Seoul National University

Lecture Note

Introduction to Stochastic Differential Equations

Lecture by Seo Insuk
Notes taken by Lee Youngjae

December 12, 2018

Contents

Chapter 0

Introduction

E-mail: *insuk.seo@snu.ac.kr, 27-212*
Office Hour: Tuesday 15:00 - 16:00
Grading

- Mid-terms 1 (15%, 10/10 or 17)
- Mid-terms 2 (15%, 11/7)
- Fianl-term (40%)
- Assignment (20%, 8-10 times)
- Attendance (10%, absent: -2%, late: -1%)

Let X be a standard normal random variable in $\mathbb{R}$. i.e., $\mathbb{P}[X \in [a,b]] = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$.
(Central Limit Theorem) If $x_1, x_2, \cdots, x_n \in X, E(x_i) = m, Var(x_i) = \sigma^2$, then

$$\frac{\frac{x_1-m}{\sigma} + \frac{x_2-m}{\sigma} + \cdots + \frac{x_n-m}{\sigma}}{\sqrt{n}} \to X$$

In this class, we study dynamic version of this theorem. If $(W_t)_{t\geq 0}$ be a fluctuation, then $(W_t)_{t\geq 0}$ be a random variable in $C[0,T]$

Example. $\frac{dX_t}{dt} = rX_t; dX_1 = rX_t dt$. Then, $X_t = X_0 e^{rt}$ (unrisky assets, bank)
$dX_t = rX_t dt + \sigma X_t dW_t$, σ : volatility (risky assets, stock)

We will study:

1. Probability Space
2. Random Variable
3. Expectation

Textbooks:

1. Stochastic Calculus for Finance II (Shreve), covering chapter 1-3 or 4
2. Introduction to Stochastic Integration (Hui-Hsiung Kuo)

Part I

Stochastic calculus for finance

Chapter 1

General Probability Theory

1.1 Infinite Probability Spaces

There are three elements consisting probability space:

- S: Sample space
- $\mathcal{E}$: Family of events $\mathcal{E} \subseteq 2^S$ (σ-algebra in measure theory)
- $\mathbb{P}$: probability $\Rightarrow \mathbb{P}(E)$ is defined for all $E \in \mathcal{E}$ (μ with $\mu(S) = 1$)

Example.

1. Toss a coin twice (H for Head, T for Tail)
   Then, $S = \{HH, HT, TT, TT\}$
2. Uniform random variable in $[0, 1]^3$
   Then, $S = [0, 1]^3$. If $E = [0, \frac{1}{2}]^3$, then $\mathbb{P}(E) = Vol(E) = \frac{1}{8}$

How to define $\mathcal{E}$?
In example 2, let $\mathcal{E}$ = family of all subsets of $[0, 1]^3$ naively. But Banach-Tarski Paradox says there are disjoint sets E, F with $\mathbb{P}(E \cup F) \neq \mathbb{P}(E) + \mathbb{P}(F)$ in this $\mathcal{E}$. Therefore we cannot naively set $\mathcal{E}$ (Use measure theory)

In example 1, suppose that we cannot see the second flip. If $\{HH\} \notin \mathcal{E}$ and $\{HT, HH\} \in \mathcal{E}$, then $\mathcal{E} = \{\phi, \{HH, HT\}, \{TH, TT\}, \{HH, HT, TH, TT\}\}$

Definition 1.1.1 (Measure)
Let Ω be a non-empty set and $\mathcal{F}$ be family of subsets of Ω with

1. $\phi \in \mathcal{F}$

2. $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$

3. $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

We say $\mathcal{F}$ as **σ-algebra** or **σ-field**, $A \in \mathcal{F}$ as **measurable**, and Ω as **measurable space**.

Exercises.

1) $\Omega \in \mathcal{F}$

2) $A_1, A_2, \cdots \in \mathcal{F}$, then $A_1 \cap A_2 \cdots \in \mathcal{F}$

3) $A_1, A_2, \cdots \in A_n \in \mathcal{F}$, then $A_1 \cup \cdots \cup A_n, A_1 \cap \cdots \cap A_n \in \mathcal{F}$.

4) $A, B \in \mathcal{F}$, then $A - B \in \mathcal{F}$

Definition 1.1.2 (Topological Space)
(See Rudin: *Real and Complex Analysis, Chapter 1.*) Let Θ be non-empty set and τ be family of subsets of Θ with

1. $\phi, \Theta \in \tau$

2. $V_1, \cdots V_n \in \tau \Rightarrow V_1 \cap \cdots \cap V_n \in \tau$

3. $V_\alpha \in \tau \ \ \forall \alpha \in I \Rightarrow \bigcup_{\alpha \in I} V_\alpha \in \tau$.

We say $V \in \tau$ be an **open set**, and (Θ, τ) be a **topological space**.

Definition 1.1.3 (Measurable Function)
$f : (\Omega, \mathcal{F}) \to (\Theta, \tau)$ is **measurable** if $f^{-1}(V) \in \mathcal{F} \ \ \forall V \in \tau$

Definition 1.1.4 (Positive Measure)
Let Ω be non-empty set and $\mathcal{F}$ be σ-algebra. Then $\mu : \mathcal{F} \to [0, \infty]$ is called **measurable** if

1. $A_1, A_2, \cdots$: disjoint members of $\mathcal{F} \Rightarrow \mu(A_1 \cup A_2 \cup \cdots) = \sum_{i=1}^{\infty} \mu(A_i)$

2. $\mu(A) < \infty$ for some $A \in \mathcal{F}$,

and $(\Omega, \mathcal{F}, \mu)$ is called a **measure space**.

Definition 1.1.5 (probability space, random variable)

1. $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **probability space** if $\mathbb{P}(\Omega) = 1$.

2. X is called a **random variable** if it is a function from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\mathbb{R}$

<u>Next Class</u>

- Borel sets on $\mathbb{R}$ or $\mathbb{R}^d$
- Lebesgue Measure
- Lebesgue Integral (Define Expectation of random variable)

Last class, we define a sample space Ω, a σ-algebra $\mathcal{F}$, and a (positive) measure $\mu : \mathcal{F} \to [0, \infty]$.

Exercises.

- $A_1 \subseteq A_2 \subseteq \cdots \Rightarrow \mu(\bigcup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mu(A_n)$
- $A_1 \supseteq A_2 \supseteq \cdots, \mu(A_1) < \infty \Rightarrow \mu(\bigcap_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mu(A_n)$

Theorem 1.1.6 (Rudin 1.10)
Let $\mathcal{F}_0$ be a collection of subset of Ω. Then, $\exists! \mathcal{F}^*$ minimal σ-algebra containing $\mathcal{F}_0$.

Proof. Let $\{\mathcal{F}_\alpha, \alpha \in I\}$ be a family of σ-algebra containing $\mathcal{F}_0$. Then, $\mathcal{F}^* = \bigcap_{\alpha \in I} F_\alpha$ satisfies the three condition: 1) contain $\mathcal{F}_0$ 2) σ-algebra 3) minimal (trivial, $\mathcal{F}^* \subseteq \mathcal{F}_\alpha$) $\square$

Definition 1.1.7 (Borel measurable)
$\mathcal{B}$ is a **Borel σ-algebra** on the topological space (Θ, τ) if $\mathcal{B}$ is a minimal σ-algebra containing τ, and B is a **Borel measurable** if $B \in \mathcal{B}$.

Remark (Completion of measure space, Rudin 1.15).
Consider an extension $(\Omega, \mathcal{F}, \mu) \to (\Omega, \overline{\mathcal{F}}, \mu)$ where

1. $\overline{\mathcal{F}} = \{A \cup N : A \in \mathcal{F}, N \subseteq A_0 \in \mathcal{F}, \mu(A_0) = 0\}$
2. $\mu(A \cup N) = \mu(A)$

Then, (Check!)

1. (well-definedness) $A_1 \cup N_1 = A_2 \cup N_2 \Rightarrow \mu(A_1) = \mu(A_2)$
2. $\mu : \overline{\mathcal{F}}$ is σ-algebra.
3. $\mu : \overline{\mathcal{F}} \to [0, \infty]$ is a measure

Example.

1) $\mathbb{R}$

$$\mathcal{F}_0 = \tau \xrightarrow{1.10} \mathcal{B} \xrightarrow{\text{completion}} \overline{\mathcal{B}}$$

$$\mathcal{L} \xrightarrow{\text{Rudin CH 2}} \mathcal{L} \xrightarrow{\text{completion}} \mathcal{L}$$

2) $C[0,T] = \Omega = \{f; f : [0,T] \to \mathbb{R}, \text{continuous}\}$.
Define $\mathcal{F}_0 = \{\bigcup_{t_1,t_2,\cdots,t_k}(A_1, A_2, \cdots, A_k) : 0 \le t_1 < t_2 < \cdots < t_k \le T; A_1, \cdots A_k \in \overline{\mathcal{B}}\}$. We call $\{f \in C[0,T] : f(t_1) \in A_1, f(t_2) \in A_2, \cdots, f(t_k) \in A_k\}$ as **cylindrical set**. Consider

$$\mathcal{F}_0 \xrightarrow{1.10} \quad \mathcal{B} \xrightarrow{\text{completion}} \quad \overline{\mathcal{B}}$$

$$\mathbb{P}_{\text{BM}} \xrightarrow{\text{KET}} \mathbb{P}_{\text{BM}} \xrightarrow{\text{completion}} \mathbb{P}^*_{\text{BM}}$$

(KET refers Kolmogorov's Extension Thm)

1.2 Random Variables and Distributions

Definition 1.2.1
$f : \Omega \to \mathbb{R}$ is measurable if $f^{-1}(V) \in \mathcal{F}$ for any open set $V \subseteq \mathbb{R}$.

Remark. $\mathcal{B}(\mathbb{R})$ = Borel σ-algebra in $\mathbb{R}$.

Remark. If f: measurable, then $f^{-1}(B) \in \mathcal{F}$ for any $B \in \mathcal{B}(\mathbb{R})$.

Proof. Let $G = \{A \subseteq \mathbb{R} : f^{-1}(A) \in \mathcal{F}\}$. Then, $\tau \subseteq G$, G : σ-algebra (check!), hence $\mathcal{B}(\mathbb{R}) \subseteq G$. □

Definition 1.2.2

- $(\Omega, \mathcal{F}, \mathbb{P})$ is a **probability space** if $(\mathbb{P}(\Omega) = 1$.
- X is a **random variable** if $X : \Omega \to \mathbb{R}$ is measurable.

Example.

1. Toss a coin Twice.
$\Omega = \{HH, HT, TH, TT\}$, $\mathcal{F} = 2^\Omega = \{\text{all subsets of } \Omega\}$, $\mathbb{P}(A) = \frac{1}{4}|A|$, $A \in \mathcal{F}$.
Then, X = the number of H's is a random variable with $X(HH) = 2, X(HT) = X(TH) = 1, X(TT) = 1$.

2. Uniform random variable in $[0,1]$
$\Omega = [0,1]$, $\mathcal{F} = \{B \in \mathcal{B}(\mathbb{R}) : B \subseteq [0,1]\}$, $\mathbb{P}(B) = \mathcal{L}(B)$ $(\mathbb{P}([0,1]) = \mathcal{L}([0,1]) = 1)$.
Then, $X : [0,1] \to \mathbb{R}$ with $X(x) = 1$ is a (uniform) random variable in $[0,1]$.

Remark. $\mathcal{L}$: Lebesgue measure on $\mathbb{R}$. i.e.,$\mathcal{L}(a,b) = b - a$. Then, $\mathcal{L}(\{a\}) = 0$
$(\because \{a\} = \bigcap_{i=1}^{\infty}(a - \frac{1}{n}, a + \frac{1}{n}) \Rightarrow \mathcal{L}(\{a\}) = \lim_{n\to\infty} \mathcal{L}((a - \frac{1}{n}, a + \frac{1}{n})) = 0)$
Similarly, $\mathcal{L}([a,b]) = \mathcal{L}([a,b)) = \mathcal{L}((a,b]) = b - a$, $\mathcal{L}(\mathbb{Q}) = \sum_{q\in\mathbb{Q}} \mathcal{L}(\{q\}) = 0$.

Return to uniform random variable,

$$\mathbb{P}[X \in (a,b)] = \mathbb{P}[\{x : X(x) \in (a,b)\}] = \mathbb{P}[(a,b)] = b - a.$$

Definition 1.2.3 (Distribution measure on X)
X is a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$. μ_X is a **distribution measure** on X if μ_X is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\mu_X(B) = \mathbb{P}[X \in B] = \mathbb{P}[\{\omega : X(\omega) \in B\}] = \mathbb{P}[X^{-1}(B)]\ \forall B \in \mathcal{B}(\mathbb{R})$$

Definition 1.2.4 (Probability density function)
f is a **probability density function** of X if $\mu_X((a,b)) = \int_a^b f(x)dx$

Remark. There is a measure with no pdf: Dirac measure

Remark. Lebesgue-Radon-Nikodym decomposition implies that any measure can be decomposed as density part and singular part.

Example (Standard Normal random variable).
Let $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Define $F : (0,1) \to \mathbb{R}$ by $F(x) = N^{-1}(x)$ for $N(X) = \int_{-\infty}^{x} \phi(y)dy$.
Let $\Omega = (0,1), \mathcal{F} = \{B \in \mathcal{B}(\mathbb{R}) : B \subseteq (0,1)\}, \mathbb{P}(A) = \mathcal{L}(A) : A \in \mathcal{B}(\mathbb{R})$.
Then, $Y : \Omega \ni x \mapsto F(x) \in \mathbb{R}$ is a random variable with

$$\begin{aligned}\mathbb{P}[Y \in (a,b)] &= \mathbb{P}[\{x : Y(x) \in (a,b)\}] \\ &= \mathbb{P}[\{x \in (N(a), N(b))\}] \\ &= N(b) - N(a) = \int_a^b \phi(x)dx,\end{aligned}$$

and a density function is ϕ.

Previous Question: In the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variable $X : \Omega \to \mathbb{R}$, the random element or random realization $\omega \in \Omega$ is a element of events in sample space. For example, $\omega = HHTTH$ is a random element in tossing a coin five times, and $X(\omega) = 3$. ($X(\omega)$ = # of Heads)
In the previous example(Standard Normal random variable), define $(\Omega, \mathcal{F}, \mathbb{P}) = ((0,1), \mathcal{B}(0,1), \mathbb{P})$, $\mathbb{P}((a,b)) = b - a$, $F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$, $X : (0,1) \ni \omega \mapsto F^{-1}(\omega) \in \mathbb{R}$. Then, X is called a standard normal random variable.

1.3 Expectations

In the following, let $\Omega = (\Omega, \mathcal{F}, \mathbb{P})$. Let $X : \Omega \to \mathbb{R}$. Then the expection $\mathbb{E}(X)$ is a mean of $X(\omega)$ with respect to the randomness of ω (given by $\mathbb{P}$)

Definition 1.3.1 (Lebesgue Integration)
$(\Omega, \mathcal{F}, \mu)$ is a measure space, and $f : \Omega \to \mathbb{R}$ is a measurable function.

(1) $f : \Omega \to [0, \infty)$
Let $0 = y_0 < y_1 < y_2 < \cdots \to \mathbb{R}$ be a partition of $[0, \infty)$,
$\Pi = \{y_0, y_1, y_2, \cdots\} : \|\Pi\| = \sup_{i \geq 1} |y_i - y_{i-1}|$, and
$\text{LS}_\pi = \sum_{i=0}^{\infty} y_i \mu[f^{-1}([y_i, y_{i+1}))]$.
In Rudin's book, $\lim_{\|\Pi\| \to 0} \text{LS}_\Pi$ converges to an element belonging to $[0, \infty]$.
Now, $\int f d\mu := \lim_{\|\Pi\| \to 0} \text{LS}_\Pi$ is called a **Lebesgue Integral**.

(2) $f : \Omega \to \mathbb{R}$
Let $f^+ = \max\{f, 0\} \geq 0$, and $f^- = -\min\{f, 0\} \geq 0$. Then, $f = f^+ - f^-$, and $|f| = f^+ + f^-$. If $\int f^+ d\mu < \infty$ and $\int f^- d\mu < \infty$, then we say f is Lebesgue integrable and $f \in L^1(\mu)$. The Lebesgue integral of $f = \int f d\mu$ is defined as $\int f^+ d\mu - \int f^- d\mu$

Remark.

1. $\int f^+ d\mu < \infty$ and $\int f^- d\mu = \infty$, then $\int f d\mu = -\infty$. The others are defined similarly.

2. $f \in L^1(\mu) \Leftrightarrow \int |f| d\mu < \infty$.

Example (Riemann vs Lebesgue Integral (p19-22)).

- $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{L})$ Lebesgue measure where $\mathcal{L}((a, b)) = b - a$.

- $f : \mathbb{R} \to \mathbb{R} \in L^1(\mathcal{L})$

- (Def) $A \subseteq \mathbb{R}$, $\int_A f d\mu := \int f \mathbb{1}_A d\mu$, where $\mathbb{1}_A(x) = 1$ if $x \in A$, and 0 otherwise.

If f is Riemann integrable, then $\int_{[a,b]} f d\mathcal{L} = \int_a^b f dx$.
Riemann integral is a limit of approximation by a partition of x-axis. On the other hand, Lebesgue integral is a limit of approximation by a partition of y-axis with preimage. Partition of x-axis is sensitive to fluctuation and restricted to Euclidean space, while partition of y-axis is not. For example, $f(x) = \mathbb{1}_{\mathbb{Q}}(x)$ is Lebesgue integrable, but it is not Riemann integrable since it is sensitive to fluctuation.

Definition 1.3.2 (Almost everywhere, 1.1.5 in Textbook)
$P(x)$ is a property at $x \in \mathbb{R}$. We say P holds **almost everywhere** (or a.e.) in $\mathbb{R}$ if and only if $\mathcal{L}(\{x : P(x) \text{ does not hold }\} = 0$.

Example. $f(x) = [x]$ is continuous almost everywhere.

Theorem 1.3.3
f is Riemann integrable if and only if f is continuous a.e.

Exercises. $f = g$ a.e. $\Rightarrow \int f d\mathcal{L} = \int g d\mathcal{L}$.

Definition 1.3.4 (Almost surely)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. The event $A (\in \mathcal{F})$ occurs **almost surely** (a.s.) if $\mathbb{P}(A) = 1$.

Example. Let X be a uniform random variable in $(0,1)$. Let $A = \{X(\omega) \neq \frac{1}{2}\}$; $\mathbb{P}(A) = 1$.

Definition 1.3.5 (Expectation, 1.3.3. in Textbook)
Expectation of $X : \Omega \to \mathbb{R}$ is defined by

$$\mathbb{E}(X) := \int_\Omega X d\mathbb{P} \quad \text{if} \quad \int_\Omega |X| d\mathbb{P} < \infty$$

Theorem 1.3.6 (1.3.4 in Textbook)

1. *X takes finite number of values $\{x_1, x_2, \cdots, x_n\} \Rightarrow \mathbb{E}(X) = \sum_{i=1}^n x_i \mathbb{P}(X = x_i)$*
2. *X, Y: random variables, $E(|X|), E(|Y|) < \infty$,*
   (i) *$X \leq Y$ a.s. (i.e. $\mathbb{P}[\{X(\omega) \leq Y(\omega)\}] = 1$), then $\mathbb{E}(X) \leq \mathbb{E}(Y)$*
   (ii) *$X = Y$ a.s. $\Rightarrow \mathbb{E}(X) = \mathbb{E}(Y)$*
3. *X, Y: random variables, $\mathbb{E}(|X|), \mathbb{E}(|Y|) < \infty \Rightarrow \mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E}(X) + \beta \mathbb{E}(Y)$.*
4. *Jensen's Inequality: $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function $\Rightarrow \phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X))$*
   (cf. $\phi(t) = t^2$)

Proof of 4. Define $S_\phi = \{(a,b) \in \mathbb{R}^2 : a + bt \leq \phi(t) \ \forall t\}$. Then $\forall t \in \mathbb{R}, \phi(t) = \sup_{(a,b) \in S_\phi} \{a + bt\}$. In fact, it is a equivalent condition. Now,

$$\begin{aligned}
\phi(\mathbb{E}[X]) &= \sup_{a,b \in S_\phi} \{a + b\mathbb{E}[X]\} \\
&= \sup_{a,b \in S_\phi} \mathbb{E}[a + bX] \\
&\leq \mathbb{E}[\sup_{a,b \in S_\phi} (a + bX)] = \mathbb{E}[\phi(X)] \quad \text{(Check!)}
\end{aligned}$$

□

Example (Dirac Measure in $\mathbb{R}$). $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \delta_y)$ $(y \in \mathbb{R})$ is a probability space with $\delta_y(A) = 1$ if $y \in A$, and 0 otherwise. Then, $\int_{\mathbb{R}} f d\delta_y = f(y)$ (Check!)

Consider modeling: X: random variable such that probability of $x_i = p_i$ with $\sum_{i=1}^n p_i = 1$. Then, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ with $\mu = \sum_{i=1}^n p_i \delta_{x_i}$ is a probability space, and $P(X = x_i) = p_i$ for $X : \mathbb{R} \ni \omega \mapsto \omega \in \mathbb{R}$: Example of thm 1.3.4.

Summary:

- Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$
- Random variables: $X : \Omega \to \mathbb{R}$
- Expectation: $E(X) = \int X d\mathbb{P}$

1.4 Convention of Integrals

We will use this section when we define the Brownian motion.

Definition 1.4.1

(1) Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and $f, f_1, f_2, \cdots$ be measurable ($\Omega \to \mathbb{R}$). Then, $f_n \to f$ **almost everywhere** (a.e.) if

$$\mu[\{\omega : (f_n(\omega))_{n=1}^\infty \text{ does not converge to } f(\omega)\}] = 0$$

(2) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $X, X_1, X_2, \cdots$ be random variables. Then, $X_n \to X$ **almost surely** (a.s.) if

$$\mathbb{P}[\{\omega : (X_n(\omega))_{n=1}^\infty \text{ does not converge to } X(\omega)\}] = 0$$

Question: $f_n \to$ a.e. Then, $\int f_n d\mu \to \int f d\mu$? $X_n \to X$ a.s. Then, $\mathbb{E}[X_n] \to \mathbb{E}[X]$?

Theorem 1.4.2 (Monotone Convergence Theorem. 1.4.5 in Textbook)
$0 \leq f_1 \leq f_2 \leq \cdots$ *(or decreasing), and* $f_n \to f$ *a.e. Then,* $\int f_n d\mu \to \int f d\mu$.

Theorem 1.4.3 (Dominated Convergence Theorem. 1.4.9 in Textbook)
$\exists g \in L^1(\mu)$ *such that* $|f_n| \leq g$ *for all* n*, and* $f_n \to f$ *a.e. Then,* $\int f_n d\mu \to \int f d\mu$.

Corollary 1.4.4
$\exists Y \in L^1(\mathbb{P})$ such that $|X_n| \leq Y$ for all n, and $X_n \to X$ a.s. Then, $\mathbb{E}[X_n] \to \mathbb{E}[X]$.

Example. Let $f_n(x) = \begin{cases} n^2 x & \text{if } 0 \leq x \leq \frac{1}{n}, \\ -n^2 x + n & \text{if } \frac{1}{n} < x \leq \frac{2}{n}, \\ 0 & \text{otherwise.} \end{cases}$ Then, $f_n \to 0$ a.e. and $\int f_n dx = 1$.

1.5 Computation of Expectations

Notation: $(X : \Omega \ni \omega \mapsto X(\omega) \in \mathbb{R})$

- $\mathbb{E}[X] = \int X d\mathbb{P} = \int_\Omega X(\omega) d\mathbb{P}(\omega)$
- $\int_B X(\omega) d\mathbb{P}(\omega) := \int \mathbb{1}_B(\omega) X(\omega) d\mathbb{P}(\omega)$

Recall: X: random variable, μ_X: distribution measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $\mu_X(B) = \mathbb{P}(X \in B)$.

Theorem 1.5.1
$g \in L^1(\mu_X)$. Then, $\mathbb{E}[g(X)] = \int_\mathbb{R} g(x) d\mu_X(x) (:= \int g d\mu_X)$.

Example. $g(x) = x$. $\int |x| d\mu_X(x) < \infty \Rightarrow \mathbb{E}[X] = \int x d\mu_X(x)$.

Proof. First, prove the thm holds for $g \geq 0$, then prove for general g by $g = g^+ - g^-$.

(1) $g = \mathbb{1}_B$
By thm 1.3.4. (1), $E[\mathbb{1}_B(X)] = 1 \cdot \mathbb{P}[\mathbb{1}_B(X) = 1] = \mathbb{P}(X \in B) = \mu_X(B) = \int \mathbb{1}_B(x) d\mu_X(x)$.

(2) $g = \sum_{k=1}^n \alpha_k \mathbb{1}_{B_k}$
Trivial by linearity.

(3) $g \geq 0$
By MCT. See *Rudin* chapter 1 for details.

□

Recall: X : random variable, X has density function f_X if

$$\mu_X((a,b)) = \int_a^b f_X(x) dx \quad \forall a, b.$$
$$\mu_X(B) = \int_B f_X d\mathcal{L} = \int_B f_X(x) d\mathcal{L}(x) = \int_B f_X(x) dx.$$

Theorem 1.5.2
$g \in L^1(\mu_X)$. Then, $E(g(X)) = \int_\mathbb{R} g(x) f_X(x) dx$.

Example. Let X be standard normal. i.e., $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (regardless what X be). Then, $E(X^4) = \int_\mathbb{R} x^4 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = 3$.

Chapter 2

Information and Conditioning

2.1 Information and σ-algebras

Example. Toss a coin Three times. $\Omega = \{HHH, HHT, \cdots, TTT\}$.
$A_H = \{HHH, HHT, HTH, HTT\}$, $A_T = \{THT, THT, TTH, TTT\}$.
Let $\mathcal{F}(1) = \{\phi, \Omega, A_H, A_T\}$ so that it is a σ-algebra containing the randomness up to time 1.
Similarly, define $A_{HH}, A_{HT}, A_{TH}, A_{TT}$.
Let $\mathcal{F}(2) = \{\phi, \Omega, A_{HH}, A_{HT}, A_{TH}, A_{TT}, A_{HH} \cup A_{HT}, \cdots, A_{TT}^C\}$ so that it is a σ-algebra containing the randomness up to time 2, and define $\mathcal{F}(0)$ similarly, and let $\mathcal{F}(0) = \{\phi, \Omega\}$.
Then, $\mathcal{F}(0) \subseteq \mathcal{F}(1) \subseteq \mathcal{F}(2) \subseteq \mathcal{F}(3)$. Let X_t = # of heads until time t. Then, X_t is $\mathcal{F}(t)$-measurable for each t.
Now, $\{X_1 = 1\} = \{\omega : X_1(\omega) = 1\} = A_H$, and $\{X_1 = 0\} = \{\omega : X_1(\omega) = 0\} = A_T$.

Definition 2.1.1 (σ-algebra generated by X)
Ω is a set, $X : \Omega \to \mathbb{R}$. $\sigma(X) = \{A \subseteq \Omega : A = X^{-1}(B) \text{ for some } B \in \mathcal{B}(\mathbb{R})\}$. Then, $\sigma(X)$ is a $\sigma-$algebra(exercise) and it is called a **σ-algebra generated by** X.

Remark. X is a random variable in $(\Omega, \sigma(X))$.
X is a random variable in $(\Omega, \mathcal{F})$, then $\sigma(X) \subseteq \mathcal{F}$ (exercise)

Definition 2.1.2 ($\mathcal{F}$-measurable)
$(\Omega, \mathcal{F})$: measure space. $X : \Omega \to \mathbb{R}$. X is called **$\mathcal{F}$-measurable** if $\sigma(X) \subseteq \mathcal{F}$. i.e., X: measurable with respect to $(\Omega, \mathcal{F})$.

In example, $X(t)$ is $\mathcal{F}(t)$-measurable $\forall t$ (check!)
cf. $X(t) : \Omega \to \mathbb{R}$. $(X(t))^{-1}(B) \in \mathcal{F}(t) \ \ \forall B \in \mathcal{B}(\mathbb{R})$.

Enough to check $(X(t))^{-1}(\{0\}), (X(t))^{-1}(\{1\}), \cdots, (X(t))^{-1}(\{t\})$.
$\mathcal{F}(t)$ has enough information to determine $X(t)$ in the sense that $\{\omega : (X(t))(\omega) \in B\} \in \mathcal{F}(t) \ \forall B \in \mathcal{B}(\mathbb{R})$.

Definition 2.1.3 (Filtration, Stochastic Process)
Ω: non-empty set, $T > 0$.

1. If $\mathcal{F}(t)$ is a σ-algebra $\forall t \in [0, T] \in T$ and $s < t \Rightarrow \mathcal{F}(s) \subseteq \mathcal{F}(t)$, then $(\mathcal{F}(t) : t \in [0, T])$ is called a **filtration**
2. If $X(t) : \Omega \to \mathbb{R}$ is $\mathcal{F}(t)$-measurable $\forall t \in [0, T]$, then $(X(t) : t \in [0, T])$ is called **Stochastic Process adopted to the filtration** $\mathcal{F}(t)$.

2.2 Independence

$X : \Omega \to \mathbb{R}$, $\mathcal{F}$: σ-algebra on Ω.

1. $\mathcal{F}$ has full information to determine $X \Rightarrow X$ is $\mathcal{F}$-measurable. (2.1)
2. $\mathcal{F}$ has no information to determine $X \Rightarrow X$ is independent to $\mathcal{F}$. (2.2)
3. $\mathcal{F}$ has a partition information to determine $X \Rightarrow$ (2.3)

Definition 2.2.1 (independent)
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $A, B \in \mathcal{F}$ is **independent** if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Question: X, Y are random variables in $(\Omega, \mathcal{F}, \mathbb{P})$. If X, Y are independent, then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, but the converse does not hold.

Definition 2.2.2
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ are sub σ-algebras of $\mathcal{F}$. $X, Y : \Omega \to \mathbb{R}$ are random variables in $(\Omega, \mathcal{F}, \mathbb{P})$.

1. $\mathcal{G}, \mathcal{H}$: independent iff $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \ \forall A \in \mathcal{G}, B \in \mathcal{H}$.
2. X, Y: independent iff $\sigma(X), \sigma(Y)$ are independent.
3. $X, \mathcal{G}$: independent iff $\sigma(X), \mathcal{G}$ are independent.

Definition 2.2.3
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.
$\mathcal{G}_1, \mathcal{G}_2, \cdots \mathcal{G}_n, \cdots$: sub σ-algebra of $\mathcal{F}$. $X_1, X_2, \cdots, X_n, \cdots$: random variable in $(\Omega, \mathcal{F}, \mathbb{P})$.

1. $\mathcal{G}_1, \cdots, \mathcal{G}_2$ are independent iff $\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n)$ for $A_1 \in \mathcal{G}_1, \cdots, A_n \in \mathcal{G}_n$.

2. $X_1, \cdots, X_n$ are independent iff $\sigma(X_1) \sim \sigma(X_n)$ are independent.

3. $\mathcal{G}_1, \mathcal{G}_2, \cdots$ are independent iff $\mathcal{G}_1 \sim \mathcal{G}_n$ are independent $\forall n$.

4. $X_1, X_2, \cdots$ are independent iff $X_1 \sim X_n$ are independent $\forall n$.

Example. Toss a coin three times.

1. $X(2), X(3)$ are not independent.
   $\mathbb{P}(\{X(2) = 2\} \cap \{X(3) = 1\}) \neq \mathbb{P}(X(2) = 2)\mathbb{P}(X(3) = 1)$.

2. $X(2), X(3) - X(2)$ are independent.
   Why: $X(2)$ is an information at tossing first, second times, and $X(3)$ is an information at tossing third time.

Definition 2.2.4 (Joint distribution)
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. X, Y are random variables in Ω. $(X, Y) : \Omega \ni \omega \mapsto (X(\omega), Y(\omega)) \in \mathbb{R}^2$

1. Joint Distribution Measure in $\mathbb{R}^2$
   $\mu_{X,Y}(C) = \mathbb{P}((X, Y) \in C)$ for $C \in \mathcal{B}(\mathbb{R}^2)$.
   (Note: We checked that $\{\omega : (X(\omega), Y(\omega)) \in C\} \in \mathcal{F}$ in real analysis.)

2. Joint Cumulative Distribution Function
   $F_{X,Y}(a, b) = \mathbb{P}(X \leq a, Y \leq b) = \mu_{X,Y}((-\infty, a] \times (-\infty, b])$ (check!)

3. Joint Probability Distribution Function
   If $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ is Borel-measurable and satisfies $\mu_{X,Y}(A \times B) = \int_B \int_A f_{X,Y}(x, y)dxdy$ for all $A, B \in \mathcal{B}(\mathbb{R})$, then $f_{X,Y}$ is called a joint probability density function (jpdf)

Theorem 2.2.5
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, X, Y are random variables in Ω. Then, the followings are equivalent.

(i) *X, Y are independent*

(ii) $\mu_{X,Y}(A \times B) = \mu_X(A)\mu_Y(B) \;\; \forall A, B \in \mathcal{B}(\mathbb{R})$

(iii) $F_{X,Y}(a, b) = F_X(a)F_Y(b) \;\; \forall a, b \in \mathbb{R}$

(iv) $\mathbb{E}[e^{uX+vY}] = \mathbb{E}[e^{uX}]\mathbb{E}[e^{vY}]$

Remark. If JPDF $f_{X,Y}$ exists, then (i) to (iv) $\Leftrightarrow f_{X,Y}(x,y) = f_X(x)f_Y(y)$ a.e.

Theorem 2.2.6
X, Y are independent if and only if $f, g : \mathbb{R} \to \mathbb{R}$ Borel-measurable, $\mathbb{E}[|f(X)g(Y)|] < \infty$ implies that $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$.

Remark. $f(x) = g(x) = x : \mathbb{E}[|XY|] < \infty \Rightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

Proof. Details are exercises.

(1) $f = \mathbb{1}_A, g = \mathbb{1}_B$

(2) f, g are simple functions

(3) $f, g \geq 0$

(4) f, g are general.

□

<u>Review</u>
$\mathcal{G}, \mathcal{H}$ are independent if $\forall A \in \mathcal{G}, B \in \mathcal{H} \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
X, Y are independent if $\sigma(X), \sigma(Y)$ are independent.
* $\sigma(X) = \{A \in \Omega : A = X^{-1}(B)$ for some $B \in \mathcal{B}(\mathbb{R})\}$.
* $\mu_{X,Y}(C) = \mathbb{P}((X,Y) \in C) \ \ \forall C \in \mathcal{B}(\mathbb{R}^2)$.

Thm. T.F.A.E.C:

1. X, Y are independent
2. $\mu_{X,Y}(A \times B) = \mu_X(A)\mu_Y(B)$
3. $\mathcal{F}_{X,Y}(x,y) = \mathcal{F}_X(x)\mathcal{F}_Y(y)$
4. (If JPDF $f_{X,Y}$ exists) $f_{X,Y}(x,y) = f_X(x)f_Y(y)$

Theorem 2.2.7
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. X, Y are independent random variables, $f, g : \mathbb{R} \to \mathbb{R}$ are Borel measurable. Then, $f(X), g(Y)$ are independent.

Proof. $A \in \sigma(f(X))$; $A = (f \circ X)^{-1}(B)$ for some $B \in \mathbb{R} = X^{-1}(f^{-1}(B)) \in \sigma(X)$.
$\therefore \sigma(f(X)) \subseteq \sigma(X)$, $\sigma(g(Y)) \subseteq \sigma(Y) \Rightarrow \sigma(f(X)), \sigma(g(Y))$ are independent. □

Corollary 2.2.8
$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X))\mathbb{E}(g(Y)]$.

Definition 2.2.9
X, Y are random variables in $(\Omega, \mathcal{F}, \mathbb{P})$.

1. $Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
2. $std(X) = \sqrt{Var(X)}$
3. $Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
4. $corr(X, Y) = cov(X, Y)/(std(X)std(Y))$

Example.

- X: standard normal random variable $(N(0, 1^2))$
- $\mathbb{P}(Z = 1) = \mathbb{P}(Z = -1) = \frac{1}{2}$ (X, Z are independent)
- $Y = XZ$. Then
  1) Y is standard normal,
  2) $corr(X, Y) = 0$.
  3) X, Y are not independent.

Definition 2.2.10 (Jointly normal)
X, Y are **jointly normal** with mean $m = (m_X, m_Y)$, $Var(C) = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$ if

$$f_{X,Y}(z) = \frac{1}{\sqrt{(2\pi)^2 \det C}} e^{-\frac{1}{2}(z-m)C^{-1}(z-m)^\intercal}$$

Theorem 2.2.11
X, Y are jointly normal and uncorrelated $(C_{12} = C_{21} = 0)$. *Then, they are independent.*

2.3 Conditional Expectation

$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $\int_A X d\mathbb{P} := \int \mathbb{1}_A X d\mathbb{P} = \int \mathbb{1}_A(\omega) X(\omega) d\mathbb{P}(\omega)$.

Lemma. $\int_A X d\mathbb{P} = \int_A Y d\mathbb{P}$ for all $A \in \mathcal{F}$ if and only if $X = Y$ a.s.

Proof. $A_n = \{\omega : X(\omega) - Y(\omega) > \frac{1}{n}\}, B_n = \{\omega : X(\omega) - Y(\omega) < -\frac{1}{n}\}$. Then,

$$0 = \int_{A_n} (X - Y) d\mathbb{P} \geq \int_{A_n} \frac{1}{n} d\mathbb{P} = \frac{1}{n} \int \mathbb{1}_{A_n} d\mathbb{P} = \frac{1}{n} \mathbb{P}(A_n)$$

Thus, $\mathbb{P}(A_n) = 0 \ \forall n$. Similarly, $\mathbb{P}(B_n) = 0$. Now, $\{\omega : X(\omega) \neq Y(\omega)\} = (\bigcup_{n=1}^\infty A_n) \cup (\bigcup_{n=1}^\infty B_n) \Rightarrow$ measure 0. $\square$

Intuition. $(\Omega, \mathcal{F}, \mathbb{P})$ is given, X : $\mathcal{F}$-measurable random variable, $\mathcal{G} \subseteq \mathcal{F}$ is a sub σ-algebra. If we know nothing, then we expect X as $\mathbb{E}[X]$. If we know $\mathcal{F}$, then we expect X as X. Now, if we know $\mathcal{G}$, then we expect X as $\mathbb{E}[X|\mathcal{G}]$ (what is it?)

Definition 2.3.1 (Conditional Expectation)
$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $X \in L^1(\mathbb{P})$ is a random variable. $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$. We define $\mathbb{E}[X|\mathcal{G}]$ as

1. $\mathcal{G}$-measurable random variable

2. $\int_A \mathbb{E}[X|\mathcal{G}](\omega) d\mathbb{P}(\omega) = \int_A X(\omega) d\mathbb{P}(\omega)$.

Question. $\mathbb{E}[X|\mathcal{G}]$ exists? (Yes! proof skip). unique? (Yes! up to a.s.)

Remark. Lemma implies determine X (a.s.) is equivalent to know $\int_A X d\mathbb{P} \ \ \forall A \in \mathcal{F}$.
In this sense, conditional expectation $Y = \mathbb{E}[X|\mathcal{G}]$ is knowing $\int_A Y d\mathbb{P} = \int_A X d\mathbb{P} \ \ \forall A \in \mathcal{G}$.

Example. Toss a coin three times.
$\mathcal{F}(0) \subseteq \mathcal{F}(1) \subseteq \mathcal{F}(2) \subseteq \mathcal{F}(3)$. $X(t)$ is a number of heads until t times; $X(t)$ is $\mathcal{F}(t)$-measurable. If $\mathcal{F}(1) = \{\phi, \Omega, A_H, A_T\}$, then $\mathbb{E}[X(2)|\mathcal{F}(1)] = X(1) + \frac{1}{2}$, since we know the information of 1st flip.

Proof. Want: $\int_A (X(1) + \frac{1}{2}) d\mathbb{P} = \int_A X(2) d\mathbb{P}$ for all $A \in \mathcal{F}(1)$ (*c.f.* $\mathbb{P}(\omega) = \frac{1}{8} \ \forall \omega \in \Omega$).
For $A = A_H$, $\int \mathbb{1}_{A_H}(\omega)(X(1)(\omega) + \frac{1}{2}) d\mathbb{P}(\omega) = \frac{3}{2}\mathbb{P}(A_H) = \frac{3}{4}$.
$\int \mathbb{1}_{A_H}(\omega)(X(2))(\omega) d\mathbb{P}(\omega) = \sum_{\omega \in A_H} (X(2))(\omega)\mathbb{P}(\omega) = \frac{1}{8}(2+2+1+1) = \frac{3}{4}$.

□

Remark. $\mathcal{G} = \sigma(Y)$; $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X|\sigma(Y)] := \mathbb{E}[X|Y]$

Theorem 2.3.2
X, Y are independent random variable in $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$.

1. $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$

2. *X is $\mathcal{G}$-measurable. Then,* $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$.

3. *$\mathcal{H}$ is a sub σ-algebra of $\mathcal{G}$. Then,* $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.

4. *$X, \mathcal{G}$ are independent, then* $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.

Proof.
1. Exercise

2. We only need to show that $X \geq 0, Y \geq 0$ implies 2.

(a) $X = \mathbb{1}_B$
Want: $\mathbb{E}[\mathbb{1}_B Y|\mathcal{G}] = \mathbb{1}_B \mathbb{E}[Y|\mathcal{G}]$ for $B \in \mathcal{G}$.

(b) $X = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{B_i}$
Use linearity.

(c) $X \geq 0$
Use MCT

3. Want: $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$
Let $A \in \mathcal{H}$. Then, $\int_A \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}](\omega)d\mathbb{P}(\omega) = \int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{H}]d\mathbb{P}$

4. Can be shown similarly as in 2. Check $X = \mathbb{1}_B$ case. (Hint: $A \in \mathcal{G} \Rightarrow A, B$ are independent.)

$\square$

Example (Revisit).

$$\begin{aligned}
\mathbb{E}[X(2)|\mathcal{F}(1)] &= \mathbb{E}[X(2) - X(1) + X(1)|\mathcal{F}(1)] \\
&= \mathbb{E}[X(2) - X(1)|\mathcal{F}(1)] + X(1) \\
&= \mathbb{E}[X(2) - X(1)] + X(1) \\
&= \frac{1}{2} + X(1)
\end{aligned}$$

Review

- $(\Omega, \mathcal{F}, \mathbb{P})$: probability space, $\mathcal{G} \subseteq \mathcal{F}$.
- X : $\mathcal{F}$-measurable random variable.
- $Y = \mathbb{E}[X|\mathcal{G}]$ if Y is $\mathcal{G}$-measurable.
- $\int_A Y(\omega)d\mathbb{P}(\omega) = \int_A X(\omega)d\mathbb{P}(\omega) \ \ \forall A \in \mathcal{G}$.

Remark. $\mathbb{E}[X|\mathcal{G}]$ is an expectation of X when we know $\mathcal{G}$.

Remark. $Y = Z$ a.s. and Z is $\mathcal{G}$-measurable, then $Z = \mathbb{E}[X|\mathcal{G}]$.

Remark. $(X(t))_{t\in[0,T]}$ is stochastic process adapted to $(\mathcal{F}(t))_{t\in[0,T]}$. In this, $(X(t))_{t\in[0,T]}$ is a random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ and $X(t)$ is $\mathcal{F}(t)$-measurable. $\mathcal{F}(s) \subseteq \mathcal{F}(t)$ for all $s < t$ and $\mathcal{F}(0) = \{\phi, \Omega\}$.

Remark. We can define $\mathcal{F}$ as $\mathcal{F}(t) = \bigcup_{s:s\leq t} \sigma(X(s))$

Definition 2.3.3 (Martingale, Markov Process)

1. Martingale $X(t)$
   $\mathbb{E}[X(t)|\mathcal{F}(s)] = X(s)$ for all $s < t$.
2. Markov Process $X(t)$
   For any borel measurable f, there exists some borel measurable g such that $\mathbb{E}[f(X(t))|\mathcal{F}(s)] = g(X(s))$

Remark. In Martingale, if we know all the previous value, then the expectation of the future is as same as the expectation of the present.

Remark. Markov process is a generalization of Markov chain. We only have to know the present value.

Chapter 3

Brownian Motions

3.1 Introduction

To study Brownian Motions, we will study:

1. Random Walks
2. Definition of Brownian Motions and its basic property (We will change the textbook!)
3. Constuction of Brownian Motions

3.2 Scaled Random Walks

Definition 3.2.1 (Random Walk)

- Let $X_i = \begin{cases} 1 & \text{prob } \frac{1}{2} \\ -1 & \text{prob } \frac{1}{2} \end{cases}; X_1, X_2, \cdots$ independent.
- $M_n = X_1 + \cdots + X_n$ is called **random walk**
- $W_n(t) = \frac{1}{\sqrt{n}} M_{nt} (:= \frac{1}{\sqrt{n}} M_{[nt]})^v, \;\; t \in \{\frac{1}{n}k, k \in \mathbb{Z}_+\}$ is called **scaled random walk**

Proposition 3.2.2
Random walk holds the following properties:

1. Independent Increament
2. Martingale

3. Quadratic Variation

Proof.

1. See Def 3.2.3

2. Let $\mathcal{F}(n) = \sigma(X_1, X_2, \cdots, X_n)$ = smallest σ-algebra making$X_1 \sim X_n$ measurable. Then,

   - $M_n = X_1 + \cdots X_n$ is $\mathcal{F}(n)$-measurable
   - $(M_n)_{n\in\mathbb{N}}$ is stochastic process adapted to $(\mathcal{F}(n))_{n=0}^{\infty}$
   - $k < l \Rightarrow \mathbb{E}[M_l|\mathcal{F}(k)] = \mathbb{E}[M_l - M_k|\mathcal{F}(k)] + \mathbb{E}[M_k|\mathcal{F}(k)] = \mathbb{E}[M_l - M_k] + M_k = \mathbb{E}[X_{k+1} + \cdots + X_l] + M_k = \mathbb{E}[X_{k+1}] + \cdots + \mathbb{E}[X_l] + M_k = M_k$.

3. $\sum_{i=1}^{n}(M_i - M_{i-1})^2 = n$

□

Definition 3.2.3 (Independent increament)
M_n is **independent increament** if $M_{k_1}, M_{k_2} - M_{k_1}, \cdots, M_{k_m - k_{m-1}}$ are independent for any $k_1 < k_2 < \cdots < k_m$. Here, $M_{k_l} - M_{k_{l-1}}$ is called increament. If M_n is a random walk, then $M_{k_1} = \sum_{i=1}^{k_1} X_i, M_{k_2 - k_1} = \sum_{i=k_1}^{k_2} X_i, \cdots$ are independent.

Remark. Proposition 3.2.2 holds for scaled random variable $W_n(t) = \frac{1}{n}M_{nt}$ ($t \in \frac{1}{n}\mathbb{Z}_+$).

Proof.

1. Independent Increament
   For $t_1 < t_2 < \cdots < t_m$, $W_n(t_1) - W_n(0), W_n(t_2) - W_n(t_1), \cdots, W_n(t_m) - W_n(t_{m-1})$ are independent, since its increaments $W_n(t_{n+1}) - W_n(t_l) = \frac{1}{n}(M_{nt_{l+1}} - M_{nt_l})$ are independent by independent increament property of M_n.

2. Martingale
   Let $\mathcal{F}_n(t) = \sigma(X_1, X_2, \cdots, X_{nt})$. Then, $W_n(t) = \frac{1}{n}(X_1 + X_2 + \cdots + X_{nt})$ is $\mathcal{F}_n(t)$-measurable. Therefore, $(W_n(t))$ is stochastic process adapted to $(\mathcal{F}_n(t))$. With some computations as before, $\mathbb{E}[W_n(t)|\mathcal{F}_n(s)] = \cdots = W_n(s)$ for $s < t$.

3. Quadratic Variation
   $\sum_{i=1}^{nt}\left(W_n(\frac{i}{n}) - W_n(\frac{i-1}{n})\right)^2 = \sum_{i=1}^{nt}\left[\frac{1}{\sqrt{n}}(M_i - M_{i-1})\right]^2 = \sum_{i=1}^{nt}\frac{1}{n}\cdot 1 = t$

□

Example. Let $f \in C^1([0,t])$. Then,

$$\begin{aligned}
\sum_{i=1}^{nt} \left(f(\frac{i}{n}) - f(\frac{i-1}{n}) \right)^2 &= \sum_{i=1}^{nt} \left[\frac{1}{n} f'(\frac{x_i}{n}) \right]^2 \\
&= \frac{1}{n}\frac{1}{n} \sum_{i=1}^{nt} \left(f'(\frac{x_i}{n})^2 \right) \quad (\to \int_0^t [f'(x)]^2 dx) \\
&\leq \frac{c}{n} \quad (\to 0)
\end{aligned}$$

It is the most different property between random process and deterministic function: Q.V. of random variable is constant but Q.V. of C^1 function is zero.

Theorem 3.2.4 (Central Limit Theorem)
Let $Y_1, Y_2, \cdots$ are independent and identically distributed (called i.i.d.) with mean 0 and variation 1 ($\mathbb{E}(Y_i) = 0, Var(Y_i) = \mathbb{E}(Y_i^2) = 1$). Then,

$$\frac{1}{\sqrt{n}} [Y_1 + \cdots Y_n] \to N(0, 1^2) \tag{$\bigstar$}$$

Remark. Meaning of $\bigstar$:

$$\mathbb{P}\left[\frac{1}{n}(Y_1 + \cdots + Y_n) \in [a,b] \right] \to \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

$W_n(t) = \frac{1}{n} M_{nt} = \frac{1}{n}(X_1 + \cdots + X_{nt}) = \sqrt{t}\frac{1}{\sqrt{nt}}(X_1 + \cdots + X_{nt}) \sim N(0,t)$
cf. $N(\mu, \sigma^2)$ is a normal random variable with mean μ and variation σ^2. Using the above,

$$\lim_{n\to\infty} \mathbb{P}\left[W_n(t) \in [a,b] \right] = \int_a^b \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} dx.$$

$$W_n(t) = \frac{1}{n^{\frac{1}{2}+\alpha}} M_{nt} \begin{cases} \alpha < 0 & |W_n(t)| \to \infty \\ \alpha > 0 & |W_n(t)| \to 0 \end{cases}$$

Remark. $\frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$ is a heat kernel in PDE.

Summary

1. Independent Increament
2. Martingale
3. Markov Process
4. $W_n(t) \sim N(0,t)$
   $W_n(t) - W_n(s) \sim N(0, t-s)$

5. Q.V. in $[0, t] = t$.

Review

$X_1, X_2, \cdots$ are i.i.d. and $X_i = \begin{cases} \pm 1 & 1/2 \\ -1 & 1/2 \end{cases}$.

Random walk: $\mu_n = X_1 + \cdots + X_n$.

Scaled random walk: $W_n(t) = \frac{1}{\sqrt{n}} M_{nt}$. Then,

1. $W_n(0) = 0)$
2. Independent Increament
   $t_1 < t_2 < \cdots < t_n$, then $W_n(t_1), W_n(t_2) - W_n(t_1), \cdots, W_n(t_n) - W_n(t_{n-1})$ are independent.
3. Asymptotic Normal
   $W_n(t) - W_n(s) \sim N(0, t-s)$ as $n \to \infty$.

Part II

Introduction to stochastic integral

Chapter 2

Brownian Motion

2.1 Definition of Brownian Motion

Definition 2.1.1 (Stochastic Process)

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.
- $[0, \infty)$ with Borel σ-algebra
- $X : [0, \infty) \times \Omega \to \mathbb{R}$, measurable.

Then, X is a **stochastic process** if

1. $X(t, \cdot) : \Omega \to \mathbb{R}$ is random variable
2. $X(\cdot, \omega) : [0, \infty) \to \mathbb{R}$ is measurable.

Remark. $X(t, \cdot) \Rightarrow X(t)$: random variable in Ω. $X(t) : \omega \mapsto [X(t)](\omega) = X(t, \omega)$.
For each $t \in [0, \infty)$ there exists random variable $X(t) : \Omega \to \mathbb{R}$. If we pick $\omega \in \Omega$, then each $X(t_i)$ is determined simultaneously by $X(t_i)(\omega)$.

Remark. We can work in $[0, T]$ instead of $[0, \infty)$. In fact, we can define in $[0, T]$ and extend to $[0, \infty)$, but it is extremly difficult.

Definition 2.1.2 (Brownian Motion in $[0, \infty)$)

- $t \in [0, \infty)$, $\omega \in \Omega((\Omega, \mathcal{F}, \mathbb{P})$: probability space)
- Stoch. Process $B(t, \omega)$

B is called **Brownian Motion** if

1. $B(0, \omega) = 0$ a.s. (i.e.,$\mathbb{P}[\{\omega : B(0, \omega) = 0\}] = 1$)

2. $B(\cdot, \omega) : [0, \infty) \to \mathbb{R}$ is a continuous function a.s.

3. $\forall 0 \leq s < t$, $B(t) - B(s) \sim N(0, t-s)$

4. Independent Increament

Remark. $B(0, \omega)$ is a measurable function.

Remark. $B(t) \sim N(0, t)$ by 3 with $s = 0$.

Remark. $(B(t))_{t \geq 0} : \Omega \to \mathbb{R}$

- $B(t)$ itself is a normal distribution
- $B(t) - B(s) : \Omega \to \mathbb{R}$ is normal distribution with variance $t - s$.

Remark. Brownian motion is a continuous version of random walk: random walk has property 1,4 and has property 3 with $n \to \infty$.

Theorem 2.1.3

1. *$s < t$: $\mathbb{E}[B(s)B(t)] = s$*

2. *$t_1 < t_2 < \cdots < t_n \Rightarrow (B(t_1), B(t_2), \cdots, B(t_n))$ is jointly normal with $\mu = (0, 0, \cdots, 0$ and $Var = C$. ($C_{ij} = t_{\min(i,j)}$ $\forall i, j$).*

Proof.

1. $\mathbb{E}[B(s)B(t)] = \mathbb{E}[B(s)(B(t) - B(s))] + \mathbb{E}[B(s)^2] = \mathbb{E}(B(s))\mathbb{E}(B(t) - B(s)) + s = s.$

2. Let $\vec{v} = (B(t_1), B(t_2) - B(t_1), \cdots, B(t_m) - B(t_{m-1}))$. Then,

$$\text{PDF of } \vec{v} = \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{x_1^2}{2t_1}} \cdot \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{x_2^2}{2(t_2 - t_1)}} \cdots \frac{1}{\sqrt{2\pi(t_m - t_{m-1})}} e^{-\frac{x_m^2}{2(t_m - t_{m-1})}}.$$

   Therefore, $\vec{v}$ is jointly normal with $\mu = 0$ and $Var = \text{diag}(t_1, t_2 - t_1, \cdots, t_n - t_{n-1}) = D$, and,

$$\vec{W} = (B(t_1), \cdots, B(t_m)) = \vec{v} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \vec{v}E.$$

   Thus, $\vec{W}$ is jointly normal with $\mu = (0, 0, \cdots, 0)$ and $Var = EDE^T = C$.

$\square$

Definition 2.1.4 (Filtration for Brownian Motion)

$$\begin{aligned}\mathcal{F}_t &= \sigma(B(s) : s \leq t)\\ &= \text{smallest } \sigma\text{-algebra containing } \{\omega : (B(s))(\omega) \in A\} \ \ \forall s \in [0,t], A : \text{Borel}\\ &= \text{smallest } \sigma\text{-algebra making } \forall B_s, s \in [0,t] \text{ measurable}\end{aligned}$$

Remark.

1. $(B(t))_{t\geq 0}$: Stochastic process adapted to the filtration $(\mathcal{F}_t)$.

2. $(B(t), \mathcal{F}(t))$: Martingale.

Lemma. $B(t) - B(s)$ is independent of $\mathcal{F}_s$ $(s < t)$.

Proof of lemma.

1. $B(t) - B(s)$ is independent of $\sigma(B(s_1), B(s_2), \cdots, B(s_n))$ for $0 < s_1 < \cdots < s_n \leq s$, and check that $\sigma(B(s_1), B(s_2), \cdots, B(s_n)) = \sigma(B(s_1), B(s_2) - B(s_1), \cdots, B(s_n) - B(s_{n-1}))$

2. Let $\mathcal{H} = \bigcup_{m=1}^{\infty} \bigcup_{0<s_1<\cdots<s_n\leq s} \sigma(B(s_1), \cdots, B(s_n))$. Then, $\mathcal{H}$ is a closed under finite intersection. i.e., $A_1, A_2, \cdots, A_n \in \mathcal{H} \Rightarrow A_1 \cap A_2 \cap \cdots \cap A_n \in \mathcal{H}$.

3. $B(t) - B(s)$ is independent of $\mathcal{H}$ by 1. Then, $B(t) - B(s)$ is independent of $\overline{\mathcal{H}} = \mathcal{F}_t$: smallest σ-algebra containing $\mathcal{H}$. (Midterm 1 problem 2)

□

Proof. 1. By construction

2. $s < t \Rightarrow \mathbb{E}[B(t)|\mathcal{F}_s] = \mathbb{E}[B(t) - B(s)|\mathcal{F}_s] + \mathbb{E}[B(s)|\mathcal{F}_s] = \mathbb{E}[B(t) - B(s)] + B(s)$.

□

Chapter 3

Constuction of Brownian Motion

There are three ways to construct Brownian motion. One is by Wiener, one is by Kolmogorov, and one is by Leby. Wiener's method gives the existence of Brownian motion in natural way. Kolmogorov's method gives the property of Brownian motion with sample path with awful ω. Levy's method gives an instruction for Brownian motion with wierd ω.

3.1 Wiener Space

Let $C = C_0[0,1] = \{f : f \text{ is continuous on } [0,1] \text{ and } f(0) = 0\}$. We give a norm to C by $\|f\| = \sup_{0\leq x\leq 1}|f(x)| = \max_{0\leq x\leq 1}|f(x)|$, and distance $d(f,g) = \|f-g\|$. Thus, there is an open ball $B_r(x) = \{y : d(x,y) < r\}$ and topology(open set) of C. Now, there is Borel σ-algebra = smallest σ-algebra containing all open sets.

Notation
From now, let $\mathcal{B}(C)$ be Borel σ-algebra in C.

Definition 3.1.1 (Cylindrical Sets)
A cylindrical sets $\mathcal{R}$ is a collection of subsets of C of the form

$$A = \{f \in C : (f(t_1), f(t_2), \cdots, f(t_m)) \in U, 0 < t_1 < t_2 < \cdots < t_m \leq 1, U \in \mathcal{B}(\mathbb{R}^n)\},$$

and A is called **cylindrical set**.

cf. For $m = 1$, $\{f : f(t_1) \in U_k\} = A_k \in \mathcal{R}$, then $\bigcup_{k=1}^{\infty} A_k = \{f : f(t_1) \in \bigcup_{k=1}^{\infty} U_k\} \in \mathcal{R}$.

Remark. $\mathcal{R}$ is not a σ-algebra.

Example. $\{(f(t_1), f(t_2)) \in (-1,1) \times (2,3)\} = \{f : f(t_1) \in (-1,1), f(t_2) \in (2,3)\}$.

Definition 3.1.2
Let $\mu : \mathcal{R} \to [0,1]$ such that

$$\mu(A) = \iint \cdots \int \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{x_1^2}{2t_1}} \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{(x_2-x_1)^2}{2(t_2-t_1)}} \cdots \frac{1}{\sqrt{2\pi(t_m t_{m-1})}} e^{-\frac{(x_m-x_{m-1})^2}{2(t_m-t_{m-1})}} dx_1 \cdots dx_m,$$

and it is a natural definition since $B(t_i)$'s are jointly normal.

Theorem 3.1.3 (Wiener)
μ is a countably additive(σ-additive) function on $\mathcal{R}$. In other words, $A_1, A_2, \cdots$ are disjoint members of $\mathbb{R}$ and $\bigcup_{k=1}^{\infty} A_k \in \mathcal{R}$, then $\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$.

Summary

- $C = C_0[0,1] \Rightarrow \mathcal{B}(C)$ is Borel σ-algebra
- $\mathcal{R}$: collection of subsets of C
- $\mu : \mathcal{R} \to [0,1]$ is σ-additive (Wiener)

Fact: $\mathcal{R}$ is a **Ring** in the sense that

1. $\phi \in \mathcal{R}$
2. $A, B \in \mathcal{R} \Rightarrow A \cup B \in \mathcal{R}$
3. $A, B \in \mathcal{R} \Rightarrow A \backslash B \in \mathcal{R}$

Theorem 3.1.4 (Caratheodory Extension Theorem)

1. *$\mu : \mathcal{R} \to [0,1]$ is σ-additive*
2. *$\mathcal{R}$ is a ring*
3. *$\overline{\mathcal{R}}$ is a smallest σ-algebra containing $\mathcal{R}$.*

Then, there exists unique extension of μ to $\overline{\mathcal{R}}$ which is a measure.

Remark. $\mu : \overline{\mathcal{R}}$ is a probability measure called Wiener measure.

Remark. We can check that $\mathcal{B}(C) \subseteq \overline{\mathcal{R}}$. (it suffices to check that $B \in \overline{\mathcal{R}}$ for all open ball B), and $\mathcal{B}(C) = \overline{\mathcal{R}}$.

Conclusion: $(C, \overline{\mathcal{R}}, \mu)$ is a probability space and says **Wiener space**.

Theorem 3.1.5
$B : [0,1] \times C \to \mathbb{R}$, and let $B(t, \omega) = \omega(t)$. Then, B is a Brownian motion.

Proof. To prove that B is a Brownian motion, we have to check

1. $B(0,\omega) = \omega(0) = 0$ (obvious $\because \omega \in C = C_0[0,1]$).
2. $B(\cdot,\omega) = \omega(\cdot)$ is continuous (obvious by construction)
3. $s < t$: $B(t) - B(s) = N(0, t-s)$.
4. Independent increament i.e., $t_1 < \cdots < t_n \Rightarrow B(t_1), B(t_2) - B(t_1), \cdots, B(t_n) - B(t_{n-1})$ are independent.

Proof of 3, 4.

$$\begin{aligned}
&\mu\left(B(t_1) \in A, B(t_2 - B_1) \in A_2, \cdots, B(t_m) - B(t_{n-1}) \in A_m\right) \\
&= \mu\left[\{\omega : (B(t_1))(\omega) \in A_1, \cdots, (B(t_m))(\omega) - (B(t_{m-1}))(\omega) \in A_m\}\right] \\
&= \mu\left[\{\omega : \omega(t_1) \in A_1, \omega(t_2) - \omega(t_1) \in A_2, \cdots, \omega(t_m) - \omega(t_{m-1}) \in A_m\}\right] \\
&= \mu\left[A_{m,t_1,\cdots,t_m,U}\middle| \text{By def } \mu\right] \quad U = \{(x_1, \cdots, x_m) \in \mathbb{R}^m : x_1 \in A_1, \cdots, x_m - x_{m-1} \in A_m\} \\
&= \iint \cdots \int \left[\prod_{i=1}^{m} \cdots\right] du_1 \cdots du_m \\
&= \iint \cdots \int \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} e^{-\frac{y_i^2}{2(t_i - t_{i-1})}} dy_n \cdots dy_1 \\
&= \prod_{i=1}^{m} \left[\int_{A_i} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} e^{-\frac{y_i^2}{2(t_i - t_{i-1})}} dy_i\right].
\end{aligned}$$

Therefore,

(i) $A_1, A_3, \cdots, A_m = \mathbb{R} \Rightarrow \mu(B(t_2) - B(t_1) \in A_2) = \int_{A_2} \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{y_2^2}{2(t_2 - t_1)}} dy_2$: Proof of 3.

(ii) By (i), $\mu(B(t_1) \in A_1, \cdots, B(t_m) - B(t_{m-1}) \in A_m) = \mu(B(t_1) \in A_1) \cdots \mu(B(t_m) - B(t_{m-1}) \in A_m)$: Proof of 4.

□

□

Remark (Invariance Principle (Donsker, 1952)).
In random walk, for $X_1, X_2, \cdots$ be i.i.d. and $\mathbb{P}(X_i = 1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}$, we defined $S_0 = X_1 \cdots + X_n$ as random walk, $W_n(t) = \frac{1}{\sqrt{n}} S_{nt}$ ($t = 0, \frac{1}{n}, \cdots, \frac{n-1}{n}, 1$) as scaled random walk.
Define $\widehat{W_n}(t) = \frac{1}{\sqrt{n}} S_{[nt]} + \frac{1}{\sqrt{n}}(nt - [nt])(S_{[nt]+1} - S_{[nt]}) \in C[0,1]$. Then, we can define probability measure μ_n in $(C, \mathcal{B}(c))$ by

$$\mu_n(A) := \mathbb{P}\left[\widehat{W_n}(\cdot) \in A\right] \quad \forall A \in \mathcal{B}(C)$$

Theorem 3.1.6
$\widehat{W_n}(\cdot)$ converges to Brownian motion in the sense that $\mu_n \to \mu$ weakly.

Definition 3.1.7 (Weakly Convergence)
Let Ω be topological space, $\mathcal{B}(\Omega)$ be Borel σ-algebra. For each $n \in \mathbb{N}$ let $\mathbb{P}_n, \mathbb{P}$ be probability measure in $(\Omega, \mathcal{B}(\Omega))$. We say $\mathbb{P}_n \to \mathbb{P}$ **weakly** if $\int f d\mathbb{P}_n \to \int f d\mathbb{P}$ for every continuous and bounded function $f : \Omega \to \mathbb{R}$.

Remark. We defined a Brownian motion B in $B : [0,1] \times C \to \mathbb{R}$. To extend to $[0,\infty) \times C \to \mathbb{R}$, define $B(t) = \begin{cases} B_1(t) & t \in [0,1] \\ B_1(1) + B_2(t-1) & t \in [1,2] \\ B_1(1) + B_2(1) + B_3(t-2) & t \in [2,3] \\ \cdots \end{cases}$

3.2 Borel-Cantelli Lemma and Chebyshev Inequality

To Be Later...

3.3 Kolmogorov's Extension and Continuity Theorems

- $\Omega = \mathbb{R}^{[0,\infty)} = \{f : \text{ s.t. } f : [0,\infty) \to \mathbb{R}, f(0) = 0\}$
- $\mathcal{R} = \{A_{m,t_1,t_2,\cdots,t_m,U} : m \in \mathbb{N}, 0 \le t_1 < \cdots < t_m, U \in B(\mathbb{R}^m)\}$
- $\mathcal{F} = \overline{\mathcal{R}}$ (Cylindrical σ-algebra)
- $\mu_{t_1,t_2,\cdots,t_m}$: probability measure in $\mathcal{B}(\mathbb{R}^m)$ by

$$\mu_{t_1,t_2,\cdots,t_m}(A) = \iint \cdots \int_A \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} e^{-\frac{(u_i - u_{i-1})^2}{2(t_i - t_{i-1})}} du_1 \cdots du_m$$

(Note: In section 3.1 we define $0 < t_1 < t_2 < \cdots$ since Brownian motion is fixed at 0, but in this section we define $0 \le t_1 < t_2 < \cdots$.)

Definition 3.3.1 (Marginal distribution of $0 \le t_1 < \cdots < t_n$)
$\mu_{t_1,t_2,\cdots,t_n} = \mathbb{P}((X(t_1), \cdots, X(t_n)) \in U)$ $(:= \mathbb{P}[\{\omega : (X(t_1,\omega), \cdots, X(t_n,\omega)) \in U\}])$is called **marginal distribution of** $0 \le t_1 < \cdots < t_n$.

Question: Can we construct a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ such that

$$\mathbb{P}[A_{n,t_1,\cdots,t_n,U}] = \mu_{t_1,\cdots,t_n}(U) \text{ ?}$$

If it is possible, then $Y : [0, \infty) \times \Omega \to \mathbb{R}$ has marginal $\mu_{t_1,\cdots,t_n}$.
(Why?)

$$\begin{aligned}
&\mathbb{P}[\{\omega : (Y(t_1, \omega), Y(t_2, \omega), \cdots, Y(t_n, \omega)) \in U\}] \\
&= \mathbb{P}[\{\omega : (\omega(t_1), \cdots, \omega(t_n)) \in U\}] \\
&= \mathbb{P}[A_{n,t_1,\cdots,t_n,U}] = \mu_{t_1,\cdots,t_n}(U)
\end{aligned}$$

<u>Observation</u> (Consistency condition)

$$\begin{aligned}
&\mu_{t_1,\cdots,t_n} : \text{marginal distribution of } X(t,\omega) \\
&\Rightarrow \mu_{t_1,\cdots,t_n}(A_1 \times \cdots \times A_{i-1} \times \mathbb{R} \times A_{i+1} \times \cdots \times A_n) \\
&= \mu_{t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_n}(A_1 \times \cdots \times A_{i-1} \times A_{i+1} \times \cdots \times A_n)
\end{aligned}$$

Theorem 3.3.2 (Kolmogorov's Extension Theorem)

1. *$\mu_{t_1,\cdots,t_n}$: Borel probability measure in $\mathbb{R}^2$*

2. *$\{\mu_{t_1}, \cdots, t_n : 0 \le t_1 < \cdots < t_n\}$ satisfies the consistency condition*

Then, $\exists \mathbb{P}$ on $(\Omega, \mathcal{F})$ such that

$$\mathbb{P}[A_{n,t_1,\cdots,t_n,U}] = \mu_{t_1,\cdots,t_n}(U)$$

Kolmogorov's method is easy to construct Brownian motion with definition 1,3,4, but it is hard to probe definition 2: $B(\cdot, \omega)$ is continuous a.s.
Kolmogorov's extension theorem can be applied to every stochastic process.

Example. $0 \le t_1 < t_2 < \cdots t_n$.

$$\mu_{t_1,t_2,\cdots,t_n}(U) = \begin{cases} \iint \cdots \int_U \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} e^{-\frac{(u_i - u_{i-1})^2}{2(t_i - t_{i-1})}} du_1 \cdots du_n & \text{if } t_1 = 0 \\ \iint \cdots \int_U \prod_{i=2}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} e^{-\frac{(u_i - u_{i-1})^2}{2(t_i - t_{i-1})}} \delta_0(u_1) du_1 \cdots du_n & \text{if } t_1 \neq 0, \end{cases}$$

and it is easy to show that $\mu_{t_1,\cdots,t_n}$ satisfies consistency condition. By Kolmogorov's extension theorem, we have $(\Omega, \mathcal{F}, \mathbb{P})$ associates to the above. Let $B : [0, t] \times \Omega \ni (t, \omega) \mapsto \omega(t) \in \mathbb{R}$ (coordinate mapping process).
Question: Is B the Brownian motion? We have to check

1. $B(0) = 0$

2. $B(\cdots, \omega)$ is continuous a.s.

3. $B(t) - B(s) : N(0, t-s)$

4. Independent increament,

and 1, 3, 4 are easy. To prove 2, we need the following theorem

Theorem 3.3.3 (Kolmogorov's Continuity Theorem)
$X : [0,1] \times \Omega \to \mathbb{R}$ is a stochastic process satisfying KET. We say $(X(t))$ is separable if

1. $\forall t \in [0,1], \epsilon > 0 \;\; \lim_{s \to t} \mathbb{P}[|X(s) - X(t)| > \epsilon] = 0$

2. $\exists \alpha, \beta, C > 0$ *such that* $\mathbb{E}[|X(t) - X(s)|^{1+\alpha}] \le C|t-s|^{\beta}$ *for all* $s, t \in [0,1]$.

Then, there exists stochastic process $\widetilde{X}(t, \omega)$ such that

1. $\widetilde{X}(\cdot, \omega)$ *is continuous* $\forall \omega \in \Omega$

2. $\Omega_0 = \{\omega : X(t, \omega) = \widetilde{X}(t, \omega) \;\; \forall t \in [0,1] \;\;, \mathbb{P}(\Omega) = 1$

Return to example, by KCT, there exists $\widetilde{B}(t, \omega)$ such that $\widetilde{B}$ satisfies condition 2 of Brownian motion. Condition 3 and 4 are obvious, and condition 1 is also true since $\widetilde{B}(0, \omega) = 0$ on Ω_0, hence a.s.
We will skip the proof of KET and KCT (see Google)
Note: Kolmogorov's method is more general than Wiener's method since Kolmogorov's mtehod can be applied to any stochastic process. However, the space in Wiener's method (Banach space) and Kolmogorov's method are different, hence Kolmogorov's construction does not contain Wiener's construction.

3.4 Levy's Interpolation Method

Let $D_n = \{\frac{k}{2^{n-1}} : k = 0, 1, \cdots, 2^{n-1}\}$. We will construct $X_1(t), X_2(t), \cdots, X_n(t), \cdots$ such that $X_n(t) \to$ BM.

Definition 3.4.1
Let $\phi_1, \cdots$ be i.i.d of $N(0,1)$. Define $X_n(0) = 0, X_1(1) = \phi_1$ and

$$X_{n+1}(t) = \begin{cases} X_n(t) & \text{if } t \in D_n \\ X_n(t) + \frac{1}{2^{\frac{n+1}{2}}}\phi_i & \text{if } t \in D_{n+1} - D_n, \end{cases}$$

and interpolate by the line. Then, for each $t \in [0,1], (X_n(t))_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\Omega) = \{X : \Omega \to \mathbb{R}; \int X^2 d\mathbb{P} < \infty\}$ with distance $d(X_m(t), X_n(t)) = \sqrt{\mathbb{E}[|X_m(t) - X_n(t)|^2]}$. Since $L^2(\Omega)$ is complete, $X_n(t) \to X(t)$ for some random variable $X(t)$.

2.3 Wiener Integral

Review: L^2 space
$(\Omega, \mathcal{F}, \mathbb{P}) = ([a,b], \mathcal{B}([a,b]), m)$.

- $L^2(\Omega) := \{X : \text{ random variable in } \Omega, \int X^2 d\mathbb{P} < \infty\}$
- $L^2([a,b]) := \{f : [a,b] \to \mathbb{R}, \int f^2 dx < \infty\}$

1. Convergence in L^2
   - $(X_n)_{n=1}^{\infty} \in L^2(\Omega); X_n \to X$ in $L^2(\Omega)$ if $\int |X_n - X|^2 d\mathbb{P} \to 0$
   - $(f_n)_{n=1}^{\infty} \in L^2([a,b]); f_n \to f$ in $L^2([a,b])$ if $\int |f_n - f|^2 dx \to 0$
   - $d(X,Y) = \sqrt{\int |X-Y|^2 d\mathbb{P}}$: distance $\Rightarrow$ metric space
2. L^2: complete:
   - $(X_n)_{n=1}^{\infty}$: Cauchy in $L^2(\Omega) \Longrightarrow \exists X \in L^2(\Omega)$ such that $X_n \to X$ in $L^2(\Omega)$
   - $(f_n)_{n=1}^{\infty}$: Cauchy in $L^2([a,b]) \Longrightarrow \exists f \in L^2([a,b])$ such that $f_n \to f$ in $L^2([a,b])$
3. Simple functions are dense in $L^2(\Omega)$
   - $X \in L^2(\Omega) \Rightarrow \exists$ simple functions $(X_n)_{n=1}^{\infty}$ such that $X_n \to X$ in $L^2(\Omega)$
   - $f \in L^2([a,b]) \Rightarrow \exists$ simple functions $(f_n)_{n=1}^{\infty}$ such that $f_n \to f$ in $L^2([a,b])$
4. $X_n \to X$ in $L^2(\Omega) \cdots (\bigstar)$
   - $X_n \to X$ a.s. if $\mathbb{P}\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = 1 \cdots (\spadesuit)$

$(\bigstar) \not\Rightarrow (\spadesuit)$ and $(\spadesuit) \not\Rightarrow (\bigstar)$. But if $(\bigstar)$ there exists a subsequence such that $(\spadesuit)$.

Question 1. What is $\int_a^b f(t)dB(t)$?
Question 2. What is $\int_a^b f(B(t))dB(t)$?

Review: Stieltjes Integral
Let $a = t_0 < t_1 < \cdots < t_n = b$, $t_i = a + \frac{b-a}{n} i$.
$\lim_{n\to} \sum_{k=1}^{n} f(t_k)(g(t_k) - g(t_{k-1})) =: \int_a^b f dg$ if exists.

Remark.

1. $g \in C^1 \Rightarrow \int_a^b f dg = \int_a^b f g' dx$
2. Well-defined if g is absolutely continuous

3. (Summation by parts) $\sum_{k=1}^{n} f(t_k)(g(t_k) - g(t_{k-1})) = \sum_{k=0}^{n-1} g(t_k)(f(t_k) - f(t_{k+1})) + f(b)g(b) - f(a)g(a)$. If f is absolutely continuous, then it converges to $-\int gdf + f(b)g(b) - f(a)g(a)$. Therefore, if f is absolutely continuous, then $\int fdg = f(b)g(b) - f(a)g(a) - \int gdf$

Example (Example to Questions).

- If f is absolutely continous, then $\int_0^1 fdf = \int_0^1 ff'dx = \frac{f(1)^2 - f(0)^2}{2}$
- However, $\int_0^1 B(t)dB(t) = \frac{B(1)^2 - B(0)^2}{2} - \frac{1}{2}$ (by quadratic variation)

Consider a random variable $B(t) : \omega \mapsto B(t, \omega)$. Then,

$$\int_a^b f(t)dB(t) : \Omega \ni \omega \mapsto \int_a^b f(t)dB(t, \omega) \in \mathbb{R}$$

is well-defined if f is absolutely continuous. However, what if $f \in L^2([a, b])$?

Step 1 f is simple, $f(x) = \sum_{i=1}^{k} \alpha_i \mathbb{1}_{[a_i, b_i]}(x) \in L^2([a, b])$
Let $I(f) = \sum_{i=1}^{k} \alpha_i (B(b_i) - B(a_i)) \in L^2(\Omega)$. W.L.O.G. let $[a_i, b_i]$ be disjoint. Note that

(a) $\int \mathbb{1}_{[a_i, b_i]} dB(t) = \int_{a_i}^{b_i} dB(t) = B(b_i) - B(a_i)$, hence $I(f) = \int_a^b fdB(t)$
(b) $I(f) : N(0, \int_a^b f^2 dx)$, since $\mathbb{E}[I(f)^2] = \sum \alpha_i^2 (b_i - a_i) = \int_a^b f^2 dx$
(c) I is linear map. i.e. f, g are simple $\Rightarrow I(f + g) = I(f) + I(g)$ and $I(cf) = cI(f)$

Step 2 Recall (3: simple functions are dense in L^2)

- $f \in L^2([a, b]) \Rightarrow \exists$ simple functions $(f_n)_{n=1}^{\infty}$ such that $f_n \to f$ in L^2.
- If $I(f_m), I(f_n) \in L^2(\Omega)$, then

$$\int_a^b |I(f_m) - I(f_n)|^2 d\mathbb{P} = \int I(f_m - f_n)^2 d\mathbb{P} = \mathbb{E}[I(f_m - f_n)^2] = \int_a^b (f_m - f_n)^2 dx.$$

Therefore (f_n) is a Cauchy sequence in $L^2([a, b]) \Rightarrow (I(f_n))_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\Omega)$. Let $Z \in L^2(\Omega)$ be $I(f_n) \to Z$, and define $I(f) := Z$.

Exercise: Well-definedness
i.e., $f_n, g_n \to f$, $I(f_n) \to Z$, $I(g_n) \to Z'$, then $Z = Z'$ a.s.?
Exercise: $(Z_n : N(\mu_n, \sigma_n^2), Z_n \to Z$ in $L^2)$ Then, 1) $\mu_n \to mu, \sigma_n \to sigam$, $Z : N(\mu, \sigma^2)$.

Theorem 2.3.1

1. $I(f) : N(0, \int_a^b f^2 dx) \quad (f \in L^2([a, b])$

2. $\mathbb{E}[I(f)I(g)] = \int_a^b fgdx \quad (f, g \in L^2([a,b])$

3. *If f is absolutely continous (or $f \in C^1([a,b])$), then* $I(f) = \int_a^b f(t)dB(t)(= f(b)B(b) - f(a)B(a) - \int_a^b B(t)df)$

Definition 2.3.2 ($\int_a^b f(t)dB(t)$)
Let $f \in L^2([a,b])$. $\int_a^b f(t)dB(t) := I(f) \in L^2(\Omega)$. Then,

- $I(f) : \omega \mapsto \int_a^b f(t)dB(t,\omega)$
- $I : L^2([a,b]) \to L^2(\Omega)$ isometry

Proof of thm.

1. Let $f_n \to f$ in $L^2([a,b])$ (f_n : *simple*) and $I(f_n) \to I(f)$ in $L^2(\Omega)$. Since $I(f_n) \sim N(0, \int_a^b f_n^2 dx)$, it converges to $N(0, \int_a^b f^2(x)dx)$. Therefore, $\int_a^b f_n^2 dx \to \int_a^b f^2 dx$.

2. By 1. $\mathbb{E}[I(f)^2] = \int_a^b f^2 dx$. Moreover, $\mathbb{E}[I(f+g)^2] = \int_a^b (f+g)^2 dx$, $\mathbb{E}[I(f)^2] = \int_a^b f^2 dx$, $\mathbb{E}[I(g)^2] = \int_a^b g^2 dx$. Check that $I(f+g) = I(f) + I(g)$.

3. Let $a = t_0 < t_1 < \cdots < t_n = b$ and $t_i = a + \frac{b-a}{n} i$. Let $f_n(t) = \sum_{i=1}^n \mathbb{1}_{[t_{i-1}, t_i)}(t) f(t_{i-1})$. Then, $f_n \to f$ in L^2 if f_n's are absolutely continuous. Therefore, $I(f) = \lim_{n\to\infty} I(f_n) = \lim_{n\to\infty} \sum_{k=1}^n f(t_{i-1})(B(t_i) - B(t_{i-1}))$ (= Definition of Stieltjes integral $\int_a^b dB(t)$)

□

Example.

1. $\int_0^1 B(t)dt$?
Let $f(x) = x - 1, df(x) = dx$. Then,

$$
\begin{aligned}
\int_0^1 B(t)dt &= B(t)df(t) \\
&= B(1)f(1) - B(0)f(0) - \int_0^1 f(t)dB(t) \\
&= -\int_0^1 f(t)dB(t) \\
&= \int_0^1 (1-t)dB(t) \\
&\Rightarrow N(0, \int_0^1 (1-t)^2 dt) = N(0, \frac{1}{3})
\end{aligned}
$$

2. By theorem 1 of 2.3.1, $X = \int_0^1 tdB(t) = N(0, \frac{1}{3})$ and $Y = \int_0^1 t^2 dB(t) = N(0, \frac{1}{5})$. Moreover, by theorem 2 of 2.3.1, $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \int_0^1 tt^2 dt = \frac{1}{4}$.

Recall (Martingale)

$(M_t, \mathcal{F}_t)_{t\geq 0}$ is a stochastic process in $(\Omega, \mathcal{F}, \mathbb{P})$ with

1. $\mathcal{F}_t$: filtration; $\mathcal{F}_s \subseteq \mathcal{F}_t$ if $s \leq t$; M_t: $\mathcal{F}_t$-measurable (adapted) $\forall t$.
2. $\mathbb{E}[|M_t|] < \infty$
3. $\mathbb{E}[M_t|\mathcal{F}_s] = M_s \quad \forall s < t$.

To show that Wiener integral is Martingale, let $M_t = \int_0^t f(u)dB(u)$, and $\mathcal{F}_t = \sigma(B_u : u \in [0, s]$. Then, $\mathbb{E}[M_t|\mathcal{F}_s] = \mathbb{E}[M_t - M_s|\mathcal{F}_s] + \mathbb{E}[M_s|\mathcal{F}_s] = \mathbb{E}[\int_s^t f(u)dB(u)|\mathcal{F}_s] + M_s = M_s$?

Theorem 2.3.3

$(M_t, \mathcal{F}_t)$ is Martingale if $f \in L^2([0, \infty)$.

Proof. 1. By constuction

2. $\mathbb{E}[M_t^2] = \int_0^t f(u)^2 du < \infty \Rightarrow \mathbb{E}[|M_t|] \leq \mathbb{E}[M_t^2]^{1/2} < \infty$ by Jensen's inequality.

3. WTS: $\mathbb{E}[\int_s^t f(u)dB(u)|\mathcal{F}_s] = 0$.

Proof. Let $f \in L^2[s, t]$. Then, $\exists$ simple functions $f_n \in L^2([s, t])$ such that $f_n \to f$ in $L^2([s, t])$. Then,

$$\begin{aligned}
f_n(x) &= \sum_{k=1}^{m} \alpha_k \mathbb{1}_{[t_k-1, t_k)}(x) \\
&\Rightarrow \int_s^t f_n(u)dB(u) = \sum_{k=1}^{m} \alpha_k (B(t_k) - B(t_{k-1})) \\
&\Rightarrow \mathbb{E}[\int_s^t f_n(u)dB(u)|\mathcal{F}_s] = 0
\end{aligned}$$

since $\mathbb{E}[B(t_i) - B(t_{i-1})|\mathcal{F}_s] = 0$. Now,

$$\begin{aligned}
X^2 &:= \mathbb{E}\left[\int_0^t f(u)dB(u)|\mathcal{F}_s\right]^2 \\
&= \mathbb{E}\left[\int_0^t (f(u) - f_n(u))dB(u)|\mathcal{F}_s\right]^2 \\
&\leq \mathbb{E}\left[\left(\int_s^t (f(u) - f_n(u))\right)^2 |\mathcal{F}_s\right]
\end{aligned}$$

Next, we take $\mathbb{E}$ both sides

$$\begin{aligned}
\mathbb{E}[X^2] &\leq \mathbb{E}\left[\mathbb{E}\left[\left(\int_s^t (f(u) - f_n(u))dB(u)\right)^2 |\mathcal{F}_s\right]\right] \\
&= \mathbb{E}\left[\left(\int_s^t (f(u) - f_n(u)dB(u)\right)^2\right] \\
&= \int_s^t (f(u) - f_n(u))^2 \to 0
\end{aligned}$$

$\square$

$\square$

A key idea is that $\int_0^t f(u)dB(u) = M_t$ is a Martingale.

Chapter 4

Stochastic Integrals

4.1 Background and Motivation

Motivation

- What is $\int_a^b f(B(u)dB(u))$?
- What is $\int_a^b f(u, B(u))dB(u)$?

Example. $\int_0^T B(t)dB(t) = \lim_{n\to\infty} \sum_{i=1}^n B(s_i)\left[B(t_i) - B(t_{i-1})\right] = ?$
Let $L_n = \sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1}))$, and $R_n = \sum_{i=1}^n B(t_i)(B(t_i) - B(t_{i-1}))$. Then,

1. $L_n + R_n = \sum_{i=1}^n (B(t_i)^2 - B(t_{i-1})^2) = B(t_n)^2 - B(t_0)^2 = B(T)^2$.
2. $R_n - L_n = \sum_{i=1}^n [(B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1})] + \sum_{i=1}^n (t_i - t_{i-1}) = (\sum_{i=1}^n X_i) + T$.
   Note that $\mathbb{E}[(\sum_{i=1}^n X_i)^2] = \sum_{i=1}^n \mathbb{E}[X_i^2] + 2\sum_{i<j} \mathbb{E}[X_i X_j]$,
   (a) $i < j \Rightarrow \mathbb{E}[X_i X_j] = \mathbb{E}[B(t_i) - B(t_{i-1}), B(t_j) - B(t_{j-1})] = \mathbb{E}[X_i]\mathbb{E}[X_j] = 0$
   (b) $\mathbb{E}[X_i]^2 = \mathbb{E}\left[(B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1})^2\right] = 2(t_i - t_{i-1})^2 = \frac{T^2}{n^2}$ (exercise).

   Therefore, $\mathbb{E}((\sum_{i=1}^n X_i)^2) = \frac{T^2}{n}$, and $\sum_{i=1}^n X_i \to 0$ in $L^2(\Omega)$.

Therefore, $R_n + L_n \to B(T)^2$ and $R_n - L_n \to T$
$\Rightarrow R_n \to \frac{B(T)^2+T}{2}, L_n \to \frac{B(T)^2-T}{2}$.

Remark. In Riemann integral, $\lim_{n\to\infty} L_n = \lim_{n\to\infty} R_n$.

Remark. L_n is called **Itô Integral**, and $M_n = \frac{L_n+R_n}{2}$ is called **Stratonovich Integral**

Remark. We use L_n in many cases since $\mathbb{E}[L_n] = 0$ and $M_t = \int_0^t B(s)dB(s)$ is Martingale in the sense of Itô.

Definition 4.1.1 (Itô Integral)
$\int_a^b f(t, B(t))dB(t) \to \lim_{n\to\infty} \sum_{i=1}^n f(t_{i-1}, B(t_{i-1}))(B(t_i) - B(t_{i-1}))$

Review
We'll check the facts that the textbook assumes but not proves.
$f \in L^2([a,b]) \to I(f) := \int_a^b f(s)dB(s) \in L^2(\Omega)$ and we've showed that

1. $I(f) \sim N(0, \int_a^b f^2(t)dt)$
2. $\mathbb{E}[I(f)I(g)] = \int_a^b f(t)g(t)Dt$

Remark. $I(f), I(g)$ are jointly normal.

Let X be random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\phi_X(t) = \mathbb{E}[e^{itX}] = \int_\Omega e^{itX(\omega)} d\mathbb{P}(\omega)$.

Example. Let $X \sim N(\mu, \sigma^2)$. Then,

$$\phi_X(t) = \int e^{itx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = e^{it\mu - \frac{1}{2}t^2\sigma^2} dx.$$

Conversely, $\phi_X(t) = e^{it\mu - \frac{1}{2}t^2\sigma^2}$ $\forall t$, then $X \sim N(\mu, \sigma^2)$.

Sketch of proof. Let f_X be a pdf of X. Then, $\int_{-\infty}^{\infty} e^{itx} f_X(x)dx = \phi_X(t)$, and by inverse Fourier transform, $f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\phi_X(t)dt$. □

Example. Let $\vec{X} = (X_1, X_2, \cdots, X_n)$ be jointly normal with mean $\vec{\mu} = (\mu_1, \cdots, \mu_n)$ and variance $\Sigma = (\Sigma_{ij})_{i,j=1}^n$. Then,

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^n \det \Sigma} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^\intercal},$$

$$\phi_{\vec{X}}(\vec{t}) = \mathbb{E}\left[e^{i(\vec{X}\cdot\vec{t})}\right] = e^{i\vec{\mu}\vec{t} - \frac{1}{2}\vec{t}\Sigma\vec{t}^\intercal}.$$

Conversely,

$$\phi_{\vec{X}}(\vec{t}) = \mathbb{E}\left[e^{i(\vec{X}\cdot\vec{t})}\right] = e^{i\vec{\mu}\vec{t} - \frac{1}{2}\vec{t}\Sigma\vec{t}^\intercal},$$

then $\vec{X} \sim N(\vec{\mu}, \Sigma)$.

Theorem 4.1.2

$$(I(f), I(g)) \sim N\left((0,0), \begin{bmatrix} \int_a^b f(t)^2 dt & \int_a^b f(t)g(t)dt \\ \int_a^b f(t)g(t)dt & \int_a^b g(t)^2 dt \end{bmatrix}\right)$$

Proof. Consider $uI(f) + vI(g) = I(uf + vg) \sim N\left(0, \int_a^b (uf(t) + vg(t))^2 dt\right)$. Then,

$$\begin{aligned} \mathbb{E}\left[e^{i(uI(f)+vI(g))}\right] &= \exp\left[-\frac{1}{2}\left(u^2\int_a^b f^2dt + v^2\int_a^b g^2dt + 2uv\int_a^b fgdt\right)\right] \\ &= \exp\left[-\frac{1}{2}(u,v)\Sigma\begin{pmatrix} u \\ v\end{pmatrix}\right]. \end{aligned}$$

Therefore, $(I(f), I(g)) \sim N(\vec{0}, \Sigma)$. □

Remark. The same holds for n functions:

$$f_1, \cdots, f_n \in L^2([a,b]) \Rightarrow (I(f_1), \cdots, I(f_n)) \sim N(\vec{0}, \Sigma), \quad \Sigma_{ij} = \int_a^b f_i(t)f_j(t)dt.$$

Let $f \in L^2([0,t])$ for all $t \geq 0$. (*cf.* $f(t) = t \notin L^2([0,\infty])$). Set $X(t) = \int_0^t f(s)dB(S)$ $(X(t) : \omega \mapsto \int_0^t f(s)dB(s,\omega))$. We've proved that X is a Martingale.

Theorem 4.1.3
$X(t)$ is independent increament.

Proof. Let $s < t$ and we will show that $X(s), X(t) - X(s)$ are independent. (proof for $X(t_1), X(t_2) - X(t_1), \cdots, X(t_m) - X(t_{m-1})$ is identical and exercise)

$$X(s) = \int_0^s f(u)dB(u) = \int_0^t g(u)dB(u), \qquad g(u) = \begin{cases} f(u) & u \leq s \\ 0 & u > s \end{cases}$$

$$X(t) - X(s) = \int_s^t f(u)dB(u) = \int_0^t h(u)dB(u), \qquad h(u) = \begin{cases} 0 & u \leq s \\ f(u) & u > s \end{cases}$$

In other words, $X(s) = I(g), X(t) - X(s) = I(h)$ (jointly normal). Therefore, $\mathbb{E}[I(g)I(h)] = \int_0^t g(u)h(u)du = 0$. Since $I(g), I(h)$ are jointly normal and uncorrelated (supports of g, h are disjoint), they are independent. □

Example. $\mathbb{E}[B(1)^2B(2)^2] = ?$

Method 1. using $(B(1), B(2)) \sim N\left((0,0), \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\right)$, calculate

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2y^2\frac{1}{2\pi}e^{-\frac{1}{2}(2x^2-2xy+y^2)}dxdy.$$

Method 2. $B(t)-B(s)$ are independent of $\mathcal{F}_s$ and $\mathcal{F}_u$ for all $(u \le s)$. In other words, give $\mathcal{F}_s$, $B(t)-B(s) \sim N(0,t-s)$. Therefore,

$$
\begin{aligned}
\mathbb{E}[B(1)^2B(2)^2] &= \mathbb{E}[\mathbb{E}[B(1)^2B(2)^2|\mathcal{F}_1]] \\
&= \mathbb{E}[B(1)^2\mathbb{E}[B(2)^2|\mathcal{F}_1]] \\
&= \mathbb{E}\left[B(1)^2\mathbb{E}[(B(2)-B(1))^2+2(B(2)-B(1))B(1)+B(1)^2|\mathcal{F}_1]\right] \\
&= \mathbb{E}\left[B(1)^2\left(\mathbb{E}[(B(2)-B(1))^2|\mathcal{F}_1]+2B(1)\mathbb{E}[(B(2)-B(1))]+B(1)^2\right)\right] \\
&= \mathbb{E}\left[B(1)^2\left(1+B(1)^2\right)\right] \\
&= 1+3=4
\end{aligned}
$$

(Fact: $\mathbb{E}[X^{2n}]=(2n-1)!!$ for standard normal random variable X)

Example. Let $X(t)=\int_0^t(a+b\frac{u}{t})dB(u)$. Then, for which a,b, $X(t)$ be Brownian motion? To say that X is a Brownian motion, we have to show that

1. $X(t)-X(s) \sim N(0,t-s)$
2. Independent increament
3. $X(0)=0$
4. continuity

1. Suppose $s<t$. Then,

$$
\begin{aligned}
X(t)-X(s) &= \int_0^t(a+b\frac{u}{t})dB(u)-\int_0^s(a+b\frac{u}{s})dB(u) \\
&= \int_0^s bu(\frac{1}{t}-\frac{1}{s})dB(u)+\int_s^t(a+b\frac{u}{t})dB(u) \\
\Rightarrow \mathbb{E}[(X(t)-X(s))^2] &= \int_0^s(bu)^2(\frac{1}{t}-\frac{1}{s})^2du+\int_s^t(a+b\frac{u}{t})^2du \\
&= \frac{b^3}{3}(\frac{1}{t}-\frac{1}{s})^2+\frac{1}{3}\frac{t}{b}\left((a+b)^3-(a+b\frac{s}{t})^3\right) \\
&= \frac{b^2}{3}\frac{s^3}{t^2}-\frac{2b^2}{3}\frac{s^2}{t}+\frac{b^2}{3}s+\frac{t}{3b}\left(3a^2b(1-\frac{s}{t})+3ab^2(1-\frac{s^2}{t^2})-\frac{s^3}{t^3}b^3+b^3\right) \\
&= (\frac{b^2}{3}-a^2)s+t(a^2+3ab+\frac{b^2}{3}) \quad \left(-\frac{2b^2}{3}-ab=0\right)
\end{aligned}
$$

Therefore, $\frac{b^2}{3}-a^2=-1, a^2+ab+\frac{b^2}{3}=1 \Rightarrow (a,b)=(1,0),(2,-3)$

2. Suppose $t_0=0<t_1<t_2<\cdots<t_m$. Then, $X(t_1)-X(t_0),X(t_2)-X(t_1),\cdots,X(t_m)-X(t_{m-1})$ are jointly normal ($\because X(t_{i+1})-X(t_i)=I(f_i)$ for some f_i). Moreover, $\mathbb{E}[I(f_i)I(f_j)]=0$ for $(a,b)=(1,0),(2,-3)$. Therefore, they are independent.

4.2 Filtration for Brownian Motion

Let $(X_t, \mathcal{F}_t)$ be stochastic process and filtration.

Proposition 4.2.1
If $X_t - X_s$ is independent of $\mathcal{F}_s$ $\forall s < t$, then (X_t) is independent increament.

Proof. $Z_1, Z_2, \cdots Z_n$ are independent if and only if (by inverse transform)

$$\mathbb{E}\left[\prod_{k=1}^{n} e^{i\theta_k Z_k}\right] = \prod_{i=1}^{n} \mathbb{E}\left[e^{i\theta_k Z_k}\right] \quad \forall \theta_1 \sim \theta_n \in \mathbb{R}.$$

Claim: $X(s) - X(t) - X(s)$ are independent.

$$\begin{aligned}
\mathbb{E}\left[e^{i\theta_1 X(s) + i\theta_2 (X(t) - X(s))}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{i\theta_1 X(s) + i\theta_2 (X(t) - X(s))}\right] | \mathcal{F}_s\right] \\
&= \mathbb{E}\left[e^{i\theta_1 X(s)} \mathbb{E}\left[e^{i\theta_2 (X(t) - X(s))} | \mathcal{F}_s\right]\right] \\
&= \mathbb{E}\left[e^{i\theta_1 X(s)}\right] \mathbb{E}\left[e^{i\theta_2 (X(t) - X(s))}\right].
\end{aligned}$$

□

4.3 Stochastic Integrals

Let $B(t)$ be Brownian motion. $\mathcal{F}_t$ is a **Brownian Filtration** if

1. $B(t)$ is $\mathcal{F}_t$-measurable $\forall t$
2. $B(t) - B(s)$ is independent of $\mathcal{F}_s$ $\forall s < t$.

Example. $\mathcal{F}_t = \sigma(B(s) : s \leq t)$ works.

Brownian filtration is essential in Itô's integral since in Wiener's integral, we concern

$$\int_a^b f(s) dB(s) : \omega \mapsto \int_a^b f(s) dB(s, \omega),$$

while in Itô's integral, we concern

$$\int_a^b f(s) dB(s) : \omega \mapsto \int_a^b f(s, \omega) dB(s, \omega),$$

and $f : [a, b] \times \Omega \to \mathbb{R}$.

Example. $\int_a^b g(B(s)) dB(s)$ and $\int_a^b g(s, B(s)) dB(s)$.

Definition 4.3.1
$f \in L^2_{\text{ad}}([a, b] \times \Omega)$ (admissible and adapted) if

1. $f(t)$ (random variable such that $(f(t))(\omega) = f(t,\omega)$) is $\mathcal{F}_t$-measurable $\forall t$. (i.e. $(f(t))$: adapted to $(\mathcal{F}_t)$)

2. $f \in L^2([a,b] \times \Omega)$ i.e.

$$\int_a^b \int_\Omega f^2(t,\omega) d\mathbb{P} dt < \infty (\Leftrightarrow \int_a^b \mathbb{E}[f(t)^2] dt < \infty).$$

Purpose of today: Define $\int_a^b f(t) dB(t)$ for $f \in L^2_{\text{ad}}$.

Step 1) Step Stochastic Process $f(t)$

Let $f(t) = \sum_{i=1}^n X(t_{i-1}) \mathbb{1}_{[t_{i-1},t_i)}(t)$, where $a = t_0 < t_1 < \cdots < t_n = b$, and $X(t_i)$: $\mathcal{F}_t$-measurable $\forall t$. Define

$$I(f) := \sum_{i=1}^n X(t_{i-1})(B(t_i) - B(t_{i-1}))$$

Lemma. $\mathbb{E}\left[I(f)^2\right] = \int_a^b \mathbb{E}\left[f(t)^2\right] dt (= \|f\|^2_{L^2([a,b]\times\Omega)})$

Proof.

$$\begin{aligned}
\mathbb{E}\left[I(f)^2\right] = & \sum_{i=1}^n \mathbb{E}\left[X(t_{i-1})^2 (B(t(_i) - B(t_{i-1}))^2)\right] \\
& + 2\sum_{i<j} \mathbb{E}\left[X(t_{i-1})X(t_{j-1})(B(t_i) - B(t_{i-1}))(B(t_j - B(t_{j-1})))\right],
\end{aligned}$$

and

$$\begin{aligned}
\sum_{i=1}^n \mathbb{E}\left[X(t_{i-1})^2 (B(t(_i) - B(t_{i-1}))^2)\right] &= \mathbb{E}\left[\mathbb{E}\left[X(t_{i-1})^2 (B(t(_i) - B(t_{i-1}))^2)\right] | \mathcal{F}_{t_{i-1}}\right] \\
&= \mathbb{E}\left[X^2_{t_{i-1}} \mathbb{E}\left[(B(t_i) - B(t_{i-1}))^2\right] | \mathcal{F}_{t_{i-1}}\right] \\
&= (t_i - t_{i-1}) \mathbb{E}\left[X^2_{t_{i-1}}\right],
\end{aligned}$$

and

$$2\sum_{i<j} \mathbb{E}\left[X(t_{i-1})X(t_{j-1})(B(t_i) - B(t_{i-1}))(B(t_j - B(t_{j-1})))\right] = 0.$$

Therefore, $\mathbb{E}\left[I(f)^2\right] = \sum_{i=1}^n (t_i - t_{i-1}) \mathbb{E}\left[X^2_{t_{i-1}}\right] = \int_a^b \mathbb{E}\left[f(t)^2\right] dt.$ $\square$

Remark. $\mathbb{E}\left[I(f)\right] = 0$ but it is possible that $I(f) \not\sim N\left(0, \int_a^b \mathbb{E}\left[f(t)^2\right] dt\right)$.

Step 2) Approximation: $f \in L^2_{\text{ad}} \Rightarrow \exists$ step stochastic process $f_n \in L^2_{\text{ad}}$ such that

$$\int_a^b \mathbb{E}\left[(f(t) - f_n(t))^2\right] dt \to 0 \quad \text{as } n \to \infty.$$

Proofs: later

Step 3) $f \in L^2_{\text{ad}} \Rightarrow \exists f_n$ steps such that $\|f_n - f\|_{L^2([a,b]\times\Omega)} \to 0$. Therefore, (f_n): convergent. i.e., Cauchy in $L^2([a,b]\times\Omega)$, and then by Lemma, $(I(f_n))$ is Cauchy in $L^2(\Omega)$. ($\because$ $\mathbb{E}\left[(I(f_m) - I(f_n))^2\right] = \mathbb{E}\left[I(f_m - f_n)^2\right] = \|f_m - f_n\|^2_{L^2([a,b]\times\Omega)}$). Therefore, $I(f_n)$ is convergent in $L^2(\Omega)$. (By the completeness of $L^2(\Omega)$) We denote by $I(f)$ this limit.

Definition 4.3.2 (Itô's integral)
$\int_a^b f(t)dB(t) := I(f)$ (Note: We have to check well-definedness)

Theorem 4.3.3

1. *$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$ (linear)*
2. $\mathbb{E}[I(f)] = 0$
3. $\mathbb{E}[I(f)^2] = \int_a^b \mathbb{E}[f(t)^2]dt$
4. $\mathbb{E}[I(f)I(g)] = \int_a^b \mathbb{E}[f(t)g(t)]dt$

Sketch of proof.

- 1, 3 $\Rightarrow$ 4 (insert $f + t$ to f at 3)
- 1,2,3: For step stochastic process $f_n \to f$ in $L^2([a,b]\times\Omega)$ show that $\mathbb{E}[I(f_n)] \to \mathbb{E}[I(f)]$ and $\mathbb{E}[I(f_n)^2] \to \mathbb{E}[I(f)^2]$.

□

Comment on Step 2) Let $\alpha(s,t) = \mathbb{E}[f(t)f(t)]$, $\alpha : [a,b]^2 \to \mathbb{R}$.

1. α is continuous in $[a,b]^2$
2. f: bounded
3. $f \in L^2_{\text{ad}}$

$X_t = \int_0^t f(s)dB(s); f \in L^2_{\text{ad}}([0,t]\times\Omega)\ \forall t > 0$.

Theorem 4.3.4
$(X_t, \mathcal{F}_t)$: Martingale

Theorem 4.3.5
X_t is continuous a.s. i.e., $\exists \Omega_0 \subseteq \Omega$ such that $\mathbb{P}(\Omega_0) = 1$, and $\omega \in \Omega_0 \Rightarrow h(t) = X_t(\omega)$ is continuous in t.

Proof. Approximations via step stochastic process. □

Theorem 4.3.6

Let $f \in L^2_{ad}$ and $\alpha(s,t) = \mathbb{E}[f(s)f(t)]$ is continous in $[0,\infty)^2$. Define $\Delta_n = \{a = t_0 < t_1 < \cdots < t_n = b\}$, $\|\Delta_n\| = \max\limits_{i=0\sim n-1}(t_{i+1} - t_i)$. Then, $lim_{\|\Delta_n\|\to 0} \sum_{i=1}^{n} f(t_{i-1},\omega)(B(t_i,\omega) - B(t_{i-1},\omega)) \to \int_a^b f(t,\omega)dB(t,\omega)$ (a.s.).

Chapter 5

An Extension of Stochastic Integral

Motivation

In chapter 4, we assumed that $\int_a^b \mathbb{E}[f(t)^2]dt < \infty$ and play on L^2_{ad}. However it is too much strong condition. Instead, in this chapter, we assume that $\int_a^b f(t,\omega)^2 dt < \infty$ a.s and play on $\widetilde{L}^2_{\text{ad}}$.
Fact: $X \geq 0, \mathbb{E}[X] < \infty \Rightarrow X < \infty$ a.s.
$\mathbb{E}[\int_a^b f(t,\omega)^2 dt] = \int_\Omega \int_a^b f(t,\omega)^2 dt d\mathbb{P}$ (By Fubini) $= \int_a^b \int_\Omega f(t,\omega)^2 d\mathbb{P} dt = \int_a^b \mathbb{E}[f(t)^2]dt$.
By Fact, condition on chapter 4 implies chdition on chapter 5, hence $L^2_{\text{ad}} \subseteq \widetilde{L}^2_{\text{ad}}$. In this chapter, we will define Itô's integral with that condition.

Goal

Define $I(f) = \int_a^b f(t)dB(t)$ for $f \in \widetilde{L}^2_{\text{ad}}$.
Idea: $f \in \widetilde{L}^2_{ad}$ then exists step stochastic process f_n such that $\int_a^b |f_n(t,\omega) - f(t,\omega)|dt \to 0$ in probability.
Note: $X_n \to X$ means

- L^1: $\mathbb{E}[|X_n - X|] \to 0$
- L^2: $\mathbb{E}[|X_n - X|^2] \to 0$
- a.s.: $\{\omega : X_n(\omega) \not\to X(\omega)\}$ measure zero.
- in probability: $\forall \delta > 0 \mathbb{P}[|X_n - X| > \delta] \to 0$.

Note

$\mathbb{P}[|X_n - X| > \delta] < \frac{\mathbb{E}[|X_n - X|^2]}{\delta^2}$. Therefore, L^2 convergence implies in probability convergence. Then, we can follow the previous strategy. (Warning: $(I(f_n))_{n=1}^\infty$ is no longer a Cauchy sequence in L^2)

Theorem 5.0.1
$X_t = \int_0^t f(t)dB(t)$ and $f \in \widetilde{L}^2_{ad}$. Then,

1. *X_t is a local martingale.*

2. *X_t has continous realization.*

Example. $f(t) = e^{B(t)^2}$. Then,

$$\mathbb{E}[f(t)^2] = \begin{cases} \frac{1}{\sqrt{1-4t}} & t < \frac{1}{4} \\ \infty & t \geq \frac{1}{4} \end{cases}$$

Hence calculating $\int_0^1 f(t)dB(t)$ is impossible in the sense of chapter 4, but it is possible in the sense of chapter 5. ($\because \int_0^1 f(t)^2 dt = \int_0^1 e^{2B^2(t)}dt < \infty$ a.s.)

Question: $X_t = \int_0^t e^{B(s)}dB(s)$. Then is it Martingale? Continuous?
<u>Note</u>: In local martingale, $\mathbb{E}[I(f)] \neq 0$.

Chapter 7

Itô Formula

Motivation

$\int f(B(t))dB(t) = ?$

Example. For f, g differentiable, $\int f'(g(t))dg(t) = \int_a^b f'(g(t))g'(t)dt = f(g(b)) - f(g(a))$.
However, $\int_a^b B(t)dB(t) = \frac{B(b)^2 - B(a)^2}{2} - \frac{b-a}{2}$.
Idea: Let $a = t_0 < t_1 < \cdots < t_n = b, f \in C^2$. Then,

$$\begin{aligned} f(B(b)) - f(B(a)) &= \sum_{i=1}^{n} [f(B(t_i)) - f(B(t_{i-1}))] \\ &\simeq \sum_{i=1}^{n} f'(B(t_{i-1}))(B(t_i) - B(t_{i-1})) + \sum_{i=1}^{n} \frac{1}{2} f''(B(t_{i-1}))(B(t_i) - B(t_{i-1}))^2 \\ &\simeq \sum_{i=1}^{n} f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})) + \frac{1}{2} \sum_{i=1}^{n} f''(B_{t_{i-1}})(t_i - t_{i-1}) \quad \text{(by lemma)} \\ &\longrightarrow \int_a^b f'(B(t))dB(t) + \frac{1}{2} \int_a^b f''(B(t))dt \quad \text{as } \|\Delta_n\| \to 0. \end{aligned}$$

7.1 Itô's Formula in the Simplest Form

Lemma. $\sum_{i=1}^{n} g(B(t_{i-1})) \left[(B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1}) \right] \to 0$ in L^2 as $\|\Delta_n\| \to 0$.

Theorem 7.1.1

$f \in C^2(\mathbb{R})$. Then,

$$f(B(b)) - f(B(a)) = \int_a^b f'(B(t))dB(t) + \frac{1}{2} \int_a^b f'(B(t))dt.$$

Proof of Lemma.

$$\begin{aligned}
\mathbb{E}[()^2] &= \sum_{i=1}^{n} \mathbb{E}\left[g(B(t_{i-1}))^2\left[(B(t_i)-B(t_{i-1}))^2-(t_i-t_{i-1})\right]^2\right] \\
&\quad + 2\sum_{i<j}\mathbb{E}\left[g(B(t_{i-1}))g(B(t_{j-1}))\left((B(t_i)-B(t_{i-1}))^2-(t_i-t_{i-1})\right)\right. \\
&\qquad \left.\times\left((B(t_j)-B(t_{j-1}))^2-(t_j-t_{j-1})\right)\right] \\
&= \mathbb{E}[g(B(t_{i-1}))^2\mathbb{E}[()^2|\mathcal{F}_{t_{i-1}}]] \\
&= \frac{2(b-a)^2}{n^2}\mathbb{E}[g(B(t_{i-1}))^2]
\end{aligned}$$

$\therefore \mathbb{E}[()^2] \leq \frac{2(b-a)^2}{n^2}\sum\frac{2(b-a)^2}{n^2}\sum_{i=1}^{n}\mathbb{E}[g(B(t_{i-1}))^2] \leq cn$ □

Remark (Theorem 4.7.1). Suppose $f \in L^2_{\text{ad}}$, $h(t,s) = \mathbb{E}[f(t)f(s)]$ is continuous in $\mathbb{R}^2$. Then,

$$\int_a^b f(t)dB(t) = \lim_{n\to\infty} f(t_{i-1})[B(t_i)-B(t_{i-1})]$$

Remark. In the above, h is continuous if (1) $f(t) = g(B(t))$ (2) $f(t) = g(t, B(t))$ for continuous g.

Theorem 7.1.2

$f \in C^2(\mathbb{R})$. Then,

$$f(B(b)) - f(B(a)) = \int_a^b f'(B(t))dB(t) + \frac{1}{2}\int_a^b f''(B(t))dt$$

Proof. Idea: If $g \in C^2$, then by intermediate theorem $g(y)-g(x) = (y-x)g'(x) + \frac{1}{2}(y-x)^2g''(\alpha)$ for some $\alpha x + \lambda(y-x), \lambda \in [0,1]$. Then,

$$\begin{aligned}
f(B(b)) - f(B(a)) &= \sum_{i=1}^{n}\left(f(B(t_i)) - f(B(t_{i-1}))\right) \\
&= \sum_{i=1}^{n} f'(B(t_{i-1}))(B(t_i)-B(t_{i-1})) \quad (1)\\
&\quad + \frac{1}{2}\sum_{i=1}^{n} f''\left(B(t_{i-1}) + \lambda_i(B(t_i)-B(t_{i-1}))\right)(B(t_i)-B(t_{i-1}))^2 \quad (2)
\end{aligned}$$

and (1) converges to $\int_a^b f'(B(t))dB(t)$ by theorem 4.7.1. Thus, it suffices to show that (2) converges to $\frac{1}{2}\int_a^b f''(B(t))dt$ for some subsequence $(n_k)_{k=1}^{\infty} \to \infty$.

STEP 1 $\sum_{i=1}^{n}\left[f''(B_{t_{i-1}}) + \lambda_i(B(t_i)-B(t_{i-1})) - f''(B(t_{i-1}))\right](B(t_i)-B(t_{i-1})) \to 0$

Let $Y_{i,n} = \sum_{i=1}^{n}\left[f''(B_{t_{i-1}}) + \lambda_i(B(t_i)-B(t_{i-1})) - f''(B(t_{i-1}))\right]$. Then,

$$\sum |Y_{i,n}(B(t_i)-B(t_{i-1}))^2| \leq Y_n \sum (B(t_i)-B(t_{i-1}))^2,$$

where $Y_n = \max\{|Y_{1,n}|, \cdots, |Y_{n,n}|\}$ and note that $\sum(B(t_i) - B(t_{i-1}))^2$ converges to $b - a$. Moreover, $Y_n \to 0$ a.s.

($\because$ let $M = \max\{B(t) : t \in [a,b]\}$ and $m = \min\{B(t) : t \in [a,b]\}$. Then

$$Z(\delta) = \sup_{x,y \in [m,M], |x-y| < \delta} |f''(y) - f''(x)| \to 0$$

as $\delta \searrow 0$ by uniform continuity. Moreover $D_n := \sup_{i=1 \sim n} |B(t_i) - B(t_{i-1})| \to 0$ as $n \to \infty$ since Brownian motion is a.s. continous. Then, $Y_n \leq Z(D_n) \to 0$ a.s. as $n \to 0$.)

STEP 2 $\sum_{i=1}^n f''(B(t_{i-1})) \left[(B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1})\right] \to 0$

We skip the detailed case. If f'' bounded then $\mathbb{E}[\sum_{i=1}^n ()]^2 = \sum_{i=1}^n \mathbb{E}[()^2] + 2\sum_{i<j} \mathbb{E}[()()]$. The first term is $\mathbb{E}[\mathbb{E}[()^2 | \mathcal{F}_{t_{i-1}}]] = \mathbb{E}[f''(B(t_{i-1}))^2 2(t_i - t_{i-1})^2] \leq c/n^2$, and the second term is $\mathbb{E}[\mathbb{E}()() | \mathcal{F}_{t_{j-1}}] = 0$

For the non-bounded case, take a bound L so that $E_L = [-L, L]$ and $\mathbb{1}_{E_L} f''$ be bounded. Then, $f'' = \mathbb{1}_{E_L} f''$ for sufficient large L.

STEP 3 $\sum_{i=1}^n f''(B(t_{i-1}))(t_i - t_{i-1}) \to \int_a^b f''(B(t))dt$

By definition of Riemann integral.

The proof ends by adding the three steps. $\square$

Remark. In the right side of the equation, the first term is defined as Itô's integral (but we can interpret it as Riemann sense), and the second term is defined in the sense of Riemann.

Example. $e^B(t) - 1 = \int_0^t e^{B(s)} dB(s) + \frac{1}{2} \int_0^t e^{B(s)} ds$.
The first term is Martingale and the second term increases in t. Set $e^{B(t)} = X(t)$. Then,

$$X(t) - X(0) = \int_0^t X(s)dB(s) + \frac{1}{2} \int_0^t X(s)ds,$$

and

$$dX(t) = X(t)dB(t) + \frac{1}{2} X(t)dt \tag{SDE}$$

Remark. The first term is noise in stock price, and the second term is increasing in long-term.

7.3 Itô's Formula Slightly Generalized

Let $f(t,x) \in C^{1,2}$ (i.e. f_t, f_x, f_{xx} are continuous) Then,

$$f(b, B(b)) - f(a, B(a)) = \int_a^b f_x(s, B(s))dB(s) + \int_a^b (f_t + \frac{1}{2} f_{xx})(s, B(s))ds$$

Proof. $f(b, B(b)) - f(a, B(a)) = \sum_{i=1}^{n} (f(t_i B_{t_i}) - f(t_{i-1}, B_{t_{i-1}}))$, and

$$\begin{aligned} f(t_i, B_{t_i} - f(t_{i-1}, B_{t_{i-1}}) &= [f(t_i, B_{t_i}) - f(t_{i-1}, B_{t_i})] + [f(t_{i-1}, B_{t_i}) - f(t_{i-1}, B_{t_{i-1}})] \\ &= f(s_i, B_{t_i})(t_i - t_{i-1}) + \text{ same as before} \\ &\to \int_a^b f_t(s, B(s))ds \end{aligned}$$

□

Example. $f(t, x) = e^{x - \frac{1}{2}t}$ (or $e^{cx - \frac{c^2}{2}t}$)
Then, $f(t, B(t)) - f(0, B(0)) = \int_0^t f(s, B(s))dB(s)$ and

1. $X(t) = e^{B(t) - \frac{1}{2}t}$ is Martingale
2. $dX(t) = X(t)dB(t)$
3. $\mathbb{E}[X(t)] = \mathbb{E}[X(0)] = 1$

7.4 Itô's Formula in the General Form

$$X(t) = X(0) + \int_0^t f(s)dB(s) + \int_0^t g(s)ds,$$

where $f \in L^2_{\text{ad}}$; $\int_0^t \mathbb{E}[f(s)^2]ds < \infty \ \forall t$ and $g \in L^2_{\text{ad}}$; g is Riemann integrable a.s.

Theorem 7.4.1
$\theta \in C^{1,2}(\mathbb{R} \times \mathbb{R})$. *Then,*

$$\begin{aligned} \theta(b, X(b)) - \theta(a, X(a)) = &\int_a^b \theta_x(s, X(s))f(s)dB(s) \\ &+ \int_a^b \left(\theta_t + g(s)\theta_x(s, X(s)) + \frac{1}{2}f^2\theta_{xx} \right)(s, X(s))ds \end{aligned}$$

Why? $d\theta(t, B(t)) = \theta_t dt + \theta_x dB_t + \frac{1}{2}\theta_{xx}(dB_t)^2$ *and* $(dB_t)^2 = dt$ *(Itô formula),* $(dB_t)^3 = 0$. *Thus,* $d\theta(t, X(t)) = \theta_t dt + \theta_x dX(t) + \frac{1}{2}\theta_{xx}(dX(t))^2 = \theta_t dt + f dB(t) + g dt + f^2 dt$.

Review
g: differentiable, $f \in C^{1,1} \Rightarrow df(t, g(t)) = f_t(t, g(t))dt + f_x(t, g(t))dg(t)$ Itô's formula is a Brownian version of the above $(f, (t, B(t)))$.
Version 1. $df(B(t)) = f'(B(t))dB(t) + \frac{1}{2}f''(B(t))dt$
Version 2. $df(t, B(t)) = f_t dt + f_x dB(t) + \frac{1}{2}f_{xx}dt = f_x dB(t) + (f_t + \frac{1}{2}f_{xx})dt$

Example (Heat Equation).

$$f(x) = \begin{cases} f_t = \frac{1}{2} f_{xx} \\ f(0,x) = h(x) \end{cases}$$

Question. What is $f(T, x_0) =$?
Answer: $g(t,x) = f(T-t, x+x_0) \Rightarrow g_t = -f_t(T-t, x+x_0), g_{xx} = f_{xx}(T-t, x+x_0)$
$\therefore f_t = \frac{1}{2} f_{xx} \Rightarrow g_t + \frac{1}{2} g_{xx} = 0$
$\therefore g(T, B(T)) = -g(0, B(0)) = \int_0^T g_x(t, B(t)) dB(t)$

Remark. $X_t = \int_0^t f(s) dB(s) \Rightarrow X_t$ is Martingale and continuous path $\Rightarrow \mathbb{E}[X_t] = \mathbb{E}[X_t | \mathcal{F}_0] = X_0 = 0$.

$\therefore \mathbb{E}[g(T, B(T))] = -g(0,0) = f(T, x_0) = 0$.
$\therefore f(T, x_0) = \mathbb{E}[f(0, x_0 + B(T))] = \mathbb{E}[h(x_0 + B(T))] = \int_{-\infty}^{\infty} h(x_0 + y) \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2T}} dy$
$\therefore f(T, x_0) = \int_{-\infty}^{\infty} h(x_0 - y) \frac{1}{\sqrt{2\pi T}} e^{-\frac{y^2}{2T}} dy = (h * P_T)(x_0)$. ($P_t$: Heat Kernel)

Remark. It is a canonical relation of heat equation(PDE) and Brownian motion(probability).

Version 3. $\theta(t,x) \in C^{1,2}$. A **integral form** is

$$X_t = X_0 + \int_0^t f(s) dB(s) + \int_0^t g(s) ds,$$

where f, g are stochastic process: $f \in L^2_{\text{ad}}, g \in L^1_{\text{ad}}$. A **stochastic differential** is

$$dX_t = f(t) dB(t) + g(t) dt$$

Formal computation: $\begin{cases} dtdt = 0 \\ dB(t)dt = 0 \\ dB(t)dB(t) = dt \end{cases}$

Theorem 7.4.2
Version 1.

$$df(B_t) = f'(B_t) dB_t + \frac{1}{2} f''(B_t)(dB_t)^2$$

Version 2.

$$df(t, B_t) = f_t dt + f_x dB_t + \frac{1}{2} f_{xx} (dB_t)^2$$

Version 3.

$$d\theta(t, X_t) = \theta_t dt + \theta_x dX_t + \frac{1}{2} \theta_{xx} (dX_t)^2$$

(Note: $dX_t = f dB_t + g dt \Rightarrow (dX_t)^2 = (f dB_t + g dt)^2 = f^2 dt$)
Version 3.*

$$d\theta(t, X_t) = [\theta_t + g\theta_x + \frac{1}{2} f^2 \theta_{xx}] dt + f \theta_x dB_t,$$

and its integral form is

$$\theta(t, X_0) - \theta(0, X_0) = \int_0^t \theta_t + g\theta_x + \frac{1}{2}f^2\theta_{xx}dt + \int_0^t f\theta_x dB_t$$

Example (Geometric Brownian Motion).

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t).$$

Set $\theta(t, x) = \log x$.

$$\begin{aligned}
d\theta(t, S(t)) &= \theta_x(t, S(t))dS(t) + \frac{1}{2}\theta_{xx}(t, S(t))(dS(t))^2 \\
&= \frac{1}{S(t)}\left[\mu S(t)dt + \sigma S(t)dB(t)\right] + \frac{1}{2}(-\frac{1}{S(t)^2})\sigma^2 S(t)^2(dB(t))^2 \\
&= (\mu - \frac{1}{2}\sigma^2)dt + \sigma dB(t).
\end{aligned}$$

It's integral form is

$$\theta(t, S(t)) - \theta(0, S(0)) = (\mu - \frac{1}{2}\sigma^2)t + \sigma B(t).$$

Therefore $\log S(t) - \log S(0) = (\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)$ and $S(t) = S(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)}$

Example (Vasicek's Model for interest rate).

$$dr_t = (\alpha - \beta r_t)dt + \sigma dB(t).$$

(Mean reverting to α/β, and $\sigma dB(t)$ is noise)
Set $\theta(t, x) = e^{\beta t}x$.

$$\begin{aligned}
d\theta(t, r_t) &= \theta_t(t, r_t)dt + \theta(t, r_t)\left[(\alpha - \beta r_t)dt + \sigma dB_t\right] + \frac{1}{2}\theta_{xx}(t, r_t)\sigma^2 dt \\
&= \beta e^{\beta t}r_t dt + e^{\beta t}x[(\alpha - \beta r_t)dt + \sigma dB_t] \\
&= \alpha e^{\beta t}dt + \sigma e^{\beta t}dB_t.
\end{aligned}$$

It's integral form is

$$\theta(t, r_t) - \theta(0, r_0) = \int_0^t \alpha e^{\beta s}ds + \int_0^t \sigma e^{\beta s}dB(s).$$

Therefore, $e^{\beta t}r_t - r_0 = \alpha\frac{e^{\beta t-1}}{\beta} + N(0, \sigma^2\frac{e^{2\beta t}-1}{2\beta})$ and

$$r_t = e^{-\beta t}r_0 + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + N\left(0, \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})\right).$$

Furthermore, $r_t \to \frac{\alpha}{\beta} + N(0, \frac{\sigma^2}{2\beta})$ as $t \to 0$.

Multidimensional Brownian Motion

Brownian motion in $\mathbb{R}^d$ is a d-independent Brownian motion $B(t) = (B_1(t), B_2(t), \cdots, B_d(t))$.

$$\begin{aligned} B(t) : \Omega &\longrightarrow \mathbb{R}^d \\ \omega &\longmapsto (B_1(t,\omega), \cdots, B_d(t,\omega)) \end{aligned}$$

Remark. In d-dimensional Brownian motion $B(t)$, consider a d-dimensional ball A. If $d = 1$, then obvious $\phi(d) = \mathbb{P}[\exists t \text{ s.t. } B(t) \in A] = 1$. If $d = 2$, then $\phi(d) = 1$, and $\phi(d) < 1$ for $d > 2$ (not zero)

Goal of the today: Multidimensional Itô's formula (chapter 7) and SDE (chapter 10).
Let $B_1(t), B_2(t), \cdots, B_n(t)$ be independent Brownian motions.

$$\begin{aligned} dX_t^{(1)} &= f_{11}(t)dB_1(t) + f_{12}(t)dB_2(t) + g_1(t)dt \\ dX_t^{(2)} &= f_{21}(t)dB_1(t) + f_{22}(t)dB_2(t) + g_2(t)dt, \end{aligned} \tag{$\bigstar$}$$

$f_{11}, f_{12}, f_{21}, f_{22} \in L^2_{\text{ad}}, g_1, g_2 \in L^1[a,b]$ a.s. ($\bigstar$) can be written as

$$d\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} d\begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix} + \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} dt$$

or

$$dY_t = F(t)dB(t) + Gdt.$$

Y_t is diffusivity, $B(t)$ is 2-dim Brownian motion, G is drift.

<u>Note</u>

- $dtdt = dB_1dt = dB_2dt = dB_1dB_2 = 0$
- $(dB_1)^2 = (dB_2)^2 = dt$

Theorem 7.4.3

$\theta(t,x,y)$, $\theta : [a,b] \times \mathbb{R}^2 \to \mathbb{R} \in C^{1,2}([a,b] \times \mathbb{R}^2)$. Then,

$$d\theta(t, X_t^{(1)}, X_t^{(2)}) = \theta_t dt + \theta_x dX_t^{(1)} + \theta_y dX_t^{(2)} + \frac{1}{2}\theta_{xx}\left(dX_t^{(1)}\right)^2 + + \frac{1}{2}\theta_{yy}\left(dX_t^{(2)}\right)^2 + \theta_{xy} dX_t^{(1)} dX_t^{(2)}$$

Example.

1. $f(t, B_1(t), B_2(t))$

$$\begin{aligned} df(t, B_1(t), B_2(t)) &= f_t dt + f_x dB_1 + f_y dB_2 + \frac{1}{2}f_{xx}dt + \frac{1}{2}f_{yy}dt \\ &= \left(f_t + \frac{1}{2}(f_{xx} + f_{yy})\right)dt + f_x dB_1 + f_y dB_2 \end{aligned}$$

Using this, we can solve the heat equation $g_t = \frac{1}{2}(g_{xx} + g_{yy}) = \frac{1}{2}\Delta g$ by putting $f(t,x,y) = g(T-t, x+x_0, y+y_0)$.

2. $X_t, Y_t; \theta(t,x,y) = xy$.

$$d(X_tY_t) = Y_t dX_t + X_t dY_t + dX_t dY_t \qquad \text{(product rule)}$$

3. $\theta(t,x,y) = \frac{x}{y}$. Find $d(\frac{X_t}{Y_t})$ (exercise!)

4. $dX_t = aX_t dB_1(t) + bX_t dB_2(t) + \alpha X_t dt$, $dY_t = cT_y dB_1(t) + dY_t dB_2(t) + \beta Y_t dt$. (Two noise in the stock price)

$$\begin{aligned} d(X_tY_t) &= X_t dY_t + Y_t dX_t + dX_t dY_t \\ &= X_tY_t((a+c)dB_1(t) + (b+d)dB_2(t)) + X_tY_t(\alpha + \beta + (ac+bd))dt \end{aligned}$$

Put $Z_t = X_tY_t$. Then,

$$\begin{aligned} dZ_t &= Z_t\left((a+c)dB_1(t) + (b+d)dB_2(t)\right) + (\alpha + \beta + (ac+bd))Z_t dt \\ &= \sqrt{(a+c)^2 + (b+d)^2} dB(t) + +(\alpha + \beta + (ac+bd))Z_t dt \\ &= \sigma Z_t dB_t + \mu Z_t dt \\ &\Rightarrow Z_t \text{ is a generalized Brownian motion.} \end{aligned}$$

Note: $\frac{\alpha B_1 + \beta B_2}{\sqrt{\alpha^2 + \beta^2}}$ is a Brownian motion (exercise)

Chapter 10

Stochastic Differential Equations

- $\sigma(t, x), f(t, x) : [a, b] \times \mathbb{R} \to \mathbb{R}$ (measurable)
- $B(t)$: Brownian motion
- $\mathcal{F}_t$: Brownian Filtration (i.e., $B(t) - B(s)$ is independent of $\mathcal{F}_t$ for all $s < t$)
- ξ: $\mathcal{F}_a$-measurable random variable.

SDE (Stochastic Differential Euqation)

$$dX_t = \sigma(t, X_t)dB(t) + f(t, X_t)dt; X_a = \xi.$$

is equivalent to **SIE** (Stochastic Integral Equation)

$$X_t = \int_a^t \sigma(s, X_s)dB(s) + \int_a^t f(s, X_s)ds + \xi.$$

10.1 Some Examples

Definition 10.1.1 (Solution of SDE)
X_t: Stochastic Process on $t \in [a, b]$. X_t is a solution of SDE if

1. $\sigma(t, X_t) \in L^2_{\text{ad}}([a, b], L^2(\Omega))$ $\forall t$ (*cf.* Ch 5.)
2. $f(t, X_t) \in L^1([a, b])$ a.s.
3. SIE holds $\forall t \in [a, b]$ a.s.

Remark. 1 implies that the first term of SIE is defined, and 2 implies that the second term of SIE is defined. Thus, the important condition is 3.

Question. 1. Exist? 2. Unique?

Example.

1. (Non-Existence) $dX_t = X_t^2 dB_t + X_t^3 dt, X_0 = 1.$

   $Y_t = \frac{1}{X_t} = f(X_t), f(x) = \frac{1}{x}$. Then,

$$
\begin{aligned}
dY_t &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 \\
&= \left(-\frac{1}{X_t^2}\right)(X_t^2 dB_t + X_t^3 dt) + \frac{1}{2}\left(\frac{2}{X_t^3}\right)X_t^4 dt \\
&= -dB_t
\end{aligned}
$$

   $\therefore Y_t = Y_0 - B = 1 - B_t$ and $X_t = \frac{1}{1-B_t}$, which makes an explosion on $B_t = 1$

2. (Non-Uniqueness) $dX_t = 3X_t^{\frac{2}{3}} dB_t + 3X_t^{\frac{1}{3}} dt, X_0 = 0.$

   $f_a(x) := \begin{cases} (x-a)^3 & x \geq a \\ 0 & x < a \end{cases}$ $(a > 0) \in C^2(\mathbb{R})$. Let $Y_t = f_a(B_t)$. Then,

$$
\begin{aligned}
dY_t &= f_a'(B_t)dB_t + \frac{1}{2}f_a''(B_t)dt \\
&= 3f_a(x)^{2/3}(B_t)dB_t + 3f_a(x)^{1/3}(B_t)dt \\
&= 3Y_t^{2/3}dB_t + 3Y_t^{1/3}dt.
\end{aligned}
$$

   Therefore, Y_t is invariant of a, and the solution of Y_t is not unique.

10.2 Bellman-Gronwall Inequality

We only concern Gronwall's Inequality. It is related to uniqueness.

Theorem 10.2.1 (Gronwall's Inequality)
Let $\phi, f \in L^1([a,b]), \beta > 0$. If $\phi(t) \leq f(t) + \beta\int_a^t \phi(s)ds\ \forall t \in [a,b]$, then

$$
\phi(t) \leq f(t) + \beta\int_a^t f(s)e^{\beta(t-s)}ds.
$$

Remark.

- f can be negative-valued.
- $f = 0$ is an interesting case. (If ϕ is non-negative, then $\phi = 0$.)

Proof. Let $g(t) = \beta\int_a^t \phi(s)ds$. Then, $g'(t) = \beta\phi(t), \frac{1}{\beta}g'(t) \leq f(t) + g(t)$, and $g'(t) - \beta g(t) \leq \beta f(t)$. Let $h(t) = e^{-\beta t}g(t)$ so that $h'(t) = e^{-\beta t}[g'(t) - \beta g(t)] \leq \beta e^{-\beta t}f(t)$. Therefore, $h(t) \leq \int_a^t \beta e^{-\beta s}f(s)ds$, $e^{-\beta t}g(t) \leq \int_a^t \beta e^{-\beta s}f(s)ds$, and $g(t) \leq \int_a^t \beta e^{\beta(t-s)}f(s)ds$. $\square$

10.3 Existence and Uniqueness Theorem

SDE is a differential equation of the form $dX_t = \sigma(t, X_t)dB_t + f(t, X_t)dt$ and $X_a = \xi : \mathcal{F}_a$-measurable, $\mathbb{E}[\xi^2] < \infty$. To show the uniqueness and existence of SDE, we need

Theorem 10.3.1
Let $dX_t = \sigma(t, X_t)dB_t + f(t, X_t)dt$.

1. *If σ, f are Lipschitz (uniformly in time), then the solution is unique*

2. *If σ, f are Lipschitz and Linear Growth, then the solution uniquely exists.*

Note

- Lipschitz (uniformly in time)

$$|g(t,x) - g(t,y)| \leq c|x - y| \text{ for some } c > 0$$

- Linear Growth

$$|g(t,x)| \leq c(|x| + 1) \text{ for some } c > 0$$

- X_t (SP) is continuous if $h(t) = X_t(\omega)$ is continous in t for a.s. $\omega \in \Omega$.

$$\begin{aligned} X :& [a,b] \times \Omega \quad \rightarrow \quad \mathbb{R} \\ & (t, \omega) \qquad \mapsto X(t,\omega) = X_t(\omega) \end{aligned}$$

Next Classes
1,2,3 $\longrightarrow$ prove uniqueness and existence
4 $\longrightarrow$ Markov Property of (X_t)
5 $\longrightarrow$ Martingale (12/12)

Proof of uniqueness. Assume the solution exists. Let there are two solutions X_t, Y_t of SDE. To show the uniqueness, we have to show

$$\mathbb{P}[X_t = Y_t \ \ \forall t \in [a,b]] = 1.$$

Write X_t and Y_t as SIE form:

$$X_t = \xi + \int_a^t \sigma(s, X_s)dB_s + \int_a^t f(s, X_s)ds,$$

$$Y_t = \xi + \int_a^t \sigma(s, Y_s)dB_s + \int_a^t f(s, Y_s)ds,$$

and the difference $Z_t = X_t - Y_t$ can be written as $\int(1)dB_s + \int(2)ds$.
Part (1): $M_t = \int_a^t [\sigma(s, X_s) - \sigma(s, Y_s)]dB_s$

$$\begin{aligned}
\mathbb{E}[M_t^2] &= \int_a^t \mathbb{E}\left[(\sigma(s, X_s) - \sigma(s, Y_s))^2\right] ds \\
&= \int_a^b \mathbb{E}\left[L^2(X_s - Y_s)^2\right] ds && \text{(By Lipschitz)} \\
&= L^2 \int_a^t \mathbb{E}\left[Z_s^2\right] ds.
\end{aligned}$$

Part (2): $N_t = \int_a^b [f(s, X_s) - f(s, Y_s)]\, ds$.

$$\begin{aligned}
N_t^2 &\le \int_a^t 1ds \int_a^t [f(s, X_s) - f(s, Y_s)]^2\, ds \\
&\le (b-a) \int_a^t L^2(X_s - Y_s)^2 ds && \text{(By Lipschitz)} \\
&= L^2(b-a) \int_a^t Z_s^2 ds.
\end{aligned}$$

$\therefore Z_t = M_t + N_t, Z_t^2 = (M_t + N_t)^2 \le 2M_t^2 + 2N_t^2$, and

$$\phi_t \le 2L^2 \int_a^t \phi(s)ds + 2L^2(b-a) \int_a^b \phi(s)ds.$$

Apply Gronwall's inequality with $Z_t = X_t - Y_t, \phi(t) = \mathbb{E}[Z_t^2], \phi(t) \le 2L^2(1 + b - a) \int_a^t \phi(s)ds$. Then, $\phi(t) \le 0 \Rightarrow \phi(t) = 0 \;\; \forall t \in [a, b]$. Since the expectation of square is zero, $Z_t = 0 \;\; \forall t \in [a, b]$ a.s.
Let $\{r_1, r_2, \cdots\}$ be enumeration of $\mathbb{Q} \cap [a, b]$. Define $\Omega_i = \{\omega : Z_{r_i}(\omega) = 0\}$, and $\mathbb{P}(\Omega_i) = 1$. Then, $\Omega_0 = \bigcap_{i=1}^{\infty} \Omega_i$ is to be $\mathbb{P}(\Omega_0) = 0$. Now, $\omega \in \Omega_0, Z_t(\omega) = 0$ for all $t \in \mathbb{Q} \cap [a, b]$. Let $\Omega_X = \{\omega : X_t(\omega) \text{ is continuous in t}\}$, and $\Omega_Y = \{\omega : Y_t(\omega) \text{ is continuous in t}\}$. Then, $\mathbb{P}(\Omega_X) = \mathbb{P}(\Omega_Y) = 1$, and $\Omega^* = \Omega_0 \cap \Omega_X \cap \Omega_Y$ is to be $\mathbb{P}(\Omega^*) = 1$.

$$\therefore \omega \in \Omega^* \Rightarrow \begin{cases} Z_t(\omega) = 0 \;\; \forall t \in \mathbb{Q} \cap [a, b] \\ Z_t(\omega) \text{ is continuous in } t. \end{cases} \Rightarrow Z_t(\omega) = 0 \;\; \forall t \in [a, b].$$

$\therefore \mathbb{P}\left[\{\omega : X_t(\omega) - Y_t(\omega) \;\; \forall t \in [a, b]\}\right] = 1$. In other words, $X_t = Y_t$. $\square$

To show the existence, we need

1. Bellman-Gronwall's Inequality
2. Borel-Cantelli's Lemma (3.2.1)
3. Doob's (Sub)Martingale Inequality (4.5.1)

Theorem 10.3.2 (Doob's Martingale inequality)
If M_t is Martingale in $[a,b]$, continuous (i.e., $\mathbb{E}[M_t|\mathcal{F}_s] = M_s \ \forall a \le s < t \le b$), then

$$\mathbb{P}\left[\sup_{a\le t\le b} |M_t| \ge \epsilon\right] \le \frac{\mathbb{E}[|M_b|]}{\epsilon}.$$

(cf. $\mathbb{E}[|M_t|]$ is increasing in t, so is $\mathbb{E}[\phi(M_t)]$ for convex ϕ.)

Remark. Applying Chebyshev's inequality,

$$\mathbb{P}\left[\sup_{a\le t\le b} |M_t| \ge \epsilon\right] \le \frac{\mathbb{E}[\sup_{a\le t\le b}|M_b|]}{\epsilon} \le \frac{\sup_{a\le t\le b}\mathbb{E}[|M_t|]}{\epsilon} = \frac{\mathbb{E}[|M_t|]}{\epsilon}$$

Theorem 10.3.3 (Borel-Cantelli's Lemma)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For $A_1, A_2, \cdots$,

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n] < \infty \Rightarrow \mathbb{P}[A_n\text{'s occur infinitely often}] = 0.$$

Remark. A_n's occur infinitely often means the number of n such that $\omega \in A_n$ is infinite.

Proof. $\{\omega : \omega \in A_n \text{ for infinitely many } n\} = \bigcap_{n=1}^{\infty}\left[\bigcup_{k=n}^{\infty} A_k\right] =: \bigcap_{n=1}^{\infty} B_n$, and $B_1 \supseteq B_2 \cdots \to B$.

$$\begin{aligned}
\therefore \mathbb{P}\left[A_n\text{'s occur infinitely often}\right] &= \mathbb{P}[\bigcap_{n=1}^{\infty} B_n] \\
&= \lim_{n\to\infty} \mathbb{P}(B_n) \\
&\le \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k) \\
&= 0.
\end{aligned}$$

□

Ah pa seo bba jim bb.
Bellman's inequality (Review)

$$\begin{cases} \Theta_n(t) \le \alpha + \beta \int_a^t \Theta_n(s)ds \\ \Theta_1(t) = c. \end{cases} \Rightarrow \Theta_{n+1}(t) \le \alpha e^{\beta(t-a)} + c\frac{\beta^n(t-a)^n}{n!}.$$

Existence and Uniqueness

- $dX_t = \sigma(t, X_t)dW_t + f(t, X_t)dt.$
- $X_\infty = \xi.$

If $\mathbb{E}[\xi^2] < \infty$ and σ, f are Lipschitz and Linear growth, then there is a unique solution $X \in L^2_{\text{ad}}([a,b] \times \Omega)$ of SDE.

Idea. Let $X_t^{(1)} = \xi \ \forall t$, and let $X_t^{(n+1)} = \xi + \int_a^t \sigma(s, X_s^{(n)}) dB_s + \int_a^t f(s, X_s^{(n)}) ds$. Then, $X_t^{(n)} \to X_t$.

1. $\int_a^b \mathbb{E}[(X_t^{(n)})^2] dt < \infty \ \forall n$.

   (a) $n = 1$: trivial.

   (b) $n \to n+1$.
   $\mathbb{E}[(X_t^{(2)})^2] =: \Theta_n(t)$. Then,

$$\begin{aligned}
\Theta_{n+1}(t) &= \mathbb{E}\left[(X_t^{n+1})^2\right] \\
&\le 3\mathbb{E}[\xi^2] + 3\mathbb{E}\left[\left(\int_a^t \sigma(s, X_s^{(n)}) dB_s\right)^2\right] + 3\mathbb{E}\left[\left(\int_a^t f(s, X_s^{(n)}) ds\right)^2\right] \\
&\le 3\mathbb{E}[\xi^2] + 3\int_0^t \mathbb{E}\left[\sigma(s, X_s^{(n)})^2\right] ds + 3\mathbb{E}\left[\int_a^t 1 ds \int_a^t f(s, X_s^{(n)})^2 ds\right] \\
&\le c_1 + c_2 \int_a^t \left(1 + \mathbb{E}\left[\left(X_s^{(n)}\right)^2\right]\right) + c_3 \int_a^t \left(1 + \mathbb{E}[X_s^{(n)}]^2\right) ds \\
&\le c_4 + c_5 \int_a^t \mathbb{E}\left[(X_s^{(n)})^2\right] ds \\
&= c_4 + c_5 \int_a^t \Theta_n(s) ds.
\end{aligned}$$

   $\therefore \Theta_{n+1}(t) \le c_4 + c_5 \int_a^t \Theta_n(s) ds$ and $\Theta_1(t) = \mathbb{E}[\xi^2]$. By Bellman's inequality,

$$\Theta_{(n+1)}(t) \le c_4 e^{c_5(t-a)} + \mathbb{E}[\xi^2] \frac{c_5^n (t-a)^n}{n!}.$$

Therefore, $X_t^{(n+1)}$ is well-defined. (the second term is Itô's integral in the sense of Ch4.)
The difference of $X_t^{(n)}$ is

$$X_{t+1}^{(n)} - X_t^{(n)} = \int_a^t \left[\sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)}) dB_s\right] + \int_a^t \left[f(s, X_s^{(n)}) - f(s, X_s^{(n-1)}) dB_s\right]$$

and denote it $X_{t+1}^{(n)} - X_t^{(n)} = \bigstar + \spadesuit$.

2. $\mathbb{E}\left[|X_t^{(n+1)} - X_t^{(n)}|^2\right] \le 2\mathbb{E}[\bigstar^2] + 2\mathbb{E}[\spadesuit^2]$.

   (a) $\mathbb{E}[\bigstar^2] = \int_a^t \mathbb{E}\left[\left(\sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)})\right)^2\right] ds \le c \int_a^t \mathbb{E}\left(|X_s^{(n)} - X_s^{(n+1)}|^2\right)$

   (b) $\mathbb{E}[\spadesuit^2] = \mathbb{E}\left[\int_a^t 1 ds \int_a^t \mathbb{E}\left[\left(f(s, X_s^{(n)}) - f(s, X_s^{(n-1)})\right)^2\right] ds\right] \le (b-a) \int_a^t \mathbb{E}\left(|X_s^{(n)} - X_s^{(n+1)}|^2\right)$

Let $\varphi_n(t) = \mathbb{E}\left[\left(X_t^{(n)} - X_t^{(n-1)}\right)^2\right]$. Then, $\varphi_{n+1}(t) \leq M \int_a^t \varphi_n(s)ds$.

Remark. $X_t^{(1)} = \xi$, and $X_t^{(2)} = \xi + \int_a^t \sigma(s,\xi)dB_s + \int_a^t f(s,\xi)ds$. Then,

$$\varphi_2(t) = \mathbb{E}[(X_t^{(2)} - X_t^{(1)})^2] \leq c(1 + \mathbb{E}[\xi^2]) := \rho.$$

By Bellman's inequality,

$$\varphi_{n+1}(t) \leq \rho \frac{M^{n-1}(t-a)^{n-1}}{(n-1)!} \leq \rho \frac{M^{n-1}(b-a)^{n-1}}{(n-1)!} := \alpha_n \searrow 0.$$

3. $\mathbb{P}\left[\sup_{a \leq t \leq b} \left|X_t^{(n+1)} - X_t^{(n)}\right| \geq \frac{1}{n^2}\right] \leq cn^4 \alpha_{n-1}$.

$$\begin{aligned}
\mathbb{P}\left[\sup_{a \leq t \leq b} \left|X_t^{(n+1)} - X_t^{(n)}\right| \geq \frac{1}{n^2}\right] &\leq \mathbb{P}\left[\left\{\sup_{a \leq t \leq b} |\bigstar| \geq \frac{1}{2n^2}\right\} \cup \left\{\sup_{a \leq t \leq b} |\spadesuit| \geq \frac{1}{2n^2}\right\}\right] \\
&\leq \mathbb{P}\left[\sup_{a \leq t \leq b} |\bigstar| \geq \frac{1}{2n^2}\right] + \mathbb{P}\left[\sup_{a \leq t \leq b} |\spadesuit| \geq \frac{1}{2n^2}\right] \\
&= p_1 + p_2.
\end{aligned}$$

(a) By Doob's Martingale inequality, i.e. $\mathbb{P}[\sup_{a \leq t \leq b} |M_t| \geq \alpha] \leq \dfrac{\mathbb{E}[M_b^2]}{\alpha^2}$

$$\begin{aligned}
p_1 &\leq \frac{1}{(\frac{1}{2n^2})^2} \mathbb{E}\left[\left\{\int_a^b \left(\sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n)})\right) dB_s\right\}\right]^2 \\
&\leq 4n^4 c \int_a^b \mathbb{E}\left[\left(X_s^{(n)} - X_s^{(n-1)}\right)^2\right] ds \\
&\leq 4n^4 c(b-a)\alpha_{n-1} \\
&= cn^4 \alpha_{n-1}.
\end{aligned}$$

(b)

$$\begin{aligned}
|\spadesuit|^2 &\leq \int_a^t 1ds \int_a^t \left(f(s, X_s^{(n)}) - f(s, X_s^{(n-1)})\right)^2 ds \\
&\leq (b-a) \int_a^b \left(f(s, X_s^{(n)}) - f(s, X_s(n-1))\right)^2 ds \\
&\leq c \int_a^b \left(X_s^{(n)} - X_s^{(n-1)}\right)^2 ds.
\end{aligned}$$

$\therefore \sup_{a \leq t \leq b} |\spadesuit| \leq \sqrt{c \int_a^b \left(X_s^{(n)} - X_s^{(n-1)}\right)^2 ds}$, and

$$\mathbb{P}\left[\sup_{a \leq t \leq b} |\spadesuit| \geq \frac{1}{2n^2}\right] \leq \frac{\mathbb{E}\left[\sup_{a \leq t \leq b} |\spadesuit|\right]}{(\frac{1}{2n^2})^2} \leq 4n^4 c(b-a)\alpha_{n-1}.$$

Since $\sum_{n=1}^{\infty} cn^4 \alpha_{n-1} < \infty$, by Borel-Cantelli's lemma

$$\mathbb{P}\left[\sup_{a\le t\le b}\left|X_t^{(n+1)} - X_t^{(n)}\right| \ge \frac{1}{n^2} \text{ infinitely often}\right] = 0$$

Therefore, on $\left\{\sup_{a\le t\le b}\left|X_t^{(n+1)} - X_t^{(n)}\right| \ge \frac{1}{n^2} \text{ finitely often}\right\}$ (with probability 1)

$$X_t = \xi + \sum_{k=1}^{n-1}\left(X_t^{(k+1)} - X_t^{(k)}\right)$$

converges uniformly to X_t by Weierstrass M-test.

□

Remark. We've showed $\int_a^b \mathbb{E}\left[X_t^2\right] dt < \infty$, and

$$X_t^{(n+1)} = \xi + \int_a^t \sigma(s, X_s^{(n)}) dB_s + \int_a^t f(s, X_s^{(n)}) ds$$

converges to

$$X_t = \xi + \int_a^t \sigma(s, X_s) dB_s + \int_a^t f(s, X_s) ds$$

by term by term (in L^2).

10.5 Markov Property

Prob → BM → Itô integral → SDE (→ Markov Property)
Notation (10.5.1 - 105.2)
$\mathbb{E}[X|Y_1, Y_2, \cdots, Y_n] = \mathbb{E}[X|\sigma(Y_1, Y_2, \cdots, Y_n)] = \theta(Y_1, Y_2, \cdots, Y_n)$ for some Borel θ. Note that $\mathbb{E}[Xg(Y_1, \cdots, Y_n)] = \mathbb{E}[\theta(Y_1, \cdots, Y_n)g(Y_1, \cdots, Y_n)]$ for some Borel g. What we want to do is calculating

$$\mathbb{E}[X|Y_1 = y_1, \cdots, Y_n = y_n] := \theta(y_1, \cdots, y_n).$$

(Note that $\mathbb{E}[X|A] = \mathbb{E}[X|\sigma(A)]$ does not makes sense since RHS is zero in this sense.)
$\mathbb{P}[A|\mathcal{G}] = \mathbb{E}[\mathbb{1}_A|\mathcal{G}]$. In particular, we have

$$\mathbb{P}[X \le x|Y_1, Y_2, \cdots, Y_n] = \mathbb{E}[\mathbb{1}_{X\le x}|Y_1, \cdots, Y_n],$$

$$\mathbb{P}[X \le x|Y_1 = y_1, Y_2 = y_2, \cdots, Y_n = y_n] = \mathbb{E}[\mathbb{1}_{X\le x}|Y_1 = y_1, \cdots, Y_n = y_n].$$

Definition 10.5.1 (Markov process)
$(X_t)_{t\in[a,b]}$ is **Markov process** if $\forall a \le t_1 < t_2 \cdots < t_n < t \le b$

$$\mathbb{P}\left[X_t \le x | X_{t_1}, X_{t_2}, \cdots, X_{t_n}\right] = \mathbb{P}[X_t \le x | X_{t_n}],$$

or equivalently

$$\mathbb{P}\left[X_t \le x | X_{t_1} = y_1, X_{t_2} = y_2, \cdots, X_{t_n} = y_n\right] = \mathbb{P}[X_t \le x | X_{t_n} = y_n],$$

$\forall y_1 \sim y_n, x \in \mathbb{R}$.

Example.

1. Suppose $X, Y_1, Y_2, \cdots, Y_n)$ has a joint PDF $f(x, y_1, y_2, \cdots, y_n)$. Then the joint PDF of $Y_1, Y_2, \cdots, Y_n$ is

$$h(y_1, y_2, \cdots, y_n) = \int_{-\infty}^{\infty} f(x, y_1, y_2, \cdots, y_n) dx.$$

We can take

$$\mathbb{P}[X \le x | Y_1 = y_1, \cdots, Y_n = y_n] = \frac{1}{h(y_1, \cdots, h_n)} \int_{-\infty}^{x} f(u, y_1, \cdots, y_n) du.$$

Proof. Let $\pi(y_1, \cdots, y_n) = \frac{1}{h(y_1, \cdots, h_n)} \int_{-\infty}^{x} f(u, y_1, \cdots, y_n) du$. Then,

$$\mathbb{E}[\mathbb{1}_{\{X \le x\}} g(Y_1, \cdots, Y_n)] = \int_{-\infty}^{x} \iint \cdots \int g(y_1, \cdots y_n) f(u, y_1, \cdots, y_n) du dy_1 \cdots dy_n,$$

$$\mathbb{E}[\pi(Y_1, \cdots, Y_n) g(Y_1, \cdots, Y_n)] = \iint \cdots \int (\pi g h)(y_1, \cdots, y_n) dy_1 \cdots dy_n.$$

□

2. Brownian Motions B_t

Let $0 = t_0 < t_1 < t_2 < \cdots t_n < t_{n+1} = t$. Then,

$$(B_{t_1}, \cdots, B_t, B_{t_{n+1}}) \sim \prod_{i=0}^{n} \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} e^{-\frac{(x_{i+1} - x_i)^2}{2(t_{i+1} - t_i)}},$$

and

$$(B_{t_1}, \cdots, B_t) \sim \prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} e^{-\frac{(x_{i+1} - x_i)^2}{2(t_{i+1} - t_i)}}.$$

Therefore by 1.

$\mathbb{P}[B_t \le x | B_{t_1} = y_1, \cdots, B_{t_n} = y_n]$

$$= \frac{1}{\prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} e^{-\frac{(y_{i+1} - y_i)^2}{2(t_{i+1} - t_i)}}} \int_{-\infty}^{\infty} \left[\prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} e^{-\frac{(y_{i+1} - y_i)^2}{2(t_{i+1} - t_i)}}\right] \frac{1}{\sqrt{2\pi(t - t_n)}} e^{-\frac{(u - y_n)^2}{2(t - t_n)}} du$$

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi(t - t_n)}} e^{-\frac{(u - y_n)^2}{2(t - t_n)}} du$$

$= \mathbb{P}\left[B_t \le x | B_{t_n} = y_n\right] \quad$ (Check!)

Therefore, Brownian motion is Markov process.

Lemma. Let $(X_t)_{t\in[a,b]}$ be stochastic process adapted to $(\mathcal{F}_t)_{t\in[a,b]}$, and $\mathbb{P}[X_t \le x|\mathcal{F}_s] = \mathbb{P}[X_t \le x|X_s]\ \ \forall s < t, x \in \mathbb{R}$. Then, (X_t) is a Markov process.

Proof.

$$\begin{aligned}
\mathbb{P}\left[X_t \le x|X_{t_1}, \cdots, X_{t_n}\right] &= \mathbb{E}[\mathbb{1}_{\{X_t\le x\}}|\sigma(X_{t_1}, \cdots, X_{t_n})] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{X_t\le x\}}|\mathcal{F}_{t_n}\right]|\sigma(X_{t_1}, \cdots, X_{t_n})\right] \\
&= \mathbb{E}[\mathbb{P}(X_t \le x|\mathcal{F}_{t_n})|\sigma(X_{t_1}, \cdots, X_{t_n})] \\
&= \mathbb{E}[\mathbb{P}(X_t \le x|X_{t_n})|\sigma(X_{t_1}, \cdots, X_{t_n})] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{X_t\le x\}}|\sigma(X_{t_n})\right]|\sigma(X_{t_1}, \cdots X_{t_n})\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\{X_t\le x\}}|\sigma(X_{t_n})\right] \\
&= \mathbb{P}[X_t \le x|X_{t_n}].
\end{aligned}$$

□

Theorem 10.5.2
σ, f is Lipschitz and Linear growth. ξ is $\mathcal{F}_a$-measurable and $\mathbb{E}[\xi^2] < \infty$. Then, X_t is a unique solution to

$$X_t = \xi + \int_a^t \sigma(s, X_s)dB_s + \int_a^t f(s, X_s)ds,$$

and therefore X_t is a Markov process.

Sketch of proof. Let $s \in [a, b]$ and Z be $\mathcal{F}_s$-measurable and $E[Z^2] < \infty$. Then,

$$Y_t = Z + \int_s^t \sigma(u, Y_u)dB_u + \int_s^t f(u, Y_u)du.$$

Then, there is a solution of the above SDE $\left(X_t^{s,Z}\right)_{t\in[s,b]}$. If $Z = x$ (non-random), write $X_t^{s,Z} = X_t^{s,x}$, and $\mathbb{P}[X_t^{s,x} \le y] =: f_{s,t,y}(x)$. Then,

$$\mathbb{P}\left[X_t^{s,Z} \le y|\mathcal{F}_s\right] = f_{s,t,y}(Z) \tag{$\bigstar$}$$

Note that proving ($\bigstar$) is difficult. Note that $t > s$, then $X_t = X_t^{s,X_s}$ (uniqueness) since

$$X_t = X_s + \int_s^t \sigma(u, X_u)dB_u + \int_s^t f(u, X_u)du.$$

Therefore if $s < t$

$$\mathbb{P}[X_t \le y|\mathcal{F}_s] = \mathbb{P}[X_t^{s,X_s} \le y|\mathcal{F}_s] = f_{s,t,y}(X_s).$$

Now $\mathbb{P}[X_t \le y|X_s] = \mathbb{E}[\mathbb{1}_{\{X_t\le y\}}|X_s] = \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{X_t\le y\}}|\mathcal{F}_s]|X_s] = \mathbb{E}[\mathbb{1}_{\{X_t\le y\}}|\mathcal{F}_s] = \mathbb{P}[X_t \le y|\mathcal{F}_s]$. The left is proving that $f_{s,t,y}$ is Borel measurable, and it is true. Therefore (X_t) is a Markov process by lemma. □

Note: Alternative way to define a Markov process.
Let $(X_t, \mathcal{F}_t)$ be a stochastic process. Then, X_t is a Markov process if

$$\forall f, \ \ s < t \ \ \exists g \text{ s.t. } \mathbb{E}[f(X_t)|\mathcal{F}_s] = g(X_s).$$

Next: Martingale (Does not cover final)

Appendix A

TA Session

Example (1.2.2). Let $(\Omega_\infty, \mathcal{F}_\infty, \mathbb{P})$ be the independent, infinite coin-toss space. Define stock price by

$$S_0(\omega) = 4 \quad \text{for all } \omega \in \Omega_\infty$$

$$S_1(\omega) = \begin{cases} 8 & \text{if } \omega_1 = H \\ 2 & \text{if } \omega_1 = T \end{cases}$$

$$S_2(\omega) = \begin{cases} 16 & \text{if } \omega_1 = \omega_2 = H \\ 4 & \text{if } \omega_1 \neq \omega_2 \\ 1 & \text{if } \omega_1 = \omega_2 = T \end{cases}$$

and in general

$$S_{n+1}(\omega) = \begin{cases} 2S_n(\omega) & \text{if } \omega_{n+1} = H \\ \frac{1}{2}S_n(\omega) & \text{if } \omega_{n+1} = T \end{cases}$$

Then, $S_0, S_1, \cdots$, are random variable.
For example, $\mathbb{P}(S_2 = 4) = \mathbb{P}(A_{HT} \cup A_{TH}) = 2pq$

Example (2.2.2). Let Ω be a three independent coin-toss space. Stock price random variables $S_0, S_1, \cdots$, are the same as the previous example. Let the probability measure $\mathbb{P}$ be given by

$$\mathbb{P}(HHH) = p^3, \mathbb{P}(HHT) = p^2 q, \cdots, \mathbb{P}(TTT) = q^3.$$

Assume $0 < p < 1$. Then, the random variables S_2 and S_3 are not independent.
$\because$ Consider the sets $\{S_3 = 32\} = \{HHH\}$ and $\{S_2 = 16\} = \{HHH, HHT\}$ whose probabilities are $\mathbb{P}(S_3 = 32) = p^3$ and $\mathbb{P}(S_2 = 16) = p^2$. In order to have Independence, $p^3 = \mathbb{P}(S_3 = 32) = \mathbb{P}(S_2 = 16 \text{ and } S_3 = 32) = \mathbb{P}(S_2 = 16)\mathbb{P}(S_3 = 32) = p^5 \Rightarrow\!\Leftarrow$.

The random variables S_2 and S_3/S_2 are independent. The σ-algebra generated by S_2 comprises ϕ, Ω, the atoms
$\{S_2 = 16\} = \{HHH, HHT\}, \{S_2 = 4\} = \{HTH, HTT, THH, THT\}, \{S_2 = 1\} = \{TTH, TTH\}$, and their unions.

The σ-algebra generated by S_3/S_2 comprises ϕ, Ω and
$\{S_3/S_2 = 2\} = \{HHH, HTH, THH, TTH\}, \{S_3/S_2 = \frac{1}{2}\} = \{HHT, HTT, THT, TTT\}$

For $A \in \sigma(S_2), B \in \sigma(S_3/S_2)$, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
ex) $p^3 = \mathbb{P}(S_2 = 16 \text{ and } S_3/S_2 = 2\} = \mathbb{P}(S_2 = 16)\mathbb{P}(S_3/S_2 = 2)) = p^2 p = p^3$

Example (2.2.10 Uncorrelated, dependent normal random variables).
Let X, Z be random variable satisfying

$$\begin{aligned} &X: \text{ standard normal random variable} \\ &Z: \text{ independent of } X,\ \mathbb{P}(Z=1) = \frac{1}{2},\ \mathbb{P}(Z=-1) = \frac{1}{2} \end{aligned}$$

Define $Y = ZX$. Show

1. Y is standard normal random variable
2. X and Y are uncorrelated but they are dependent.

Proof.

1. $$\begin{aligned} F_Y(b) &= \mathbb{P}(Y \le b) \\ &= \mathbb{P}(Y \le b \text{ and } Z = 1) + \mathbb{P}(Y \le b \text{ and } Z = -1) \\ &= \mathbb{P}(X \le b \text{ and } Z = 1) + \mathbb{P}(X \ge -b \text{ and } Z = -1) \\ &= \mathbb{P}(X \le b)\mathbb{P}(Z = 1) + \mathbb{P}(X \ge -b)\mathbb{P}(Z = -1) \\ &= \frac{1}{2}N(b) + \frac{1}{2}N(b) \\ &= N(b) \end{aligned}$$

2. Since $\mathbb{E}X = \mathbb{E}Y = 0$,

$$Cov(X, Y) = \mathbb{E}[XY] = \mathbb{E}[ZX^2] = \mathbb{E}[Z]\mathbb{E}[X^2] = 0$$

$\therefore$ X and Y are uncorrelated.
If X and Y are independent, $|X|$ and $|Y|$ are independent. But $\mathbb{P}(|X| \le 1, |Y| \le 1) = \mathbb{P}(|X| \le 1) = N(1) - N(-1)$, and $\mathbb{P}(|X| \le 1, |Y| \le 1) = \mathbb{P}(|X| \le 1)\mathbb{P}(|Y| \le 1) = (N(1) - N(-1))^2 \Rightarrow\Leftarrow$

$\square$

Let $\mu_{X,Y}$ be a joint distribution measure of (X, Y). Since $|X| = |Y|$, (X, Y) takes values only in the set $C = \{(x, y) : x = \pm y\}$.
It follows that for any measurable function f,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{1}_C(x,y) f_{X,Y}(x,y)dydx = 0$$

$\therefore$ There is no joint density $f_{X,Y}$ for (X, Y).

$$\begin{aligned}
F_{X,Y}(a,b) &= \mathbb{P}(X \le a, Y \le b) \\
&= \mathbb{P}(X \le a, X \le b, Z = 1) + \mathbb{P}(X \le a, -X \le b, Z = -1) \\
&= \frac{1}{2}\mathbb{P}(X \le a \wedge b) + \frac{1}{2}\mathbb{P}(-b \le X \le a) \\
&= \frac{1}{2}N(a \wedge b) + \frac{1}{2}((N(a) - N(-b)) \vee 0)
\end{aligned}$$

Example (2.2.12). Let (X, Y) be jointly normal with the density

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right)$$

Define $W = Y - \frac{\rho\sigma_2}{\sigma_1}X$. Then, X and W are independent.
Note that linear combination of jointly normal random variables are jointly normal (i.e., (X, W) is jointly normal).
Thus it suffices to show that $Cov(X, W) = 0$ (by Thm 2.2.9)

$$\begin{aligned}
Cov(X,W) &= \mathbb{E}[(X - \mathbb{E}X)(W - \mathbb{E}W)] \\
&= \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] - \mathbb{E}[\frac{\rho\sigma_2}{\sigma_1}(X - \mathbb{E}X)^2] \\
&= Cov(X,Y) - \frac{\rho\sigma_2}{\sigma_1}\sigma_1^2 \\
&= 0
\end{aligned}$$

Let $f_{X,W}$ be joint density of X and W.

$$\begin{aligned}
\mathbb{E}[W] &= \mu_2 - \frac{\rho\sigma_2\mu_1}{\sigma_1} =: \mu_3 \\
\mathbb{E}[(W - \mathbb{E}W)^2] &= \mathbb{E}[(Y - \mathbb{E}Y)^2] - \frac{2\rho\sigma_2}{\sigma_1}\mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] + \frac{\rho^2\sigma_2^2}{\sigma_1^2}\mathbb{E}[(X - \mathbb{E}X)^2] \\
&= \sigma^2 - \frac{2\rho\sigma_2}{\sigma_1}\rho\sigma_1\sigma_2 + \frac{\rho^2\sigma_2^2}{\sigma_1^2}\sigma_1^2 \\
&= (1-\rho^2)\sigma_2^2 =: \sigma_3^2
\end{aligned}$$

$\therefore f_{X,W}(x,w) = \frac{1}{2\pi\sigma_1\sigma_3} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(w-\mu_3)^2}{2\sigma_3^2}\right)$.
Note that we have decomposed Y into the linear combination $Y = \frac{\rho\sigma_2}{\sigma_1}X + W$ of a pair of independent normal random variables X and W.

Example (2.3.3). Let $\mathcal{G} = \sigma(X)$. Observe estimate Y based on X and error.

$$\begin{aligned} \mathbb{E}[Y|X] &= \frac{\rho\sigma_2}{\sigma_1} + \mathbb{E}[W] = \frac{\rho\sigma_2}{\sigma_1}(X - \mu_1) + \mu_2. \\ Y - \mathbb{E}[Y|X] &= W - \mathbb{E}[W] \end{aligned}$$

Note that the error is random variable with expected value zero and independent of the estimation $\mathbb{E}[Y|X]$.

Appendix B

Special lecture on Martingale

B.1 Doob-Meyer Decomposition

Let $(X_t, \mathcal{F}_t)$ be stochastic process and filtration. We call $(X_t, \mathcal{F}_t)$ **Martingale** if

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \ \ \forall s < t,$$

and **Sub Martingale** if

$$\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s \ \ \forall s < t.$$

Theorem B.1.1 (Doob-Meyer Decomposition)
Let RCLL denote right continuous and left limit. If $(X_t, \mathcal{F}_t)$ is a RCLL sub martingale, then there exists a unique decomposition $X_t = L_t + A_t$ such that L_t is RCLL Martingale and $(A_t, \mathcal{F}_t)$ is stochastic process, $A_0 = 0$, A_t is increasing in t RCLL.

Lemma. If M_t is a Martingale and φ is convex, then $\varphi(M_t)$ is submartingale.

Proof. By Jensen's inequality,

$$\mathbb{E}[\varphi(M_t)|\mathcal{F}_s] \geq \varphi(\mathbb{E}[M_t|\mathcal{F}_s]) = \varphi(M_s)$$

□

Remark. If M_t is a Marginal, then by lemma M_t^2 is submartingale. By theorem, $M_t^2 = L_t + A_t$; and call $A_t = \langle M \rangle_t$ quadratic variation.

Example.

1. $M_t = B_t$.

   $B_t^2 = (B_t^2 - t) + t$. $(B_t^2 - t)$ is a Martingale (by exam 1) and $t = \langle M \rangle_t$ is increasing.

2. $M_t = \int_0^t f(s)dB_s$ (f is a deterministic function).

   $\langle M\rangle_t = \int_0^t f(s)^2 ds$.

3. $M_t = B_t^2 - t$.

   $\langle M\rangle_t = 4\int_0^t B_s^2 ds$. Note that QV of M is a random variable.

B.2 Itô's Integral for Martingale

We want to find $\int_0^t f(s)dM_s$. Assume

- f satisfies $\mathbb{E}[\int_0^t f^2(s)d\langle M\rangle_s] < \infty$
- f is predictable (adapted if $M_t = B_t$)

Then, $I(f) = \int_0^t f(s)dM_s$ is defined as before.
cf. M_t is RCLL (Assumption)

Theorem B.2.1
$X_t = \int_0^t f(s)dM_s$. Then,

1. *X_t is RCLL Martingale. (Remark that even continuous M_t generates only RCLL X_t)*
2. $\mathbb{E}[X_t^2] = \mathbb{E}[\int_0^t f(s)^2 d\langle M\rangle_s]$

Sketch of proof. Let $f(s,\omega) = \sum_{i=0}^n \mathbb{1}_{[t_i,t_{i+1})}(s)f_i(w)$ for $t_0 = 0 < t_1 < \cdots < t_{n+1} = t$. Then,

$$\int f(s)dM_s = \sum f_i(\omega)\left(M_{t_{i+1}} - M_{t_i}\right) =: I(f),$$

$$\mathbb{E}[I(f)^2] = \sum_{i=1}^n \mathbb{E}\left(f_i^2(M_{t_{i+1}} - M_{t_i})^2\right) + 2\sum_{i<j}\mathbb{E}\left[f_i f_j(M_{t_{i+1}} - M_{t_i})(M_{t_{j+1}} - M_{t_j})\right].$$

Note that the second term is canceled, and

$$\mathbb{E}[f_i^2(M_{t_{i+1}} - M_{t_i})^2] = \mathbb{E}[\mathbb{E}[\cdots|\mathcal{F}_i]] = \mathbb{E}[f_i^2\mathbb{E}[(M_{t_{i+1}} - M_{t_i})^2|\mathcal{F}_t]], \tag{B.1}$$

$$\begin{aligned}\mathbb{E}[(M_t - M_s)^2|\mathcal{F}_s] &= \mathbb{E}[M_t^2|\mathcal{F}_s] - 2M_s\mathbb{E}[M_t|\mathcal{F}_s] + M_s^2 \\ &= \mathbb{E}[M_t^2 - M_s^2|\mathcal{F}_s] \qquad \text{(B.2)} \\ &= \mathbb{E}[\langle M\rangle_t - \langle M\rangle_s + L_t - L_s|\mathcal{F}_s].\end{aligned}$$

(B.1) and (B.2) finish the proof of 2. □

B.3 Levy's Characterization

Question: Why SDE chooses Brownian motion as randomness?

Theorem B.3.1
Let M_t be continuous Martingale and $\langle M\rangle_t = t$ for every t. Then, M_t is a Brownian motion.

Example. Let $\mathrm{sgn}(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}$, and $X_t = \int_0^t \mathrm{sgn}(B_t)dB_t$. Then, $\langle X\rangle_t = \int_0^t \mathrm{sgn}(B_t)^2 dt = \int_0^1 1dt = t$, and X_t is a Brownian motion.

Remark that this is a counterexample for many cases. For example, X_t and B_t are uncorrelated but they are not independent.

Theorem B.3.2
M_t is continuous Martingale defined in $(\Omega, \mathcal{F}, \mathbb{P})$, and $\langle M\rangle_t$ is absolutely continuous with respect to t almost surely. Then, there is a exntension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\exists$ Brownian motion B_t in $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ satisfying

$$M_t = \int_0^t X_s dB_s \ \ \forall t \text{ a.s. for some } X_s \in L^2_{ad}.$$

Remark. $dM_t = X_t dB_t$ in this sense. Thus, we don't have to concern Martingale, only Brownian motion.

Theorem B.3.3
M_t is a continuous Martingale, $\langle M\rangle_t \to \infty$ as $t \to \infty$, and $T(s) = \inf\{t : \langle M\rangle_t > s\}$ (generalized inverse). Then,

1. *$M_{T(s)} = B_s (\to$ Brownian motion).*
2. *$M_t = B_{\langle M\rangle_t}$.*

Example. $X_t = B_{t^3+2t}$ is a Marginal since $t^3 + 2t$ is increasing. $X_t = B_{e^t-1}$ is a Martingale since $e^t - 1$ is increasing. Moreover, every Martingale is such types.