Stochastic Differential Equations and Applications

Mentor: Dr. Mrinmay Biswas

Summer - 2023

Contents

1 Lecture-wise Summary

Lec No.	Date	Topic
1	May 15	Probability space and Random Variable
2	May 17	Random variable and its Distribution
3	May 19	Expectation of Random Variable
4	May 22	Independence and Conditional Expectation
5	May 24	Measure theory
6	May 26	Measure theory
7	May 29	Measure theory
8	May 31	Measure theory and stochastic process
9	Jun 5	Brownian Motion and Martingales
10	Jun 7	Ordinary Differential Equations
11	Jun 12	Wiener Integral
12	Jun 14	Itô integral
13	Jun 19	Metric Space and Stochastic Process
14	Jun 21	Stochastic Differential Equations
15	Jun 23	Stochastic Differential Equations
16	Jun 26	Generalized Itô Chain Rule
17	Jun 28	Examples

2 Measure Theory

2.1 Metric Space

Metric Space

Metric space (Y, d), $d\colon Y \times Y \to [0, \infty)$ satisfying

1. $d(x, y) = 0$ iff $x = y$
2. $d(y, x) = d(x, y)\ \forall\ x, y \in Y$
3. $d(x, y) \leq d(x, z) + d(z, y)\ \forall\ x, y, z \in Y$

Cauchy Sequence

A sequence $\{x_n\} \subseteq Y$ is said to be Cauchy if for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon\ \forall\ n, m \geq N$

$$d(x_n, x_m) \to 0 \text{ as } n, m \to \infty$$

Completeness

A metric space (Y, d) is called Complete if any Cauchy sequence converges in Y.

2.2 σ-algebra

Let Ω be a set. A σ-algebra on Ω is a collection $\mathcal{F}$ of subsets of Ω satisfying the following properties:

1. $\Omega \in \mathcal{F}$.
2. $A \subseteq \Omega$ and $A \in \mathcal{F}$, $\Rightarrow A^C \in \mathcal{F}$.
3. If $A_1, A_2, \cdots \in \mathcal{F}$, $\Rightarrow \bigcup_{i=1}^{\infty} A_k \in \mathcal{F}$.

Remark 1. Here are some following remarks:

1. We note that $\emptyset \in \mathcal{F}$.
2. If $A_1, A_2, \cdots \in \mathcal{F}$, $\Rightarrow \bigcap_{i=1}^{\infty} A_k \in \mathcal{F}$

Proposition 1. Let $\{\mathcal{F}_\alpha\}_{\alpha \in \wedge}$ be a collection of σ-algebras on Ω,then $\bigcap_{\alpha \in \wedge} \mathcal{F}_\alpha$ is a σ-algebra on Ω.

Proof. Let $\zeta = \bigcap_{\alpha \in \wedge} \mathcal{F}_\alpha$

(i)Since $\mathcal{F}_\alpha$ is a σ-algebra for each $\alpha \in \wedge$,
$\Rightarrow \Omega \in \mathcal{F}_\alpha,\ \forall\ \alpha \in \wedge,$
$\Rightarrow \Omega \in \bigcap_{\alpha\in\wedge} \mathcal{F}_\alpha = \zeta$
(ii) $A \in \zeta \Rightarrow A \in \bigcap_{\alpha\in\wedge} \mathcal{F}_\alpha \Rightarrow A \in \mathcal{F}_\alpha,\ \forall\ \alpha \in \wedge$
$\Rightarrow A^c \in \mathcal{F}_\alpha,\ \forall\ \alpha \in \wedge$ (because $\mathcal{F}_\alpha$ is a σ-algebra)
$\Rightarrow A^c \in \bigcap_{\alpha\in\wedge} \mathcal{F}_\alpha \Rightarrow A^c \in \zeta$
(iii) $A_i \in \zeta\, \forall\ \alpha \in \wedge$
$\Rightarrow A_i \in \bigcap_{\alpha\in\wedge} \mathcal{F}_\alpha,\ \forall\ \alpha \in \wedge, i \in \mathbb{N}$
$\Rightarrow \forall \alpha \in \wedge, A_i \in \mathcal{F}_\alpha\ \forall\ i \in \mathbb{N}$
$\Rightarrow \bigcup_{i\in\mathbb{N}} A_i \in \mathcal{F}_\alpha,\ \forall\ \alpha \in \wedge$
$\Rightarrow \bigcup_{i\in\mathbb{N}} A_i \in \bigcap_{\alpha\in\wedge} \mathcal{F}_\alpha$
$\Rightarrow \bigcup_{i\in\mathbb{N}} A_i \in \zeta$

Therefore $\{\mathcal{F}_\alpha\}_{\alpha\in\wedge}$ is a σ-algebra.

□

Proposition 2. $2^\Omega = \mathcal{P}(\Omega) :=$ power set of Ω = set of all subsets of Ω, then 2^Ω is a σ-algebra.

Proof. (i) $\Omega \in 2^\Omega$ (2^Ω contains all subsets of Ω)
(ii) $A \in 2^\Omega \Rightarrow A^c \subseteq \Omega \Rightarrow A^c \in 2^\Omega$
(iii) $A_i \in 2^\Omega \Rightarrow \bigcup_{i\in\mathbb{N}} A_i \in \Omega \Rightarrow \bigcup_{i\in\mathbb{N}} A_i \in 2^\Omega$

Therefore, 2^Ω is a σ-algebra.

□

Proposition 3. Let S be any collection of subsets of Ω, then there exists a smallest σ-algebra containing S (It is denoted by σ(S)).

Proof. We will be proving it by an example.

Example 1. When we toss two coins, the outcomes are $\{HH, HT.TH, TT\}$
$S = \{\Omega, \{HH\}, \emptyset\}$ which is not σ-algebra.
$\mathcal{F}_1 = \{\Omega, \{HH\}, \emptyset, \{HT, TH, TT\}\}$, $S \subset \mathcal{F}_1$ and also $S \subset 2^\Omega$
$\sigma(\mathrm{S}) = \mathcal{F}_1 \cap 2^\Omega = \mathcal{F}_1$ ■

$\square$

2.3 Measure on $(\Omega, \mathcal{F})$

Definition 2. $(\Omega, \mathcal{F})$ is a measurable space
Let measure $\mu\colon \mathcal{F} \to [0, \infty)$ *on* Ω is a function satisfying

1. $\mu(\emptyset) = 0$
2. For any $\{A_i\}_{i\in\mathbb{N}} \subseteq f$ with $A_j \cap A_k = \emptyset$ for j $\neq$ k, j,k $\in \mathbb{N}$

$$\mu(\bigcup_{i\in\mathbb{N}} A_i) = \Sigma_{i\in\mathbb{N}}\mu(A_i)$$

$\blacksquare$

- The triplet $(\Omega, \mathcal{F}, \mu)$ is called measure space.
- If $\mu(\Omega) < \infty$ then μ is called finite measure.
- If $\bigcup_{i\in\mathbb{N}} \Omega_i = \Omega, \Omega_i \in \mathcal{F}$ and $\mu(\Omega_i) < \infty$, then μ is called σ-finite measure.
- $\mu(\Omega) = 1$, we use $\mu = \mathbb{P}$ (Probability measure $\mathbb{P}\colon \mathcal{F} \to [0, 1]$)

Example 3. 1. $\Omega = \mathbb{N}, f = 2^{\mathbb{N}}, \mu =$ counting measure
$E \in \mathcal{F}, \mu(E) =$ number of elements in set E
$E_1 = \{1, 2, 3, \ldots, 10\} \Rightarrow\ \mu(E_1) = 10$
$E_2 = \{2, 4, 6, \ldots, 100\} \Rightarrow\ \mu(E_2) = 50$

2. $\Omega = \mathbb{R}, f = 2^{\mathbb{R}}, \mu =$ Dirac measure centred at $0 = \delta_0$
$E \in 2^{\mathbb{R}}$

$$\delta_0(E) = \begin{cases} 0 & 0 \notin E \\ 1 & 0 \in E \end{cases}$$

$E_1 = (1, 2) \Rightarrow\ \delta_0(E_1) = 0$
$E_2 = (-1, 1) \Rightarrow\ \delta_0(E_2) = 1$

3. $\Omega = \mathbb{R}, f = \mathcal{B}$
$S = \{(a, b) | a < b\},\ \mathcal{B} = \sigma(S) = \sigma$-algebra containing S, $\mu =$ Lebesgue measure
$\mu((a, b)) = b - a$, It is same for all $\mu([a, b]),\ \mu([a, b))$ and $\mu((a, b])$.

4. $\Omega = \{1, 2, \ldots, 6\}, \mathcal{F} = 2^{\Omega}$
   measure $= \mathbb{P}(\{i\}) = \frac{1}{6}, \ \forall \ i = 1, \ldots, 6$

$\blacksquare$

Theorem 1. Continuity of measure from below:
Let $A_i \in \mathcal{F}, \ i \in \mathbb{N}$ *and* $A_i \subseteq A_{i+1}$, then

$$\mu(\bigcup_{i \in \mathbb{N}} A_i) = Lim_{i \to \infty} \mu(A_i)$$

Theorem 2. Continuity of measure from above:
Let $A_i \in \mathcal{F}, \ i \in \mathbb{N}, \ \mu(\Omega) < \infty$ *and* $A_{i+1} \subseteq A_i$, then

$$\mu(\bigcap_{i \in \mathbb{N}} A_i) = Lim_{i \to \infty} \mu(A_i)$$

2.4 Measurable Functions

Definition 4. Let $(\Omega, \mathcal{F})$ and $(\Omega^1, \mathcal{F}^1)$ be a measurable space. A function $f \colon \Omega \to \Omega^1$ is said to be measurable if $f^{-1}(B) \in \mathcal{F}$ for any $B \in \mathcal{F}^1$. $\blacksquare$

- Let $(\Omega, \mathcal{F})$ be a measurable space. A function $f \colon \Omega \to \mathbb{R}$ is said to be measurable if f is $(\mathcal{F}, \mathcal{B})$ measurable.
  $B \in \mathcal{B} \Rightarrow f^{-1}(B) \in \mathcal{F}$, $\mathcal{B}$ is Borel σ-algebra on $\mathbb{R}$.

Equivalent criteria for measurability of real-valued function:

Let $f \colon \Omega \to \mathbb{R}$ is measurable and $\alpha \in \mathbb{R}$ if

$$f^{-1}((\alpha, \infty)) = \{x \in \mathbb{R} \mid f(x) > \alpha\} \in \mathcal{F}$$

$$f^{-1}([\alpha, \infty)) = \{x \in \mathbb{R} \mid f(x) \geq \alpha\} \in \mathcal{F}$$

$$f^{-1}((-\infty, \alpha)) = \{x \in \mathbb{R} \mid f(x) < \alpha\} \in \mathcal{F}$$

$$f^{-1}((-\infty, \alpha]) = \{x \in \mathbb{R} \mid f(x) \leq \alpha\} \in \mathcal{F}$$

Theorem 3. Let $(\Omega, \mathcal{F})$ be a measurable space.

1. $f, g \colon \Omega \to \mathbb{R}$ be measurable then $f + g, fg, cf$ and $f/g[g \neq 0]$ are also measurable.

2. f^n is measurable.

3. $|f|$ is measurable.

2.5 Measurable simple functions

Let $(\Omega, \mathcal{F})$ be a measurable space. A function $f\colon \Omega \to \mathbb{R}$ is said to be a simple function if the range of $f(\Omega)$ is a finite set.

- $f(\Omega) = \{a_1, a_2, \ldots, a_k\}$, where $k \in \mathbb{N}$.
- $f^{-1}(a_i) = E_i$ for $i = 1, 2, \ldots, k$.

$$f(x) = a_1\mathbf{1}_{E_1}(x) + a_2\mathbf{1}_{E_2}(x) + \ldots a_k\mathbf{1}_{E_k}(x)$$
$$f(x) = \sum_{i=1}^{k} a_i\mathbf{1}_{E_i}(x)$$

Let $x \in \Omega, \Rightarrow f(x) \in f(\Omega)$
$\Rightarrow f(x) = a_j \text{for some} j$
$\Rightarrow x \in f^{-1}(\{a_j\}) = E_j$
$\Rightarrow x \notin E_1, E_2 \ldots E_{j-1}, E_{j+1}, \ldots E_k$

Characteristic Function: Let $E \subseteq \Omega$ be any set then characteristic function of E is

$$\mathbf{1}_E(x) = \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

Note: $\delta_0\colon \mathcal{F} \to [0,1]$ and $\mathbf{1}_E(x)\colon \Omega \to [0,1]$ are different!

Example 5. Let

$$\mathbf{1}_{[1,2]}(x) = \begin{cases} 1, & \text{if } 1 \leq x \leq 2 \\ 0, & \text{if otherwise.} \end{cases}$$

$\star$ If $E \in \mathcal{F}$, then $\mathbf{1}_E$ is measurable. ■

Remark 2.

- Step functions are also simple functions. But simple functions may not be step functions.
- For Step functions E_i are intervals.

- A simple function f is measurable if all $E_i \; \forall \; i = 1, 2, \ldots, k$ are measurable.

Example 6.

$$f(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1 \\ 2, & \text{if } 1 \leq x \leq 2. \end{cases}$$

$$g(x) = \begin{cases} 1, & \text{if } 0 \leq x < 0.5, x = 1 \\ 2, & \text{if } 1 < x \leq 2 \\ 0, & \text{if otherwise.} \end{cases}$$

Here, $f(x)$ is both a step function and a simple function, but $g(x)$ is a simple function, not a step function. ■

2.6 Integration of Non-negative simple function

Let $f \colon \Omega \to [0, \infty)$ be a non-simple function, then f is given by $f(x) = \sum_{i=1}^{k} a_i \mathbf{1}_{E_i}(x)$.

$$\int_\Omega f(x)\, d\mu(x) = \sum_{i=1}^{k} a_i \mu(E_i)$$

$$\int_\Omega \mathbf{1}_E(x)\, d\mu = \mu(E)$$

for some $k \in \mathbb{N}$
$a_1, a_2, \ldots a_k \in [0, \infty)$ and
$E_i = f^{-1}(a_i)$

Example 7. Let $\Omega = \{HH, HT, TH, TT\}$ and $X \colon \Omega \to \mathbb{R}$
$\mathbb{P}(\{HH\}) = 1/4$
$\mathbb{P}(\{HT\}) = 1/4$
$\mathbb{P}(\{TH\}) = 1/4$
$\mathbb{P}(\{TT\}) = 1/4$

$$X(\omega) = \begin{cases} 2, & \text{if } \omega = HH \\ 1, & \text{if } \omega = HT \text{or} TH \\ 0, & \text{if } \omega = TT. \end{cases}$$

$\mathbb{E}(X) = \int_\Omega X(\omega)\, d\mathbb{P}(\omega)$

$= 2 \cdot \mathbb{P}(\{HH\}) + 1 \cdot \mathbb{P}(\{TH\}\{HT\}) + 0 \cdot \mathbb{P}(\{TT\})$
$= 2/4 + (1/4 + 1/4) + 0$
$= 1$ ■

Let $X\colon \Omega \Rightarrow [0, \infty)$ be a non-negative simple function then

$$\int_\Omega X(\omega)\, d\mathbb{P}(\omega) = \sum_{i=1}^{k} a_i \mathbb{P}(E_i)$$

$$= \sum_{i=1}^{k} a_i f_i$$

$$X(\omega) = \sum_{i=1}^{k} a_i \mathbf{1}_{E_i}(\omega)$$

$E_i = X^{-1}(a_i)$
$\mathbb{P}(E_i) = \mathbb{P}(X^{-1}(a_i) = \mathbb{P}(X = a_i) = f_i$
Now, let $\mathcal{S}$ = Set of simple measurable functions from $\Omega \to \mathbb{R}$
$\mathcal{S}_+$ = Set of non-negative simple function
$S \in \mathcal{S}_+ \Rightarrow s\colon \Omega \to [0, \infty)$ and $S(\Omega)$ is finite.

$$S(x) = \sum_{i=1}^{k} a_i \mathbf{1}_{E_i}(x)$$

$$\int_\Omega S(x)\, d\mu(x) = \sum_{i=1}^{k} a_i \mu(E_i)$$

Let $\mathcal{L}$ = Set of real valued measurable function.
$= f\colon \Omega \to \mathbb{R}$ is measurable
$\mathcal{L}_+$ = non-negative measurable function.
Let $f \in \mathcal{L}_+$

$$\int f(x)\, d\mu(x) = \sup\left\{\int S(x)\, d\mu(x) : 0 \le S(x) \le f(x), \text{for a.e} x \in \Omega, S(x) \in \mathcal{S}_+\right\}$$

Meaning of almost everywhere(a.e): Let (Ω,$\mathcal{F}$,μ) be a measurable space.
A property $\mathfrak{f}$ holds a.e means if $Q = \{x \in \Omega \mid \mathfrak{f}(x) \text{does not hold}\} \Rightarrow \mu(Q) = 0, Q \in \mathcal{F}$

Theorem 4. Let $f\colon \Omega \to [0, \infty]$ measurable on $f \in \mathcal{L}_+$.

Then $\exists$ a sequence $S_n \in \mathcal{S}_+$
$0 \leq S_n(x) \leq S_{n+1}(x) \leq f(x)$ such that $\lim_{n\to\infty} S_n(x) = f(x)$, a.e$x \in \Omega$.

Proof. Let $n \in \mathbb{N}$

$$E_i = f^{-1}([\frac{i-1}{2^n}, \frac{i}{2^n}]), i = 1, 2, \ldots n2^n$$

$$F_n = f^{-1}([n, \infty))$$

$$S_n(x) = \sum_{i=1}^{n2^n} (\frac{i-1}{2^n})\mathbf{1}_{E_i}(x) + n\mathbf{1}_{E_n}(x)$$

$0 \leq S_n(x) \leq S_{n+1}(x) \leq f(x)$ and $S_n(x) \leq f(x) \leq S_n(x) + 1/2^n$

□

Let $f \in \mathcal{L}$
Define $f_+(x) = \max f(x), 0$ and $f_-(x) = -\min f(x), 0$

$$f_+(x) = \begin{cases} f(x), & \text{if } f(x) \geq 0 \\ 0, & \text{if } f(x) < 0 \end{cases}$$

$$f_-(x) = \begin{cases} f(x), & \text{if } -f(x) \leq 0 \\ 0, & \text{if } f(x) > 0 \end{cases}$$

as we know, $\max(a, b) = \frac{a+b+|a-b|}{2}$ and $\max(a, b) = \frac{a+b+|a-b|}{2}$
$f_+(x) = \frac{f(x)+|f(x)|}{2}$
$f_-(x) = \frac{|f(x)|-f(x)}{2}$

$f_+(x) + f_-(x) = |f(x)|$
$f_+(x) + f_-(x) = f(x)$
$f(x) = f_+(x) - f_-(x) \text{where} f_+(x) \leq 0, f_-(x) \leq 0$
So, we have seen that any function can be represented by the difference of two non-negative functions. Also, we can define $\int_\Omega f_+(x)\, d\mu(x)$ and $\int_\Omega f_-(x)\, d\mu(x)$

$$\int_\Omega f(x)\, d\mu(x) = \int_\Omega f_+(x)\, d\mu(x) - \int_\Omega f_-(x)\, d\mu(x)$$

provided one of them is finite.

$$\int_\Omega |f(x)|\, d\mu(x) = \int_\Omega f_+(x)\, d\mu(x) - \int_\Omega f_-(x)\, d\mu(x)$$

Notation: $\mathcal{L}^1 = \{f \in \mathcal{L} | \int_\Omega |f(x)| \, d\mu(x) < \infty$
$\mathcal{L}^p = \{f \in \mathcal{L} | \int_\Omega |f(x)|^p \, d\mu(x) < \infty$

Example 8. $f_n(x) = nx^n, x \in (0, 1)$

$$\lim_{n \to \infty} f_n(x) = 0 = f(x)$$

$$\int_0^1 f(x) \, dx = 0$$

$$\int_0^1 f_n(x) = n \int_0^1 x^n \, dx = \frac{1}{1 + \frac{1}{n}} \to 1$$

let $\frac{1}{x} = 1 + b, b > 0$

$$\Rightarrow \frac{1}{x^n} = (1 + b)^n$$

$$= 1 + nb + \frac{n(n-1)b^2}{2} + \dots b^n$$

$$\frac{1}{x} > \frac{n(n-1)b^2}{2}$$

$$0 < nx^n < \frac{2}{b^2(n-1)}$$

$lim_{n \to \infty} nx^n = 0$ (by sandwitch theorm)

■

2.7 Monotone Convergence Theorem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measurable space. Let $f_n \in \mathcal{L}_+$ sequence such that $o \leq f_n(x) \leq f_{n+1}(x)$ and $\lim_{n \to \infty} f_n(x) = f(x)$ a.e $x \in \Omega$. Then,

$$\lim_{n \to \infty} \int_\Omega f_n(x) \, d\mu(x) = \int_\omega f(x) \, d\mu(x)$$

Remark 3. Monotone convergence also holds for decreasing sequences provided $|int_\Omega f_1(x) \, d\mu(x) < \infty$.

2.8 Fatous Lemma

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measurable space and $f_n \in \mathcal{L}_+$. Then,

$$\int_\Omega \lim_{n\to\infty} \inf f_n(x)\, d\mu(x) \leq lim_{n\to\infty} \int_\Omega f_n(x)\, d\mu(x)$$

2.9 Dominated Convergence Theorem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measurable function. Let $f\colon f_n \in \mathcal{L}$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ a.s $x \in \Omega$ and $\exists g \in \mathcal{L}_+^1$ such that $|f_n(x)| \leq g(x)$ provided $\int_\Omega g(x)\, d\mu(x) < \infty$a.e $x \in \Omega \forall n \geq 1$. Then,

$$\lim_{n\to\infty} \int_\Omega f_n(x)\, d\mu(x) = \int_\omega f(x)\, d\mu(x)$$

3 Probability space

In probability theory, a probability space or a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ is a mathematical construct that provides a formal model of a random process or "experiment". For example, one can define a probability space that models throwing a die or tossing a coin. A probability space consists of three elements:

1. A sample space, Ω, which is the non-empty finite or countably infinite set of all possible outcomes.
2. The σ-algebra $\mathcal{F}$ is a collection of all the events we would like to consider. An event space is a set of events, $\mathcal{F}$, an event being a set of outcomes in the sample space.
3. A probability function or measure defined on $\mathcal{F}$, $\mathbb{P}$, which assigns each event in the event space a probability, which is a number between 0 and 1.

3.1 Probability function

Definition 1. Given a collection $\mathcal{F}$ of subsets of Ω, we refer to functions $f : \mathcal{F} \to \mathbb{R}$ as a Set Function.

Definition 2. A real valued set function $\mathbb{P}$ defined on $\mathcal{F}$ is said to be a probability function/measure if

1. $\mathbb{P}(\Omega) = 1$.
2. $\mathbb{P}(A) \geq 0 \quad \forall \quad A \in F$.
3. <u>Countable Additivity</u>: given a sequence of events A_1, A_2, $\ldots$ which are pairwise disjoint $(A_i \cap A_j \; \forall \; i \neq j)$ then

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$$

.

Some properties of probability function/measure:

1. $\mathbb{P}(\emptyset) = 0$.

   Proof.

$$A_1 = \Omega$$

$$A_2 = \emptyset, n \geq 2$$

$$\bigcup_{n=1}^{\infty} A_n = \Omega$$

by using countable additivity,

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) + \dots$$

$$\Rightarrow \mathbb{P}(\emptyset) = 0$$

□

2. Finite Additivity: Let $A_1, A_2, \dots, A_n$ be a finite collection of pairwise disjoint events. Then

$$\mathbb{P}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i)$$

Proof. given is

$$A_1, A_2, \dots A_n$$

$$A_m = \emptyset, m > n$$

□

3. Complementarity: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Proof.

$$A \cup A^c = \Omega$$

$$\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega)$$

$$\mathbb{P}(A) + \mathbb{P}(A^c) = 1$$

hence,

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$$

□

4. $\forall A \in \mathcal{F}, \mathbb{P}(A) \leq 1$

Proof.

$$\mathbb{P}(A^c) \geq 0$$

then

$$\mathbb{P}(A) \leq \mathbb{P}(A) + \mathbb{P}(A^c)$$

$$\mathbb{P}(A) \leq 1$$

□

5. <u>Monotonicity</u>: If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

   Proof.

$$B = A \cup (B \backslash A)$$

$$\Rightarrow \mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \backslash A)$$

$$\Rightarrow \mathbb{P}(B) \geq \mathbb{P}(A)$$

□

6. <u>Inclusion-Exclusion</u>: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

   Proof.

$$A = (A \cap B) \cup (A \backslash B)$$

$$B = (A \cap B) \cup (B \backslash A)$$

$$\mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A \backslash B) + \mathbb{P}(A \cap B) + \mathbb{P}(B \backslash A)$$

$$= \mathbb{P}(A \cap B) + \mathbb{P}(A \cup B)$$

□

Example 9. Let $\Omega = \{H, T\}$, $U = \{\emptyset, H, T, \Omega\}$ and $\mathbb{P} : U \to [0, 1]$
Let $\mathbb{P}(\{H\}) = \frac{1}{3}$ and $\mathbb{P}(\{T\}) = \frac{2}{3}$
Here, we can find $\mathbb{P}(A)$ for any set A such that $A \subseteq \Omega$. Now, let $\Omega = \{HH, HT, TH, TT\}$ and same $\mathbb{P} : U \to [0, 1]$
$\mathbb{P}(\{HH\}) = \frac{1}{9}$,$\mathbb{P}(\{TH\}) = \frac{2}{9}$,$\mathbb{P}(\{HT\}) = \frac{2}{9}$ and $\mathbb{P}(\{TT\}) = \frac{4}{9}$. ■

Example 10. Let $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ be a finite set, and suppose we are given numbers $0 \leq p_j \leq 1$ for $j = 1, \ldots, N$ satisfying $\sum_{j=1}^{N} p_j = 1$.

We take U to comprise all subsets of Ω.

For each set $A = \{\omega_{j_1}, \omega_{j_2}, \ldots, \omega_{j_m}\} \in U$, with $1 \leq j_1 < j_2 < \ldots < j_m \leq N$,

we define $\mathbb{P}(A) := \mathbb{P}(\omega_{j1}) + \mathbb{P}(\omega_{j2}) + \cdots + \mathbb{P}(\omega_{jm})$
$= p_{j_1} + p_{j_2} + \ldots + p_{j_m}$. ■

4 Random Variable

A random variable X is a measurable function $X : \Omega \to E$ from a sample space Ω as a set of possible outcomes to a measurable space E. The technical axiomatic definition requires the sample space Ω to be a sample space of a probability triple $(\Omega,\mathcal{F},\mathbb{P})$.

Definition 3. The Borel subsets of $\mathbb{R}^{\mathbf{n}}$, denoted $\mathcal{B}$, comprise the smallest σ-algebra of subsets of $\mathbb{R}^{\mathbf{n}}$ containing all open sets. We may henceforth informally just think of $\mathcal{B}$ as containing all the "nice, well-behaved" subsets of $\mathbb{R}^{\mathbf{n}}$.

Definition 4. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space. A mapping $X : \Omega \to \mathbb{R}^{\mathbf{n}}$ is called an n-dimensional random variable if for each $B \in \mathcal{B}$, we have $X^{-1}(B) \in \mathcal{F}$. We equivalently say that X is $\mathcal{U}$-measurable.

4.1 Distribution Function

Let $F(x)$ be the distribution function defined as follows:

$$F(x) = \mathbb{P}(X \leq x),$$

where X is a random variable and $\mathbb{P}(\cdot)$ denotes the probability.

The distribution function $F(x)$ satisfies the following properties:

1. Non-decreasing: $F(x_1) \leq F(x_2)$ if $x_1 \leq x_2$.
2. Right-continuous: $\lim_{h\to 0^+} F(x + h) = F(x)$.
3. Limits at infinities: $\lim_{x\to -\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

4.2 Discrete random variable and its support

Let $X : (\Omega,\mathcal{F},\mathbb{P}) \to (\mathbb{R},\mathcal{B},\mathbb{P} \circ X^{-1})$ be a Random variable, then there exist a finite or countably infinite set $S \subset R$ such that $\mathbb{P} \circ X^{-1}(S) = 1$, $\sum_{x\in S} \mathbb{P} \circ X^{-1}(x) = 1$, $\sum_{x\in S} P(X = x) = 1$, then we say X is a discrete random variable and S is known as its support

4.3 Probability mass function

X is a discrete random variable, then the function $f_x : \mathbb{R} \rightarrow [0,1]$ defined as

$$f_x(x) = \begin{cases} \mathbb{P}(X = x) & x \in S \\ 0 & x \in S^c \end{cases}$$

, then f_x is the probability mass function

For example: When X follows Poisson distribution then probability mass function is given as

$$f_x(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2.....$$

Example 11. Let E be an event of tossing a fair coin twice independently, X be a random variable that counts the number of heads, then what is its probability mass function f_x and Distribution function F_x?

Sample space of X be Ω
$\Omega = \{$HH,HT,TH,TT$\}$
let's define X
X(HH) $= 2$ $\quad$ X(HT)=X(TH) $= 1$ $\quad$ X(TT) $= 0$ and its support $S = \{0, 1, 2\}$
$f_x = \mathbb{P}(X = x)$, f_x is pmf of X

$$f_x = \begin{cases} \frac{1}{4} & x = 0 \\ \frac{1}{2} & x = 1 \\ \frac{1}{4} & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

F_x is the distribution function of X

$$F_x = \begin{cases} 0 & x \leq 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{1}{4} + \frac{1}{2} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

■

4.4 Absolutely continuous random variable

Let X : $(\Omega,\mathcal{F},\mathbb{P}) \rightarrow (\mathbb{R},\mathcal{B},P \circ X^{-1})$ be a Random variable with distribution function F_x if there exist a function $f_x\colon R{\rightarrow}[0,\infty)$ which is integrable then $F_x = \int_{-\infty}^{x} f(t)\,dt$, then X is absolutely continuous random variable and f is probability density function of X.

Properties of pdf:
1. $f\colon R{\rightarrow}[0,\infty)$
2. $\int_{-\infty}^{\infty} f(t)\,dt = 1$

Then set $S = \{x \in R | \forall h > 0, \mathbb{P} \circ X^{-1}((x-h, x+h)) > 0$ is called its support.
For example: X follows Normal distribution which has pdf as f

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in R$$

Another example: X follows uniform distribution which has pdf as f

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b, \\ 0 & \text{otherwise,} \end{cases}$$

5 Expectation of Random Variable

Let X be a discrete/continuous random variable with pmf/pdf f_x
(i) If X is discrete random variable and $\sum_{x \in S_x} |x| \cdot f_x(x) < \infty$ then we say expectation of X exists and equals to

$$\mathrm{E}[X] = \sum_{x \in S_x} x \cdot f_x(x)$$

(ii) If X is continuous random variable and $\int_{-\infty}^{\infty} |x| \cdot f_x(t)\,dt < \infty$ then we say expectation of X exists and equals to

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x \cdot f_x(t)\,dt$$

If the "absolute convergence condition" does not hold, we say expectation does not exist

Example 12. Let X be a discrete random variable with probability mass function f_x

$$f(x) = \begin{cases} \frac{1}{4} & \text{if } x \in \{\pm 1, \pm 2\} \\ 0 & \text{otherwise,} \end{cases}$$

find expectation of X

Solution:

$$\sum_{x \in S_x} |x| \cdot f_x(x) = |-1| \times \frac{1}{4} + |-2| \times \frac{1}{4} + 1 \times \frac{1}{4} + 2 \times \frac{1}{4} = \frac{3}{2} < \infty$$

$$\mathrm{E}[X] = \sum_{x \in S_x} x \cdot f_x(x) = \frac{1}{4} \times (-1 - 2 + 1 + 2) = 0$$

■

Example 13. Let X be a discrete random variable with probability mass function f_x

$$f(x) = \begin{cases} \frac{3}{\pi^2} \frac{1}{x^2} & \text{if } x \in \{\pm 1, \pm 2, \dots\} \\ 0 & \text{otherwise,} \end{cases}$$

find expectation of X

Solution:

$$\sum_{x \in S_x} |x| \cdot f_x(x) = \frac{3}{\pi^2} \times 2 \sum_{x=1}^{\infty} \frac{|x|}{x^2} = \frac{6}{\pi^2} \sum_{x=1}^{\infty} \frac{1}{x} = \infty$$

As $\sum_{x \in S_x} |x| \cdot f_x(x) = \infty$ so we can say expectation of X does not exist. ■

1. Suppose X is a discrete/continuous random variable with pmf/pdf f_x such that $E[X]$ exists, then

$$\mathrm{E}[X] = \int_0^{\infty} \mathbb{P}(X > x)\, dx - \int_{-\infty}^{0} \mathbb{P}(X < x)\, dx$$

Proof. when X is continuous

LHS =

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x \cdot f_x(x)\, dx$$

$$= \int_0^{\infty} x \cdot f_x(x)\, dx + \int_{-\infty}^{0} x \cdot f_x(x)\, dx$$

$$= \int_{x=0}^{\infty} \int_{y=0}^{x} f_x(x)\,dy\,dx - \int_{x=-\infty}^{0} \int_{y=x}^{0} f_x(x)\,dy\,dx$$

Then

$$\mathrm{E}[X] = \int_{0}^{\infty} \mathbb{P}(X > x)\,dx - \int_{-\infty}^{0} \mathbb{P}(X < x)\,dx$$

□

2. If $\mathbb{P}(X \geq 0) = 1$ we usually say X is a non-negative random variable. Then $\mathbb{P}(X < 0) = 0 \; \forall \; x < 0$
   If $\mathrm{E}[X]$ exists, $\mathrm{E}[X] = \int_0^\infty \mathbb{P}(X > x)\,dx$

3. Suppose X is a discrete random variable with $\mathbb{P}(X \in \{0, 1, 2, \dots\}) = 1$.
   If

$$\begin{aligned} \mathrm{E}[X] &= \int_{0}^{\infty} \mathbb{P}(X > x)\,dx \\ &= \sum_{n=0}^{\infty} \int_{n}^{n+1} \mathbb{P}(X > x)\,dx \\ &= \sum_{n=0}^{\infty} \int_{n}^{n+1} \mathbb{P}(X \geq n+1)\,dx \\ &= \sum_{n=0}^{\infty} \mathbb{P}(X \geq n+1) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) \end{aligned}$$

4. <u>Function of random variable</u>:
   Let X be any random variable then $Y = h(X)$ be a function of X
   (i) If X is discrete random variable then Y is also discrete RV and if $\sum_{x \in S_x} |h(x)| \cdot f_x(x) < \infty$ then $\mathrm{E}(Y)$ exists and equals $\sum_{x \in S_X} h(x) \cdot f_x(x)$
   (ii) If X is continuous random variable then Y is also continuous RV and if $\int_{-\infty}^{\infty} |h(x)| \cdot f_x(x) < \infty$ then $\mathrm{E}[Y]$ exists and equals $\int_{-\infty}^{\infty} h(x) \cdot f_x(x)\,dx$

5. $h(X) = aX + b, x \;\forall\; R$
   then $\mathrm{E}(h(X)) = a\mathrm{E}(X) + b$

   Proof. Let for discrete case, $E(X)$ exists
   then $\sum_{x \in S_x} |aX + b| \cdot f_x(x) \leq |a| \sum_{x \in S_x} x \cdot f_x(x) + |b| \sum_{x \in S_x} f_x(x)$

$= |a| \sum_{x \in S_x} f_x(x) < \infty$
$\mathrm{E}[h(X)] = \sum_{x \in S_x} aX + b \cdot f_x(x)$
$= a \sum_{x \in S_x} x \cdot f_x(x) + b \sum_{x \in S_x} f_x(x)$

$$\mathrm{E}[h(X)] = a\mathrm{E}[X] + b$$

If X is continuous, replace $\sum_{x \in S_x}$ by $\int_{-\infty}^{\infty} \ldots \, dx$. □

6. Let X be a discrete/continuous with pmf/pdf f_x
(i) $h_i \colon \mathbb{R} \to \mathbb{R}$ be any real-valued function $i = 1, 2, \ldots, n$ and $a_1, a_2, \ldots, a_n \in \mathbb{R}$
Then
$$\mathrm{E}[\sum_{i=1}^{n} a_i \cdot h_i(X)] = \sum_{i=1}^{n} a_i \mathrm{E}[h_i(X)]$$
provided the expectation exists.
(ii) if $h_1, h_2 \colon \mathbb{R} \to \mathbb{R}$ such that $h_1(x) \leq h_2(x) \; \forall x \in S_x$
then
$$\mathrm{E}[h_1(X)] \leq \mathrm{E}[h_2(X)]$$

7. we call $\mathrm{E}[X - \mu_1^1]^r$ as rth central moment of X.

5.1 Variance

Definition 14. The variance is the mean squared difference between each data point and the center of the distribution measured by the mean. ■

when $r = 2$, then second central moment of X is said to variance.

$$\mu_2 = var(X) = \mathrm{E}(X - \mu_1^1)^2$$

$$\begin{aligned}
\mu_2 &= \mathrm{E}(X - \mu_1^1)^2 \\
&= \mathrm{E}[X - \mathrm{E}(X)]^2 \\
&= \mathrm{E}[X^2 + (\mathrm{E}(X))^2 - 2 \cdot X \cdot \mathrm{E}(X)] \\
&= \mathrm{E}[X^2] + (\mathrm{E}(X))^2 - 2 \cdot \mathrm{E}[X] \cdot \mathrm{E}[X] \\
&= \mathrm{E}[X^2] - (\mathrm{E}(X))^2
\end{aligned}$$

Hence,

$$var(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = \mu_2^1 - (\mu_1^1)^2$$

1. Since $var(X) \geq 0$, then $(\mathrm{E}(X))^2 \leq \mathrm{E}(X^2)$
2. If $var(X) = 0$, then there exists $c \in \mathbb{R}$ such that $\mathbb{P}(X = c) = 1$
3. $var(aX + b) = a^2 var(X) \quad a, b \in R$
4. $Y = \frac{X - \mathrm{E}(X)}{\sqrt{var(X)}}$ then
   $\mathrm{E}[Y] = 0$ and $var(Y) = 1$

Example 15. Let $X \sim \text{Poisson}(\lambda)$ then find $\mathrm{E}(X)$ and $var(X)$
Solution:
We know that $\sum_{x \in S_x} |x| \cdot f_x(x) < \infty$ so expectation exists.

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{x=0}^{\infty} x \cdot f_x(x) \\
&= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{(x)!} \\
&= \lambda e^{-\lambda} e^{\lambda} \\
\mathrm{E}(X) = \lambda
\end{aligned}$$

$$\begin{aligned}
Var(X) &= \sum_{k=0}^{\infty} (k - \lambda)^2 \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} \\
&= \sum_{k=0}^{\infty} (k^2 - 2k\lambda + \lambda^2) \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} \\
&= \sum_{k=0}^{\infty} k^2 \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} - 2\lambda \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} + \lambda^2 \sum_{k=0}^{\infty} \frac{e^{-\lambda} \cdot \lambda^k}{k!}
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{E}(X^2) - 2\lambda \mathrm{E}(X) + \lambda^2 \\
&= \lambda + \lambda^2 - 2\lambda^2 + \lambda^2 \\
&= \lambda
\end{aligned}
$$

hence, $var(X) = \lambda$ ■

Example 16. Let $X \sim \text{Normal}(\mu, \sigma^2)$ then find $\mathrm{E}(X)$ and $var(X)$
Solution:
We know that $\int_{-\infty}^{\infty} |x| \cdot f_x(t)\, dt < \infty$ so expectation exists.

$$
\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} x \cdot f_x(t)\, dt \\
&= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \int_{-\infty}^{\infty} (x-\mu) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx + \int_{-\infty}^{\infty} \mu \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} [\int_{-\infty}^{\infty} t \cdot e^{\frac{-t^2}{2\sigma^2}}\, dt + \mu \cdot \sqrt{2\pi\sigma^2}] \\
&= 0 + \mu
\end{aligned}
$$

Hence

$$\mathrm{E}(X) = \mu$$

$$
\begin{aligned}
\mathrm{Var}(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx \\
&= \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) f(x)\, dx \\
&= \int_{-\infty}^{\infty} x^2 f(x)\, dx - 2\mu \int_{-\infty}^{\infty} x f(x)\, dx + \mu^2 \int_{-\infty}^{\infty} f(x)\, dx \\
&= \mathrm{E}(X^2) - 2\mu \mathrm{E}(X) + \mu^2 \\
&= \sigma^2
\end{aligned}
$$

■

5.2 Chebysheu's inequality

Definition 17. Let X be a random variable with a finite mean denoted as μ and a finite non-zero variance denoted as σ^2 for any real number, $c > 0$. ■

$$\mathbb{P}(|X - \mathrm{E}X| \geq c) \leq \frac{\mathrm{E}\,|X - \mathrm{E}X|^2}{c^2}$$

Proof. Before proofing Chebysheu's inequality, we will state and prove Markov's inequality

Markov's inequality:

Definition 18. Let X be a random variable with $\mathbb{P}(X \geq 0) = 1$ $[\mathbb{P}(X < 0) = 0]$, then there exist $c > 0$

$$\mathbb{P}(X \geq c) \leq \frac{\mathrm{E}X}{c}$$

■

Proof. In continuous case:

$$\begin{aligned}
\mathrm{E}(x) &= \int_{-\infty}^{\infty} x \cdot f_x(t)\,dt \\
&= \int_{0}^{\infty} x \cdot f_x(t)\,dt \\
&\geq \int_{c}^{\infty} x \cdot f_x(t)\,dt \\
&\geq c\int_{c}^{\infty} f_x(t)\,dt = x\mathbb{P}(X \geq c)
\end{aligned}$$

Hence

$$\mathbb{P}(X \geq c) \leq \frac{\mathrm{E}X}{c}$$

Corollary 1. 1. Let $h \colon \mathbb{R} \to [0, \infty)$, then any Y be any non-negative random variable of function of X, $Y = h(X)$, then

$$\mathbb{P}(Y \geq c) = \mathbb{P}(h(X) \geq c) \leq \frac{\mathrm{E}X}{c}$$

2. If h is strictly increasing function, then

$$\mathbb{P}(X \geq c) = \mathbb{P}(h(X) \geq h(c)) \leq \frac{\mathrm{E}h(X)}{h(c)}$$

3. If $h(X) = |X|$,then $\mathbb{P}(|X| \geq c) \leq \frac{\mathrm{E}|X|}{c}$
4. If $h(X) = |X|^r \;\; \forall r > 0$,then $\mathbb{P}(|X|^r \geq c^r) \leq \frac{\mathrm{E}|X|^r}{c^r}$

□

For proving Chebysheu's inequality, we will be considering Markov's inequality and corollary-2

Let $Y = |X - \mathrm{E}X|\,,\; h(Y) = Y^2$

$$\mathbb{P}(Y \geq c) = \mathbb{P}(h(Y) \geq h(c)) = \mathbb{P}(Y^2 \geq c^2) \leq \frac{\mathrm{E}Y^2}{c^2}$$

Hence

$$\mathbb{P}(|X - \mathrm{E}X| \geq c) \leq \frac{\mathrm{E}\,|X - \mathrm{E}X|^2}{c^2}$$

□

5.3 Modes of Convergence

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{X_n\}$ be a sequence of random variable $[X_n \colon \Omega \to \mathbb{R}]$

1. converges almost everywhere to random variable $X \colon \Omega \to \mathbb{R},$ if $\mathbb{P}(\omega \colon X_n(\omega) \to X(\omega), n \to \infty) = 1$ or $\mathbb{P}(\omega \mid Lim_{n\to\infty} X_n(\omega) = X(\omega)) = 1$
2. converges in probability to X, if every $\epsilon > 0$ $\mathbb{P}(\omega \mid |X_n(\omega) - X(\omega)| > \epsilon) \to 0$ as $n \to \infty \mid Lim_{n\to\infty}\mathbb{P}(\omega \mid |X_n(\omega) - X(\omega)| > \epsilon) = 0$
3. converges in pth mean to X if $\mathrm{E}\,|X_n - X|^p \to 0$ as $n \to \infty / Lim_{n\to\infty}\mathrm{E}[|X_n - X|^p] = 0[1 \leq p < \infty]$
4. converges in distribution to X if $F_{X_n}(x) \to F_X(x)$ for all point of continuity of $F_X(x)$

Theorem 5. Let X and $\{X_n\}$ be sequence of random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ then

1. $X_n \xrightarrow{\text{a.e.}} X \Rightarrow X_n \xrightarrow{\text{in probability}} X$
2. $X_n \xrightarrow{\text{pth mean}} X \Rightarrow X_n \xrightarrow{\text{in probability}} X$
3. $X_n \xrightarrow{\text{in probability}} X \Rightarrow X_n \xrightarrow{\text{in distribution}} X$
4. $X_n \xrightarrow{\text{in probability}} X \Rightarrow$ there exists subsequence $\{X_{n_k}\}$ such that $X_{n_k} \xrightarrow{\text{a.e.}} X$

Proof. $X_n \xrightarrow{\text{pth mean}} X \Rightarrow X_n \xrightarrow{\text{in probability}} X$
Using Chebysheu's inequality, $0 \leq \mathbb{P}(|X_n - X| > \epsilon) \leq \frac{\mathrm{E}[|X_n - X|^p]}{\epsilon^p}$
As $n \to \infty$ then $\mathrm{E}[|X_n - X|^p] \to 0$ and $\mathbb{P}(|X_n - X| > \epsilon) \to 0$ for fixed ϵ □

5.4 Independent Random variable

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\omega \Rightarrow$ sample space and $\mathcal{F} \Rightarrow$ event space

1. Independent events: Two events $A_1, A_2 \in \mathcal{F}$ are independent, if

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2)$$

   Any finite number of events $A_1, A_2, A_3, \ldots, A_k \in \mathcal{F}$ are independent

$$\mathbb{P}(\bigcap_{i=1}^{k} A_i) = \prod_{i=1}^{k} \mathbb{P}(A_i)$$

2. Independent σ-algebra: Two σ-algebras $\mathcal{F}_\infty, \mathcal{F}_\in \subseteq \mathcal{F}$ are independent, if

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2)$$

   for any $A_1 \in \mathcal{F}_\infty, A_2 \in \mathcal{F}_\in$
   $\sigma(X)$:
   As $\mathcal{B}$ = smallest σ of open subsets of R.
   Let $X\colon (\Omega, \mathcal{F}) \to (\mathcal{R}, \mathcal{B})$ be a random variable.
   σ(X) = σ-algebra generated by X.

   Example 19. $\Omega = \{HH, HT, TH, TT\}$
   $\mathcal{F} = \mathbb{P}(\Omega) = 2^4 = 16$
   $X\colon \Omega \to \mathbb{R}$
   $S_X = \{\Omega, \emptyset, \{HH\}, \{HT\}, \{TH\}\}$ ■

3. Independent Random Variable: Two random variable X_1, X_2 are independent if $\sigma(X_1)$ and $\sigma(X_2)$ are independent, that is

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2)$$

if $A_1 \in \sigma(X_1), A_2 \in \sigma(X_2)$

$$\Rightarrow F_X, Y(x, y) = F_X(x) \cdot F_Y(y)$$

Example 20. Let X and Y be random variables that output the number of heads of different independent coins when tossed, and find $\mathbb{P}((X < 2) \cap (Y > 1))$
Solution:
As given X and Y are independent discrete random variable

$$\mathbb{P}(X \cap Y) = \mathbb{P}(X) \cdot \mathbb{P}(Y)$$

$$\begin{aligned}\mathbb{P}((X < 2) \cap (Y > 1)) &= \mathbb{P}(X < 2) \cdot \mathbb{P}(Y > 1) \\ &= [P_X(0) + P_X(1)]P_Y(2) \\ &= (\frac{1}{4} + \frac{1}{2})\frac{1}{4} \\ &= \frac{3}{16}\end{aligned}$$

■

5.5 Conditional Expectation

Let X and Y be any random variable, then the conditional expectation is given by $\mathrm{E}[X|Y]$
$\mathrm{E}[X|Y]$ is average value of X when the value of Y is known.

$$\mathrm{E}[X \mid Y = y] = \begin{cases} \sum_{x \in X} x \cdot f_{X|Y}(X = x \mid y) & X, Y \textit{ are discrete random variable} \\ \int_{-\infty}^{\infty} x \cdot f_{X|Y}(X = x \mid y) & X, Y \textit{ are continuous random variable} \end{cases}$$

Properties:

1. $\mathrm{E}[\mathrm{E}[X \mid \mathcal{G}]] = \mathrm{E}[X]$

2. If X is $(\mathcal{G},\mathcal{B})$ measurable, then $\mathrm{E}[X \mid \mathcal{G}] = X$
3. X and $\mathcal{G}$ are independent random variable, $\mathrm{E}[X \mid \mathcal{G}] = \mathrm{E}[X]$
4. If $Z\colon \Omega \to \mathbb{R}$ be $(\mathcal{G},\mathcal{B})$ measurable and $\mathrm{E}[|XZ|] < \infty$, then $\mathrm{E}[XZ \mid \mathcal{G}] = Z \cdot \mathrm{E}[X \mid \mathcal{G}]$
5. Let $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, $\mathrm{E}[\mathrm{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathrm{E}[X \mid \mathcal{H}]$
6. $\mathrm{E}[a \mid X] = a$
7. $\mathrm{E}[aX + bZ \mid Y] = a\mathrm{E}[X \mid Y] + b\mathrm{E}[Z \mid Y]$
8. If $X \geq 0$, then $\mathrm{E}[X \mid Y] \geq 0$
9. $X, Z\colon \Omega \to \mathbb{R}$, $\mathrm{E}\,|X| < \infty$, $\mathrm{E}\,|Z| < \infty$. If $X \leq Z \Rightarrow \mathrm{E}[X \mid \mathcal{G}] \leq \mathrm{E}[Z \mid \mathcal{G}]$
10. $|\mathrm{E}[X \mid \mathcal{G}]| \leq \mathrm{E}[|X| \mid \mathcal{G}]$
11. <u>Conditional Fatous Lemma:</u>
    Let $X_n \in \mathcal{L}^1_+(\Omega), \int_\Omega X_n dP, \infty$ then $\mathrm{E}[Lim_{n\to\infty} X_n \mid \mathcal{G}] \leq Lim_{n\to\infty}\mathrm{E}[X_n \mid \mathcal{G}]$
12. <u>Conditional Monotone Convergence Theorem:</u>
    Let $X, X_n \in \mathcal{L}^1_+$, $0 \leq X_n \leq X_{n+1} \leq X$, $Lim_{n\to\infty} X_n = X$, then

Example 21. Let X and Y be discrete random variable

X/Y	Y=0	Y=1
X=0	$\frac{1}{5}$	$\frac{2}{5}$
X=1	$\frac{2}{5}$	0

and find $E[X \mid Y]$

Solution:
For finding $\mathrm{E}[X \mid Y]$,
$\mathrm{E}[X \mid Y] = \sum_{x \in X} x \cdot f_{X|Y}(X = x \mid y)$

$$f_X(0) = \frac{1}{5} + \frac{2}{5} = \frac{3}{5} \; f_X(1) = \frac{2}{5} + 0 = \frac{2}{5} \; f_Y(0) = \frac{1}{5} + \frac{2}{5} = \frac{3}{5} \; f_Y(1) = \frac{2}{5} + 0 = \frac{2}{5}$$

$$f_{X|Y=0} = \frac{f_X, Y(x,0)}{f_Y(0)} \begin{cases} \frac{\frac{1}{5}}{\frac{3}{5}} = \frac{1}{3} & , x = 0 \\ \frac{\frac{2}{5}}{\frac{3}{5}} = \frac{2}{3} & , x = 1 \end{cases}$$

$$f_{X|Y=1} = \frac{f_X, Y(x,1)}{f_Y(1)} \begin{cases} \frac{\frac{2}{5}}{\frac{2}{5}} = 1 & , x = 0 \\ \frac{0}{\frac{2}{5}} = 0 & , x = 1 \end{cases}$$

$$\mathrm{E}[X \mid Y = 0] = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$

$$\mathrm{E}[X \mid Y = 1] = 0 \cdot 1 + 1 \cdot 0 = 0$$

$$\mathrm{E}[X \mid Y = y] = \begin{cases} \frac{2}{3} & , y = 0 \\ 0 & , y = 1 \end{cases}$$

■

6 Brownian Motion

6.1 Stochastic Process

- A collection $\{\mathbb{X}(t)|t \geq 0\}$ of random variables is called a Stochastic process.
- For each point $\omega \in \Omega$, the mapping $t \mapsto \mathbb{X}(t, \omega)$ is the corresponding sample path.

6.2 Brownian Motion

- It is a continuous time-space Stochastic process.
- It is a random variable or collection of random variables that represent a continuous and random movement of a particle or system over time.
- It is denoted by $(\mathbb{W}_t)_{t \geq 0}$ and is a real-valued function.

Properties of Brownian motion:

1. $(\mathbb{W}_0) = 0$
2. Independent increments:
   The random variable $(\mathbb{W}_v) - (\mathbb{W}_u)$ and $(\mathbb{W}_t) - (\mathbb{W}_s)$ are independent whenever $u \leq v \leq s \leq t$. (u, v) and (s, t) are disjoint random variable.
3. Normal increments:
   $(\mathbb{W}_{t+s}) - (\mathbb{W}_s) \sim N(0, t)$ or $(\mathbb{W}_t) - (\mathbb{W}_0) \sim N(0, t)$ where $t > 0, x \in \mathbb{R}$

$$f(x; 0, t) = \frac{1}{(2\pi t)^{1/2}} e^{\frac{-(x)^2}{2t}}$$

4. Continuous Sample space:
   With probability 1, the function $t \mapsto \mathbb{W}(t, \omega)$ is continuous almost sure and it doesn't have any jumps or discontinues.
5. Markov property:
   This property of Brownian motion states that the future behavior of the process depends only on its current state and is independent of past history.
6. Self-similarity:
   It is a property of Brownian motion where the statistical properties of the process are similar at different time scales.

- $\mathrm{E}[\mathbb{W}_t] = 0$

- $\mathrm{E}[\mathbb{W}_t^2] = t$
- $\mathrm{E}[\mathbb{W}_s(\mathbb{W}_t - \mathbb{W}_s)] = 0$ where $0 \leq s \leq t$
- $\mathrm{E}[\mathbb{W}_t \mathbb{W}_s] = min\{t, s\}$

6.3 Martingale

- Filtration:
  Increasing family of sub σ-algebras $\{f_t\}$ of f
  $f_0 \subseteq f_s \subseteq f_t \subseteq f$ if $0 \leq s \leq t$
- Right-Continuous:
  A filtration is said to be right-continuous if

$$\bigcap_{t>s} f_t = f_s$$

Definition 22. A stochastic process $\{X_t\}_{t\geq 0}$ is called a Martingale with respect to a filtration if it satisfies

- $\mathrm{E}[|X_t|] < \infty \ \forall \ t \geq 0$
- X_t is f_t measurable(adaptedness)
- $\mathrm{E}[X_t | f_s] = X_s, \ 0 \leq s \leq t$

■

Remark 4. The expectation of Martingale is constant with respect to the family of the index.

$$\mathrm{E}[\mathrm{E}[X_t | f_s]] = \mathrm{E}[X_s]$$

$$\mathrm{E}[X_t] = \mathrm{E}[X_s]$$

7 Ordinary Differential Equation (ODE)

Example:

$$\tfrac{d}{dt}y(t) = a \cdot y(t)\ ,\ y(0) = y_0\ ,\ y\colon [0,\infty) \to R$$

$$\begin{aligned} \frac{dy}{dt} = ay \to (1) \Rightarrow \int \frac{dy}{y} &= \int a \cdot dt + c_1 \\ ln\,|y| &= a \cdot t + c_1 \\ y(t) &= e^{at} \cdot c_2\ ,\ c_2 = e^{c_1} \\ \text{At } t = 0\ ,\ y(0) &= y_0 \\ y(t) &= y_0 \cdot e^{at}\ ,\ t > 0 \to (2) \end{aligned}$$

7.1 General form of first-order ODE

$$\frac{dy}{dt} = f(t, y(t))$$

Examples of $f(t, y(t))$:

1. $f(t, y(t)) = t + e^{y(t)}$
2. $f(t, y(t)) = \sin(y) + \cos(y)$
3. $f(t, y(t)) = ln\,|y(t)|$
4. $f(t, y(t)) = a \cdot y(t)$

Remark 5. : If $f(t, y(t)) = f(y(t))$ then $ff(t, y(t))$ is autonomous, otherwise it is non-autonomous.

Definition 23. Solution of (2) is a function $y\colon [0,\infty) \to R$ such that y is a continuous and differentiable in $(0, T)$, satisfies (1) for all $t \in (0, T)$ and $y(0) = y_0$

$$\frac{dy}{dt} = f(t, y(t))\ ,\ y(0) = y_0\ ,\ t \in (0, T)\ ,\ T > 0$$

■

7.2 Existence and Uniqueness of solution

Theorem 6. Let $f\colon [0,T] \mathsf{x} [y_0 - R, y_0 + R] \to \mathbb{R}$ be continuous and f satisfies Lipschitz condition [i.e $\exists\, \mathsf{L}_R > 0$ such that $|f(y_1,t) - f(y_2,t)| \geq \mathsf{L}_R \cdot (y_1 - y_2)$, $\forall t \in [0,T]$, $y_1, y_2 \in [y_0 - R, y_0 + R]$]. Then there exists $T \geq \delta > 0$ such that there has a unique solution.

Proof.

$$\int_0^t \frac{dy}{f(t,y)} = \int_0^t dt \Rightarrow \int_0^t \frac{dy(s)}{ds} ds = \int_0^t f(s,y(s))ds$$
$$y(t) - y(0) = \int_0^t f(s,y(s))ds$$
$$y(t) = y_0 + \int_0^t f(s,y(s))ds$$

$$\frac{dy(t)}{dt} = f(t,y(t)) \text{ is similar } y(t) = y_0 + \int_0^t f(s,y(s))ds$$

$y(t) = y_0 + \int_0^t f(s,y(s))ds$ is the integral form.
$\frac{dy(t)}{dt} = f(t,y(t))$ is the differentiable form. $\square$

We will be constructing sequences of integral form function

$$y_1(t) = y_0 + \int_0^t f(s,y_0)ds$$

$$y_2(t) = y_0 + \int_0^t f(s,y_1(s))ds$$

. . . .

$$y_n(t) = y_0 + \int_0^t f(s,y_{n-1}(s))ds$$

If $y_n(t) \to y(t)$ as $n \to \infty$ is said **Picard's sequence**

Example:

$$\tfrac{d}{dt}y(t) = a \cdot y(t)\ ,\ y(0) = y_0\ ,\ y\colon [0,\infty) \to R$$

Integral form of $\frac{dy(t)}{dt} = f(t, y(t))$ is $y(t) = y_0 + \int_0^t f(s, y(s))ds$

$$y_1(t) = y_0 + a\int_0^t y_0 ds = y_0 + ay_0 t = (1 + at)y_0$$
$$y_2(t) = y_0 + a\int_0^t y_1 ds = y_0 + a\int_0^t (1 + at)y_0 ds = y_0 + ay_0 t + a^2t^2/2 = (1 + at + \frac{(at)^2}{2})y_0$$

$\cdots\cdots \quad \cdots$

$$y_n(t) = y_0[1 + at + \frac{(at)^2}{2!} + \cdots + \frac{(at)^n}{n!}]$$

As $n \to \infty$, $y_n(t) \to y(t) = y_0 e^{at}$
$f(t, y(t))$ for (1)

1. continuous
2. It satisfies Lipschitz condition
   $|f(y_1, t) - f(y_2, t)| = |ay_1 - ay_2| = a \cdot |y_1 - y_2|$

- $\frac{dy(t)}{dt} = ay(t)$ is the rate of the change in $y(t)$ is a times the present state.
- $a = (r + "e")$ $\rightarrow e$= error in measurements follows probabilistic theory
- $\frac{dy(t)}{dt} = r \cdot y(t) + "e" \cdot y(t)$
- In Brownian motion , $"e" = \frac{dW_t}{dt}$

$$\frac{dy}{dt} = r \cdot y(t) + \frac{dW_t}{dt} \cdot y(t)$$
$$dy = r \cdot y(t) + y(t) \cdot dW_t$$

General stochastic differentiable process:

$$dy(t) = b(t, y(t))dt + \sigma(t, y(t))dW_t \ \ , \ y(0) = y_0$$

Equivalent Integral form:

$$y(t) = y_0 + \int_0^t b(s, y(s))ds + \int_0^t \sigma(s, y(s))dW_s$$

The random variable $W_t(\omega)$ can be defined as

- $t \mapsto W_t(\omega)$ is the function from $[0, T]$ to $\mathbb{R}$ when ω is fixed.

- $\omega \mapsto W_t(\omega)$ is the function from Ω to $\mathbb{R}$ when t is fixed.

8 Stochastic Differential Equation (SDE)

8.1 Wiener integral

$\int_0^t \sigma(s, y(s))dW_s$, the integral depends on s, ω.
Before we compute this integral, we will start with a simple case

$$\text{Simple Case: } \int_a^b f(s)dW_s$$

$$f(s) = \Sigma_{i=1}^k a_i \mathbf{1}_{[t_{i-1},t_i]}$$
$$\int_a^b f(s)ds = \Sigma_{i=1}^k a_i \mu([t_{i-1}, t_i]) = \Sigma_{i=1}^k a_i(t_i - t_{i-1}) \rightarrow \text{Riemann Integration}$$

g is a function of Bounded variation(B.V)
g: $[a, b] \rightarrow \mathbb{R}$ and Let $\mathbf{P} = \{a = t_0 < t_1 < t_2 < \cdots < t_k = b\}$ be partition of $[a, b]$
$V(g, \mathbf{P}) = \Sigma_{i=1}^k |g(t_i) - g(t_{i-1})|$. If $\sup_{\mathbf{P}} V(g, \mathbf{P}) < \infty$, then g is function of Bounded Variation(B.V)

$$\int_a^b f(s)dg(s) = \Sigma_{i=1}^k a_i(g(t_i) - g(t_{i-1}))$$

- Particular Case: If g is increasing, then g is function of bounded variation (B.V)

$$\begin{aligned} V(g, \mathbf{P}) &= \Sigma_{i=1}^k (g(t_i) - g(t_{i-1})) \\ &= g(t_n) - g(t_0) = g(b) - g(a) < \infty \end{aligned}$$

Riemann Stieltjes:

$$\int_a^b f(s)dg(s) = Lim_{n\rightarrow\infty, \|\Delta_n\|\rightarrow 0} \Sigma_{i=1}^n f(x_i)[g(t_i) - g(t_{i-1})]$$

where $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ are partitions of $[a, b]$ and $x_i \in [t_{i-1}, t_i]$(called as tag), $\|\Delta_n\| = max_{1\leq i\leq n}\{(t_i - t_{i-1})\}$

- As $s \mapsto W_s(.)$ is not a function of Bounded variation, we cannot use **Riemann Stieltjes**.

- As for different tags (x_i) have different values in Integral.

$$\int_a^b f(s)dW_s = Lim_{n\to\infty,\|\Delta_n\|\to 0}\Sigma_{i=1}^n f(x_i)[W_{t_i} - W_{t_{i-1}}]$$

$$x_i = t_{i-1}, \text{then integral is called Ito integral.} \tag{1}$$

$$x_i = \frac{(t_i + t_{i-1})}{2}, \text{then integral is called Strotonovich integral.} \tag{2}$$

- Ito integral doesn't hold the Fundamental theorem of calculus and holds Markov property.
- Strotonovich integral holds the Fundamental theorem of calculus.
- $\int_a^b f(s)dW_s = Lim_{n\to\infty,\|\Delta_n\|\to 0}\Sigma_{i=1}^n f(t_{i-1})[W_{t_i} - W_{t_{i-1}}]$ $\to$ Itô
- $\int_a^b f(s) \circ dW_s = Lim_{n\to\infty,\|\Delta_n\|\to 0}\Sigma_{i=1}^n f(\frac{(t_i+t_{i-1})}{2})[W_{t_i} - W_{t_{i-1}}]$ $\to$ Strotonovich integral.

Observations:

1. $E[\int_a^b f(s)dW_s] = \Sigma_{i=1}^k a_i[E[W_{t_i}] - E[W_{t_{i-1}}]] = 0$
2. $Var(\int_a^b f(s)dW_s) = \Sigma_{i=1}^k a_i^2 Var(W_{t_i} - W_{t_{i-1}}) = \Sigma_{i=1}^k a_i^2(t_i - t_{i-1})$
3. $E[(\int_a^b f(s)dW_s)^2] = \int_a^b f(s)^2 ds$

If $f\colon [a,b] \to \mathbb{R}$ be a step function(simple function) , $f(s) = \Sigma_{i=1}^k a_i \mathbf{1}_{[t_{i-1},t_i]}$, then Ito integral of f $[\int_a^b f(s)dW_s]$ is a **Normal Random variable** with mean 0 and variance $\int_a^b f(s)^2 ds$.

- $\mathcal{L}^p(\Omega, f, \mathbb{P}) = \{f\colon \Omega \to \mathbb{R}/ \int_\Omega |f|^p \, d\mu(x) < \infty\}$
- $\mathrm{L}^p(\Omega, f, \mathbb{P}) = \mathrm{L}^p = \{f\colon \Omega \to \mathbb{R}/ \int_\Omega |f|^p \, d\mu(x) < \infty\}$ and $f = g$ in L^p if $f(x) = g(x)$ a.e. $x \in \Omega$
- $\mathrm{L}^2(\Omega) = \mathrm{L}^2(\Omega, f, \mathbb{P}) = \{f\colon \Omega \to \mathbb{R}/ \int_\Omega |f|^2 \, d\mu(x) < \infty\}$
- $\mathrm{L}^2(a,b) = \mathrm{L}^2([a,b], \mathcal{B}, \mu) = \{f\colon \Omega \to \mathbb{R}/ \int_a^b |f|^2 \, d\mu(x) < \infty\}$ and $\mu =$ Lebesgue measure.

8.2 Hölder's inequality

If $f \in \mathrm{L}^p, g \in \mathrm{L}^q$ then $f \cdot g \in \mathrm{L}^1$ if $\frac{1}{p} + \frac{1}{q} = 1$

$$\int_\Omega |f(x) \cdot g(x)| \, d\mu(x) \leq [\int_\Omega |f(x)|^p \, d\mu(x)]^{\frac{1}{p}} \cdot [\int_\Omega |g(x)|^q \, d\mu(x)]^{\frac{1}{q}}$$

1. If $p = q = 2$

$$\int_\Omega |f(x) \cdot g(x)| \, d\mu(x) \leq [\int_\Omega |f(x)|^2 \, d\mu(x)]^{\frac{1}{2}} \cdot [\int_\Omega |g(x)|^2 \, d\mu(x)]^{\frac{1}{2}}$$

2. If $g(x) = 1$ and $p = q = 2$

$$\int_\Omega |f(x) \cdot g(x)| \, d\mu(x) \leq [\int_\Omega |f(x)|^2 \, d\mu(x)]^{\frac{1}{2}} \cdot [\mu(\Omega)]^{\frac{1}{2}}$$

3. If $g(x) = 1, f = X, \mu = \mathbb{P}$ and $p = q = 2$

$$\int_\Omega |X(\omega)| \, d\mathbb{P}(\omega) \leq [\int_\Omega |X(\omega)|^2 \, d\mathbb{P}(\omega)]^{\frac{1}{2}}$$

$$E\,|X| \leq [E\,|X|^2]^{\frac{1}{2}}$$

- $\mathrm{I}(f)$ defined for f(simple or step function) $[f\colon [a,b] \to \mathbb{R}]$
- Itô isometery $= E[(\mathrm{I}(f))^2] = \int_a^b |f(x)|^2 \, dx < \infty$ and $f \in \mathrm{L}^2([a,b])$

<u>Procedure:</u>

To define $\mathrm{I}(f) = \int_a^b f(s) dW_s$ for $f \in \mathrm{L}^2([a,b])$

Let $f_n \in \mathrm{L}^2([a,b])$ be sequences of step functions such that $Lim_{n\to\infty} \int_a^b |f_n - f|^2 \, dx \to 0$ $[f_n \xrightarrow{\mathrm{L}^2} f]$

$$f_n(s) = \Sigma_{i=1}^{k_n} a_i^n \mathbf{1}_{[t_{i-1}^n, t_i^n]}$$

$$\mathrm{I}(f_n) = \int_a^b f_n(s) dW_s = \Sigma_{i=1}^{k_n} a_i^n [\mathbb{W}_{t_i^n} - \mathbb{W}_{t_{i-1}^n}] \to \mathrm{I}(f) = \int_a^b f(s) dW_s = \Sigma_{i=1}^{k} a_i [\mathbb{W}_{t_i} - \mathbb{W}_{t_{i-1}}]$$

$$\mathrm{I}(f_n) \to \mathrm{I}(f) \text{ as } n \to \infty$$

8.3 Cauchy Sequence Convergence

- $\mathrm{I}(f) = \int_a^b f(t) dW_t$, $f(t) = \Sigma_{i=1}^k a_i \mathbf{1}_{[t_{i-1}, t_i]}(t)$
    1. $\mathrm{I}(f) \sim N(0, \sigma^2)$, $\sigma^2 = \int_a^b [f(t)]^2 dt$

2. $\mathrm{I}(af + bg) = a\mathrm{I}(f) + b\mathrm{I}(g),\ a, b \in \mathbb{R}, f, g\colon [a, b] \to \mathbb{R}$

- Let $f \in \mathrm{L}^2(a, b)$, there exists sequence of step functions $f_n \in \mathrm{L}^2(a, b)$ such that $f_n \to f$ in $\mathrm{L}^2(a, b)$
$$\int_a^b |f_n(t) - f(t)|^2 \, dt \to 0$$
$$\mathrm{I}(f_n) \to \mathrm{I}(f) = lim_{n\to\infty}\mathrm{I}(f_n) \text{ in } \mathrm{L}^2(\Omega)$$
$$\mathrm{I}\colon \mathrm{L}^2(a, b) \to \mathrm{L}^2(\Omega) \Rightarrow \text{square integrable random variable}$$
$$f \mapsto \mathrm{I}(f)\colon \Omega \to \mathbb{R}$$

- We take $Y = \mathrm{L}^2(\Omega), d_2(f, g) = [\int_\Omega |f(\omega) - g(\omega)|^2 \, d\mathbb{P}(\omega)]^{\frac{1}{2}}$
In general $Y = \mathrm{L}^p(\Omega), d_p(f, g) = [\int_\Omega |f(\omega) - g(\omega)|^p \, d\mathbb{P}(\omega)]^{\frac{1}{p}},\ 1 \le p < \infty$
$f_n \to f$ in $\mathrm{L}^2(a, b), \mathrm{I}\colon \mathrm{L}^2(a, b) \to \mathrm{L}^2(\Omega)$
$$\mathrm{I}(f_n) - \text{Cauchy in } \mathrm{L}^2(\Omega)$$

- $(\mathrm{L}^2(\Omega), d_2)$ is complete metric space
In general $(\mathrm{L}^p(\Omega), d_p)$ is complete metric space

$$\begin{aligned}
d_2(\mathrm{I}(f_n), \mathrm{I}(f_m)) &= (\int_\Omega |\mathrm{I}(f_n - f_m)|^2 \, d\mathbb{P}(\omega))^{\frac{1}{2}} \\
&= (\mathrm{E}\, |\mathrm{I}(f_n - f_m)|^2)^{\frac{1}{2}} \\
&= [\int_a^b |f_n(t) - f_m(t)|^2 \, dt]^{\frac{1}{2}} \\
&= [\int_a^b |f_n(t) - f(t) + f(t) - f_m(t)|^2 \, dt]^{\frac{1}{2}} \\
&\le [\int_a^b |f_n(t) - f(t)|^2 \, dt]^{\frac{1}{2}} + [\int_a^b |f(t) - f_m(t)|^2 \, dt]^{\frac{1}{2}}
\end{aligned}$$

$d_2(\mathrm{I}(f_n), \mathrm{I}(f_m)) \to 0$ as $n, m \to \infty$

Therefore $\mathrm{I}(f_n)$ is a Cauchy sequences in $\mathrm{L}^2(\Omega)$, so it converges (as $\mathrm{L}^2(\Omega)$ is complete)

Definition 24. $\mathrm{I}(f) = Lim_{n\to\infty}\mathrm{I}(f_n)$ in $\mathrm{L}^2(\Omega)$, there exists sub-sequence $\{\mathrm{I}(f_{n_k})\}$ such that $\mathrm{I}(f)(\omega) = Lim_{k\to\infty}\mathrm{I}(f_{n_k})(\omega)$ almost surely in $\mathbb{P}$. ■

Property:
$\mathrm{I}(f) \sim N(0, \sigma^2),\ \sigma^2 = \int_a^b |f(t)|^2\, dt$

To define $\mathrm{I}(f)(\omega) = \int_a^b f(t,\omega) dW_t(\omega)$
Fix a Brownian motion, $\{W_t\}_{a \leq t \leq b}$ and a filtration $\{\mathcal{F}_t\}_{a \leq t \leq b}$ satisfying

- For each $t,\ W_t$ is $\mathcal{F}_t$ measurable
- For any $a \leq s \leq t \leq b$, the random variable $W_t - W_s$ is independent of $\mathcal{F}_s$

Let $\mathrm{L}^2([a,b] \times \Omega)$ set of all stochastic processes, $f\colon [a,b] \times \Omega \to \mathbb{R}$ satisfying

- $f(t, \boldsymbol{.})$ is $\mathcal{F}_t$ adapted(measurable)
- $\mathrm{E}[\int_a^b |f(t,\omega)|^2\, dt] < \infty$

Step-1:-

f is a step process in $\mathrm{L}^2([a,b] \times \Omega)$

$$f(t,\omega) = \Sigma_{i=1}^k a_i(\omega) \mathbf{1}_{[t_{i-1}, t_i]}(t)$$

$$\mathrm{I}(f)(\omega) = \int_a^b f(t,\omega) dW_t(\omega) = \Sigma_{i=1}^k a_i(\omega)[W_{t_i} - W_{t_{i-1}}]$$

Property

1. $\mathrm{E}[\mathrm{I}(f)] = 0$
2. $\mathrm{Var}(\mathrm{I}(f)) = \mathrm{E}[\mathrm{I}(f)]^2 = \mathrm{E}[\int_a^b f^2(t,\omega) dt] = \int_\Omega [\int_a^b |f(t,\omega)|^2\, d\mathbb{P}(\omega)]$

Step-2:-

Let $f \in \mathrm{L}^2_{ad}([a,b] \times \Omega)$, we can choose a sequence of step function $f_n \in \mathrm{L}^2_{ad}([a,b] \times \Omega)$ such that $\mathrm{E}[\int_a^b |f_n(t,\omega) - f(t,\omega)|^2\, dt] \to 0$
As above $\mathrm{I}\colon \mathrm{L}^2_{ad}([a,b] \times \Omega) \to \mathrm{L}^2(\Omega),\ f \mapsto \mathrm{I}(f)$
$\{\mathrm{I}(f_n)\}$ is a Cauchy sequence in $\mathrm{L}^2(\Omega)$

Property

1. $\mathrm{E}[\mathrm{I}(f)] = \mathrm{E}[\int_a^b f(t) dW_t] = 0$
2. $\mathrm{E}[|\mathrm{I}(f)|^2] = \mathrm{E}[\left|\int_a^b f(t,\omega) dW_t(\omega)\right|^2] = \mathrm{E}[\int_a^b |f(t)|^2\, dt]$

- $\int_0^T \sigma(t, y(t,\omega)) dW_t(\omega)$ is a random variable with
  1. Mean $= \mathrm{E}[\int_0^T \sigma(t, y(t,\omega)) dW_t(\omega)] = 0$

2. Variance $= \mathrm{E}[\left|\int_0^T \sigma(t, y(t,\omega))dW_t\right|^2] = \mathrm{E}[\int_0^T |\sigma(t, y(t,\omega))|^2\, dt]$

Theorem 7. For $f \in \mathrm{L}^2(a,b)$, Let us define $X(t,\omega) = \int_a^t f(s)dW_s(\omega)$, then the stochastic process $\{X(t)\}_{a\leq t\leq b}$ is a continuous in t and Martingale with respect to filtration $\mathcal{F}_t = \sigma\{W_s \mid a \leq s \leq t\}$

Proof. 1. $\mathrm{E}\,|X(t)| < \infty$

$$\begin{aligned}
\mathrm{E}\,|X(t)| &\leq [\mathrm{E}\,|X(t)|^2]^{\frac{1}{2}} \\
&= (\mathrm{E}[\left|\int_a^t f(s)dW_s\right|^2])^{\frac{1}{2}} \\
&= (\int_a^t |f(s)|^2)^{\frac{1}{2}} \\
&\leq (\int_a^b |f(s)|^2)^{\frac{1}{2}} < \infty
\end{aligned}$$

2. $X(t)$ is $\mathcal{F}_t$ adapted (measurable)
   Case-1:- f is step function, $f(t) = \Sigma_{i=1}^k a_i \mathbf{1}_{[t_{i-1},t_i]}(t)$
   $X(t,\omega) = \Sigma_{i=1}^k a_i(W_{t_i\wedge t} - W_{t_{i-1}\wedge t})$

$$X(t,\omega) \text{is a } \mathcal{F}_t \text{ measurable}$$

   Case-2:- $f \in \mathrm{L}^2(a,b),\ f_n \to f$ in $\mathrm{L}^2(a,b)$
   $X_n(t) = \int_a^b f_n(s)dW_s,\ X_n(t)$ is $\mathcal{F}_t$ measurable.
   $X_n(t) = \int_a^b f_n(s)dW_s \xrightarrow{n\to\infty} \int_a^b f(s)dW_s = X(t)$ in $\mathrm{L}^2(\Omega)$
   $X_n(t)$ are $\mathcal{F}_t$ measurable, $X(t)$ is $\mathcal{F}_t$ measurable.

3. $\mathrm{E}[X(t) \mid \mathcal{F}_s] = X(s), a \leq s \leq t$

$$\begin{aligned}
X(t) &= \int_a^t f(r)dW_r \\
&= \int_a^s f(r)dW_r + \int_s^t f(r)dW_r \\
&= X(s) + \int_s^t f(r)dW_r \\
\mathrm{E}[X(t) \mid \mathcal{F}_s] &= \mathrm{E}[X(s) + \int_s^t f(r)dW_r \mid \mathcal{F}_s] \\
&= \mathrm{E}[X(s) \mid \mathcal{F}_s] + \mathrm{E}[\int_s^t f(r)dW_r \mid \mathcal{F}_s]
\end{aligned}$$

$$= X(s) + \mathrm{E}[\int_s^t f(r)dW_r \mid \mathcal{F}_s]$$

To prove $\mathrm{E}[\int_s^t f(r)dW_r \mid \mathcal{F}_s] = 0$

<u>Case-1</u>:- f is step(simple function)

$\int_s^t f(r)dW_r = \Sigma_{i=1}^k a_i(W_{t_i} - W_{t_{i-1}}), s = t_0 < t_1 < t_2 \cdots < t_n = t$ and

$f(r) = \Sigma_{i=1}^k \mathbf{1}_{[t_{i-1},t_i)}(r)$

$\mathrm{E}[\int_s^t f(r)dW_r \mid \mathcal{F}_s] = \Sigma_{i=1}^k a_i \mathrm{E}[W_{t_i} - W_{t_{i-1}} \mid \mathcal{F}_s] = \Sigma_{i=1}^k a_i \mathrm{E}[W_{t_i} - W_{t_{i-1}}] = 0$

<u>Case-2</u>:- $f \in \mathrm{L}^2(a,b)$, there exists sequences of step function $f_n \in \mathrm{L}^2(a,b)$ such that $f_n \to f$ in $\mathrm{L}^2(a,b)$

$$\begin{aligned}
\mathrm{E}[\int_s^t f_n(r)dW_r - \int_s^t f(r)dW_r \mid \mathcal{F}_s](\omega) &\to 0 \\
(\mathrm{E}[\int_s^t f_n(r)dW_r - \int_s^t f(r)dW_r \mid \mathcal{F}_s])^2 &\leq \mathrm{E}[(\int_s^t f_n(r)dW_r - \int_s^t f(r)dW_r)^2 \mid \mathcal{F}_s] \\
\mathrm{E}[(\mathrm{E}[\int_s^t f_n(r)dW_r - \int_s^t f(r)dW_r \mid \mathcal{F}_s])^2] &\leq \mathrm{E}[\mathrm{E}[(\int_s^t f_n(r)dW_r - \int_s^t f(r)dW_r)^2 \mid \mathcal{F}_s]] \\
&\leq \mathrm{E}[(\int_s^t f_n(r)dW_r - \int_s^t f(r)dW_r)^2] \\
&= \int_s^t |f_n(r) - f(r)|^2 \, dr \to 0 \text{ as } n \to \infty \\
\mathrm{E}[\int_s^t f_n(r)dW_r \mid \mathcal{F}_s] &\to \mathrm{E}[\int_s^t f(r)dW_r \mid \mathcal{F}_s] \text{ in } \mathrm{L}^2(\Omega)
\end{aligned}$$

Upto a subsequence

$$\mathrm{E}[\int_s^t f_{n_k}(r)dW_r \mid \mathcal{F}_s](\omega) \to \mathrm{E}[\int_s^t f(r)dW_r \mid \mathcal{F}_s](\omega) \text{ a.s. } \mathbb{P}$$

$\therefore X(t,\omega) = \int_a^t f(s)dW_s(\omega)$ is Martingale.

$\square$

Theorem 8. For $f \in \mathrm{L}^2_{ad}([a,b] \times \Omega)$, let us define $X(t,\omega) = \int -a^t f(s,\omega)dW_s(\omega), a \leq t \leq b$, then the stochastic process $\{X(t)\}_{a \leq t \leq b}$ is continuous in t and Martingale with respect to filtration given $\mathcal{F}_t$.

Remark 6.

$$F(t) = \int_a^t f(s)ds$$

If f is Riemann Integrable, F is Lipschitz continuous
If f is continuous, F is differentiable $[F^{'}(t) = f(t)]$
$t \mapsto X(t,\omega)$ is continuous only

$$dX(t) = f(t,\omega) \cdot dW_t$$

For $f \in \mathrm{L}^2_{ad}([a,b] \times \Omega)$, $\int_a^b f(t,\omega)dW_t(\omega)$
$X(t,\omega) = \int_a^t f(s,\omega)dW_s(\omega)$ is martingale and continuous.

- $g, f \in \mathrm{L}^2(a,b)$ then $\mathrm{I}(f) = \int_a^b f(t)dW_t$ satisfies
    1. $\int_a^b (c_1 f(t) + c_2 g(t))dW_t = c_1 \int_a^b f(t)dW_t + c_2 \int_a^b g(t)dW_t$
    2. $\mathrm{E}[\int_a^b f(t)dW_t] = 0$
    3. $\mathrm{E}[(\int_a^b f(t)dW_t)^2] = \int_a^b |f(t)|^2 \, dt = \sigma_f^2$
    4. define $X(t,\omega) = \int_a^t f(s)dW_s(\omega)$ is a continuous martingale.
- $f, g \in \mathrm{L}^2_{ad}([a,b] \times \Omega)$
    1. $\int_a^b (c_1 f(t,\omega) + c_2 g(t,\omega))dW_t = c_1 \int_a^b f(t,\omega)dW_t + c_2 \int_a^b g(t,\omega)dW_t$
    2. $\mathrm{E}[\int_a^b f(t,\omega)dW_t] = 0$
    3. $\mathrm{E}[(\int_a^b f(t,\centerdot)dW_t)^2] = \int_a^b |f(t,\centerdot)|^2 \, dt$
    4. define $X(t,\omega) = \int_a^t f(s,\omega)dW_s(\omega)$ is a continuous martingale.
- $\mathcal{L}_{ad}(\Omega, \mathrm{L}^2([a,b]))$ = set of all stochastic processes $f\colon [a,b] \times \Omega$ satisfying
    1. $f(t)$ is $\mathcal{F}_t$ measurable
    2. $\int_a^b |f(t)|^2 \, dt < \infty$ a.s. $\mathbb{P}$

Goal to define $\mathrm{I}(f)(\omega) = \int_a^b f(t,\omega)dW_t(\omega)$
Procedure: $\mathrm{L}^2_{ad}([a,b] \times \Omega) \subset \mathcal{L}_{ad}(\Omega, \mathrm{L}^2([a,b]))$

$$f \in \mathrm{L}^2_{ad}([a,b] \times \Omega) \to \mathrm{E}[\int_a^b |f(t,\centerdot)|^2 \, dt] < \infty$$

$$\Rightarrow \int_a^b |f(t,\omega)|^2 \, d\mathbb{P}(\omega) \text{ a.s. } \mathbb{P} \text{ by Fubini theorem}$$

$$\Rightarrow f \in \mathcal{L}_{ad}(\Omega, \mathrm{L}^2([a,b]))$$

Lemma 1. Let $f \in \mathcal{L}_{ad}(\Omega, \mathrm{L}^2([a,b]))$, then there exists a sequences $\{f_n(t)\} \subset \mathrm{L}^2_{ad}([a,b] \times \Omega)$ of step processes such that
$Lim_{n\to\infty} \int_a^b |f_n(t) - f(t)|^2 \, dt = 0$ in probability
i.e For any $\epsilon > 0, lim_{n\to\infty}\mathbb{P}(\omega \mid \int_a^b |f_n(t,\omega) - f(t,\omega)|^2 \, dt > \epsilon) = 0$
For each n, we define $\mathrm{I}(f_n) = \int_a^b f_n(t,\omega)dW_t$ and $(f_n(t,\omega) = \sum_{i=1}^k \xi_i^n(\omega)\mathbf{1}_{[t_{i-1}^n, t_i^n]}(t))$

Lemma 2. Let $\mathrm{I}(f_n)$ is defined above, then for every $\epsilon > 0$,
$\mathbb{P}(\omega \mid |\mathrm{I}(f_n) - \mathrm{I}(f_m)| > \epsilon) \to 0$ as $n, m \to \infty$.Hence there exists a random variable $\mathrm{I}(f)\colon \Omega \to \mathbb{R}$ such that for every $\epsilon > 0, lim_{n\to\infty}\mathbb{P}(\omega \mid |\mathrm{I}(f_n) - \mathrm{I}(f_m)| > \epsilon) = 0$

$$\int_a^b f(t,\omega)dW_t = Lim_{n\to\infty}\mathrm{I}(f_n)$$

$$= Lim_{n\to\infty} \int_a^b f_n(t,\omega)dW_t = \mathrm{I}(f) \text{ in probability}$$

Theorem 9. Let $f \in \mathcal{L}_{ad}(\Omega, \mathrm{L}^2([a,b]))$, define $X(t,\omega) = \int_a^t f(s,\omega)dW_s(\omega)$. Then $\{X(t)\}_{a\leq t\leq b}$ is a local martingale and as continuous realization.

8.4 Itô chain rule and product rule

Theorem 10. **Chain Rule**:-
Let $X\colon [0,T] \times \Omega \to \mathbb{R}$ be a stochastic process satisfying
$X(r) = X(s) + \int_s^r F(t)dt + \int_s^r G(t)dW_t$ for $0 \leq s \leq r \leq T$ or
$dX(t) = F(t)dt + G(t)dW_t,\ 0 \leq t \leq T$ (for Itô process)
with $F \in \mathrm{L}^1([0,T] \times \Omega), G \in \mathrm{L}^2([0,T] \times \Omega)$

Let $u\colon [0,T] \times \mathbb{R} \to \mathbb{R}$ such that $(t,x) \to u(t,x)$, $\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}$ are continuous. Then, $Y(t) = u(t, X(t)), 0 \leq t \leq T satisfies$

$$dY(t) = d\,(ut, X(t))$$

$$dY(t) = \frac{\partial u}{\partial t}(t, X(t))dt + \frac{\partial u}{\partial X}(t, X(t))dX(t) + \frac{1}{2}\frac{\partial^2 u}{\partial X^2}(t, X(t))G^2 dt$$

$$
\begin{aligned}
d\,f(x,y) &= \frac{\partial f}{\partial x}(x,y).dx + \frac{\partial f}{\partial y}(x,y).dy + \frac{1}{2}.\frac{\partial^2 f}{\partial x^2}dx^2 + \frac{\partial^2 f}{\partial x \partial y}.dxdy + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}dy^2 + \ldots \\
&= \frac{\partial f}{\partial x}(x,y).dx + \frac{\partial f}{\partial y}(x,y).dy
\end{aligned}
$$

higher power terms goes to 0

$$
\begin{aligned}
d\,Y(t) &= d\,u(t, X(t)) \\
&= \frac{\partial u}{\partial t}(t, X(t)).dt + \frac{\partial u}{\partial X_t}(t, X(t)).dX_t + \frac{1}{2}.\frac{\partial^2 u}{\partial t^2}(t, X(t)).dt^2 + \frac{\partial^2 u}{\partial t \partial X_t}(t, X(t)).dX_t dt + \frac{1}{2}.\frac{\partial^2 u}{\partial X_t^2}(t, X \\
&= \frac{\partial u}{\partial t}(t, X(t)).dt + \frac{\partial u}{\partial X_t}(t, X(t)).dX_t + \frac{1}{2}.\frac{\partial^2 u}{\partial X_t^2}(t, X(t)).dX_t^2
\end{aligned}
$$

- $dt^2 = 0$, $[dt << 1]$
- $dt.dW_t = 0$
- $dW_t^2 = dt$

Then, $dX_t = Fdt + GdW_t$
$dX_t^2 = F^2dt^2 + FGdW_tdt + G^2dW_t^2 = G^2dW_t^2$
$dX_t^2 = G^2dt$

$$
d\,Y(t) = \frac{\partial u}{\partial t}(t, X(t)).dt + \frac{\partial u}{\partial X_t}(t, X(t)).dX_t + \frac{1}{2}\frac{\partial^2 u}{\partial X_t^2}(t, X(t)).G^2dt
$$

$$
= \frac{\partial u}{\partial t}(t, X(t)).dt + \frac{1}{2}\frac{\partial^2 u}{\partial X_t^2}(t, X(t)).G^2dt + \frac{\partial u}{\partial X_t}(t, X(t)).(Fdt + GdW_t)
$$

$$
= [\frac{\partial u}{\partial t}(t, X(t)) + \frac{\partial u}{\partial X_t}(t, X(t)).F + \frac{1}{2}\frac{\partial^2 u}{\partial X_t^2}(t, X(t)).G^2].dt + \frac{\partial u}{\partial X_t}(t, X(t)).GdW_t
$$

Taking limits from r to s

$$
Y(s) = Y(r) + \int_r^s [\frac{\partial u}{\partial t}(t, X(t)) + \frac{\partial u}{\partial X_t}(t, X(t)).F + \frac{1}{2}\frac{\partial^2 u}{\partial X_t^2}(t, X(t)).G^2].dt + \int_r^s \frac{\partial u}{\partial X_t}(t, X(t)).GdW_t
$$

<u>Particular case:</u> $u(t, x) = u(x)$
$u(X(s)) = u(X(0)) + \int_0^s [\frac{\partial u(X(t))}{\partial x}F(t) + \frac{1}{2}\frac{\partial^2 u(X(t))}{\partial x^2}G^2(t)]dt + \int_0^s \frac{\partial^2 u(X(t))}{\partial x}G(t)dW_t$
$X(t) = W_t, u(x) = x^2$ so $dX_t = dW_t$, then $F = 0, G = 1$

$$W_s^2 = W_0^2 + \int_0^s [2W_t.F + 2G^2]dt + \int_0^s 2W_t.GdW_t$$

$$W_s^2 = W_0^2 + s + 2\int_0^s W_t dW_t$$

$$\int_0^s W_t dW_t = \frac{1}{2}W_s^2 - \frac{1}{2}s$$

When we integrate $\int_0^s ydy$ is equal to $\frac{1}{2}[y^2(s) - y^2(0)]$, but when we integrate $\int_0^s W_t dW_t$ we get $\frac{1}{2}W_s^2 - \frac{1}{2}s$, In this $(-\frac{1}{2}s)$ is extra term

In general, $\int_0^s f(W_t)dW_t$, where $F = 0, G = 0$

$$\int_0^s \frac{\partial u}{\partial x}(W_t)dW_t = u(W_s) - u(W_0) - \frac{1}{2}\int_0^s \frac{\partial^2 u}{\partial x^2}(W_t)dt$$

If $\frac{\partial u}{\partial x} = f$, then $u(x) = F(x) = \int f(x)dx$ and $\frac{\partial^2 u}{\partial x^2}(x) = f^{'}(x)$

$$\int_0^s f(W_t)dW_t = F(W_s) - F(W_0) - \int_0^s f^{'}(W_t)dt$$

- <u>Example-1</u>:-
  $\int_0^s e^{W_t}dW_t = e^{W_s} - \frac{1}{2}\int_0^s e^{W_t}dt$
  as $f(x) = F(x) = f^{'}(x) = e^x$

- <u>Example-2</u>:-
  $\int_0^s W_t dW_t = \frac{1}{2}W_s^2 - \frac{1}{2}s$
  as $f(x) = x, F(x) = \frac{1}{2}x^2, f^{'}(x) = 1$

- <u>Example-3</u>:-
  $\int_0^s W_t e^{W_t}dW_t = W_s e^{W_s} + 1 - e^{W_s} - \frac{1}{2}\int_0^s (e^{W_s} + W_s e^{W_s})ds$
  as $f(x) = xe^x, F(x) = xe^x - e^x, f^{'}(x) = xe^x + e^x$

- <u>Example-4</u>:-
  $\int_0^s W_t^m dW_t = \frac{W_s^{m+1}}{m+1} - \frac{1}{2}\int_0^s mW_s^{m-1}ds$
  as $f(x) = x^m, F(x) = \frac{x^{m+1}}{m+1}, f^{'}(x) = mx^{m-1}$

8.5 Itô formula

Let $dX_t = Fdt + GdW_t, (0 \le t \le T)$ be a Itô process and $u\colon [0,T] \times \mathbb{R} \to \mathbb{R}$ such that $\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}$ are continuous. Then, the process $Y(t) = u(t, X(t))$ satisfies

$$dY(t) = d(u(t, X(t)))$$

$$dY(t) = [\frac{\partial u}{\partial t}(t, X(t)) + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t, X(t)).G^2(t) + \frac{\partial u}{\partial x}(t, X(t)).F(t)]dt + \frac{\partial u}{\partial x}(t, X(t)).G(t)dW_t$$

Case:-

$X(t) = W_t$ then $dX(t) = dW_t \Rightarrow F = 0, G = 1$

$$d(u(t, X(t))) = [\frac{\partial u}{\partial t}(t, W_t) + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t, W_t)]dt + \frac{\partial u}{\partial x}(t, W_t)dW_t$$

Particular Case:-

$u(t, x) = u(x)$

$$d(u(t, X(t))) = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(W_t)dt + \frac{\partial u}{\partial x}(W_t)dW_t$$

Integrate on both sides on limits from 0 to t

$$\int_0^t \frac{\partial u}{\partial x}(W_s)dW_s = u(W_t) - u(W_0) - \frac{1}{2}\int_0^t \frac{\partial^2 u}{\partial x^2}(W_s)ds$$

$f(x) = \frac{\partial u}{\partial x}, F(x) = u(x), f'(x) = \frac{\partial^2 u}{\partial x^2}$

$$\int_0^t f(W_s)dW_s = F(W_t) - F(W_0) - \frac{1}{2}\int_0^t f'(W_s)ds$$

If $u(t, X(t)) = u(t, W_t)$

$$\frac{\partial u}{\partial x}(t, W_t)dW_t = d(u(t, W_t)) - [\frac{\partial u}{\partial t}(t, W_t) + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t, W_t)]dt$$

Integrate on both sides on limits from 0 to s

$$\int_0^s \frac{\partial u}{\partial x}(t, W_t)dW_t = u(s, W_s) - u(0, W_0) - \int_0^s [\frac{\partial u}{\partial t}(t, W_t) + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t, W_t)]dt$$

8.6 Generalized Itô Chain rule

Let $\{X^i(t)\}_{0\leq t\leq T}$ be $i = 1, 2, 3, \dots, n$ number of Itô processes given by $dX^i(t) = F^i(t).dt+G^i(t).dW_t; i = 1, 2, 3, \dots, n$ and $u\colon [0,T]\times\mathbb{R}^n \to \mathbb{R}$ be such that $\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x^i}, \frac{\partial^2 u}{\partial x^i \partial x^j}$ are continuous $[i, j = 1, 2, \dots, n]$.Then

$$d(u(t, X^1(t), X^2(t), \dots, X^n(t)))$$
$$= \frac{\partial u}{\partial t}(t, X^1(t), X^2(t), \dots, X^n(t))dt+\sum_{i=1}^{n}\frac{\partial u}{\partial x^i}(t, X^1(t), X^2(t), \dots, X^n(t)).dX^i(t)+\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 u}{\partial x^i\partial x^j}G^iG^j dt$$

Example 25. Differentiate $u(t, x) = x_1x_2$

Solution: $dX_1(t) = G_1(t).dW_t, dX_2(t) = G_2(t).dW_t$

$$\begin{aligned}
d(u(t, x)) &= d(x_1(t).x_2(t))\\
&= x_2(t).dx_1(t) + x_1(t).dx_2(t) + \frac{1}{2}[dx_1(t).dx_2(t) + dx_1(t).dx_2(t) + \frac{\partial x_2}{\partial x_1}.dx_2 + \frac{\partial x_1}{\partial x_2}.dx_1]\\
&= x_2(t).dx_1(t) + x_1(t).dx_2(t) + \frac{1}{2}[2.dx_1(t).dx_2(t) + 0 + 0]\\
&= x_2(t).dx_1(t) + x_1(t).dx_2(t) + dx_1(t).dx_2(t)\\
&= x_2(t).dx_1(t) + x_1(t).dx_2(t) + G_1(t).G_2(t)dt
\end{aligned}$$

■

8.7 Generalized Itô Chain rule in higher dimensions

$\mathbf{W}_t\colon \Omega \to \mathbb{R}^m$, $X(t)\colon \Omega \to \mathbb{R}^n$ and $m, n \in \mathbb{N}$

Let $\mathbf{W}_t = [W_t^1, W_t^2, \dots, W_t^m]^T$ be a m-dimensional Brownian motion (column vector) and $\{\mathcal{F}_t\}_{0\leq t\leq T}$ be a filtration such that

- W_t is $\{\mathcal{F}_t\}$ adapted
- $W_t - W_s$ is independent of $\{\mathcal{F}_s\}$, $0 \leq s \leq t \leq T$

$F\colon [0,T]\times\Omega \to \mathbb{R}^n$ $F(t) = [F^1(t), F^2(t), \dots, F^n(t)]^T$ and $G\colon [0,T]\times\Omega \to \mathbb{R}^{n\times m}$

$$G(t) = \begin{bmatrix} G^{11}(t) & G^{12}(t) & \dots & G^{1m}(t)\\ G^{21}(t) & G^{22}(t) & \dots & G^{2m}(t)\\ \dots & \dots\dots\dots & \dots & \\ G^{n1}(t) & G^{n2}(t) & \dots & G^{nm}(t)\end{bmatrix}$$

$F\colon [0,T]\times\Omega\to\mathbb{R}^n$ such that $\mathrm{E}[\int_0^T |F^i(t)|\,dt] < \infty [\in \mathrm{L}^1((0,T)\times\Omega\ , i = 1,2,\dots,n]$ and $G\colon [0,T]\times\Omega\to\mathbb{R}^{n\times m}$ such that $\mathrm{E}[\int_0^T |G^{ij}(t)|^2\,dt] < \infty[\in \mathrm{L}^2((0,T)\times\Omega\ , i = 1,2,\dots,n\ , j = 1,2,\dots,m]$

Let $X\colon [0,T]\times\Omega\to\mathbb{R}^n$ be a stochastic process given by $dX(t) = F(t).dt + G(t).dW_t$ or $X(t) = X_0 + \int_0^t F(s).ds + \int_0^t G(s).dW_s$

$$dX(t) = \begin{bmatrix} dX^1(t) \\ dX^2(t) \\ \cdots \\ \cdots \\ \cdots \\ dX^n(t) \end{bmatrix}_{n\times 1} \qquad F(t) = \begin{bmatrix} F^1(t) \\ F^2(t) \\ \cdots \\ \cdots \\ \cdots \\ F^n(t) \end{bmatrix}_{n\times 1} \qquad G(t)_{n\times m}.dW(t)_{m\times 1} = \begin{bmatrix} \sum_{j=1}^m G^{1j}.dW_t^j \\ \sum_{j=1}^m G^{2j}.dW_t^j \\ \cdots \\ \cdots \\ \cdots \\ \sum_{j=1}^m G^{nj}.dW_t^j \end{bmatrix}_{n\times 1}$$

$$dX^i(t) = F^i(t).dt + \sum_{j=1}^m G^{ij}.dW_t^j\ ;\ i = 1,2,\dots,n$$

Let $u\colon [0,T]\times\mathbb{R}^n\to\mathbb{R}$ be such that $\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x^i}, \frac{\partial^2 u}{\partial x^i.\partial x^k}\ ;\ i,k = 1,2,\dots,n$ are continuous. Then,

$$d(u(t,X(t))) = \frac{\partial u}{\partial t}(\centerdot)dt + \sum_{i=1}^n \frac{\partial u}{\partial x^i}(\centerdot)dX^i(t) + \frac{1}{2}\sum_{i,k=1}^n \frac{\partial u}{\partial x^i \partial x^k}(\centerdot)\sum_{j=1}^m G^{ij}(t).G^{kj}(t)dt$$

8.8 Stochastic Differential equation

$$dX(t) = f(t,X(t)).dt + \sigma(t,X(t)).dW_t\ ; X(0) = X_0,\ t\in(0,T) \to (1)$$

Definition 26. A stochastic process $X\colon [0,T]\times\Omega\to\mathbb{R}$ is called a solution of the Stochastic Differential equation (1) if

1. $X(\centerdot)$ is progressively measurable
2. $f(\centerdot, X(\centerdot)) \in \mathrm{L}^1_{ad}([0,T]\times\Omega)$
3. $\sigma(\centerdot, X(\centerdot)) \in \mathrm{L}^2_{ad}([0,T]\times\Omega)$

4. For all $t \in [0,T]$; $X(t)$ satisfies $X(t) = X_0 + \int_0^t f(s, X(s)).ds + \int_0^t \sigma(s, X(s)).dW_s$ a.s. $\mathbb{P}$

$\blacksquare$

Theorem 11. Existence and Uniqueness
Let $f\colon [0,T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma\colon [0,T] \times \mathbb{R} \to \mathbb{R}$ are continuous and satisfies

$$|f(t,x) - f(t,y)| \leq L\,|x-y|\,,\ |\sigma(t,x) - \sigma(t,y)| \leq L\,|x-y|$$

$$|f(t,x)| \leq L(1+|x|), |\sigma(t,x)| \leq L(1+|x|)\ \forall t \in [0,T], x \in \mathbb{R}$$

Let $X_0\colon \Omega \to \mathbb{R}$ be random variable, $\mathrm{E}\,|X_0|^2 < \infty$. Then, there exists a unique solution $X \in \mathrm{L}^2_{ad}([0,T] \times \Omega)$ of SDE(1).

Example 27. Let $S(t)$ denote the stock prices at time t. The evolution of $S(t)$ is given by the SDE $\frac{dS(t)}{S(t)} = \mu.dt + \sigma.dW_t; S(0) = S_0,\ \mu > 0,\ \sigma \in \mathbb{R}$ $\blacksquare$

Example 28. Brownian Bridge
$dB(t) = \frac{-B}{1-t}.dt + dW_t;\ 0 < t < 1,\ B(0) = 0$ $\blacksquare$

Example 29. Langevin's equation
$dX(t) = -bX(t).dt + \sigma.dW_t;\ X(0) = X_0$ $\blacksquare$

Example 30. Omstein-Uhlenbeck process
$\ddot{Y} = -b\dot{Y} + \sigma\xi$ $\blacksquare$

Example 31. Find the solution of $dX(t) = gX(t).dW_t;\ X(0) = X_0$ (g is function of t)
Solution: Firstly we will integrate $dx = g(t)xdt$ instead of dW_t

$$dx = g(t)xdt; x(0) = x_0$$
$$\frac{dx}{x} = g(t).dt$$

integrating on limits from 0 to t

$$ln(x(t)) = \int_0^t g(s).ds + c; x(0) = x_0$$

$$ln(x(t)) = \int_0^t g(s).ds + ln(x_0)$$
$$x(t) = x_0.e^{\int_0^t g(s).ds}$$

So, $X(t) = X_0.e^{\int_0^t g(s).dW_s}$ is half solution of the SDE. Now we need to find the extra term
For finding the extra term $X(t) = X_0.e^{Y(t)}$; $Y(t) = \int_0^t g(s).dW_s \Leftrightarrow dY(t) = g(t).dW_t$

$$\begin{aligned}
X(t) &= X_0.e^{Y(t)} \\
dX(t) &= X_0.d(e^{Y(t)}) = X_0.d(u(Y(t))) \ ; \ u(y) = e^y \\
dX(t) &= X_0[e^{Y(t)}.dY(t) + \frac{1}{2}e^{Y(t)}g^2(t).dt] \\
&= X_0e^{Y(t)}g(t).dW_t + \frac{1}{2}X_0e^{Y(t)}g^2(t).dt \\
&= X(t)g(t).dW_t + \frac{1}{2}X(t)g^2(t).dt \\
\frac{dX(t)}{X(t)} &= g(t).dW_t + \frac{1}{2}g^2(t).dt
\end{aligned}$$

So checking with the extra term is right for this solution,
$Y(t) = \int_0^t g(s)dW_s - \frac{1}{2}\int_0^t g^2(s).ds \ \Leftrightarrow dY(t) = g(t).dW_t - \frac{1}{2}g^2(t)dt$

$$\begin{aligned}
X(t) &= X_0.e^{Y(t)} \\
dX(t) &= X_0[e^{Y(t)}.dY(t) + \frac{1}{2}e^{Y(t)}g^2(t).dt] \\
&= X_0[e^{Y(t)}g(t).dW_t - e^{Y(t)}\frac{1}{2}g^2(t)dt + \frac{1}{2}e^{Y(t)}g^2(t).dt] \\
&= X_0e^{Y(t)}g(t).dW_t = X(t)g(t).dW_t
\end{aligned}$$

$\therefore X(t) = X_0.e^{(\int_0^t g(s)dW_s - \frac{1}{2}\int_0^t g^2(s).ds)}$ is solution of $dX(t) = gX(t).dW_t$; $X(0) = X_0$ ■

Example 32. Find the solution of $dX(t) = fX(t).dt + gX(t).dW_t$; $X(0) = X_0$ (f, g is function of t)
Solution:

$$\frac{dX(t)}{X(t)} = f(t).dt + g(t).dW_t$$

$X(t) = X_0e^{Y(t)}$

$Y(t) = \int_0^t g(s)dW_s - \frac{1}{2}\int_0^t g^2(s).ds + \int_0^t f(s).ds \Leftrightarrow dY(t) = f(t).dt + g(t).dW_t - \frac{1}{2}g^2(t).dt$
Checking if $X_0 e^{Y(t)}$ is correct solution of SDE

$$\begin{aligned}
dX(t) &= d(X_0 e^{Y(t)}) \\
&= X_0[e^{Y(t)}(f(t).dt + g(t).dW_t - \frac{1}{2}g^2(t).dt) + \frac{1}{2}e^{Y(t)}g^2(t).dt] \\
&= X_0 e^{Y(t)} f(t).dt + X_0 e^{Y(t)} g(t).dW_t \\
dX(t) &= X(t)f(t).dt + X(t)g(t).dW_t
\end{aligned}$$

$X(t) = X_0 e^{\int_0^t g(s)dW_s - \frac{1}{2}\int_0^t g^2(s).ds + \int_0^t f(s).ds}$ is solution of $dX(t) = fX(t).dt + gX(t).dW_t;\ X(0) = X_0$ ■

8.9 Stock prices

$S(t)$ is stock price at time t

$$\frac{dS(t)}{S(t)} = \mu.dt + \sigma.dW_t;\ S(0) = S_0,\ \mu > 0, \sigma \in \mathbb{R}$$

As above differential equation is the same as in the example-(32). So, solution of this SDE is in form of $S(t) = S_0 e^{Y(t)}$
$dY(t) = \mu.dt + \sigma.dW_t - \frac{1}{2}\sigma^2.dt$

$$\begin{aligned}
Y(t) &= \int_0^t \mu.ds + \int_0^t \sigma.dW_s - \frac{1}{2}\int_0^t \sigma^2.ds \\
Y(t) &= \mu t + \sigma W_t - \frac{1}{2}\sigma^2 t
\end{aligned}$$

$\therefore S(t) = S_0 e^{(\mu t + \sigma W_t - \frac{1}{2}\sigma^2 t)}$
$dS(t) = \mu S(t).dt + \sigma S(t).dW_t \Leftrightarrow S(t) = S_0 + \int_0^t \mu S(s).ds + \int_0^t \sigma S(s).dW_s$

Finding expectation of $S(t)$

$$\begin{aligned}
\mathrm{E}[S(t)] &= S_0 + \mu \mathrm{E}[\int_0^t S(s).ds] + \sigma \mathrm{E}[\int_0^t S(s).dW_s] \\
&= S_0 + \mu \mathrm{E}[\int_0^t S(s).ds] \\
&= S_0 + \mu \int_0^t \mathrm{E}[S(s).ds] \text{ Using Fubini theoremLet } \mathrm{E}[S(t)] = m(t) \\
m(t) &= m_0 + \mu \int_0^t m(s).ds
\end{aligned}$$

differential the above equation

$$\frac{dm(t)}{dt} = \mu m(t)$$

On integrating

$$m(t) = m_0 e^{\mu t}$$

$$\mathrm{E}[S(t)] = \mathrm{E}[S_0] e^{\mu t}$$

$$= S_0 e^{\mu t} (\text{if } S_0 \text{ is determinate})$$

$$S(t) = S_0 e^{[(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t]}$$

8.10 General Linear

$$dX(t) = [a(t) + f(t, X(t))].dt + [b(t) + \sigma(t, X(t))].dW_t \Rightarrow (2)$$

Solution of (2) can be written as

$$X(t) = X_0 e^{Y(t)} + \int_0^t e^{Y(t)-Y(s)} (a(s) - b(s).g(s)).ds + \int_0^t e^{Y(t)-Y(s)} b(s).dW_s$$

where

$$Y(t) = \int_0^t [f(r) - \frac{1}{2} g^2(r)].dr + \int_0^t g(r).dW_r$$