{P}(\omega_i) \geq 0$
- *For any event* $A \subseteq \Omega, \mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$.
- $\sum_{i=1}^{n} \mathbb{P}(\omega_i) = 1$

This is rather useless theorem imo. Everything is trivial the slides even made stupid mistake stating $A \in \Omega$. Then we use the above axioms to prove several basic probability laws like $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. The proof for

"If $A \subseteq B \subseteq \Omega$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$" is as follows

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}(A \cup (B \backslash A)) \\ &= \mathbb{P}\left(A \cup (B \cap A^c)\right) \\ &= \mathbb{P}(A) + \mathbb{P}\left(B \cap A^c\right)\end{aligned}$$

1.4 DEFINITION. A random variable X is defined as a map from sample space to the real line

$$X : \Omega \longrightarrow \mathbb{R}$$

The Ω can be a set of anything not necessarily a number.

1.5 DEFINITION. (Support of a random variable) Let X be a random variable or vector. The support of X of that of its distribution is the set of all x s.t $\forall \delta > 0$,

$$\mathbb{P}\{X \in (x - \delta, x + \delta)\} > 0$$

To differentiate discrete and continuous r.v. we look at the support. Discrete r.v. has a countable support (i.e. finite or countably infinite) while continuous random variables has uncountable infinite support.

From now on we focus on discrete r.v. first.

1.6 DEFINITION. The distribution or the law of a random variable X is characterized by its (cumulative) distribution function $F_X : \mathbb{R} \rightarrow [0,1]$ $x \rightarrow F(x)$. So, if $\mathbb{P}$ is the probability measure assigned to Ω then

$$\forall x \in \mathbb{R}, \quad F_X(x) = \mathbb{P}\Big(\{\omega \in \Omega \mid X(\omega) \leq x\}\Big)$$

Now if we reduce the Ω to a finite (e.g. $\text{card}(\Omega) < \infty$) then we could also use the pmf to characterize the distribution function (e.g. characterize we mean uniquely determine the distribution). For continuoius r.v CDF but not PDF pin down the distribution while for discrete r.v both CDF and PMF works. The following theorem shows this

1.7 THEOREM. *In the case where Card $(\Omega) < \infty$, the distribution of a random variable is also characterized by its probability mass function $f_X : \mathbb{R} \rightarrow [0,1]$ that is*

$$\forall x \in \mathbb{R}, f_X(x) = \mathbb{P}\{\omega \in \Omega \mid X(\omega) = x\}$$

Proof. The logic is to show that

$$F_X(x) \iff f_X(x)$$

given one then the other is uniquely determined. Detail see slides 1 appendix. □

1.8 DEFINITION. Two random variables X and Y are said to be equal if and only if

$$\forall \omega \in \Omega, X(\omega) = Y(\omega)$$

They are said to be equal in distribution (or in law) when they have the same distribution.

Clearly by theorem 1.7, if the r.v. is discrete, then equal in distribution can be equal in either PDF or CDF. Random variables are equal $\implies$ Equal in distribution but NOT vice versa. The first statement is stronger. Equal in distribution depends on the assigned probability measure while equal of r.v. does not depends on probability measure. Two r.v. are equal means the equality is point-wisely in the sample space.

1.9 DEFINITION. ($\sigma-$ Algebra) Let $\mathcal{F}$ be a collection of subsets of a sample space $\Omega.\mathcal{F}$ is called a σ-field (or σ-algebra) if and only if it has the following properties.

- The empty set $\emptyset \in \mathcal{F}$.
- If $A \in \mathcal{F}$, then the complement $A^c \in \mathcal{F}$.
- If $A_i \in \mathcal{F}, i = 1, 2, \ldots$, then their union $\cup A_i \in \mathcal{F}$.

Notice the union must be countable. If further Ω is finite, then the third condition is the same as

$$\bigcup_{i}^{n} A_i \in \mathcal{F}$$

since there have to be finitely many events. σ-algebra is a collection of events.

1.10 DEFINITION. (Partition) A family $\mathcal{P} = \{A_1, \ldots, A_n\}$ of events in Ω is called a finite partition of Ω if

- $\forall i \in \{1, \ldots, n\}, A_i \neq \varnothing$
- $\forall i, j \in \{1, \ldots, n\}$ such that $i \neq j, A_i \cap A_j = \varnothing$
- $\bigcup_{i=1}^{n} A_i = \Omega$

In short, it is the union of disjoint non-empty events that span the whole sample space.

1.11 DEFINITION. A σ-algebra $\mathcal{F}$ is said to be generated from the finite partition $\mathcal{P}$ if it is the smallest σ-algebra that contains all the elements of $\mathcal{P}$. In that case $\mathcal{F}$ is denoted $\sigma(\mathcal{P})$ and the elements of $\mathcal{P}$ are the atoms of $\mathcal{F}$.

All σ-algebra have atoms.

1.12 DEFINITION. The pair $(\Omega, \mathcal{F})$ is called the measurable space.

1.13 DEFINITION. (r.v. on measurable space) A random variable X constructed on the measurable space $(\Omega, \mathcal{F})$, is a real-valued function $X : \Omega \to \mathbb{R}$ such that

$$\forall x \in \mathbb{R}, \left\{\omega \in \Omega : X(\omega) \leq x\right\} \in \mathcal{F}$$

It can be shown that for those Ω s.t. $\text{Card}(\Omega) < \infty$ the definition is equivalent to

$$\forall x \in \mathbb{R}, \{\omega \in \Omega \mid X(\omega) = x\} \in \mathcal{F}$$

The random variable X is also called $\mathcal{F}$-measurable.

Notice the $\{\omega \in \Omega : X(\omega) \leq x\}$ is an event so the notation is correct (e.g. it is a set of outcomes from the sample space). The tentative proof is as follows

Proof. Let We show the $\implies$ first. Let $\{X_{(1)}, X_{(2)}, \cdots, X_{(n)}\}$ be in ascending order. Then $\forall\, x \in \mathbf{R}, i = 1, 2, ..., n$

$$\begin{aligned}\{\omega \in \Omega \,|\, X(\omega) \leq x\} &= \{\omega \in \Omega \,|\, X(\omega) \leq X_{(i)}\} \\ &= \{\omega \in \Omega \,|\, X(\omega) = X_{(1)} \text{ or } X(\omega) = X_{(2)} \text{ or } \cdots X(\omega) = X_{(i)}\}\end{aligned}$$

Since the condition is "or" so

$$\{\omega \in \Omega \,|\, X(\omega) = X_{(i)}\} \in \mathcal{F}$$

For those x s.t. $\{\omega \in \Omega \,|\, X(\omega) = x\} = \emptyset \in \mathcal{F}$ by the definition of σ-algebra. □

1.14 THEOREM. *Let $(\Omega, \mathcal{F})$, Card $(\Omega) < \infty$, be a measurable space and $\mathcal{P} = \{A_1, \ldots, A_n\}$, be the finite partition of Ω that generates $\mathcal{F}$. The function $X : \Omega \to \mathbb{R}$ is a random variable on that space (X is $\mathcal{F}$-measurable) if and only if X is constant on the atoms of $\mathcal{F}$.*

Constants can differ from events to events as long as $X(\omega)$ does not vary for $\omega \in A_i$. Proof see Chapter 1 slides page 42. One immediate result from theorem 1.14 is that for the r.v. X on the trivial σ-algebra $\{\Omega, \emptyset\}$ is constant since the only atoms is Ω (or $\emptyset$). By theorem 1.14 the r.v. can only takes 1 value.

1.15 COROLLARY. *If $\mathcal{F}$ = the set of all possible events in Ω (the power set of Ω) then any real-valued function on $\Omega (X : \Omega \to \mathbb{R})$ is $\mathcal{F}$-measurable, that is to say that X is a random variable on $(\Omega, \mathcal{F})$.*

1.16 DEFINITION. (σ-algebra generated by random variable) Let $X : \Omega \to \mathbb{R}$. The smallest σ-algebra $\mathcal{F}$ that make X a random variable on the measurable space $(\Omega, \mathcal{F})$ is called the σ-algebra generated by X and is denoted $\sigma(X)$.

This is simply the reverse order of defining an $\mathcal{F}$-measurable r.v. . Consider the finite case or Ω. Let $X(\omega) \in \{x_1, x_2, ..., x_n\}$. Then we define

$$A_i = \{\omega \in \Omega \,|\, X(\omega) = x_i\}$$

Notice $\{A_1, ..., A_n\}$ form a finite partition of Ω. We call this the "events characterizing the random variable". Then by the routine

$$\sigma(X) = \sigma(\mathcal{P}_X) = \mathcal{F}_X$$

1.17 THEOREM. *Let's assume that $\text{Card}(\Omega) < \infty$. If $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ are $\mathcal{F}$-measurable, then $\forall a, b \in \mathbb{R}$, $aX + bY$ is also $\mathcal{F}$-measurable, which is to say that any linear combination of random variables built on the same measurable space is a random variable of that space.*

Proof. Since $\text{Card}(\Omega) < \infty$, the random variables X and Y can only take a finite number of values, let's say $x_1 < \ldots < x_m$ and $y_1 < \ldots < y_n$ respectively. $\forall z \in \mathbb{R}$,

$$\begin{aligned}
&\{\omega \in \Omega \mid aX(\omega) + bY(\omega) \leq z\} \\
&= \bigcup_{ax_i + by_j \leq z} \left\{\omega \in \Omega \mid X(\omega) = x_i \text{ and } Y(\omega) = y_j\right\} \\
&= \bigcup_{ax_i + by_j \leq z} \underbrace{\{\omega \in \Omega \mid X(\omega) = x_i\}}_{\in \mathcal{F}} \cap \underbrace{\left\{\omega \in \Omega \mid Y(\omega) = y_j\right\}}_{\in \mathcal{F}} \in \mathcal{F}.
\end{aligned}$$

□

1.18 DEFINITION. (Probability Measure) Let $(\Omega, \mathcal{F})$ be a measurable space. The function $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure on $(\Omega, \mathcal{F})$ if

- $\mathbb{P}(\Omega) = 1$.
- $\forall A \in \mathcal{F}, 0 \leq \mathbb{P}(A) \leq 1$.

- $\forall A_1, A_2, \ldots \in \mathcal{F}$ such that $A_i \cap A_j = \varnothing$ if $i \neq j$,

$$\mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) = \sum_{i\geq 1} \mathbb{P}(A_i)$$

1.1 Practice

Here are some good practices.

1. Let $\mathrm{Card}(\Omega) < \infty$. Show that if $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ are $\mathcal{F}$-measurable then $\min\{X, Y\}, \max\{X, Y\}$ and XY are also $\mathcal{F}$-measurable. In addition, show that if $\forall \omega \in \Omega, Y(\omega) \neq 0$ then X/Y is $\mathcal{F}$-measurable.

   Proof. Here we only show the proof for XY. The others are trivial and in a similar manner. Let us show that XY is $\mathcal{F}$-measurable : $\forall z \in \mathbb{R}$,

$$\begin{aligned}
&\{\omega \in \Omega \mid X(\omega)Y(\omega) \leq z\} \\
&= \bigcup_{x_i y_j \leq z} \left\{\omega \in \Omega \mid X(\omega) = x_i \text{ and } Y(\omega) = y_j\right\} \\
&= \bigcup_{x_i y_j \leq z} \underbrace{\{\omega \in \Omega \mid X(\omega) = x_i\}}_{\in \mathcal{F}} \cap \underbrace{\left\{\omega \in \Omega \mid Y(\omega) = y_j\right\}}_{\in \mathcal{F}} \in \mathcal{F}.
\end{aligned}$$

   □

   The takeaway point is to transfer the conditions to the union. This works because there are countably many combination of X and Y (they are discrete so finitely many values to take).

2 STOCHASTIC PROCESS

2.1 Definition. (Stochastic Process) Let $(\Omega, \mathcal{F})$ be a measurable space. A stochastic process

$$X = \{X_t : t \in \mathcal{T}\}$$

is a family of random variable, all built on the same measurable space $(\Omega, \mathcal{F})$.

2.2 Definition. (Filtration) A family $\mathbb{F} = \{\mathcal{F}_t : t \in \mathcal{T}\}$ of σ-algebras on Ω is a filtration on the measurable space $(\Omega, \mathcal{F})$ if

- $\forall t \in \mathcal{T}, \mathcal{F}_t \subseteq \mathcal{F}$,

- $\forall t_1, t_2 \in \mathcal{T}$ such that $t_1 \leq t_2, \mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$.

Filtration is also called the information set. It contains all the information revealed by time t. Usually the $\mathcal{F}_0 = \{\emptyset, \Omega\}$. If we say the filtration $\mathbb{F}$ is generated by the stochastic process $\{X_t : t = 1, \cdots, n\}$, this means $\forall t \in \mathcal{T}$,

$$\mathcal{F}_t = \sigma(X_1, X_2, \cdots, X_t)$$

2.3 DEFINITION. A stochastic process $X = \{X_t : t \in \mathcal{T}\}$ is said to be adapted to the filtration $\mathbb{F} = \{\mathcal{F}_t : t \in \mathcal{T}\}$ if

$$\forall t \in \mathcal{T}, X_t \text{ is } \mathcal{F}_t\text{-measurable.}$$

2.4 DEFINITION. (Stopping time) Let $(\Omega, \mathcal{F})$ be a measurable space such that $\text{Card}(\Omega) < \infty$ and equipped with the filtration $\mathbb{F} = \{\mathcal{F}_t : t \in \{0, 1, \ldots\}\}$. A stopping time τ is a $(\Omega, \mathcal{F})$-random variable that takes its values in $\{0, 1, \ldots\}$ such that

$$\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t, \quad \forall\ t \in \{0, 1, \ldots\}$$

For countable set of Ω the definition is equivalent to

$$\{\omega \in \Omega : \tau(\omega) = t\} \in \mathcal{F}_t, \quad \forall\ t \in \{0, 1, \ldots\}$$

Consider stopping time is a decision or strategy to buy or sell a stock. Recall the filtration can be viewed as an information set up to time t. The take away point is that if the decision is made based on the future information that yet to happen, then it is not a stopping time (e.g. sell the stock as soon as it make a profit. This requires knowledge of the price at $t + 1$ while we know only price up to time t). This is reflected in the definition that the event $\{\omega \in \Omega : \tau(\omega) \leq t\}$ must be contained in the information set. It is reasonable to consider the stopping to be exogeneous.

2.5 THEOREM. *Given* $(\Omega, \mathcal{F})$ *and a corresponding filtration* $\mathbb{F}$*, let* τ_1 *and* τ_2 *are stopping time based on the same* $\mathbb{F}$*. Then*

$$\tau_1 \wedge \tau_2 \equiv \min\{\tau_1, \tau_2\}, \quad \tau_1 \vee \tau_2 \equiv \max\{\tau_1, \tau_2\}$$

are both stopping time as well.

Proof is trivial. Meaningless to practice.

2.6 DEFINITION. (First passage/Hitting time) Let $(\Omega, \mathcal{F})$ be a measurable space such that $\text{Card}(\Omega) < \infty$ and equipped with the filtration $\mathbb{F} = \{\mathcal{F}_t : t \in \{0, 1, \ldots\}\}$.

$X = \{X_t : t \in \{0, 1, \ldots\}\}$ represents a stochastic process adapted to that filtration. Let $B \subseteq \mathbb{R}$. We define the time until the stochastic process X first enters the set B as

$$\tau_B(\omega) = \min\{t \in \{0, 1, \ldots\} : X_t(\omega) \in B\}.$$

If it happened that the path $t \to X_t(\omega)$ never hits the set B then we define $\tau_B(\omega) = \infty$.

2.7 THEOREM. *The random variable τ_B is a stopping time.*

Proof. Since $\text{Card}(\Omega) < \infty$, then $\forall t \in \{0, 1, \ldots\}$, X_t can only take a finite number of values. Let's denote them by

$$x_1^{(t)} < \ldots < x_{m_t}^{(t)}.$$

$\forall t \in \{0, 1, \ldots\}$. Actually the m_t with subscript t make sense because each r.v. X_t has different co-domain. m only can not pin down the maximum value that X_t take. Then

$$\begin{aligned}
&\{\omega \in \Omega : \tau_B(\omega) = t\} \\
&= \{\omega \in \Omega : X_0(\omega) \notin B, \ldots, X_{t-1}(\omega) \notin B, X_t(\omega) \in B\} \\
&= \left(\bigcap_{k=0}^{t-1} \{\omega \in \Omega : X_k(\omega) \notin B\}\right) \cap \{\omega \in \Omega : X_t(\omega) \in B\} \\
&= \left(\bigcap_{k=0}^{t-1} \bigcup_{x_i^{(k)} \notin B} \left\{\omega \in \Omega : X_k(\omega) = x_i^{(k)}\right\}\right) \bigcap \left(\bigcup_{x_i^{(t)} \in B} \left\{\omega \in \Omega : X_t(\omega) = x_i^{(t)}\right\}\right) \\
&\in \mathcal{F}_t
\end{aligned}$$

□

the last "in" is due to adaption and

For practical purpose. How to set up the measure framework for a real-world scenario? Usually we should consider the variable of interest as the random variable (e.g. stock price, interest rate, balance of bank account etc.) and the state of the world as the sample sample space Ω (e.g. some factors that could possibly related or specify the value of the r.v. of interest). The state of the world should be treated as deterministic otherwise it is not possible to arrange the value of random variables. Further more, to determine the evolution of r.v. by time, the value at t should also be determined by the state of the world. This could be done through a function.

3 EXPECTATION AND CONDITIONAL EXPECTATION

3.1 DEFINITION. The random variables X and Y are both built on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathrm{Card}(\Omega) < \infty$. Let $\mathcal{P}_X = \{A_1, \ldots, A_m\}$ and $\mathcal{P}_Y = \{B_1, \ldots, B_n\}$ be two finite partitions that respectively generate the sigma-algebras $\sigma(X)$ and $\sigma(Y)$. The random variables X and Y are said to be (pariwise) independent if

$$\forall A \in \mathcal{P}_X \text{ and } \forall B \in \mathcal{P}_Y, \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

By the definition, it is sufficient to just check the pairs of atoms to determine the independence. Not necessary to check every elements in $\sigma(X)$ and $\sigma(Y)$. Also notice the probability measure play a role in Independence. The definition can be extended to multiple random variables.

3.2 DEFINITION. Let $X_1, X_2, ..., X_n$ be random variables defined on the same measure space. Then they are said to be *mutually independent* if and only if

- They are pairwise independent
- The following rule is observed

$$\mathbb{P}(\bigcap_{i \in S} X_i) = \prod_{i \in S} \mathbb{P}(X_i) \ \forall S \subseteq \{1, 2, \cdots, n\}$$

Notice it is possible for r.v. to be pairwise independent but not mutually independent.

3.3 DEFINITION. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space such that $\mathrm{Card}(\Omega) < \infty$. For all event $A \in \mathcal{F}$ with a positive probability, $\mathbb{P}(A) > 0$, the conditional probability given A, denoted $\mathbb{P}(\bullet \mid A)$, is defined as

$$\forall B \in \mathcal{F}, \mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$$

Based on conditional probability, we can define another theorem for independence.

3.4 THEOREM. *The random variables* X *and* Y *are both built on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *Card* $(\Omega) < \infty$. *Let* $\mathcal{P}_X = \{A_1, \ldots, A_m\}$ *and* $\mathcal{P}_Y = \{B_1, \ldots, B_n\}$, *be two finite partitions that respectively generate the sigma-algebras* $\sigma(X)$ *and* $\sigma(Y)$. *If*

$$\forall A \in \mathcal{P}_X \ s.t \ \mathbb{P}(A) > 0, \mathbb{P}(B \mid A) = \mathbb{P}(B), \forall B \in \mathcal{P}_Y$$

or again, if

$$\forall B \in \mathcal{P}_Y \ s.t \ \mathbb{P}(B) > 0, \mathbb{P}(A \mid B) = \mathbb{P}(A), \forall A \in \mathcal{P}_X$$

then X *and* Y *are independent.*

The proof is trivial. Notice that the probability of the conditioning event is assumed to be positive. If $\exists B$ such that $\mathbb{P}(B) = 0$ we are still able to show independence by

$$0 \leq \mathbb{P}(B \cap A) \leq \mathbb{P}(B) = 0 = \mathbb{P}(A)\mathbb{P}(B)$$

3.5 DEFINITION. (Expectation) Let X be a random variable built on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\text{Card}(\Omega) < \infty$. The expectation of X, denoted $E^{\mathbb{P}}[X]$ is

$$\mathbb{E}^{\mathbb{P}}[X] = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega)$$

For discrete random variable, the equation is equivalent to

$$\mathbb{E}^{\mathbb{P}}[X] = \sum_{i=1}^{n} x_i \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) = x_i\} = \sum_{i=1}^{n} x_i f_X(x_i)$$

where x_i are the possible values that X can take, f_X is the pmf of X.

The common properties of expectation still holds under measurable space (e.g. linearity). The proof for linearity is trivial. Remember the probability measure is defined on sample space for each ω but not the sigma field. Expectation of product is product of expectation if independent. Moment.

3.6 DEFINITION. (Conditional Expectation) The random variable X is built on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, Card $(\Omega) < \infty$. Let $\mathcal{G} \subseteq \mathcal{F}$, a sigma-algebra generated by the finite partition $\mathcal{P} = \{A_1, \ldots, A_n\}$ satisfying $\forall i \in \{1, \ldots, n\}, \mathbb{P}(A_i) > 0$. The conditional expectation of X given $\mathcal{G}$, denoted $E^{\mathbb{P}}[X \mid \mathcal{G}]$ is

$$\begin{aligned}
\mathbb{E}(X \,|\, \mathcal{G})(\omega) &\equiv \sum_{i=1}^{n} \mathbb{E}[X \,|\, A_i](\omega) * \mathbb{I}_{A_i}(\omega) \\
&= \sum_{i=1}^{n} \sum_{\omega \in \Omega} X(\omega) \frac{\mathbb{P}(\omega \cap A_i)}{\mathbb{P}(A_i)} \mathbb{I}_{A_i}(\omega) \\
&= \sum_{i=1}^{n} \frac{\mathbb{I}_{A_i}(\omega)}{\mathbb{P}(A_i)} \sum_{\omega^* \in A_i} X(\omega^*)\,\mathbb{P}(\omega^*)
\end{aligned}$$

Notice the last equality is due to $\mathbb{P}(\emptyset) = 0$ by the property of (probability) measure. Thus it can be seen conditional expectation is a random variable due to undetermined events on $\mathcal{G}$: eventually we find that the expression of $\mathbb{E}(X \,|\, \mathcal{G})(\omega)$ is a function of ω thus indeed a mapping from Ω to $\mathbb{R}$.

3.7 LEMMA. *(Important properties of conditional expectation) Let* X *and* Y *be two random variables in the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. Let sigma-algebras* $\mathcal{G}, \mathcal{G}_1$ *and* $\mathcal{G}_2$ *are respectively generated by the finite partitions* $\mathcal{P} = \{A_1, \dots, A_n\}$ *and* $\mathcal{P}_1 = \{B_1, \dots, B_m\}$ *and* $\mathcal{P}_2 = \{C_1, \dots, C_n\}$*. The following properties holds:*

1. *If* X *is* $\mathcal{G}$*-measurable, then* $E^{\mathbb{P}}[X \mid \mathcal{G}](\omega) = X(\omega)$.
2. *If* $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$ *are sigma-algebras, then*
$$E^{\mathbb{P}}\left[E^{\mathbb{P}}[X \mid \mathcal{G}_1] \mid \mathcal{G}_2\right] = E^{\mathbb{P}}[X \mid \mathcal{G}_1]$$
3. *(Tower rule or Iterated expectation) If* $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$ *are sigma-algebras, then*
$$E^{\mathbb{P}}\left[E^{\mathbb{P}}[X \mid \mathcal{G}_2] \mid \mathcal{G}_1\right] = E^{\mathbb{P}}[X \mid \mathcal{G}_1]$$
4. *Conditioning on the trivial algebra*
$$E^{\mathbb{P}}[X \mid \{\emptyset, \Omega\}] = E^{\mathbb{P}}[X]$$
5. *If* X *is* $\mathcal{G}$*-measurable then*
$$E^{\mathbb{P}}\left[E^{\mathbb{P}}[X \mid \mathcal{G}]\right] = E^{\mathbb{P}}[X]$$
6. *If* Y *is* $\mathcal{G}$*-measurable, then*
$$E^{\mathbb{P}}[XY \mid \mathcal{G}] = YE^{\mathbb{P}}[X \mid \mathcal{G}]$$
7. *If* X *and* Y *are independent, then*
$$E^{\mathbb{P}}[X \mid \sigma(Y)] = E^{\mathbb{P}}[X]$$
8. $\forall a, b \in \mathbb{R}$,
$$E^{\mathbb{P}}[aX + bY \mid \mathcal{G}] = aE^{\mathbb{P}}[X \mid \mathcal{G}] + bE^{\mathbb{P}}[Y \mid \mathcal{G}]$$

Proof. Here we show the proof for properties 1,3,6 and 7 only. Notice that it is the sample space ω that is always partitioned. □

Comparing properties 1 and 4. Intuitively, the trivial sigma algebra does not provide any information thus there is no "randomness" in terms of in which event the outcome happens. Thus conditioning on the trivial algebra the expectation is a constant. However if the sigma algebra is not the trivial set then we are not sure in which event the outcome happens so consequently it is a random variable.

4.1 DEFINITION. (Martingales) On the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F}$ is the filtration $\{\mathcal{F}_t : t \in \{0, 1, 2, \ldots\}\}$, the stochastic process

$$M = \{M_t : t \in \{0, 1, 2, \ldots\}\}$$

is a discrete-time martingale if

1. $\forall t \in \{0, 1, 2, \ldots\}, E^{\mathbb{P}}[|M_t|] < \infty$
2. $\forall t \in \{0, 1, 2, \ldots\}$, M_t is $\mathcal{F}_t-$ measurable;
3. $\forall s, t \in \{0, 1, 2, \ldots\}$ such that $s < t, \mathbb{E}^{\mathbb{P}}[M_t \mid \mathcal{F}_s] = M_s$. This is equivalent to

$$\forall t \in \{1, 2, \ldots\}, E^{\mathbb{P}}[M_t \mid \mathcal{F}_{t-1}] = M_{t-1}$$

   Can be proved by induction.

Notice that the Martingale depends on both the probability measure and the filtration $\mathbb{F}$ so it is also sometimes called $(\mathbb{F}, \mathbb{P})$-martingales. To prove a stochastic process is martingale, we just need to show the above three properties are satisfied. The intuition of the 3rd property is that the best bet for the future is to take the current value as prediction. Martingale process is considered to be a "flat line" in probabilistic view. In financial asset pricing, "the martingality of an asset is equivalent to not being able to conduct arbitrage through trades in that asset".

4.2 LEMMA. *Let* $M = \{M_t : t \in \{0, 1, 2, \ldots\}\}$ *be a martingale built on the filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. *Then*

$$\forall t \in \{1, 2, \ldots\}, E^{\mathbb{P}}[M_t] = E^{\mathbb{P}}[M_0]$$

This means on average a martingale process is constant. This does not mean that the process varies a little since the variance at each time $\mathrm{Var}^{\mathbb{P}}(M_t)$ can be infinite. Also notice here the subscript t is deterministic (e.g. for any given t).

4.3 EXAMPLE. Let $\{\xi_t : t \in \{1, 2, \ldots\}\}$ be a sequence of $(\Omega, \mathcal{F})$-independent and identically distributed random variables with respect to the measure $\mathbb{P}$ and such that

$$E^{\mathbb{P}}[\xi_t] = 0 \text{ and } E^{\mathbb{P}}\left[\xi_t^2\right] < \infty.$$

Let's define

$$\mathcal{F}_0 = \{\varnothing, \Omega\}$$

$$\forall t = \{1, 2, \ldots\}, \mathcal{F}_t = \sigma\{\xi_s : s \in \{1, \ldots, t\}\}$$

and

$$M_0 = 0, M_t = \sum_{s=1}^{t} \xi_s.$$

Then show that the stochastic process M is a martingale on the space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Proof. Indeed,

$$E^{\mathbb{P}}[|M_t|] = E^{\mathbb{P}}\left[\left|\sum_{s=1}^{t} \xi_s\right|\right] \leq \sum_{s=1}^{t} E^{\mathbb{P}}[|\xi_s|] \leq \sum_{s=1}^{t} \sqrt{E^{\mathbb{P}}\left[\xi_s^2\right]} < \infty$$

where the second inequality comes from the fact that, for any random variable,

$$0 \leq \text{Var}[|X|] = E\left[|X|^2\right] - (E[|X|])^2 \Rightarrow E[|X|] \leq \sqrt{E\left[|X|^2\right]}$$

Given the selected filtration, M is adapted (*) (which is to say that $\forall t \in \{0, 1, 2, \ldots\}$, M_t is $\mathcal{F}_t$-measurable).Lastly, $\forall s, t \in \{0, 1, 2, \ldots\}$ such that $s < t$

$$\begin{aligned}
&E^{\mathbb{P}}[M_t \mid \mathcal{F}_s] \\
=&E^{\mathbb{P}}\left[M_s + \sum_{u=s+1}^{t} \xi_u \mid \mathcal{F}_s\right] \\
=&E^{\mathbb{P}}[M_s \mid \mathcal{F}_s] + \sum_{u=s+1}^{t} E^{\mathbb{P}}[\xi_u \mid \mathcal{F}_s] \\
=&M_s + \sum_{u=s+1}^{t} \underbrace{E^{\mathbb{P}}[\xi_u \mid \mathcal{F}_s]}_{=E^{\mathbb{P}}[\xi_u]} = 0
\end{aligned}$$

The last equation is because first, M_s is $\mathcal{F}$-measurable by inception and the independence due to $\{\xi_1, \xi_2, ..., \xi_s\}$ and $\{\xi_{s+1}, \cdots, \xi_t\}$. The independence then brings out all the expectation equals $M_0 = 0$. Thus it is a martingale. □

Notes: * The meaning of "adapted" is not very clear. We could consider the fact that, for random variables based on the same measurable space $(\Omega, \mathcal{F})$ then their summation, or further any basic operations like multiplication, division etc., is still $\mathcal{F}$-measurable.

4.4 DEFINITION. (Stopped process) The stochastic process X and the stopping time τ are built on the same filtered measurable space $(\Omega, \mathcal{F}, \mathbb{F})$. The stochastic process X^τ defined by

$$X_t^\tau(\omega) = X_{t \wedge \tau(\omega)}(\omega)$$

is called a stopped process with stopping time τ.

Through time, it take the value X take until stopping time. From stopping time onward, its value freezes at the value at stopping time. Notice the value of $\tau(\omega)$ is deterministic (I believe in some sense we could treat it as exogenous) while the value for t is not.

4.5 THEOREM. *If the martingale* M *and the stopping time* τ *are built on the same filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *then the stopped process* M^{τ} *is also a martingale on that space.*

Proof. Make up later. See slides 19. □

4.6 THEOREM. *(Optional Stopping Theorem) Let* $M = \{M_t \mid t = 0, 1, \cdots\}$ *be an martingale and* τ *be a stopping time both defined on the same filtered probablility space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$*. If the following conditions hold*

- $\mathbb{E}^{\mathbb{P}}(\tau) < \infty$
- $\mathbb{E}^{\mathbb{P}}(|M_{t+1} - M_t| \,\big|\, \mathcal{F}_t) < c$ *for some constant* c *and* $\forall\, t$

then $\mathbb{E}(M_\tau) = \mathbb{E}(M_0)$ *where we say* τ *is an optional stopping time.*

The version of OST shown in the slides is a bit different

4.7 THEOREM. *(Optional Stopping time) Let* $X = \{X_t \mid t = 0, 1, \cdots\}$ *be an stochastic process and* τ *be a stopping time both defined on the same filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$*. Then M is a martingale* ***IFF***

$$\mathbb{E}^{\mathbb{P}}(X_\tau) = \mathbb{E}^{\mathbb{P}}(X_0) \forall\, \tau\, s.t.\ 0 \geq \tau(\omega) \geq c$$

for some constant c (e.g. for any τ *that is bounded).*

Proof. The proof is exactly the same as in slides just with necessary comments.

- " $\implies$ ": Assume X is martingale.

$$
\begin{aligned}
E^{\mathbb{P}}[X_\tau] &= E^{\mathbb{P}}\left[\sum_{k=0}^{b} X_k \mathbb{I}_{\{\tau=k\}}\right] \\
&= E^{\mathbb{P}}\left[\sum_{k=0}^{b} X_k \left(\mathbb{I}_{\{\tau\geq k\}} - \mathbb{I}_{\{\tau\geq k+1\}}\right)\right] \\
&= \sum_{k=0}^{b} E^{\mathbb{P}}\left[X_k \mathbb{I}_{\{\tau\geq k\}}\right] - \sum_{k=0}^{b} E^{\mathbb{P}}\left[X_k \mathbb{I}_{\{\tau\geq k+1\}}\right] \\
&= E^{\mathbb{P}}[X_0] + \sum_{k=1}^{b} E^{\mathbb{P}}\left[X_k \mathbb{I}_{\{\tau\geq k\}}\right] - \sum_{k=0}^{b-1} E^{\mathbb{P}}\left[X_k \mathbb{I}_{\{\tau\geq k+1\}}\right] \\
&= E^{\mathbb{P}}[X_0] + \sum_{k=1}^{b} E^{\mathbb{P}}\left[X_k \mathbb{I}_{\{\tau\geq k\}}\right] - \sum_{k=1}^{b} E^{\mathbb{P}}\left[X_{k-1} \mathbb{I}_{\{\tau\geq k\}}\right] \\
&= E^{\mathbb{P}}[X_0] + \sum_{k=1}^{b} E^{\mathbb{P}}\left[(X_k - X_{k-1}) \mathbb{I}_{\{\tau\geq k\}}\right] \\
&= E^{\mathbb{P}}[X_0] + \sum_{k=1}^{b} E^{\mathbb{P}}\left[E^{\mathbb{P}}\left[(X_k - X_{k-1}) \mathbb{I}_{\{\tau\geq k\}} \mid \mathcal{F}_{k-1}\right]\right] \\
&= E^{\mathbb{P}}[X_0] + \sum_{k=1}^{b} E^{\mathbb{P}}\left[\mathbb{I}_{\{\tau\geq k\}} E^{\mathbb{P}}[X_k - X_{k-1} \mid \mathcal{F}_{k-1}]\right] \\
&= E^{\mathbb{P}}[X_0]
\end{aligned}
$$

□

Comments: The second last line is because $\mathbb{I}_{\tau\geq k}$ is $\mathcal{F}_{k-1}$ measurable. This can be shown as follows. Recall $\mathbb{I}_{\tau\geq k}$ is defined as

$$
\mathbb{I}_{\{\tau\geqslant k\}} = \begin{cases} 1, & \tau \geqslant k \\ 0, & \tau < k \end{cases}
$$

So appeal to the first principle we have

$$
\begin{aligned}
\left\{\omega \in \Omega \mid I_{\{\tau\geqslant k\}} = 1\right\} &= \{\omega \in \Omega \mid \tau(\omega) < k\}^c \\
&= \left\{\bigcup_{k'=0}^{k-1} \{\omega \in \Omega \mid \tau(\omega) = k'\}\right\}^c
\end{aligned}
$$

τ_i is $\mathcal{F}_{k-1}$-measurable for all $i < (k-1)$ so its union is also $\mathcal{F}_{k-1}$-measurable. So does its complement (due to sigma-field). why the last 3rd and 4th line?

Optional stopping time is used to compute the expected return or risk of a portfolio or financial assets. Also it is a proper technique to evaluate American option.

5 DISCRETE-TIME MARKET MODELS: BINOMIAL MODEL

Notation and definition: We consider two type of assets which are riskless asset and risky assets. They are represented by $S_t^{(1)}$ and $S_t^{(2)}$ respectively. We consider $t = 0, 1$. They could take values as follows.

$$\forall\omega \in \Omega, S_1^{(1)}(\omega) = S_0^{(1)}(\omega)(1+r) \equiv (1+r)$$

$$\forall\omega \in \Omega, S_0^{(2)}(\omega) = s_0 \in \mathbb{R}_+ = (0, \infty), \quad S_1^{(2)} = \begin{cases} s_{11} \in \mathbb{R}^+ \\ s_{12} \in \mathbb{R}^+ \end{cases}$$

where r is periodic interest rate (constant through out the period); $S_1^{(2)}$ can take either s_{11} or s_{12} where $s_{11} < s_{12}$. Assume $S_0^{(1)} = 1$ for all ω and $S_0^{(2)}$ is known for certain as s_0.

The measurable space is build in a back-tracking way. Notice we want to see how an asset evolves over time, so a trajectory should be a series of price at each point of time. Define the two assets in a single vector

$$\mathbf{S}_t = \left(S_t^{(1)}, S_t^{(2)}\right)', \quad t = 0, 1$$

as t changes we have a stochastic process. Then the process is

	$\left(S_0^{(1)}, S_0^{(2)}\right)^\top$	$\left(S_1^{(1)}, S_1^{(2)}\right)^\top$
trajectoire #1(ω_1)	$(1, s_0)^\top$	$(1+r, s_{11})^\top$
trajectoire #2(ω_2)	$(1, s_0)^\top$	$(1+r, s_{12})^\top$

Now each trajectory is equivalent to a $\omega \in \Omega$. This is how we inversely generate the sample space. In practice, the sample space is not known as well. Based on our assumption (e.g. s_{11}, s_{22} etc.) there can be more than two ω in Ω while they will be absorbed into either of the two trajectory which is

$$\exists \omega_1, \omega_2 \in \Omega \text{ such that } \forall t \in \{0, 1\} \mathbf{S}_t(\omega_1) = \mathbf{S}_t(\omega_2)$$

In other words we are not able to specify them. The set Ω is not unique in this sense. Thus the $\mathcal{F} = \mathcal{F}_1 = \{\emptyset, \Omega, \sigma_1, \sigma_2\}$.

5.1 DEFINITION. (Portfolio) The pair

$$\boldsymbol{\phi} = (\phi_1, \phi_2) \in \mathbb{R} \times \mathbb{R}$$

is called a portfolio. The value of a portfolio at state ω at time t is then

$$V_{\boldsymbol{\phi}}(t,\omega) = \boldsymbol{\phi}'\mathbf{S}_t(\omega) = \sum_{i=1}^{2} \phi_i S_t^{(i)}(\omega)$$

We define that if $V(0,\omega) = V(0) > 0$, we need to pay V to acquire the asset. If $V(0,\omega) = V(0) < 0$ means we will receive $-V$ amount in acquiring the asset (e.g. negative V means short selling and thus in debt).

Notice that ϕ can be negative in order to reflect possible short selling (since price can't be negative).

5.2 DEFINITION. We say that $\boldsymbol{\phi}$ is an arbitrage opportunity if

1. $\forall \omega \in \Omega, V_{\boldsymbol{\phi}}(0,\omega) = 0$
2. $\forall \omega \in \Omega, V_{\boldsymbol{\phi}}(1,\omega) \geq 0$
3. $\exists \omega \in \Omega, V_{\boldsymbol{\phi}}(1,\omega) > 0$

i.e., starting from a zero investment (1), we are certain not to incur a loss (2) and we have a positive probability to make a gain (3)

COMMENTS As long as there's price difference in the future that we know for certain, or using such a strategy that no cost is incurred, there is arbitrage opportunity. The existence of arbitrage has nothing to do with the price of asset today. The (1) condition is trying to make sure we are getting money from nowhere by short selling. The logic is actually quite simple: At $t = 0$ we short sell the asset having (relative) lower price at $t = 1$.

ARBITRAGE IN ONE PERIOD MODEL Condition (1) implies

$$\phi_1 = -\phi_2 s_0 \rightarrow \boldsymbol{\phi} = (-\phi_2 s_0 \;\; \phi_2)'$$

Condition (3) assures that $(0,0)$ is not an arbitrage opportunity and thus it is safe to assume that $\phi_2 \neq 0$. Conditional (2) together with above portfolio then implies that

$$V_{\phi}(1,\omega_1) = \begin{cases} \phi_2 (s_{11} - s_0(1+r)) \geq 0, & \omega = \omega_1 \\ \phi_2 (s_{12} - s_0(1+r)) \geq 0, & \omega = \omega_2 \end{cases}$$

Thus we conclude

$$\phi_2 > 0 \iff s_{12} > s_{11} \geq s_0(1+r)$$

$$\phi_2 < 0 \iff s_{11} < s_{12} \leq s_0(1+r)$$

(Notice $s_{11} < s_{12}$ strictly is part of the assumption of the model so the "iff" make sense) To interpret: First of all we need to agree that at $t = 0$ we need to short sell the asset that has a relatively lower price at $t = 1$. Then

1. If $s_{12} > s_{11} \geq s_0(1+r)$, then $\forall\, \phi_2 > 0$ the portfolio $(-\phi_2 s_0, \phi_2)$ is an arbitrage opportunity. This means we perform the following procedure:

   - Short sell $ns_0/1$ share of riskless asset to buy $ns_0/s_0 = n$ share of risky asset. In this case $\boldsymbol{\phi} = (-ns_0 \;\; n)'$. The amount paid out is then

     $$V_\phi(0) = -nS_0 + nS_0 = 0$$

   - At $t = 1$ we repurchase the share of risk-less asset and give it back at price $(1+r)$ and sell risky asset at price $S_1^{(2)}(\omega)$ the net amount is then

     $$\begin{aligned} \Delta V_\phi &= V_\phi(1,\omega) - V_\phi(0,\omega) = V_\phi(1,\omega) \\ &= \begin{cases} ns_{11} - ns_0(1+r) = n\left(s_{11} - s_0(1+r)\right) \geq 0 & \text{if } \omega = \omega_1 \\ ns_{12} - ns_0(1+r) = n\left(s_{12} - s_0(1+r)\right) > 0 & \text{if } \omega = \omega_2 \end{cases} \end{aligned}$$

2. If $s_{11} < s_{12} \leq s_0(1+r)$, then $\forall \phi_2 < 0$ the portfolio $(-\phi_2 s_0, \phi_2)$ is an arbitrage opportunity. This means we perform the following procedure:

   - Short sell n shares of risky asset and purchase ns_0 share of riskless asset. In this case $\boldsymbol{\phi} = (ns_0 \;\; -n)'$. The amount paid out is then

     $$V_\phi(0) = ns_0 1 - ns_0 = 0$$

   - At $t = 1$ we repurchase the share of risky asset and give it back at price $S_2^{(1)}(\omega)$ then the net amount paid is then

     $$\begin{aligned} \Delta V_\phi &= V_\phi(1,\omega) - V_\phi(0,\omega) = V_\phi(1,\omega) \\ &= \begin{cases} ns_0(1+r) - ns_{11} = n\left(s_0(1+r) - s_{11}\right) > 0 & \text{if } \omega = \omega_1 \\ ns_0(1+r) - ns_{12} = n\left(s_0(1+r) - s_{12}\right) \geq 0 & \text{if } \omega = \omega_2 \end{cases} \end{aligned}$$

   Actually the ΔV is the net profit not the net amount paid.

This is saying that there is no arbitrage opportunity if

$$s_{11} < s_0(1+r) < s_{12}$$

5.3 DEFINITION. (Contingent claim) A contingent claim is a contract between two parties, a seller and a buyer, the value of which will depend on the state of the market during the contract validity period. It is like an insurance contract. Mathematically speaking, a contingent claim C is any non-negative $(\Omega, \mathcal{F})$-random variable.

In short option is a type of contingent claim. Then we use the one period model to price the option. The idea to price the contingent claim is to match the price of contingent claim and price of portfolio at $t = 1$. Why? The logic to match the value is not quite clear. Let's consider the $(\Omega, \mathcal{F})$ generated by the above assets in a one period model. To match the price we have

$$\begin{aligned}
&\forall \omega \in \Omega, V_\phi(1, \omega) = C(\omega) \\
\Leftrightarrow &\phi_1(1+r) + \phi_2 s_{11} = c_1 \text{ and } \phi_1(1+r) + \phi_2 s_{12} = c_2 \\
\Leftrightarrow &\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \frac{s_{12}c_1 - s_{11}c_2}{(s_{12}-s_{11})(1+r)} \\ \frac{c_2 - c_1}{s_{12}-s_{11}} \end{pmatrix}
\end{aligned}$$

and consequently the portfolio value at $t = 0$ is

$$V_\phi(0, \omega) = V_\phi(0) = \frac{c_1}{1+r} \underbrace{\frac{s_{12} - s_0(1+r)}{s_{12} - s_{11}}}_{q} + \frac{c_2}{1+r} \underbrace{\frac{s_0(1+r) - s_{11}}{s_{12} - s_{11}}}_{1-q}$$

and we derive the new probability measure $\mathbb{Q}$ which is

$$\mathbb{Q}(\omega_1) = q, \ \ \mathbb{Q}(\omega_2) = 1 - q$$

This make sense since c_1 and c_2 are the values of a random variable (e.g. actually the contingent claim price) at $t = 1$ in two state of the world respectively and consequently

$$V_\phi(0, \omega) = \frac{1}{1+r} \mathbb{E}^{\mathbb{Q}}[C]$$

where we use C the express the price of contingent claim at $t = 1$ and $(1 + r)^{-1}$ is the discount factor. Now we have $(\Omega, \mathcal{F}, \mathbb{Q})$. Another important implication is that *no arbitrage opportunity* $\iff$ $q \in (0, 1)$.

Risk Neutral Measure $\mathbb{Q}$ The probability measure derived form above is called risk neutral measure. The reason to be called neutral is as follows. Define the return of asset at time t to be

$$R_S(t, \omega) = \frac{S_t(\omega) - S_{t-1}(\omega)}{S_{t-1}(\omega)}$$

5.4 fact. Under any probability measure $\mathbb{P}$ the expected return of the riskless asset in a $\{0, 1\}$ period is r.

Proof.

$$\begin{aligned}
E^{\mathbb{P}}\left[R_{S^{(1)}}(t=1)\right] &= \sum_{\omega\in\Omega} \frac{S_1^{(1)}(\omega) - S_0^{(1)}(\omega)}{S_0^{(1)}(\omega)}\mathbb{P}(\omega) \\
&= \sum_{\omega\in\Omega} \frac{(1+r)-1}{1}\mathbb{P}(\omega) \\
&= \sum_{\omega\in\Omega} r\mathbb{P}(\omega) \\
&= r.
\end{aligned}$$

□

5.5 FACT. The return of a risky asset depends on the probability measure. The general form is

$$\begin{aligned}
E^{\mathbb{P}}\left[R_{S^{(2)}}(1)\right] &= \sum_{i=1}^{2} \frac{S_1^{(2)}(\omega_i) - S_0^{(2)}(\omega_i)}{S_0^{(2)}(\omega_i)}\mathbb{P}(\omega_i) \\
&= \frac{s_{11} - s_0}{s_0}\mathbb{P}(\omega_1) + \frac{s_{12} - s_0}{s_0}\mathbb{P}(\omega_2)
\end{aligned}$$

Notice this is usually different form the physical probability in the real world since the probability measure can hardly be the same as physical probability. Also as we can see the expected return for risky and riskless asset are usually not equal. However under the measure $\mathbb{Q}$ they are equal which is

$$\begin{aligned}
E^{\mathbb{Q}}\left[R_{S^{(2)}}(1)\right] =& \frac{s_{11} - s_0}{s_0}\mathbb{Q}(\omega_1) + \frac{s_{12} - s_0}{s_0}\mathbb{Q}(\omega_2) \\
=& \frac{s_{11} - s_0}{s_0}\frac{s_{12} - s_0(1+r)}{s_{12} - s_{11}} \\
&+ \frac{s_{12} - s_0}{s_0}\left(1 - \frac{s_{12} - s_0(1+r)}{s_{12} - s_{11}}\right) \\
=& r
\end{aligned}$$

What this imply is important: *On the probability space* $(\Omega, \mathcal{F}, \mathbb{Q})$, *there is no benefit associated with risk, the expected return on the risky security is the same as the one on the riskless security.*

EQUIVALENT MARTINGALE MEASURE (EMM) The measure $\mathbb{Q}$ derived above is also called a EMM.

5.6 FACT. On the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, the discounted price processes of the securities $\left\{\frac{S_t^{(i)}}{(1+r)^t} : t = 0, 1\right\}$ are martingales.

Proof. We check the third condition of martingle only. For riskless asset

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\frac{S_1^{(1)}}{1+r} \mid \mathcal{F}_0\right] &= \mathrm{E}^{\mathbb{Q}}\left[\frac{1+r}{1+r} \mid \mathcal{F}_0\right] \\ &= 1 \\ &= \frac{S_0^{(1)}}{(1+r)^0}.\end{aligned}$$

For risky asset

$$\begin{aligned}\mathrm{E}^{\mathbb{Q}}\left[\frac{S_1^{(2)}}{1+r} \mid \mathcal{F}_0\right] = \mathrm{E}^{\mathbb{Q}}\left[\frac{S_1^{(2)}}{1+r}\right]& \\ &= \sum_{i=1}^{2} \frac{S_1^{(2)}(\omega_i)}{1+r}\mathbb{Q}(\omega_i) \\ &= \frac{s_{11}}{1+r}\mathbb{Q}(\omega_1) + \frac{s_{12}}{1+r}\mathbb{Q}(\omega_2) \\ &= \frac{s_{11}}{1+r}\frac{s_{12}-s_0(1+r)}{s_{12}-s_{11}} + \frac{s_{12}}{1+r}\frac{s_0(1+r)-s_{11}}{s_{12}-s_{11}} \\ &= s_0 \\ &= \frac{S_0^{(2)}}{(1+r)^0}.\end{aligned}$$

□

In reality for most of security

$$\mathrm{E}^{\mathbb{P}}\left[\frac{S_1^{(2)}}{1+r} \mid \mathcal{F}_0\right] > s_0$$

6 REPLICATION AND RISK NEUTRAL MEASURES

This section is base on the paper from Harrison & Pliska.

Model Setup

1. We work on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where Ω is finite $|\Omega| < \infty$. Let the probability measure $\mathbb{P}$ to be the real world physical probability.

2. $\forall \omega \in \Omega$ we have $\mathbb{P}(\omega) > 0$. All participates in the market agrees that Ω includes all possible state of the world.

3. The time scale is finite $\mathcal{T} = \{0, 1, \cdots, T\}$.

4. In terms of filtration we impose the following restriction:

   (a) $\mathbb{F} = \{\mathcal{F}_t : t \in \mathcal{T}\}$ where we assume $\mathcal{F}_0 = \{\Omega, \emptyset\}$ and $\mathcal{F}_T$ is the power set of $\mathcal{F}$ the largest σ-algebra.

   (b) $\mathcal{P}_t = \{A_1^t, A_2^t, \cdots, A_{n_t}^t\}$ s.t. $\mathcal{F}_t = \sigma(\mathcal{P}_t)$.

   (c) $\mathbb{F}$ describes how information is revealed at time t. For example at time t, investors can distinguish between A_t^i and A_t^j in $\mathcal{P}_t$ while they can't distinguish between ω with in A_t^i. Also $\mathcal{F}_0 = \{\Omega, \emptyset\}$ implies the price at $t = 0$ is constant s_0 known for certain.

   (d) In total $(k+1)$ assets are modeled indexing as $0, 1, \cdots, k$. The price evolution of all security is

   $$\mathbf{S} = \{\mathbf{S}_t : \ t \in \mathcal{T}\}$$

   where $\mathbf{S}_t \in \mathbb{R}^{k+1}$. Each of the S_t^i indicates the price of unit share of security i.

5. For those price process we assume the following

   (a) $\mathbf{S}$ is $\mathbb{F}$-adapted(e.g. $\forall\ t$ the price S_t^i is $\mathcal{F}_t$ measurable).

   (b) $S_t^i > 0$ the price is strictly positive.

   (c) $S^0 = \{S_t^0 \ : t \in \mathcal{T}\}$ is considered as the risk free security. Since is risk free then it's safe to assume:

      i. $S_t^0 \leq S_{t+1}^0$ for any $t = 0, 1, \cdots, T-1$

      ii. $\forall\ \omega \in \Omega$ we assume $S_0^0(\omega) = 1$

   (d) Define the discount factor

   $$\beta_t = \frac{1}{S_t^0}$$

   as a scalar stochastic process (e.g. same across ω but vary with t).

Model

6.1 definition. (Predictable process) As stochastic process $X = \{X_t : t \in \mathcal{T}\}$ is predictable if

1. X_0 is $\mathcal{F}_0$-measurable
2. X_t is $\mathcal{F}_{t-1}$-measurable.

The second condition means that all events characterizing distribution of X_t is captured in $\mathcal{F}_{t-1}$. An example is a deterministic process that $X_0 = a$ is $\mathcal{F}_0$-measurable.

6.2 definition. A trading strategy is a predictable vector stochastic process

$$\boldsymbol{\phi} = \{\boldsymbol{\phi}_t : t \in \{1, \ldots, T\}\}$$

where the random row vector $\boldsymbol{\phi}_t = \left(\phi_t^0, \phi_t^1, \ldots \phi_t^K\right)$ represents the investor's portfolio at time t : ϕ_t^k = the number of shares of security k held at time t. Moreover we define ϕ_t^k be the portfolio held by the investor during $(t-1, t]$, $t = 1, 2, \cdots, T$

It is actually crucial to distinguish a strategy and a portfolio while in the lecture and also in this documents we still mix the use of them such that both of them are represented as $\boldsymbol{\phi}$.

The $(t-1, t]$ assumption make sense. In order to show predictable, we allow investor to chose the portfolio right after time t. So the portfolio held at time t should base on a decision made depending on $\mathcal{F}_{t-1}$. Conventionally the process $V(\boldsymbol{\phi}) = \{V_t(\boldsymbol{\phi}) : t \in \{0, 1, \ldots, T\}\}$ represents the market value of the strategy at any time.

$$V_t(\boldsymbol{\phi}) = \begin{cases} \boldsymbol{\phi}_1' \mathbf{S}_0 & \text{if } t = 0 \\ \boldsymbol{\phi}_t' \mathbf{S}_t = \sum_{i=0}^{k} \phi_t^i S_t^i & \text{if } t \in \{1, \ldots, T\} \end{cases}$$

6.3 definition. (Self-financing Strategy) We say a strategy is self-financing if no funds are added to or withdrawn from the value of the portfolio after time $t = 0$, i.e.

$$\forall t \in \{1, \ldots, T-1\}, \boldsymbol{\phi}_t' \mathbf{S}_t = \boldsymbol{\phi}_{t+1}' \mathbf{S}_t$$

Where $\boldsymbol{\phi}_t' \mathbf{S}_t$ is the amount we receive if we liquidated the portfolio $\boldsymbol{\phi}_t'$ at t and $\boldsymbol{\phi}_{t+1}' \mathbf{S}_t$ is the amount we pay to acquire the new portfolio $\boldsymbol{\phi}_{t+1}'$. This means no money is add to or withdraw from the portfolio however the composition of the portfolio can changes. It is a refinance process.

6.4 Definition. (Admissible Strategy) The set of strategy Φ is said to be admissible if

1. Each component ϕ^k of $\boldsymbol{\phi}$ is predictable
2. $\boldsymbol{\phi}$ is self-financing
3. For any $t \in \mathcal{T}$, $V_t(\boldsymbol{\phi}) \geq 0$

If change the last condition to be $V_T(\boldsymbol{\phi}) \geq$ where the value of the portfolio is positive only for T, then it's the relaxed version of admissible strategy.

The relaxed version allows the value to be negative for $t < T$ due to shortselling.

6.5 Definition. An admissible strategy $\boldsymbol{\phi}$ is an arbitrage opportunity if

$$V_0(\boldsymbol{\phi}) = 0 \ \ \& \ \ \mathbb{E}^P\left[V_T(\boldsymbol{\phi})\right] > 0$$

Notice the second condition implies that

$$\mathbb{E}^P\left[V_T(\boldsymbol{\phi})\right] = \sum_{\omega \in \Omega} \underbrace{V_T(\boldsymbol{\phi}, \omega)}_{\geq 0} \underbrace{P(\omega)}_{>0}$$

so there exists at least one ω s.t. $V_T(\boldsymbol{\phi}, \omega) > 0$. So, the strategy $\boldsymbol{\phi}$ is an arbitrage opportunity when, while investing nothing ($V_0(\boldsymbol{\phi}) = 0$), we are assured not to lose any money ($V_T(\boldsymbol{\phi}, \omega) \geq 0$ since $\boldsymbol{\phi}$ is admissible) and we have a positive probability to make a gain ($\exists \omega \in \Omega$ for which $V_T(\boldsymbol{\phi}, \omega) > 0$). Notice here the time is at big T so we only try to make sure the arbitrage works in the end of the investment period.

6.6 Definition. (Contingent Claim in multi-period model) A contingent claim X is a $(\Omega, \mathcal{F}_T)$ – non-negative random variable. Mathematically

$$\begin{aligned}\mathbf{X} &= \text{the set of contingent claims} \\ &= \left\{ X \;\middle|\; \begin{array}{c} X \text{ is a } (\Omega, \mathcal{F}_T) - \text{ random variable} \\ \text{such that } \forall \omega \in \Omega, X(\omega) \geq 0 \end{array} \right\}\end{aligned}$$

6.7 Definition. A contingent claim X is said to be attainable if there exists an **admissible** trading strategy $\boldsymbol{\phi}$ that can replicate the cash flow generated by X, i.e. $V_T(\boldsymbol{\phi}) = X$ has the same value as the contingent claim at T.

It worth notice that $V_T(\phi)$ is indeed F_T-measurable. I assume we just want to model the outcome of the cc at maturity T and we don't care about the value of cc during the period before maturity.

6.8 DEFINITION. (Price System) A price system π is a linear operator on $\mathbf{X}$ that returns non-negative values $\pi : \mathbf{X} \to [0, \infty)$ and that satisfies

$$\pi(X) = 0 \Leftrightarrow X = 0$$

and

$$\forall a, b \in \mathbb{R} \text{ and } \forall X_1 X_2 \in \mathbf{X}, \pi(aX_1 + bX_2) = a\pi(X_1) + b\pi(X_2)$$

The price system tries to attach a price to each of the possible contingent claim. Also an important fact is that

$$V_T(\phi) \in \mathbf{X}$$

for any admissible strategy ϕ. In short the value of an admissible strategy at time T (end of the maturity period) is an contingent claim. Notice $V_T(\phi)$ is an random variable. It's easy to check it's a map from $\Omega \to \mathbb{R}$.

6.9 DEFINITION. A price system is said to be consistent with the market model if the price associated with the contingent claim $V_T(\phi)$ is its market value at time $t = 0$, $V_0(\phi)$, i.e.

$$\pi(V_T(\phi)) = V_0(\phi)$$

6.10 DEFINITION. Let Π be the set of consistent price system. It is defined to be

$$\left\{ \pi : \mathbf{X} \to [0, \infty) \middle| \begin{array}{c} \pi(X) = 0 \Leftrightarrow X = 0 \\ \forall a, b \geq 0 \text{ and } \forall X_1, X_2 \in \mathbf{X}, \\ \pi(aX_1 + bX_2) = a\pi(X_1) + b\pi(X_2) \\ \forall \phi \in \Phi, \pi(V_T(\phi)) = V_0(\phi) \end{array} \right\}$$

Is π trying to price the CC at time 0?

6.11 DEFINITION. (Equivalent risk-neutral or martingale measures (EMMs)) Let $\mathbb{Q}_{\mathbb{P}}$ be the set of Equivalent martingale measure of probability measure $\mathbb{P}$. It is defined to be

$$\mathbb{Q}_{\mathbb{P}} = \left\{ Q \mid \begin{array}{c} Q \text{ is a probability measure on } (\Omega, \mathcal{F}), \\ \forall \omega \in \Omega, Q(\omega) > 0 \text{ and} \\ \forall k \in \{0, 1, \ldots, K\}, \beta S^k \text{ is a Q - martingale.} \end{array} \right\}$$

where

$$\beta S^k = \left\{\beta_t S_t^k \;:\; t \in \mathcal{T}\right\}$$

The EMM is just a convenient tool to price contingent claim.

6.12 DEFINITION. (Equivalent probability measure) Let P and Q be two probability measures that exist on the measurable space $(\Omega, \mathcal{F})$. The measures P and Q are said to be equivalent if and only if the impossible events are the same under both measures, i.e.

$$\forall A \in \mathcal{F}, P(A) = 0 \Leftrightarrow Q(A) = 0$$

In our case, since $\forall \omega \in \Omega, P(\omega) > 0$, all measures Q equivalent to P shall satisfy the condition

$$\forall \omega \in \Omega, Q(\omega) > 0$$

6.13 THEOREM. *Proposal. There is bijective correspondence between the set Π of price systems that are consistent with the market model and the set $\mathbb{P}$ of martingale measures equivalent to* P*. Such as correspondence is*

$$\pi(X) = E^Q\left[\beta_T X\right], X \in \mathbf{X}$$

$$Q(A) = \pi\left(S_T^0 \mathbb{I}_A\right), A \in \mathcal{F}$$

INTERPRETATION

1. If we know a martingale measure Q equivalent to P, then we can build a consistent price system π by setting $\forall X \in \mathbf{X}, \pi(X) = E^Q\left[\beta_T X\right]$

2. On the other hand, if a price system π consistent with the market model is available, then we can build a martingale measure Q equivalent to P by defining $\forall A \in \mathcal{F}, Q(A) = \pi\left(S_T^0 \mathbb{I}_A\right)$

3. On the other hand, if a price system π consistent with the market model is available, then we can build a martingale measure Q equivalent to P by defining $\forall A \in \mathcal{F}, Q(A) = \pi\left(S_T^0 \mathbb{I}_A\right)$

4. Such a proposal tells us that, if there exists a martingale measure equivalent to P or if there exists a price system consistent with the market model, then both exist, and the proposal establishes the link between them.

5. However, nothing allows us to show that either of such two objects exist.

6.14 LEMMA. *(Page 228) If there exists a self-financing strategy* ϕ *(not necessarily admissible) such that* $V_0(\phi) = 0$, $V_T(\phi) \geq 0$ *and* $E^P[V_T(\phi)] > 0$, *then there exists an arbitrage opportunity.*

6.15 THEOREM. *(Fundamental theorem of Asset Pricing FTAP 1) The market model contains no arbitrage opportunities if and only if there exists at least one martingale measure equivalent to* $\mathbb{P}$.

The proof is omitted. it can be found in slides Theorem 2.7. Together with theorem 6.13, they provides the following logic flow:

$$\text{No arbitrage opportunity} \to \exists \mathbb{Q} \sim \mathbb{P} \to \pi(X) = E^{\mathbb{Q}}[\beta_T X], X \in \mathbf{X}$$

If no arbitrage opportunity then by theorem 6.15 there exists equivalent probability measure $\mathbb{Q}$. Then by proposal 6.13 we are able to build a consistent price system π.

6.16 COROLLARY. *Corollary on page 228. If the market model contains no arbitrage, then there is a single price associated with any attainable contingent claim* X *and it satisfies*

$$\pi = \mathbb{E}^{\mathbb{Q}}[\beta_T X] \;\; \forall \mathbb{Q} \sim \mathbb{P}$$

The logic is that if there's no arbitrage opportunity then there can be more than one equivalent probability measure $\mathbb{Q}$ to $\mathbb{P}$ and thus there's theorem tells us there can more more than one consistent price system. What corollary 6.16 says is that the price for any attainable price agree across different price systems.

6.17 THEOREM. *(Proposal 2.8) If* ϕ *is an admissible strategy, then the process* $\beta V(\phi)$, *representing its discounted market value, is a* Q*-martingale for each measure* $\mathbb{Q} \sim \mathbb{P}$.

The proof is easy. Just remember to use the the fact in definition 6.11 that βS^k is a $\mathbb{Q}$-martingale and thus

$$\mathbb{E}^{\mathbb{Q}}[\beta_t^k S_t^k \mid \mathcal{F}_{t-1}] = \beta_{t-1}^k S_{t-1}^k$$

6.18 THEOREM. *Proposal 2.9. If* $X \in \mathbf{X}$ *is an attainable contingent claim, then*

$$\beta_t V_t(\phi) = E^{Q}[\beta_T X \mid \mathcal{F}_t]$$

for any trading strategy ϕ *which generates* X *and for each measure* $\mathbb{Q} \in \mathbb{Q}_{\mathbb{P}}$.

Interpretation If the contingent claim X can be replicated with strategy ϕ, $V_T(\phi) = X$, then the market value of such a contingent claim must be, at any time, equal to the market value of the strategy, otherwise, there is a way to create an arbitrage opportunity. This also can be shown in mathematical way as follows

placeholder

The short selling scenario is as follows. the logic why no arbitrage is not clear.

Recall we always short sell the on of the asset which has **relatively** lower price or in other words the one whose relative price drops in the future. In this case since at maturity T we know ϕ will exactly match the value (price) of the cc (e.g. assumption of the theorem that X is generated by some ϕ), so if there's price discrepancy before T there will be arbitrage opportunity. Furthermore, we could also use the discrepancy price at t to by riskless asset for profit. Not finished yet

6.19 lemma. *One of the implication form theorem 6.18 is that*

$$X_t = V_t(\phi) = \frac{1}{\beta_t}\mathbb{E}^{\mathbb{Q}}\left[\beta_T X \mid \mathcal{F}_t\right]$$

which price the contingent calim at any given time t.

6.20 definition. (Complete Market) A market is said to be complete if it contains **no arbitrage** opportunity and if all contingent claims are **attainable**.

Recall from corollary 6.16, attainable cc in risk free market have unique price across different price systems (e.g. indeed probability measure). However for those cc that are NOT attainable we can not find a portfolio replicating its cash flow. There price is then not unique.

6.21 theorem. *A market is complete if and only if there exists a single martingale measure. On the contrary, the arbitrage-free market is incomplete if there exists at least one cc that is not attainable.*

Notice there can be attainable cc in incomplete market and its price is unique. Theorem 6.21 tells that based on FTAP, if there exists only one measure equivalent to $\mathbb{P}$ then market is complete; if there's more than one then the market is not complete.

In practice we are given the price process and we want to check if the market is complete to design an arbitrage strategy. There are tow ways to do this. The first one is based on theorem 6.21. The logic is to check if the discounting process $\{S_t^k \beta_t\}_t$ with given prices over time allows unique solution to $\mathbb{Q}(\omega)$. The equations are given by

$$\begin{cases} \mathbb{E}^{\mathbb{Q}}[S_t^k \beta_t \mid \mathcal{F}_{t-1}] & = S_{t-1}^k \beta_{t-1} \ \forall k \\ \sum_{\omega \in \Omega} \mathbb{Q}(\omega) & = 1 \end{cases}$$

This set of equations can be turned into a linear system and we solve or check the solution situation. Then eventually if the solution is unique then market is complete otherwise not. Another approach is that we form a linear system as follows based on the definition 6.20. This approach is to solve a system of linear equation built based on the fact that all cc in complete market are attainable.

To set up the framework. Consider a one-period model where $\mathcal{T} = \{0, 1\}$ and $k+1$ security $\mathbf{S}_t = \{S_t^0, S_t^1, \cdots, S_t^k\}$ defined on $(\Omega, \mathcal{F})$ (we also use the set of asset $\mathbf{S}$ to represent the "market"). Let $\mathrm{Card}(\Omega) = n$. The logic is then since all cc are attainable in a complete market this means $\forall\, X \in \mathbf{X}$, $\exists\, \boldsymbol{\phi}$ s.t. $V_T(\boldsymbol{\phi}) = X_T \ \forall \omega \in \Omega$. Expand the above equation we have

$$V_T(\boldsymbol{\phi}, \omega) = \sum_{i=0}^{k} \phi_T^i S_T^i(\omega) = X(\omega)$$

since this have to work for all ω thus we set up the following system

$$\mathbf{A}\boldsymbol{\phi}_T = \begin{bmatrix} S_T^0(\omega_1) & S_T^1(\omega_1) & \cdots & S_T^k(\omega_1) \\ S_T^0(\omega_2) & S_T^1(\omega_2) & \cdots & S_T^k(\omega_2) \\ \cdots & & \cdots & \cdots \\ S_T^0(\omega_n) & S_T^1(\omega_n) & \cdots & S_T^k(\omega_n) \end{bmatrix} = \begin{bmatrix} X(\omega_1) \\ X(\omega_2) \\ \cdots \\ X(\omega_n) \end{bmatrix} = \mathbf{X}$$

where the columns range over all $k+1$ securities and the rows range over all possible state of the world. Thus in order to have all equations satisfied, the system should have unique solution thus equivalently the rank of $\mathbf{A}$ must have full rank. Noitce that in terms of the matrix, $n \le k+1$. The number of rows must be equal to the number of states since we want our equations to be met at each of every states.

To extend this to a multi-period model of $\mathcal{T} = \{0, 1, \cdots, T\}$ we just have to check that at every time $t \in \mathcal{T}$, all matrix involved have to be of full rank. This is guaranteed by lemma 6.19.

6.22 THEOREM. *Assume a one-period model* $M = \left(s^0, s^1, \ldots, s^k\right)$. *Then a cc is attainable iff*

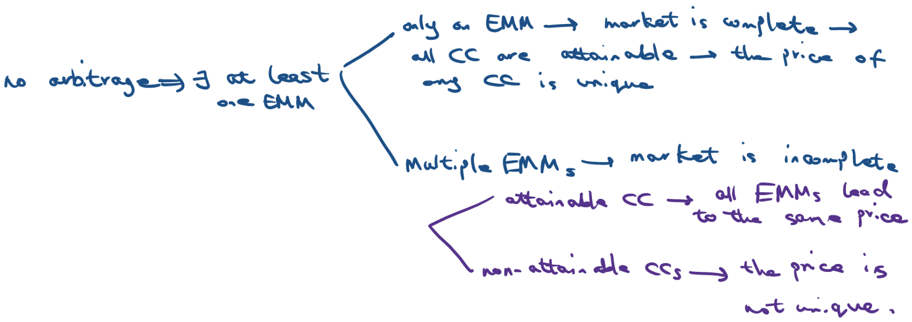
$$X_0 = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{S_T^0} X\right]$$

have the same value across all EMMs $\mathbb{Q} \sim \mathbb{P}$

Notice no matter in a complete or incomplete market, the price associated to a contingent claim are always equal across those EMMs as long as the market is arbitrage free. However for those that are not attainable, the price are not unique (e.g. differ EMM generate different price).

6.23 THEOREM. *(FTAP 2) An arbitrage-free market is complete iff there exists a unique EMM to $\mathbb{P}$.*

The two FTAP can be summarised as the graph below.



The proof in $\Longrightarrow$ does not make sense. The proof should based on the fact that it works for all cc. However the proof takes only one specific X?

7 APPLICATION: PRICE THE EUROPEAN AND AMERICAN STYLE OPTION

The model set up are exactly the same as introduced in the beginning of chapter 6.

7.1 DEFINITION. A European-style contingent claim is a non-negative random variable, $\mathcal{F}_T$-measurable since the said contingent claim can only be exercised at maturity, that is at time T.

7.2 DEFINITION. An American-style contingent claim, by contrast, can be represented as an $\mathbb{F}$-adapted stochastic process $X = \{X_t : t = 0, 1, \ldots, T\}$ where X_t represents the contingent claim value at time t if it is exercised at that time.

The holder of European option has no right to exercises until maturity date T while American option holder can chose to exercise during the life time of the option. Thus the time to exercise an American option is a stopping time. The set of those stooping time is defined to be in a set

$$\Lambda_0 = \{\tau : \Omega \longrightarrow \mathcal{T} = \{0, \ldots, T\} \mid \tau \text{ is a stopping time }\}$$

the subscript 0 indicates the set is about admissible exercising time when we are at time 0. Now consider the following definition

$$\boldsymbol{\tau} : \{\Omega\} \longrightarrow \mathcal{T}^{\mathrm{Card}(\Omega)} \subseteq \mathbb{R}^{\mathrm{Card}(\Omega)}$$

$$\boldsymbol{\tau}(\Omega) \to \left[\tau(\omega_1), \cdots, \tau(\omega_{\mathrm{Card}(\Omega)})\right]'$$

The expression is not quite rigorous Need some proper definition from function space I believe. The logic is as follows. We may not know the exact rule of a random time sequence (e.g. for example the time when the discounted expected value maximized). However, since $\mathcal{T}$ is a finite set, then there are also just finitely many possible results of those random time sequence which is $|\mathcal{T}|^{\mathrm{Card}(\Omega)}$. Thus wee can narrow down our set of all random sequences to a set of stopping time by checking the definition of stopping time.

7.3 EXAMPLE. Let's consider an American-style contingent claim, a put option with an 80-dollar strike price. The value of such a put at time t, if the put is exercised, is

$$X_t = \max\left\{80 - S_t^1, 0\right\}$$

and its discounted value at time t, if again the put is exercised, is

$$Y_t = \beta_t X_t = (1+r)^{-t}\left\{80 - S_t^1, 0\right\} = 1.115^{-t}\max\left\{80 - S_t^1, 0\right\}$$

Once again, the filtration is determined by the price process of all asset. That is, the $\mathbf{S}_{t=1}^{\mathrm{T}}$ process. The price process is given as follows. The price process is given and the values of the discounted process is given as follows.

ω	$\begin{pmatrix} S_0^0(\omega) \\ S_0^1(\omega) \end{pmatrix}$	$\begin{pmatrix} S_1^0(\omega) \\ S_1^1(\omega) \end{pmatrix}$	$\begin{pmatrix} S_2^0(\omega) \\ S_2^1(\omega) \end{pmatrix}$	$\begin{pmatrix} S_3^0(\omega) \\ S_3^1(\omega) \end{pmatrix}$	$\mathbb{Q}(\omega)$
ω_1	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 100 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 125 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 156,25 \end{pmatrix}$	0.343
ω_2	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 100 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 125 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 100 \end{pmatrix}$	0.147
ω_3	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 100 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 100 \end{pmatrix}$	0.147
ω_4	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 100 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 64 \end{pmatrix}$	0.063
ω_5	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 64 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 100 \end{pmatrix}$	0.147
ω_6	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 64 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 64 \end{pmatrix}$	0.063
ω_7	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 64 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 51.20 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 64 \end{pmatrix}$	0.063
ω_8	$\begin{pmatrix} 1 \\ 80 \end{pmatrix}$	$\begin{pmatrix} 1.115 \\ 64 \end{pmatrix}$	$\begin{pmatrix} 1.115^2 \\ 51.20 \end{pmatrix}$	$\begin{pmatrix} 1.115^3 \\ 40.96 \end{pmatrix}$	0.027

Then the discounted price process Y_t is

ω	Y_0	Y_1	Y_2	Y_3
ω_1	0	0	0	0
ω_2	0	0	0	0
ω_3	0	0	0	0
ω_4	0	0	0	$\frac{16}{1.115^3} \cong 11.5424$
ω_5	0	$\frac{16}{1.115} \cong 14.3498$	0	0
ω_6	0	$\frac{16}{1.115} \cong 14.3498$	0	$\frac{16}{1.115^3} \cong 11.5424$
ω_7	0	$\frac{16}{1.115} \cong 14.3498$	$\frac{28.80}{1.115^2} \cong 23.1656$	$\frac{16}{1.115^3} \cong 11.5424$
ω_8	0	$\frac{16}{1.115} \cong 14.3498$	$\frac{28.80}{1.115^2} \cong 23.1656$	$\frac{39.04}{1.115^3} \cong 28.1634$

Now define a random time process of interest (want to check if it's a valid stopping time) as "for each ω, a time when the option value is the greatest".

Then we find that

$$\begin{aligned}\tau(\boldsymbol{\omega}) =& \left[\tau(\omega_1), \tau(\omega_2), \tau(\omega_3), \tau(\omega_4), \tau(\omega_5), \tau(\omega_6), \tau(\omega_7), \tau(\omega_8)\right]' \\ =&(3, 3, 3, 3, 1, 1, 2, 3)'\end{aligned}$$

If we check that t = 2, then we find $\{\omega \in \Omega : \tau(\omega) = 2\} = \{\omega_7\} \notin \mathcal{F}_2$ so it is not a valid stopping time.

Now think in a way that the rule of a stopping time is not given. We consider all the possible result, that is the co-domain of a $\boldsymbol{\tau}$. Thus there will be

$$|\mathcal{T}|^{\text{Card}(\Omega)}$$

number of results. We can actually check whether each of them is a valid stopping time or not in the same manner as above example. The definition of $\boldsymbol{\tau}$ have to be revised. Λ_0 is the set of VALID stopping time. The logic discussed below is picking items from a space of all random time sequences.

Decision on American Option Exercises

7.4 DEFINITION. (Formulation of Snell's problem) Let $Y = \{Y_t : t = 0, 1, \ldots, T\}$ be a stochastic process, $\mathbb{F}$-adapted. For all $t \in \{0, 1, \ldots, T\}$, we can define the set of stopping times taking their values in the set $\{t, \ldots, T\}$:

$$\Lambda_t = \{\tau : \Omega \longrightarrow \{t, \ldots, T\} \mid \tau \text{ is a stopping time}\}$$

Note that $\Lambda_T \subseteq \Lambda_{T-1} \subseteq \ldots \subseteq \Lambda_0$ Make up a formal proof here. Logic is that if we consider the Λ_t as set of $\boldsymbol{\tau}$ under definition 7.2, then as t decreases, we simply eliminated a column of S_t so there's less choice of each of the element in the vector. Snell's problem is as follows: can we determine a stopping time $\tau^* \in \Lambda_0$ satisfying

$$E[Y_{\tau^*}] = \sup_{\tau \in \Lambda_0} E[Y_\tau]$$

In other words, we are looking to determine, for each of the $\omega \in \Omega$, the time $\tau^*(\omega)$ when we should stop the stochastic process Y in order to maximize the expected value of the random variable Y_τ.

7.5 EXAMPLE. Let's assume that $\forall t \in \{0, 1, \ldots, T\}, \mathcal{F}_t = \{\emptyset, \Omega\}$. Under such conditions, since Y_t is $\mathcal{F}_t-$ measurable, then Y_t is constant, i.e. there exists a real number y_t for which

$$\forall \omega \in \Omega, Y_t(\omega) = y_t$$

That is we can group all trajectory into one single path. In this case our stopping time is also deterministic: Consider

$$\Lambda_0 = \{\tau_0, \tau_1, \ldots, \tau_T\}$$

(the τ_t is like the column vector in example 7.3) $\{\omega : \tau(\omega) = t\} \in \mathcal{F}_t = \{\emptyset, \Omega\}$ (e.g. τ is $\{\emptyset, \Omega\}$-measurable so can only take one constant value across ω which is $\tau_t(\omega) = t$). Then this problem turns to be

$$y_{t^*} = \max_{t \in \{0,1,\ldots,T\}} y_t$$

The procedure is as follows. The basic idea is to introduce an auxiliary sequence $\{z_t : t = 0, 1, \ldots, T\}$ defined as

$$z_t = \max_{u \in \{t,\ldots,T\}} y_u$$

Note that this sequence is decreasing (actually non-increasing. There have to exists some = otherwise otherwise the whole sequence is constant) (i.e. $\forall t \in \{1, \ldots, T\}, z_t \leq z_{t-1}$) and

$$z_t = \max\{y_t, z_{t+1}\}$$

Let's set

$$t^* = \min\{t \in \{0, 1, \ldots, T\} \mid z_t = y_t\}$$

and let's show that t^* satisfies equation we want.

A General Case What we want is now just to build a correct form of decreasing sequence z_t satisfying $z_t = \max\{y_t, z_{t+1}\}$.

Let's set

$$Z_t = \begin{cases} Y_T & \text{if } t = T \\ \max\{Y_t, E[Z_{t+1} \mid \mathcal{F}_t]\} & \text{if } t \in \{0, \ldots, T-1\} \end{cases}$$

Interpretation. Z_t represents, conditionally to the information available at time t, the maximum between the discounted value of the option if it is exercised at that time and its expected discounted value if it is exercised subsequently, at a time judiciously chosen. Z_t is thus the discounted value of the American option at time t. Then there are 2 cases happens at each point of time t:

$$\begin{cases} Y_t \geq E[Z_{t+1} \mid \mathcal{F}_t] \implies Z_t = Y_t \implies \text{Exercises} \\ Y_t < E[Z_{t+1} \mid \mathcal{F}_t] \implies Z_t \neq Y_t \implies \text{Don't exercise} \end{cases}$$

Notice, when we try to specify the atoms of $\mathcal{F}_t$, it is determined by the **S** process which contains all possible assets. It is not determined by the process of Y_t.

8 BROWNIAN MOTION

We starts by introducing the continuous time process. Consider a continuous time process with the following properties:

- The value that the variable takes can change at any point in time and from moment to moment.
- It can take any real number as value.
- The value have to change continuously with no jump (e.g. if considers the process as a function of time then it is continuous in time).
- It takes value at random

No we will construct the continuous time process based on single period binomial model. Consider on single period binomial model with

$$S_t = \begin{cases} uS_{t-1}\ , & \text{upward price movement} \\ dS_{t-1}\ , & \text{downward price movement} \end{cases}$$

where $0 < d < 1 < u$. Consider on a time line from 0 to $t \in \mathbb{Z}$ and between each 1 unit of time there are n times of price movements happened (e.g. in total $nt \in \mathbb{Z}$ time of movements). Also we assume:

(a) For each binomial movement, the upward and downward factors depends on n and $\sigma > 0$, the volatility of price are the same over time which are in form

$$\begin{cases} u_n = 1 + \frac{\sigma}{\sqrt{n}} \\ d_n = 1 - \frac{\sigma}{\sqrt{n}} \end{cases}$$

(b) The risk neutral probability $q = 0.5$ and $1 - q = 1 - 0.5 = 0.5$

Then we define $N_u(\omega)$ and $N_d(\omega)$ as the number of upward and downward movements over time $[0, t]$ at a given state ω. Thus the price at time t is then given by

$$S_{t,n}(\omega) = S_0 \left(1 + \frac{\sigma}{\sqrt{n}}\right)^{N_u(\omega)} \left(1 - \frac{\sigma}{\sqrt{n}}\right)^{N_d(\omega)}$$

The following theorem shows the asymptotic behaviour of the above price.

8.1 THEOREM. *As $n \to \infty$, the distribution of $S_{t,n}$ converges to*

$$S_{t,n}(\omega) \longrightarrow_d S_t(\omega) = S_0 \exp\left\{\sigma W_t(\omega) - \frac{1}{2}\sigma^2 t\right\}$$

where $W_t(\omega)$ is a normal random variable with mean zero and variance t.

Comments: The Ω can also be uncountable defined using Borel set. Now in this case we only deal with finite Ω.

8.1 Brownian Motion

8.2 DEFINITION. A standard Brownian motion $\{W_t : t \geq 0\}$ is an adapted stochastic process, built on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ such that:

(a) $\forall \omega \in \Omega, W_0(\omega) = 0$

(b) $\forall 0 \leq t_0 \leq t_1 \leq \ldots \leq t_k$, the random variables $W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \ldots, W_{t_k} - W_{t_{k-1}}$ are independent (for any given ω)

(c) $\forall s, t \geq 0$ such that $s < t$, the random variable $W_t - W_s$ is normally distributed with expectation 0 and variance $t - s$ i.e. $W_t - W_s \sim N(0, t - s)$

(d) $\forall \omega \in \Omega$, the path $t \to W_t(\omega)$ is continuous (continuous in t)

In short, a Brownian motion is a stochastic process that is **continuous in time, starting with value 0, with independent and normally distributed increments over time**. The definition, cdf and pdf of normal r.v. is omitted. Just remember that independence implies 0 covariance but not vice versa. Also notice the BM is adapted to **F** by definition so W_t is always $\mathcal{F}_t$ measurable. The reverse is true only if both random variables are normally distributed. This is in short:

8.3 FACT. Let $X \sim N(0, \sigma_1^2)$ and $Y \sim N(0, \sigma_2^2)$. Then

$$X \perp Y \iff \text{Cov}(X, Y) = 0$$

FILTRATION The most often used filtration for Brownian motion is $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$

$$\mathcal{F}_t = \sigma(\{W_s : 0 \leq s \leq t\} \cup \mathcal{N})$$

where $\mathcal{N}$ is the collection of 0 measure events. This is also the smallest σ–algebra for which the $W_s : s \in [0, t]$ are measurable (smallest since it is exactly generated by the r.vs). What is the sigma algebra generated by a continuous r.v? The sigma algebra still meet the following properties as in discrete case:

- $\forall 0 \leq s \leq t$, $\mathcal{F}_s \subseteq \mathcal{F}_t$ (as much information as there in the later earlier time)
- (Adaptivity) For any $t \geq 0$, the Brownian motion W_t is $\mathcal{F}_t$ measurable (the information at t is sufficient to evaluate W_t at time t).
- (Independence of future increment) $\forall 0 \leq t < u$, the incremental $W_u - W_t$ is independent of $\mathcal{F}_t$ (any increment of the Brownian motion after t is independent of the information at time t).

8.4 LEMMA. *Let $\{W_t : t \geq 0\}$ be a standard Brownian motion. Then*

- $\forall s > 0, \{W_{t+s} - W_s : t \geq 0\}$ *(time homogeneity)*
- $\{-W_t : t \geq 0\}$ *(symmetry)*
- $\left\{cW_{\frac{t}{c^2}} : t \geq 0\right\}$ *(time rescaling)*
- $\left\{W_t^* = tW_{\frac{1}{t}}\mathbb{1}_{t>0} : t \geq 0\right\}$ *(time inversion)*

are also standard Brownian motions.

How to show the time inversion is also a standard BM?

Proof. Just show the time homo case. make up later □

8.5 DEFINITION. (Martingale in continuous time) On the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F}$ is the filtration $\{\mathcal{F}_t : t \geq 0\}$, the stochastic process $M = \{M_t : t \geq 0\}$ is a martingale in continuous time if

- $\forall t \geq 0, E^{\mathbb{P}}[|M_t|] < \infty$;
- $\forall t \geq 0, M_t$ is $\mathcal{F}_t$ - measurable;
- $\forall s, t \geq 0$ s.t. $s < t, E^{\mathbb{P}}[M_t \mid \mathcal{F}_s] = M_s$.

Nothing different from discrete case since the Ω is still finite.

8.6 THEOREM. *The Brownian Motion is an Martingale process.*

Proof. Easy to check the expectation of $|W_t|$ is finite. Also W_t is $\mathcal{F}_t$ measurable by construction. The last property is checked as follows.

$$\begin{aligned} E^{\mathbb{P}}[W_t \mid \mathcal{F}_s] &= E^{\mathbb{P}}[W_t - W_s + W_s \mid \mathcal{F}_s] \\ &= E^{\mathbb{P}}[W_t - W_s \mid \mathcal{F}_s] + E^{\mathbb{P}}[W_s \mid \mathcal{F}_s] \\ &= E^{\mathbb{P}}[W_t - W_s] + W_s \\ &= W_s \end{aligned}$$

□

8.7 LEMMA. *The Brownian Motion is a Markov process*

Proof in slides Brownian motion page 15.

8.8 FACT. The path $W(t)$ is nowhere differentiable.

Some good intuitions are here for non-differentiable.

8.9 DEFINITION. (Stopping time) Let $(\Omega, \mathcal{F})$ be a measurable space equipped with the filtration $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$. A stopping time τ is a function of Ω into $[0, \infty]\mathcal{F}$-measurable such that

$$\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$$

Intuitively, a stopping time tells at each given state ω when to stop and also given information until time t we should know whether to stop or not thus have to be $\mathcal{F}_t$ measurable.

Consider the first hitting time in continuous time process. It is a stopping time as shown in discrete time case. It is defined as follows in continuous time cases.

8.10 DEFINITION. Let $a > 0$. Let's define

$$\tau_a(\omega) = \begin{cases} \inf\{s \geq 0 : W_s(\omega) = a\} & \text{if } \{s \geq 0 : W_s(\omega) = a\} \neq \varnothing \\ \infty & \text{if } \{s \geq 0 : W_s(\omega) = a\} = \varnothing \end{cases}$$

the first time when Brownian motion W reaches point a. Shouldn't we consider a as an interval? $W_s(\omega)$ is a conts. r.v. how could it takes a single value?

8.11 LEMMA. *The random variable τ_a is a stopping time.*

Proof. Proof of the lemma. We must show that for all $t \geq 0$, the event $\{\omega \in \Omega : \tau_a \leq t\}$ belongs to the sigma-algebra $\mathcal{F}_t$. If $\mathbb{Q}$ represents the set of all rational numbers, then

$$\begin{aligned} &\{\omega \in \Omega : \tau_a \leq t\} \\ &= \left\{\omega \in \Omega : \sup_{0 \leq s \leq t} W_s(\omega) \geq a\right\} \\ &= \bigcap_{n=1}^{\infty} \left\{\omega \in \Omega : \sup_{0 \leq s \leq t} W_s(\omega) > a - \frac{1}{n}\right\} \\ &= \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{r \in \mathbb{Q} \cap [0,t]} \underbrace{\left\{\omega \in \Omega : W_r(\omega) > a - \frac{1}{n}\right\}}_{\mathcal{F}_r \text{ therefore } \in \mathcal{F}_t}}_{\in \mathcal{F}_t} \in \mathcal{F}_t \end{aligned}$$

where the last equality is obtained from the fact that $\sup_{0\leq s\leq t} W_s(\omega) > a - \frac{1}{n}$ if and only if there exists at least one rational number r smaller than or equal to t for which

$$W_r(\omega) > a - \frac{1}{n}$$

$\square$

yet to go through.

8.12 LEMMA. *The stopping time τ_a is finite almost surely, i.e.*

$$\mathbb{P}[\tau_a = \infty] = 0$$

or equivlently

$$\mathbb{P}(\{\omega : \tau_a(\omega) = \infty\}) = 0$$

In short the above lemmas and definitions show the following facts:

- BM will eventually reach every real value a no matter how large a is.
- BM is recurrent. It visits each of its states infinite number of times.

The proof of the theorems and lemmas can be find in slides from pages 27 to 37.

8.2 Multi-dimensional Brownian Motion

8.13 DEFINITION. Standard Brownian motion W of dimension n is a family of random vectors

$$W = \left\{ \mathbf{W}_t = \left(W_t^{(1)}, \ldots, W_t^{(n)} \right)^\top : t \geq 0 \right\}$$

where $W^{(1)}, \ldots, W^{(n)}$ represent independent Brownian motions built on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Standard BM consists of independent individual BMs. In practice they are not independent. We usually use BM to model sources uncertainty of asset price in financial market. Now we discuss techniques used to obtain a dependent multi-dimensional BMs. To do so, we consider linear combinations of those independent BMs.

8.14 LEMMA. *Let $\Gamma = \left(\gamma_{ij}\right)_{i,j\in\{1,2,\ldots,n\}}$ is a matrix of constants and $\boldsymbol{W} = \left(W^{(1)}, \ldots, W^{(n)}\right)^\top$ is a vector made up of independent Brownian motions. For all t, let's set $B_t = \Gamma W_t$. Then B_t is un random vector of dimension n, the ith component of which is*

$\boldsymbol{B}_t^{(i)} = \sum_{k=1}^{n} \gamma_{ik} \boldsymbol{W}_t^{(k)}$. *Moreover*

$$\text{Cov}\left[B_t^{(i)}, B_t^{(j)}\right] = t \sum_{k=1}^{n} \gamma_{ik}\gamma_{jk}$$

$$\text{Cor}\left[B_t^{(i)}, B_t^{(j)}\right] = \frac{\sum_{k=1}^{n} \gamma_{ik}\gamma_{jk}}{\sqrt{\sum_{k=1}^{n} \gamma_{ik}^2}\sqrt{\sum_{k=1}^{n} \gamma_{jk}^2}}$$

The proof shown in class is unnecessary and tedious. In matrix notation, we have

$$\begin{aligned}\Sigma_{\mathbf{B}_t} &= \mathbb{E}\left[\Gamma(\mathbf{W}_t - \boldsymbol{\mu}_{\mathbf{W}_t})(\mathbf{W}_t - \boldsymbol{\mu}_{\mathbf{W}_t})'\Gamma'\right] \\ &= \Gamma \Sigma_{\mathbf{W}_t} \Gamma' = t\Gamma\Gamma'\end{aligned}$$

Just remember the fact that

$$W_t = W_t - W_0 \sim N(0, \sigma^2 = t)$$

Reversely, if we've already known a correlated multi-dimensional BM with given correlation matrix $\rho \in \mathbb{R}^{n\times n} = (\rho_{ij})$, we can also restore the independent BM guaranteed by the following theorem.

8.15 THEOREM. *Let's now assume that* $B^{(1)}, \ldots, B^{(n)}$ *represent correlated Brownian motions, built on the filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *and that*

$$\forall i, j \in \{1, \ldots, n\} \text{ and } \forall t \geq 0, \text{Cor}\left[B_t^{(i)}, B_t^{(j)}\right] = \rho_{ij}$$

There exists a matrix $\mathbf{A}$ *of format* $n \times n$ *such that*

- $\mathbf{B} = \mathbf{AW}$
- $\text{Cor}\left[B_t^{(i)}, B_t^{(j)}\right] = \rho_{ij}$
- $\mathbf{W}$ *is made of independent Brownian motions.*

TO COMPUTE MATRIX A Recall the fact that the correlation matrix and covariance matrix are both symmetric and s.p.d. The above theorem tells:

- Since Each of the $B^{(i)}$ is a standard BM so

$$\rho_{ij} = \frac{\text{Cov}(B_t^{(i)}, B_t^{(j)})}{\sigma_{B_t^{(i)}}\sigma_{B_t^{(j)}}} = \frac{\text{Cov}(B_t^{(i)}, B_t^{(j)})}{\sqrt{t}\sqrt{t}} \implies \text{Cov}(B_t^{(i)}, B_t^{(j)}) = t\rho_{ij}$$

  then we have

$$\Sigma_{\mathbf{B}_t} = t\rho$$

where ρ is the known correlation matrix.

- Then the theorem says $\mathbf{B}_t = \mathbf{A}\mathbf{W}_t$ then by basic algebra

$$\Sigma_{\mathbf{B}_t} = \mathbf{A}\Sigma_{\mathbf{W}_t}\mathbf{A}' = t\mathbf{A}\mathbf{A}'$$

In short now we have

$$\Sigma_{\mathbf{B}_t} = t\mathbf{A}\mathbf{A}' = t\rho$$

Recall for a p.s.d matrix it admits the eigenvalue decomposition (spectral decomposition if in addition symmetric). So we could decompose $t\rho$

$$\Sigma_{\mathbf{B}_t} = t\rho = t\mathbf{Q}\Lambda\mathbf{Q}' = \underbrace{[\sqrt{t}\mathbf{Q}\Lambda^{1/2}]}_{\mathbf{U}}[\sqrt{t}\mathbf{Q}\Lambda^{1/2}]' = t\mathbf{A}\mathbf{A}'$$

Thus

$$\mathbf{A} = \frac{1}{\sqrt{t}}\mathbf{U}$$

9 STOCHASTIC INTEGRAL

Here we omitted the discussion about Riemann integral. The stochastic integral is aspired by the following eqation

$$I_t = \int_0^t X_s dW_s$$

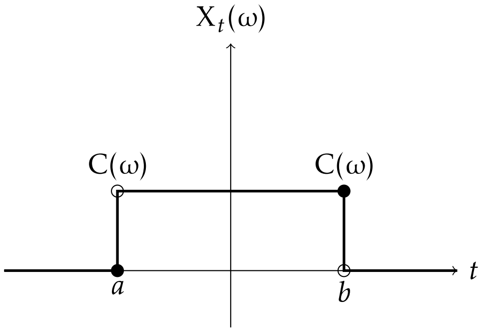
where the X_s is the number of shares of an asset we are holding, dW_s is the incremental of a Brownian motion which roughly equivalent to the variation of share price and thus the I_t is the profit (or loss) during the horizon $(0, t)$.

9.1 DEFINITION. (Basic Stochastic Process) We call X a basic stochastic process if X admits the following representation:

$$X_t(\omega) = C(\omega)\mathbb{I}_{(a,b]}(t)$$

where $a < b \in \mathbb{R}$ and C is a random variable, $\mathcal{F}_a$-measurable and square-integrable, i.e. $E^{\mathbb{P}}\left[C^2\right] < \infty$.

In orther words



It worth notice that the r.v. X_t is actually $\mathcal{F}_t$ measurable. Intuitively, consider that at time a right after the price is revealed we by $C(\omega)$ share of the asset and hold is during $(a, b]$ and then sell all of them at time b once the price is revealed. Are we using the sigma algebra generated by the price process as the filter?

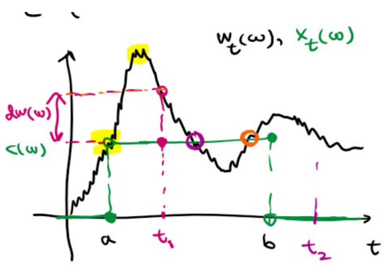
9.2 DEFINITION. (Stochastic integral of basic process) The stochastic integral of X with respect to the Brownian motion is defined as

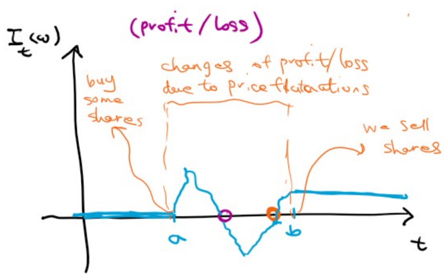
$$\left(\int_0^t X_s dW_s\right)(\omega)$$

$$=C(\omega)\left(W_{t\wedge b}(\omega) - W_{t\wedge a}(\omega)\right)$$

$$=\begin{cases} 0 & \text{if } 0 \leq t \leq a \\ C(\omega)\left(W_t(\omega) - W_a(\omega)\right) & \text{if } a < t \leq b \\ C(\omega)\left(W_b(\omega) - W_a(\omega)\right) & \text{if } b < t. \end{cases}$$

The quantity $\left(\int_0^t X_s dW_s\right)(\omega)$ is a r.v. for any t and thus $\left\{\left(\int_0^t X_s dW_s\right)(\omega) : t \geq 0\right\}$ is a stochastic process. The illustration is as follows.





where the left one shows the evolution of **change in price**. Notice the Brownian motion models the change in price $\Delta P(\omega)$ not the price it self. Page 19 exercise. Notice the above calculus result is the same as in RS integral taking the integral w.r.t W_s.

9.3 LEMMA. *If X is a basic stochastic process, then*

$$\left\{ \int_0^t X_s dW_s : t \geq 0 \right\}$$

is a $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$*-martingale. Then by the martingale properties*

$$\mathbb{E}\left[\int_0^t X_u dW_u \,\middle|\, \mathcal{F}_s \right] = \int_0^s X_u dW_u$$

See proof in slides on page 22. Also The properties indicates a constant expectation of the martingale which is

$$\mathbb{E}[I_t] = \mathbb{E}[I_t|\mathcal{F}_0] = I_0 = \int_0^0 X_s dW_s = 0$$

9.4 LEMMA. *Lemma 2. If X and Y are basic processes, then for all* $t \geq 0$*,*

$$E^{\mathbb{P}}\left[\int_0^t X_s Y_s dW_s \right] = E^{\mathbb{P}}\left[\left(\int_0^t X_s dW_s \right) \left(\int_0^t Y_s dW_s \right) \right]$$

Check the proof in slides.

9.5 DEFINITION. (Simple Stochastic process) We call X a simple stochastic process if X is a finite sum of basic processes :

$$X_t(\omega) = \sum_{i=1}^{n} C_i(\omega) \mathbb{I}_{(a_i, b_i]}(t)$$

9.6 DEFINITION. The stochastic integral of X with respect to the Brownian motion is defined as the sum of the stochastic integrals of the basic processes which constitute X :

$$\int_0^t X_s dW_s = \sum_{i=1}^{n} \int_0^t C_i \mathbb{I}_{(a_i, b_i]}(s) dW_s$$

Notice, in practice we compute the results of the integral, we take

$$C_i = \lim_{t^+ \longrightarrow a_i} X_t(\omega)$$

That is the left end point of X (or the right limit). This will be consistent with Ito's integral for general stochastic process. Further more, the stochastic integral of simple process doesn't depend on representation (e.g. the choice of basic process). The integral will give same result which is

$$\sum_{i=1}^{n} \int_0^t C_i \mathbb{I}_{(a_i,b_i]}(s) dW_s = \sum_{j=1}^{m} \int_0^t \widetilde{C}_j \mathbb{I}_{(\widetilde{a}_j,\widetilde{b}_j]}(s) dW_s$$

9.7 LEMMA. *Lemma 3. If X is a simple stochastic process, then*

$$\left\{ \int_0^t X_s dW_s : t \geq 0 \right\}$$

is a $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$*-martingale.*

Proof is simple. The integral of basic process is martingale. Then the sum of martingale process is also martingale.

9.8 LEMMA. *(Ito's isometry) If X is a simple process, then for all* $t \geq 0$*,*

$$E^{\mathbb{P}}\left[\int_0^t X_s^2 ds\right] = E^{\mathbb{P}}\left[\left(\int_0^t X_s dW_s\right)^2\right]$$

Proof see appendix of slides. Using this we can compute the variance of the process $\{I_t : t \geq 0\}$ which is

$$\begin{aligned} &\mathrm{Var}^{\mathbb{P}}\left[\int_0^t X_s dW_s\right] \\ =&E^{\mathbb{P}}\left[\left(\int_0^t X_s dW_s\right)^2\right] - (\underbrace{E^{\mathbb{P}}\left[\left(\int_0^t X_s dW_s\right)\right]}_{=0,\text{Martingale}})^2 \\ =&E^{\mathbb{P}}\left[\int_0^t X_s^2 ds\right] = \int_0^t E^{\mathbb{P}}\left[X_s^2\right] ds. \end{aligned}$$

It is possible to extend this calculation to other processes X and to establish in a similar manner a method to calculate the covariance between two stochastic integrals (see the Appendix).

GENERALIZATION OF ITO'S INTEGRAL FOR GENERAL STOCHASTIC PROCESS Purpose is to extend the class of process for which the stochastic integral w.r.t Brownian motion can be defined. We define a class of stochastic process consisting of those can be approximated by simple process. The set-up of those process is as follows.

Consider stochastic process $X_t, t \geq 0$ that allows continuous variation over time and also jumps (not restricted to simple process). We assume the following conditions:

(a) $\{X_t : t \geq 0\}$ is $\mathbb{F}$ adapted

(b) For any T, the r.v. X_t is square integrable which is

$$\mathbb{E}\left[\int_0^T X_t^2 \mathrm{dt}\right] < \infty$$

Notice the basic and simple process satisfy this property as well.

Then to approximate the process $\int_0^T X_t^2 \mathrm{d}W_t$ we approximate X by simple process. To construct such a process first divide $[0, T]$ into n partitions $[0, T] = \cup_{i=0}^{n-1}(t_i, t_{i+1}]$. Then we define the sequence

$$X_t^{(n)} = X_{t_i}, t \in (t_i, t_i + 1]$$

which gives a left continuous simple process. Then we could show $\forall t$

$$\lim_{n \longrightarrow \infty} X_t^{(n)} = X_t$$

where the distance measure for converges is in the following manner

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \left(X_t^{(n)} - X_t\right)^2 dt\right] = 0$$

Hence the stochastic integral is given by

$$I_t = \int_0^t X_s \, \mathrm{d}\, W_s = \lim_{n\to\infty} \int_0^t X_s^{(n)} \, \mathrm{d}\, W_s \; \forall t \in [0, T]$$

The followinig theorem is a summary of the above generalization.

9.9 THEOREM. *$\forall T > 0$, assume $X = \{X_t : t \geq 0\}$ is a $\mathcal{F}_t$-measurable process built on $(\omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ s.t.*

$$\mathbb{E}\left\{\int_0^T X_t^2 dt\right\} < \infty$$

Then the integral

$$I_t = \int_0^t X_s dW_s$$

can be computed using

$$I_t = \int_0^t X_s \, d \, W_s = \lim_{n \to \infty} \int_0^t X_s^{(n)} \, d \, W_s \ \forall t \in [0, T]$$

and I_t *admits the following properties:*

(a) Adaptivity: $\forall t$, I_t *is* $\mathcal{F}_t$*-measurable*

(b) Linearity: I_t *is linear in* X_s *which is*

$$\int_0^t X_s + Y_s dW_s = \int_0^t X_s dW_s + \int_0^t Y_s dW_s$$

where both X_s *and* Y_s *must hold for square integrable*

(c) Martingale: I_t *is a Martingale*

(d) Ito isometry:

$$E^{\mathbb{P}}\left[\int_0^t X_s^2 ds\right] = E^{\mathbb{P}}\left[\left(\int_0^t X_s dW_s\right)^2\right]$$

9.10 EXAMPLE. Now consider the Brownian motion and we would like to compute its integral. We assume the Brownian motion is square integrable (How to show this?). Then we using a eimple process to approximate W_s and eventually we arrive at the follwoing conclusion:

9.11 FACT. Let W_t be a standard Brownian motion. By using the simple process to approximate W_t we get

$$I_t = \int_0^T W_s dW_s = \frac{1}{2} W_T^2 - \frac{1}{2} T$$

where the 1/2 T is called the Ito correction computed using quadratic variation. Notice using RS integral the result is simply

$$\int_0^T x dx = \frac{1}{2} x^2 \Big|_0^T = \frac{1}{2} x^2$$

there's no correction. The detail computation process see written lecture note 9 page 11 - 13. Make up later.

10 STOCHASTIC DIFFERENTIAL EQUATION

Recall that the evolution of a deterministic process through time can be described using ODE while if the process is stochastic we need to introduce SDE. To compare for example

$$dX(t) = \mu(t, X(t))dt$$

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW_t$$

where the part involves W_t is stochastic Brownian motion.

10.1 Convergence Review

Recall $X = Y$ a.s./with probability 1/ a.everywhere means

$$\mathbb{P}(\omega \in \Omega : X(\omega) \neq Y(\omega)) = 0$$

this is equivalent to

$$\exists A \subseteq \Omega \text{ s.t. } \mathbb{P}(A) = 1, \ \ \forall \omega \in A \ X(\omega) = Y(\omega)$$

In our context we well dealing with the following zero-measure set

$$\mathcal{N} =: \{B \in \mathcal{F} : \ \mathbb{P}(B) = 0\} \subset \mathcal{F}_0$$

which is the collection of zero-measure events. The following lemma is useful.

10.1 LEMMA. *Let $X = Y$ almost surely and X is $\mathcal{F}_t$ measurable. Then Y is also $\mathcal{F}_t$ measurable*

Finally we will work on the filtration $\mathbb{F}$

$$\mathbb{F} := \left\{ \mathcal{F}_t, t \geq 0 : \mathcal{F}_t = \sigma(W_s : s \in [0, t]) \cup \mathcal{N} \right\}$$

10.2 Scaled Random walk (s.r.w)

10.2 DEFINITION. Consider time horizon $[0, T]$ divided evenly into n sub-intervals with length $\Delta t = T/n$. Then the s.r.w process is

$$W_k = W_{k-1} + \sqrt{\Delta t} Z_k, \ \ Z_k \sim_{i.i.d} \text{Bernoulli}((1,\text{-}1), p = 1/2) \ k = 1, 2, \cdots, n$$

Properties for s.r.w easy to shows:

(a) The process is equivalent to

$$W_k = W_0 + \sum_{j=1}^{k} \sqrt{\Delta t} Z_j$$

(b) W_k has 0 expectation and variance $n\Delta t = T$. Notice that

$$W_0 =_{a.s.} 0$$

so it has 0 variance.

(c) It has similar properties as Brownian motion which is for any disjoint intervals $(u, v) \cap (s, t) = \emptyset$

$$W_v - W_u \perp W_t - W_s$$

and thus covariance of the difference is 0. Further more

$$\begin{cases} \mathbb{E}(W_t - W_s) = 0 \\ \text{Var}(W_t - W_s) = t - s \text{shouldn't it be} (t-s)\Delta t \end{cases}$$

(d) By central limit theorem we have

$$\begin{aligned} W_n &= \sum_{i=1}^{n} \sqrt{\Delta t} Z_i \\ &= \sum_{i=1}^{n} \sqrt{\frac{T}{n}} Z_i \\ &= \sqrt{T} \left[\frac{\sum_{i=1}^{n} Z_i}{n} \sqrt{n} \right] = \sqrt{T} \left[\frac{(\overline{Z}_n - 0)\sqrt{n})}{1} \right] \longrightarrow_d \sqrt{T} Z, Z \sim N(0,1) \end{aligned}$$

recall that converges in distribution here means

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{W_n}{\sqrt{T}} < z \right) = \Phi(z)$$

so more generally random walk involves sum of i.i.d random variables not necessarily Bernoulli r.v..

10.3 Short Review of ODE

Consider $f(t)$ to be the price of an asset. Let the behaviour of the price follows

$$f(t + \Delta t) - f(t) = \mu \Delta t f(t)$$

Why this assumption make sense? The variation over $[t, t+\Delta]$ should be bigger if the interval is longer and also, the higher the price at time t the wilder the price can vary. So it make sence. Then by taking the limit $\Delta \to 0$ we have

$$\frac{f(t)'}{f(t)} = \mu$$

Then solve if we have

$$\int_0^t \frac{d}{ds}[\ln f(s)]\,\mathrm{ds} = \int_0^t \mu \mathrm{ds} \implies f(x) = ce^{\mu t}$$

eventually given initial condition $f(0) = f_0$ we can solve for c and then the whole $f(x)$ is known. In short, since the process is deterministic in time t, so once we modeled the infinitesimal behaviour, we will be able to determine the price at any time.

10.4 Stochastic differential equation

Approaching SED via ODE. Now we consider that there can be some unpredictable random components of the price. Besides the deterministic trend, we added a stochastic components so the process becomes

$$S_{t+\Delta t} - S_t = \mu S_t \Delta t + \sigma S_t \sqrt{\Delta t}\xi_t$$

The stochastic part is kind of an analogy of its deterministic counterpart: the random change is proportional to price S_t at the beginning with constant σ while the difference is that the "variance" part is not as deterministic as $\sqrt{\Delta t}$ but a r.v. ξ_t instead. We assume that

$$\xi \sim_{i.i.d} N(0,1), \quad \xi_t \perp \{S_u : u \in [0,t]\}$$

The independence actually stress on that fact that utilizing information till t inclusive (e.g. the observed price of S till t), ξ_t is not predictable. Notice S_t is $\mathcal{F}_t$ measurable. If the above process is satisfied we mean that

$$\Pr\left[\left\{\omega \in \Omega : S_{t+\Delta t}(\omega) - S_t(\omega) = \mu S_t(\omega)\Delta t + \sigma S_t(\omega)\sqrt{\Delta t}\xi_t(\omega)\right\}\right] = 1$$

Conditional mean and var   Consider the random part only. We have

$$\begin{aligned} E\left[\sigma S_t \sqrt{\Delta t}\xi_t \mid \sigma\{S_u : u \in \{0, \Delta t, \ldots, t\}\}\right] &= \sigma S_t \sqrt{\Delta t} E[\xi_t] \\ &= 0 \\ E\left[\left(\sigma S_t \sqrt{\Delta t}\xi_t\right)^2 \mid \sigma\{S_u : u \in \{0, \Delta t, \ldots, t\}\}\right] &= \sigma^2 S_t^2 \Delta t E\left[\xi_t^2\right] \\ &= \sigma^2 S_t^2 \Delta t \end{aligned}$$

Prove is kindergarten level but remember if X_t is $\mathcal{F}_t$ measurable then any real valued function $f(X_t)$ is $\mathcal{F}_t$ measurable. So the higher the Δt or σ the more dispersed the stochastic value around its conditional mean.

Consider a scaled random walk with incremental Z_k, $Z_k \sim N(0,1)$ (e.g. $W_k = W_{k-1} + Z_k$) Then we can rewrite the process as

$$S_{t+\Delta t} - S_t = \mu S_t \Delta t + \sigma S_t \sqrt{\Delta t}\xi_t$$

Instead of considering random walk as shown in the written notes. Consider that the most important property we have to keep with the ξ is that it is independent of $\{S_u : u \in [0,t]\}$. Recall $\sqrt{\Delta t}\xi_t \sim N(0, \Delta t)$ then we rewrite it by a Brownian Motion

$$S_{t+\Delta t} - S_t = \mu S_t \Delta t + \sigma S_t (W_{t+\Delta t} - W_t)$$

where the increment has exactly the same distribution and properties as ξ and also notice the whole process is then $\mathcal{F}_{t+\Delta t}$-measurable. Finally the SDE we obtained becomes

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

This SDE fit a more general form

$$dX_t = b(X_t, t)dt + a(X_t, t)dW_t \qquad (*)$$

where $b(X_t, t)$ is called the *drift coefficient* and the $a(X_t, t)$ is called the *diffusion coefficient*. Further more, recall that the BM W_t is not differentiable. So what we mean by $(*)$ is actually the integral form

$$\underbrace{\int_0^t dX_s}_{X_t - X_0} = \int_0^t b(X_s, s)\, ds + \underbrace{\int_0^t a(X_s, s)\, dW_s}_{\text{stochastic integral}}$$

The solution to an SDE is a stochastic process while for ODE is a deterministic process. To solve for SDE, we have to use Ito's lemma.

10.5 Ito's Lemma

Ito's lemma is the counterpart of fundamental theorem of calculus in stochastic calculus. Ito's lemma has different manifestation. Several of them are shown below. Ito's lemma eventually provides us with a SED that is satisfied by the given function $f(W_s)$.

10.3 THEOREM. *(Ito's lemma 1st version) Let W be a Brownian motion built on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and let $f : \mathbb{R} \to \mathbb{R}$ be a function, the first two* *derivatives of which exist and are continuous. Then* $\forall 0 \le t \le T$,

$$f(W_t) - f(W_0) \overset{\mathbb{P}-a.s.}{=} \int_0^t \frac{df}{dw}(W_s)\, dW_s + \frac{1}{2}\int_0^t \frac{d^2 f}{dw^2}(W_s)\, ds$$

In its differential form, equation (8) is written

$$df(W_t) = \frac{df}{dw}(W_t)\, dW_t + \frac{1}{2}\frac{d^2 f}{dw^2}(W_t)\, dt$$

which is the SDE that $f(W_t)$ *satisfies.*

Example see page 32 of slide 11.

10.4 THEOREM. *(Ito's lemma 2nd version) Let* W *be a Brownian motion built on the filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *and let* $f : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ *be a function, the first and second partial derivatives of which exist and are continuous. Then* $\forall 0 \le t \le T$

$$\begin{aligned} & f(W_t, t) - f(W_0, 0) \\ &= \int_0^t \left(\frac{\partial f}{\partial s}(W_s, s) + \frac{1}{2}\frac{\partial^2 f}{\partial w^2}(W_s, s) \right) ds \\ &\quad + \int_0^t \frac{\partial f}{\partial w}(W_s, s)\, dW_s \end{aligned}$$

In its differential form, we have

$$df(W_t, t) = \left(\frac{\partial f}{\partial t}(W_t, t) + \frac{1}{2}\frac{\partial^2 f}{\partial w^2}(W_t, t) \right) dt + \frac{\partial f}{\partial w}(W_t, t)\, dW_t$$

No matter for which type of solution, remember there's a 1/2 multiplied by the second derivative of $f(W_t, t)$.

10.5 EXAMPLE. (Black Scholes model) Consider the risky asset price to be

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right]$$

so $S_t = f(W_t, t)$ where W_t is a Brownian Motion. Using Ito's lemma, the SDE that S_t satisfies easy to find. We directly looking for the differential form and then restore the correct integral form which are

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

$$S_t = S_0 + \int_0^t \mu S_n \mathrm{d}n + \int_0^t \sigma S_n \mathrm{d}W_n$$

which is the SDE solved by S_t. Some comments:

* Since $W_t \sim N(0, t)$ then

$$\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t \sim N\left(\left(\mu - \frac{\sigma^2}{2}\right) t, \sigma^2 t\right)$$

  and thus S_t is log-normally distributed. Again, take logarithm of a log-normal r.v. lead to a normal r.v. so inversely take exponential of a normal r.v. lead to a log-normal r.v.

* The part S_0 is no doubt a stochastic component (e.g. $S_0 = S_0(\omega)$ depending on the states of the world. Further more we can compute the expectation and variance of S_t which are

$$\begin{aligned} E[S_t] &= E\left[s_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t}\right] \\ &= e^{\left(\mu - \frac{\sigma^2}{2}\right)t} E[S_0] E\left[e^{\sigma W_t}\right] \\ &= e^{\left(\mu - \frac{\sigma^2}{2}\right)t} E[S_0] e^{0*\sigma + \frac{1}{2}t\sigma^2} = e^{\mu t} E[S_0] \end{aligned}$$

  The 3rd equality is due to the fact about the moment generating function of a normal random variable: If $X \sim N(m, v)$, then

$$M_t(X) = \mathbb{E}\left[\exp(mt + \frac{1}{2}vt^2)\right]$$

  Similarly we can check that

$$\mathbb{E}(X_t^2) = E\left[S_0^2\right] e^{2\mu t} e^{\sigma^2 t}, \ \mathrm{Var}(S_t) = E\left[S_0^2\right] e^{2\mu t} e^{\sigma^2 t} - E[S_0]^2 e^{2\mu t}$$

10.6 DEFINITION. (Ito's process) An Itô process is defined to be a process $X = \{X_t : 0 \leq t \leq T\}$ taking its values in $\mathbb{R}$ such that:

$$X_t \equiv X_0 + \int_0^t K_s ds + \int_0^t H_s dW_s$$

where $K = \{K_t : 0 \leq t \leq T\}$ and $H = \{H_t : 0 \leq t \leq T\}$ are processes adapted to the filtration $\{\mathcal{F}_t\}$, satisfying $\mathbb{P}\left[\int_0^T |K_s|\, ds < \infty\right] = 1$ and $\mathbb{P}\left[\int_0^T |H_s|^2\, ds < \infty\right] = 1$

10.7 DEFINITION. (Total variation of a stoc. process) Let $[0, T]$ be partitioned into $\pi = \{t_0 = 0, t_1, \cdots, t_n = T\}$. Define the norm of a partition π as

$$\|\pi\| = \max_{k=1,\cdots,n} \{t_k - t_{k-1}\}$$

which is the length of the largest interval. Then the total variation of a stochastic process $X = \{X_t : t \geq 0\}$ is defined as

$$TV_T^X = \lim_{\|\pi\|\to 0} \sum_{k=1}^{n} \left|X_{t_k} - X_{t_{k-1}}\right|$$

The total variation is the limit of the sum of increment. It can be viewed as the length of a sample path $\{X_t(\omega) < t \in [0, T]\}$. It can also be view as the measure of differentiability: for example, if the function $f(t)$ is differentiable then its total variation is finite since

$$\begin{aligned}
TV_T^f &= \lim_{\|\pi\|\to 0} \sum_{i=1}^{n} \left|f_{t_k} - f_{t_{k-1}}\right| \\
&= \lim_{\|z\|\to 0} \sum_{i=1}^{n} |\, f'_{t_k^*} \underbrace{(t_{k-1} - t_k)}_{\Delta t_k > 0} | \\
&= \lim_{\|\pi\|\to 0} \sum_{T=1}^{n} \left|f'_{t_k^*}\right| \Delta t_k \\
&\leqslant \lim_{\|\pi\|\to 0} \sum_{i=1}^{n} \sup_{t_k} \left|f'_{t_k^*}\right| \Delta t_k = \lim_{\|\pi\|\to 0} \sup_{t_k} |\, f'_{t_k^*} \,| \underbrace{\sum_{k=1}^{n} \Delta t_k}_{=T} < +\infty
\end{aligned}$$

the supreme of $f'_{t_k^*}$ is finite since f is differentiable.