

# Notes on Probability

## Contents

<b>1</b>	<b>Modes of Convergence</b>	<b>3</b>
1.1	Uniform Integrability (u.i.) . . . . .	3
1.2	Convergence: Almost Sure, in Probability, and in $L^p$ . . . . .	5
1.3	Weak Convergence . . . . .	10
1.4	Examples . . . . .	15
1.5	The Moment Problem . . . . .	16
<b>2</b>	<b>Law of Large Numbers and Central Limit Theorem</b>	<b>18</b>
2.1	Strong Law of Large Numbers and Weak Law of Large Numbers . . . . .	18
2.2	Central Limit Theorem (CLT) . . . . .	22
2.3	Delta's Method . . . . .	22
<b>3</b>	<b>Discrete-Time Martingales</b>	<b>23</b>
3.1	Conditional Expectation . . . . .	23
3.2	Basic Properties . . . . .	25
3.3	Doob's Decomposition . . . . .	26
3.4	Martingale Convergence . . . . .	29
3.5	Optimal Stopping Theorem . . . . .	32
3.6	Inequalities . . . . .	37
<b>4</b>	<b>Examples of Discrete Stochastic Processes</b>	<b>40</b>
4.1	Finite Martingales . . . . .	40
4.2	Markov Chains (MC) . . . . .	40
4.3	Generating functions . . . . .	42
4.4	Branching Processes . . . . .	42
4.5	Birth-Death Processes . . . . .	43
4.6	Random Walks . . . . .	43
<b>5</b>	<b>General Stochastic Processes</b>	<b>44</b>
5.1	Finite-Dimensional Distributions (fdd) . . . . .	44
5.2	Continuity of Sample Paths . . . . .	46
5.3	Stopping Times . . . . .	47
5.4	Total Variation and Quadratic Variation . . . . .	50
<b>6</b>	<b>Discrete and Continuous Time Markov Chains</b>	<b>53</b>
6.1	Characterization . . . . .	53
6.2	Markov Chains: Classification of States . . . . .	54
6.3	Markov Chains: Classification of Chains . . . . .	55
6.4	Stationary distributions and the limit theorem . . . . .	56
6.5	Markov Chains: Examples . . . . .	56
6.6	Markov Processes: Examples . . . . .	56
<b>7</b>	<b>Poisson Processes</b>	<b>57</b>
7.1	Characterization . . . . .	57
7.2	Strong Markov Properties . . . . .	59
7.3	Compound Poisson . . . . .	59

<b>8</b>	<b>Brownian Motion</b>	<b>60</b>
8.1	Basic Properties . . . . .	60
8.2	Martingale Connection . . . . .	62
8.3	Strong Markov Properties (SMP) . . . . .	62
8.4	Hitting Time and Running Maximum . . . . .	65
8.5	Path Properties . . . . .	67
8.6	Zeros of Brownian motion . . . . .	68
<b>9</b>	<b>Basic Stochastic Calculate</b>	<b>69</b>
9.1	Definition of $\int_0^t X_s \, ds$ . . . . .	69
9.2	Itô Integral: Definitions of $\int_0^t X_s \, dB_s$ . . . . .	69
9.3	Properties of Stochastic Integral . . . . .	70
9.4	Itô Formula . . . . .	74
9.5	Applications of Itô formula . . . . .	74

# 1 Modes of Convergence

## 1.1 Uniform Integrability (u.i.)

**Definition 1.1** (Uniformly Integrable).  $\mathcal{K}$  - a family of real-value RVs  
 $\mathcal{K}$  is uniformly intergrable if

$$\begin{aligned} k(b) &= \sup_{X \in \mathcal{K}} \mathbb{E}(|X| \mathbf{1}_{\{|X| > b\}}) \\ &\rightarrow 0 \quad \text{as } b \rightarrow \infty \end{aligned}$$

**Proposition 1.2.** Let  $\psi : [0, \infty) \rightarrow [0, \infty)$  be a non-negative increasing convex function such that  $\lim_{r \rightarrow \infty} \frac{\psi(r)}{r} = \infty$ . Then

$$\sup_n \mathbb{E}(\psi(|X_n|)) < \infty \implies X \text{ is u.i.}$$

*Remark.* The converse is also true: if  $\{X_n\}$  is a u.i. family, then there exists a non-negative increasing convex function satisfying these properties.

**Possible choices of  $\psi$ :**  $r \rightarrow r^2$ ,  $r \rightarrow [(1+r) \log(1+r) - r], \dots$

*Proof.* Let  $M = \sup_n \mathbb{E}(\psi(|X_n|))$ . Notice that  $\lim_{r \rightarrow \infty} \frac{\psi(r)}{r} = \infty$  means: For all  $n > 0$ , there exists  $r_n > 0$  such that for all  $r > r_n$ ,

$$\psi(r) > nM \cdot r.$$

Then for a fixed  $n$ , we choose  $r > r_n$ :

$$\begin{aligned} M &\geq \mathbb{E}(\psi(|X_n|)) \geq \mathbb{E}(\psi(|X_n|) \mathbf{1}_{\{|X_n| > r\}}) \\ &\geq nM \mathbb{E}|X_n| \mathbf{1}_{\{|X_n| > r\}} \end{aligned}$$

$$\implies \mathbb{E}|X_n| \mathbf{1}_{\{|X_n| > r\}} \leq \frac{1}{n}. \text{ We are done.}$$

□

**Theorem 1.3.**

$\{X_k\}_{k \in \mathcal{K}}$  is u.i.  $\iff$  for all  $\epsilon > 0$ ,  $\exists \delta$ , such that for all  $A$  with  $\mathbb{P}A < \delta$ :  $\sup_{k \in \mathcal{K}} \mathbb{E}X_k \mathbf{1}_A < \epsilon$ .

## Exercises

**Exercise 1.1.1.** Let  $X$  be integrable and define  $X_n = X$  for all  $n$ . Show that  $\{X_n\}$  is u.i.

*Proof.* Compute  $k(b)$ :

$$k(b) = \sup_n \mathbb{E}(|X_n| \mathbf{1}_{\{|X_n| > b\}}) = \mathbb{E}(|X| \mathbf{1}_{\{|X| > b\}}).$$

Notice that  $|X| \mathbf{1}_{\{|X| > b\}} \leq |X|$ . By the dominated convergence theorem:

$$\lim_{b \rightarrow \infty} \mathbb{E}(|X| \mathbf{1}_{\{|X| > b\}}) = \mathbb{E}\left(\lim_{b \rightarrow \infty} |X| \mathbf{1}_{\{|X| > b\}}\right) = 0.$$

The last equality is baesd on the integrability of  $X$ . Then by the definition,  $\{X_n\}$  is u.i. □

**Exercise 1.1.2.** u.i.  $\implies \sup_n \mathbb{E}X_n^+ < \infty$

*Proof.* Notice that

$$\mathbb{E}X_n^+ \leq \mathbb{E}|X_n| \leq \mathbb{E}|X_n| \mathbf{1}_{\{|X| > b\}} + b.$$

By u.i., for all  $\epsilon > 0$ ,  $\exists b_0$  s.t.  $\forall b > b_0$ ,

$$\sup_n |X_n| \mathbf{1}_{\{|X_n| > b\}} < \epsilon.$$

$$\implies \sup_n \mathbb{E}X_n^+ \leq b_0 + 1 + \epsilon < \infty \text{ (Take } b = b_0 + 1).$$

□

**Exercise 1.1.3.** Let  $\{X_n\}$  be a sequence of RVs such that  $\mathbb{E}(\sup_n |X_n|) < \infty$ .

(a) Show that  $\{X_n\}$  is u.i.

(b) Give an example of a sequence  $\{Y_n\}$  such that  $\sup_n \mathbb{E}|Y_n| < \infty$  but  $\{Y_n\}$  is not u.i.

*Proof.* (a) Notice two facts:

$$\begin{aligned} |X_n| &\leq \sup_n |X_n| \\ \{|X_n| > b\} &\subset \{\sup_n |X_n| > b\} \end{aligned}$$

For all  $n$ , we have:

$$\begin{aligned} |X_n| \mathbf{1}_{\{|X_n| > b\}} &\leq \sup_n |X_n| \cdot \mathbf{1}_{\{\sup_n |X_n| > b\}} \\ \mathbb{E}\left(|X_n| \mathbf{1}_{\{|X_n| > b\}}\right) &\leq \mathbb{E}\left(\sup_n |X_n| \cdot \mathbf{1}_{\{\sup_n |X_n| > b\}}\right) \end{aligned}$$

Compute  $k(b)$ :

$$\begin{aligned} \sup_n \mathbb{E}\left(|X_n| \mathbf{1}_{\{|X_n| > b\}}\right) &\leq \mathbb{E}\left(\sup_n |X_n| \cdot \mathbf{1}_{\{\sup_n |X_n| > b\}}\right) \\ &\longrightarrow 0, \end{aligned}$$

since  $\sup_n |X_n|$  is integrable.

(b) Define  $Y_n(\omega) = \begin{cases} n & 0 < \omega \leq \frac{1}{n} \\ 0 & \text{o.w.} \end{cases}$

Then  $\sup \mathbb{E}Y_n = 1$ ; however,  $\sup_n \mathbb{E}Y_n \mathbf{1}_{\{Y_n > b\}} = 1$  for all  $n > b$ . Thus,  $\{Y_n\}$  is not u.i. □

**Exercise 1.1.4.** Let  $\{X_n\}$  and  $\{Y_n\}$  be two u.i. sequences. Then

(a)  $\{X_n + Y_n\}$  is u.i.

(b)  $\{X_n \vee Y_n\}$  is u.i.

*Proof.* Part (b) is trivial.

For (a), consider the following inequality:

$$|X + Y| \mathbf{1}_{\{|X+Y| \geq 2b\}} \leq 2|X| \mathbf{1}_{\{|X| \geq b\}} + 2|Y| \mathbf{1}_{\{|Y| \geq b\}}.$$

Or we have another option: Theorem 1.3. By the u.i. of  $X$  and  $Y$ , we have:

For all  $\epsilon > 0$  there exists  $\delta > 0$  such that if  $\mathbb{P}A < \delta$  then  $\sup_n \mathbb{E}|X_n| \mathbf{1}_A < \epsilon$  and  $\sup_n \mathbb{E}|Y_n| \mathbf{1}_A < \epsilon$ .

$\implies \mathbb{E}|X_n + Y_n| \mathbf{1}_A \leq \mathbb{E}|X_n| \mathbf{1}_A + \mathbb{E}|Y_n| \mathbf{1}_A < 2\epsilon$ .

Then use Theorem 1.3 again. □

**Exercise 1.1.5.** Let  $\{X_i\}$  be a sequence of i.i.d integrable RVs with mean  $\mu$ . Show that the sample mean is u.i. and  $\bar{X}_n = S_n/n \xrightarrow{L^1} \mu$  as  $n \rightarrow \infty$ .

*Proof.* Recall the SLLN:  $\{X_i\}$  are i.i.d. RVs and  $\mathbb{E}|X| = \mu < \infty \implies \bar{X}_n \xrightarrow{a.s.} \mu$ .

$\implies \bar{X}_n \xrightarrow{\mathbb{P}} \mu$ .

Because  $X_n \xrightarrow{L^1} X \iff X_n \xrightarrow{\mathbb{P}} X + \{X_n\}$  is u.i. (Theorem 1.8), it suffices to prove  $\{\bar{X}_n\}$  is a u.i. sequence.

Use the following result ( ): if  $X_n \xrightarrow{w} X$ ,

$$\{X_n\} \text{ is u.i. } \iff X_n, X \text{ are integrable } + \mathbb{E}X_n \rightarrow \mathbb{E}X.$$

$\implies \{X_n\}$  is a u.i. sequence. □

**Exercise 1.1.6.** Any u.i. sequence is tight.

*Proof.* Tool: the definition of the tightness. Let  $\{X_n\}$  be a u.i. sequence.

$$\begin{aligned} \sup_n \mathbb{P}(|X_n| > M) &= \sup_n \mathbb{E} \mathbf{1}_{\{|X_n| > M\}} \\ (M > 1) &< \sup_n \mathbb{E}\{|X_n| \mathbf{1}_{\{|X_n| > M\}}\} \\ (\text{u.i.}) &\rightarrow 0 \quad \text{as } M \rightarrow \infty \end{aligned}$$

$\implies \forall \epsilon > 0, \exists M_\epsilon$  such that

$$\sup_n \mathbb{P}(|X_n| > M_\epsilon) < \epsilon$$

$\implies \{X_n\}$  is tight. □

**Exercise 1.1.7.** Let  $B$  be a Brownian motion. Show that  $B_t$ ,  $B_t^2 - t$ , and  $e^{\lambda B_t - \frac{\lambda^2}{2}t}$  are not uniformly integrable.

*Proof.* Notice that  $\mathbb{E}|X_t| \mathbf{1}_{\{|X_t| > b\}}$  can be calculated explicitly. □

## 1.2 Convergence: Almost Sure, in Probability, and in $L^p$

**Definition 1.4.** Let  $(X_n)$  be a sequence RVs.

- $X_n \xrightarrow{a.s.} X$  if

$$\mathbb{P}\{\omega : X_n(\omega) \rightarrow X(\omega)\} = 1.$$

- $X_n \xrightarrow{\mathbb{P}} X$  if for every  $\epsilon > 0$ ,

$$\mathbb{P}(|X_n - X| > \epsilon) \rightarrow 0.$$

- $X_n \xrightarrow{L^p} X$  if

$$\mathbb{E}|X_n - X|^p \rightarrow 0.$$

**Theorem 1.5** (Continuous mapping theorem). Let  $(X_n)$  be a sequence RVs and  $f : \mathbb{R} \rightarrow \mathbb{R}$  be continuous.

- $X_n \xrightarrow{a.s.} X$

$$\implies f \circ X_n \xrightarrow{a.s.} f \circ X.$$

- $X_n \xrightarrow{\mathbb{P}} X$

$$\implies f \circ X_n \xrightarrow{\mathbb{P}} f \circ X.$$

### Almost sure convergence

**Theorem 1.6** (Borel-Canteli lemma). Let  $A_1, A_2, \dots$  be a sequence of events.

a)

$$\sum_{n=1}^{\infty} \mathbb{P}A_n < \infty \implies \mathbb{P}(\limsup_n A_n) = 0.$$

b) If  $\{A_i\}$  are independent, then

$$\sum_{n=1}^{\infty} \mathbb{P}A_n = \infty \implies \mathbb{P}(\limsup_n A_n) = 1.$$

*Remark.*

- Prove that  $X_n \xrightarrow{a.s.} X$ :

Define  $A_k = \{\omega : |X_k(\omega) - X(\omega)| \geq a_k\}$ . Choose  $\{a_k\} \downarrow 0$  such that

$$\sum_{n=1}^{\infty} \mathbb{P}A_n < \infty.$$

Or we check for every  $\epsilon > 0$ ,

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| > \epsilon) < \infty.$$

- Prove that  $X_n \not\xrightarrow{q.s.} X$ :

Define  $A_k = \{\omega : |X_k(\omega) - X(\omega)| > \epsilon\}$ . If the independence of the sequence of RVs  $\{X_n\}$  is given, it suffices to check if

$$\sum_{n=1}^{\infty} \mathbb{P}A_n = \infty.$$

**Example 1.7. Note:** There exists a sequence of events such that  $\sum \mathbb{P}A_n = \infty$  and  $\mathbb{P}(\limsup_n A_n) = 0$ . For example, take  $A_n = (0, a_n)$  with  $a_n \downarrow 0$  and  $a_n \geq \frac{1}{n}$ .

### Convergence in probability

**Proposition 1.8** (Relations to a.s. convergence).

- $X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{\mathbb{P}} X$
- $X_n \xrightarrow{\mathbb{P}} X \iff \text{Each subsequence } \{X_{n_k}\} \text{ contains a further subsequence } X_{n_{k_i}} \xrightarrow{a.s.} X.$

**Proposition 1.9** (Relations to convergence in  $L^p$ ).

- $X_n \xrightarrow{L^p} X \implies X_n \xrightarrow{\mathbb{P}} X$

### Convergence in $L^p$

**Proposition 1.10.**  $\forall p > q \geq 1$ , if  $X_n \rightarrow X$  in  $L^p$ , then  $X_n \rightarrow X$  in  $L^q$ .

*Remark.* Lyapunov's inequality:  $\|X\|_p \geq \|X\|_q$  for  $p \geq q \geq 1$ .

**Theorem 1.11.** If  $X_n \xrightarrow{\mathbb{P}} X$ , the following are equivalent:

- $\{X_n\}_{n \in \mathbb{N}}$  is uniform integrable.
- $\mathbb{E}|X_n| < \infty$  for all  $n$ ,  $\mathbb{E}|X| < \infty$ , and  $X_n \xrightarrow{L^1} X$ .
- $\mathbb{E}|X_n| < \infty$  for all  $n$ , and  $\mathbb{E}|X_n| \rightarrow \mathbb{E}|X| < \infty$ .

**Theorem 1.12.**

$$X_n \xrightarrow{L^1} X \iff \{X_n\} \text{ is u.i. and } X_n \xrightarrow{\mathbb{P}} X.$$

### Exercises

**Exercise 1.2.1.** Let  $(X_n)$  be a sequence of independent RVs such that  $\mathbb{P}(X_n < \infty) = 1$ ,  $\forall n$ .

(a) Show that

$$\sup_n X_n < \infty \iff \sum_n \mathbb{P}(X_n > A) < \infty \text{ for some } A.$$

(b) With no assumption of indep., give an example s.t.  $\sup_n X_n < \infty$  while  $\sum_n \mathbb{P}(X_n > A) = \infty$ .

*Proof.*

(a)  $\Leftarrow$  : By B-C (a),  $\mathbb{P}(X_n > A \text{ i.o.}) = 0$ . It implies

$$\mathbb{P}(X_n \leq A \text{ i.o.}) = \mathbb{P}(\sup_n X_n \leq A) = 1.$$

$\Rightarrow$  : Assume  $\sum_n \mathbb{P}(X_n > A) = \infty$  for all  $A$ . Notice that  $\{X_n > A\}$  is indep. of  $\{X_m > A\}$  because  $X_n$  is indep. of  $X_m$  when  $n \neq m$ . Then by B-C (b),

$$\mathbb{P}(X_n > A \text{ i.o.}) = 1$$

for all  $A$ . It implies  $\sup_n X_n = \infty$ . Contradiction.

(b)  $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, 1], \mathcal{B}, \text{Leb})$ . Let  $X_n(\omega) = \begin{cases} n & 1 - \frac{1}{n} < \omega \leq 1 \\ 0 & \text{o.w.} \end{cases}$ . Then

$$\sum_n \mathbb{P}(X_n > A) = \infty.$$

However,  $\limsup_n X_n = \begin{cases} \infty & \omega = 1 \\ 0 & \omega \in [0, 1) \end{cases}$ . So

$$\mathbb{P}(\limsup_n X_n < \infty) = 1.$$

□

**Exercise 1.2.2.**  $X_n \stackrel{iid}{\sim} F$  with  $F(x) < 1$  for all  $x$ . Set  $M_n = \max\{X_1, \dots, X_n\}$ . Prove that  $M_n \uparrow \infty$ .

*Proof.* Let  $N_j = \bigcap_{n=1}^{\infty} \{M_n \leq j\}$ .

$$\mathbb{P}(N_j) = \lim_{n \rightarrow \infty} \mathbb{P}\{M_n \leq j\} = \lim_{n \rightarrow \infty} F^n(j) = 0.$$

Let  $N = \bigcup_j N_j$ . It is a measurable set and  $\mathbb{P}(N) \leq \sum_j \mathbb{P}(N_j) = 0$ . Thus, if we take

$$\begin{aligned} \omega \in N^c &= \bigcap_j N_j^c \\ \Rightarrow \omega \notin N_j, \forall j \\ \Rightarrow \omega \notin \bigcap_n \{M_n \leq j\}, \forall j \\ \Rightarrow \omega \notin \lim_n \{M_n \leq j\}, \forall j \\ \Rightarrow \forall j, \exists n_0(\omega, j), \text{ such that } \forall n > n_0(\omega, j), M_n(\omega) > j \end{aligned}$$

It is equivalent to say  $\lim_n M_n = +\infty$ .

□

**Exercise 1.2.3.** (a) Show that for the non-negative RV  $X$ ,

$$\sum_{i=1}^{\infty} \mathbb{P}(X \geq i) \leq \mathbb{E}X < 1 + \sum_{i=1}^{\infty} \mathbb{P}(X \geq i).$$

(b) Show that if the integer-valued RV  $X \geq 0$ ,

$$\mathbb{E}X = \sum_{i=1}^{\infty} \mathbb{P}(X \geq i).$$

(c) Let  $X_0, X_1, \dots$  be i.i.d. continuous RVs and  $N = \inf\{n \geq 1 : X_n > X_0\}$ . Show that  $\mathbb{E}N = \infty$ .

*Proof.* (a) Define  $A_i = \{i - 1 \leq X < i\}$ . Then

$$X = \sum_{i=1}^{\infty} X \mathbf{1}_{A_i}, \text{ and} \\ (i-1)\mathbf{1}_{A_i} \leq X \mathbf{1}_{A_i} \leq i\mathbf{1}_{A_i}, \forall i.$$

$$\begin{aligned} \implies \sum_i (i-1)\mathbf{1}_{A_i} &\leq X \leq \sum_i i\mathbf{1}_{A_i}. \\ \implies \sum_i (i-1)\mathbb{P}(i-1 \leq X < i) &\leq \mathbb{E}X \leq \sum_i i\mathbb{P}(i-1 \leq X < i). \\ \implies \sum_{i=1}^{\infty} \mathbb{P}(X \geq i) &\leq \mathbb{E}X < 1 + \sum_{i=1}^{\infty} \mathbb{P}(X \geq i). \end{aligned}$$

(b) Because  $X = \sum_{i=1}^{\infty} X \mathbf{1}_{A_i}$ ,

$$\implies \mathbb{E}X = \sum_{i=1}^{\infty} \mathbb{E}(X \mathbf{1}_{A_i}) = \sum_{i=1}^{\infty} (i-1)\mathbb{P}(X = i-1) = \sum_{i=1}^{\infty} \mathbb{P}(X \geq i)$$

(c) Notice that  $\mathbb{E}N = \sum_{n=0}^{\infty} \mathbb{P}(N > n)$ . Now, we want to find  $\mathbb{P}(N > n)$ .

For  $n = 1$ , by symmetry,

$$\mathbb{P}(N > 1) = \mathbb{P}(X_1 \leq X_0) = \mathbb{P}(X_0 \leq X_1)$$

So  $\mathbb{P}(N > 1) = 1/2$ . (Note:  $\mathbb{P}(X_1 = X_0) = 0$  by continuity.)

For  $n > 1$ ,

$$\begin{aligned} \mathbb{P}(N > n) &= \mathbb{P}(X_1 \leq X_0, \dots, X_n \leq X_0) \\ &= \mathbb{P}(X_1 < X_2 < \dots < X_n < X_0) + \dots + \mathbb{P}(X_n < X_{n-1} < \dots < X_1 < X_0) \\ &= n! \mathbb{P}(X_1 < X_2 < \dots < X_n < X_0) \end{aligned}$$

And notice that  $\mathbb{P}(X_1 < X_2 < \dots < X_n < X_0) = \frac{1}{(n+1)!}$ , by symmetry. Thus, we have  $\mathbb{P}(N > n) = \frac{1}{n+1}$ ; and  $\mathbb{E}N = \sum_{n=0}^{\infty} \frac{1}{n+1} = \infty$ . □

**Exercise 1.2.4.**  $X_n \xrightarrow{\mathbb{P}} 0$  if and only if  $\mathbb{E}(|X_n|/(1 + |X_n|)) \rightarrow 0$ .

*Proof.*  $\implies$  : By continuous mapping theorem,

$$|X_n|/(1 + |X_n|) \xrightarrow{\mathbb{P}} 0.$$

Notice that  $\mathbb{E}\left(\sup_n \frac{|X_n|}{1+|X_n|}\right) \leq 1$ . So  $\{\frac{|X_n|}{1+|X_n|}\}_{n \in \mathbb{N}}$  is u.i.

Then use the following theorem: u.i.  $+\xrightarrow{\mathbb{P}} \implies \xrightarrow{L^1}$ .

$\Leftarrow$  : Notice that  $\xrightarrow{L^1} \implies \xrightarrow{\mathbb{P}}$ . Then by continuous mapping theorem,

$$|X_n|/(1 + |X_n|) \xrightarrow{\mathbb{P}} 0.$$

□

### -Applications of Borel-Canteli lemma-

**Exercise 1.2.5.** Let  $Y_n = \begin{cases} 1 & \text{w.p. } \frac{1}{n} \\ 0 & \text{w.p. } 1 - \frac{1}{n} \end{cases}$  and  $Z_n = \begin{cases} n^2 & \text{w.p. } \frac{1}{n^2} \\ 0 & \text{w.p. } 1 - \frac{1}{n^2} \end{cases}$  be two independent RVs sequences. Assume  $Y$  and  $Z$  are independent. Define  $X_n = Y_n Z_n$ . Show  $X_n \xrightarrow{a.s.} 0$ .

*Proof.* First, compute the distribution of  $X$ .

$$X_n = \begin{cases} n^2 & \text{w.p. } 1/n^3 \\ 0 & \text{o.w.} \end{cases}$$



For any  $\epsilon > 0$ , let  $A_n = \{|X_n| > \epsilon\}$ . Then notice that

$$\sum_n \mathbb{P}A_n = \sum_n \frac{1}{n^3} < \infty.$$

By Borel-Canteli (a),

$$\mathbb{P}(A_n \text{ i.o.}) = 0, \forall \epsilon > 0.$$

$$\implies X_n \xrightarrow{a.s.} 0.$$

□

**Exercise 1.2.6.** Let  $H_n = \begin{cases} 1 & \text{w.p. } \frac{1}{n} \\ 0 & \text{w.p. } 1 - \frac{1}{n} \end{cases}$  be an independent RVs sequence. Define  $X_n = (-1)^n n H_n$ .

Show that  $X_n \xrightarrow{\mathbb{P}} 0$ , but  $\limsup_{n \rightarrow \infty} X_n = \infty$  a.s. and  $\liminf_{n \rightarrow \infty} X_n = -\infty$  a.s.

*Proof.* First, directly check the definition of  $\xrightarrow{\mathbb{P}}$ .

For every  $\epsilon > 0$ ,

$$\mathbb{P}(|X_n| > \epsilon) = \mathbb{P}(nH_n > \epsilon) \leq \mathbb{P}(nH_n = n) = \frac{1}{n} \rightarrow 0.$$

$$\implies X_n \xrightarrow{\mathbb{P}} 0.$$

Then, let  $A_n = \{X_n > \frac{n}{2}\}$ .

$$\sum_{n=1}^{\infty} \mathbb{P}A_n = \sum_{n=1}^{\infty} \mathbb{P}(X_n = n) = \sum_{n=2,4,\dots} \frac{1}{n} = \infty$$

By Borel-Canteli (b),

$$\mathbb{P}(A_n \text{ i.o.}) = \mathbb{P}(X_n > \frac{n}{2} \text{ i.o.}) = \mathbb{P}(\limsup_{n \rightarrow \infty} X_n = \infty) = 1.$$

Similarly, take  $B_n = \{X_n < -\frac{n}{2}\}$ .

$$\sum_{n=1,3,\dots} \mathbb{P}B_n = \sum_{n=1,3,\dots} \frac{1}{n}.$$

By Borel-Canteli (b) again,

$$\mathbb{P}(B_n \text{ i.o.}) = \mathbb{P}(\liminf_{n \rightarrow \infty} = -\infty) = 1.$$

□

**Exercise 1.2.7.** Let  $Y_n = \begin{cases} 1 & \text{w.p. } \frac{1}{n} \\ 0 & \text{w.p. } 1 - \frac{1}{n} \end{cases}$  be an independent RVs sequence and  $X$  be another RV independent of  $\{Y_n\}$ . Define  $Z_n = X + Y_n$ . Does  $\{Z_n\}$  converge weakly? in probability? a.s.?

*Proof.*

- **in probability.** Yes, its limit is  $X$ .

For all  $\epsilon > 0$ ,

$$\begin{aligned} \mathbb{P}(|Z_n - X| > \epsilon) &= \mathbb{P}(|Y_n| > \epsilon) \\ &\leq \mathbb{P}(|Y_n| = 1) = 1/n \rightarrow 0 \end{aligned}$$

- **in distribution** Yes, its limit is  $X$ .

Implied by convergence in probability.

- **a.s.** No, it doesn't converge almost surely.

Assume  $(Z_n)$  converge almost surely. Then the limit must be  $X$ .

For  $\epsilon \in (0, 1)$ , let  $A_n = \{|Z_n - X| > \epsilon\} = \{|Y_n| > \epsilon\}$ . Because

$$\sum_n \mathbb{P}A_n = \sum_n \mathbb{P}(Y_n = 1) = \sum_n \frac{1}{n} = \infty$$

and  $\{A_n\}$  are independent (by independence of  $(Y_n)$ ), by Borel-Canteli (b),

$$\mathbb{P}(A_n \text{ i.o.}) = 1.$$

Thus,  $Z_n$  doesn't converge to  $X$ . Contradiction.

□

### -m.s. convergence-

**Exercise 1.2.8.**  $X_n \xrightarrow{L^2} X$  as  $n \rightarrow \infty$  implies that  $\mathbb{E}X_n \rightarrow \mathbb{E}X$  and  $\mathbb{E}X_n^2 \rightarrow \mathbb{E}X^2$ .

*Proof.* By Minkowski inequality,

$$\begin{aligned} \|X_n\|_2 &= \|X + (X_n - X)\|_2 \leq \|X\|_2 + \|X - X_n\|_2, \\ \|X\|_2 &= \|X_n + (X - X_n)\|_2 \leq \|X_n\|_2 + \|X - X_n\|_2. \end{aligned}$$

Therefore, we have

$$\left| \|X_n\|_2 - \|X\|_2 \right| \leq \|X - X_n\|_2 \rightarrow 0.$$

$$\implies \mathbb{E}X_n^2 \rightarrow \mathbb{E}X^2.$$

□

**Exercise 1.2.9** (Loeve's criterion).  $\{X_n\}$  converges in  $L^2$  if and only if  $\mathbb{E}X_n X_m \rightarrow C$  as  $m, n \rightarrow \infty$  for some constant  $C$ .

*Proof.*  $\Leftarrow$  : Take  $m = n$ . We are done.

$\Rightarrow$  : By Cauchy criterion for  $L^2$ -convergence,

$$\|X_n - X_m\|_2 \rightarrow 0, \text{ as } m, n \rightarrow \infty.$$

Thus, we have

$$\begin{aligned} \|X_m - X_n\|_2^2 &= \mathbb{E}X_m^2 + \mathbb{E}X_n^2 - 2\mathbb{E}X_m X_n \\ \lim_{m, n \rightarrow \infty} \|X_m - X_n\|_2^2 &= \lim_{m, n \rightarrow \infty} (\mathbb{E}X_m^2 + \mathbb{E}X_n^2 - 2\mathbb{E}X_m X_n) \\ &= 0 = \mathbb{E}X^2 + \mathbb{E}X^2 - 2 \lim_{m, n \rightarrow \infty} \mathbb{E}X_m X_n \end{aligned}$$

It implies that  $\mathbb{E}X_n X_m \rightarrow \mathbb{E}X^2$ .

□

## 1.3 Weak Convergence

**Definition 1.13.** Let  $(X_n)$  be a sequence RVs. The following definitions of  $X_n \xrightarrow{w} X$  are equivalent:

- $\mathbb{P} \circ X_n^{-1} \xrightarrow{w} \mathbb{P} \circ X^{-1}$ .
- $F_n \rightarrow F$  for all continuous points of  $F$ .

The following theorem can be used to check the weak convergence conveniently.

**Theorem 1.14.**

- **Lévy's continuity theorem**

Let  $\{F_n\}$  be a sequence of CDFs, and  $\{\phi_n\}$  be the corresponding CFs.

- (i) If  $F_n \xrightarrow{w} F$ , then  $\phi_n(t) \rightarrow \phi(t)$  for all  $t$ ; and conversely,
- (ii) suppose that  $\lim \phi_n$  exists for all  $t$ , and  $\phi$  is continuous at 0. Then  $F_n \xrightarrow{w} F$  and  $\{F_n\}$  are tight.

• **Cramer-Wald device**

$\{X_n\}$  are  $\mathbb{R}^k$ -value RVs.

$$X_n \xrightarrow{w} X \iff a^T X_n \xrightarrow{w} a^T X \quad \forall a \in \mathbb{R}^k$$

**Theorem 1.15** (Scheffé's Theorem: Convergence of PDF  $\implies$  Weak convergence). Assume there is a sequence of PDFs  $f_n \rightarrow f_\infty$  almost surely. Define the corresponding Borel measures

$$\mu_n(B) = \int_B f_n(x) \, dx.$$

Then we have

$$\sup_{B \in \mathcal{B}(\mathbb{R})} |\mu_n(B) - \mu_\infty(B)| \rightarrow 0.$$

*Remark.*

- A more general result holds for  $L^p$  space ( $p \geq 1$ ): if we have  $\lim_n \|f_n\|_p = \|f\|_p$  and  $f_n \xrightarrow{a.s.} f$ , then

$$\lim_n \|f - f_n\|_p = 0.$$

- If we take  $B = (-\infty, x]$ , we can get the weak convergence.

*Proof.* For all  $B \in \mathcal{B}(\mathbb{R})$ ,

$$\left| \int_B f_n \, dx - \int_B f_\infty \, dx \right| \leq \int_B |f_n - f_\infty| \, dx.$$

Therefore,

$$\begin{aligned} \sup_B \left| \int_B f_n \, dx - \int_B f_\infty \, dx \right| &\leq \int_{\mathbb{R}} |f_n - f_\infty| \, dx \\ &\rightarrow 0. \end{aligned}$$

For “ $\rightarrow 0$ ” part, we first notice the fact where  $\int (f_\infty - f) \, dx = 1 - 1 = 0$ , which implies that

$$\int_{\mathbb{R}} |f_n - f_\infty| \, dx = 2 \int_{\mathbb{R}} (f_\infty - f_n)^+ \, dx;$$

then, we use the dominated convergence theorem: because  $(f_\infty - f_n)^+ < f_\infty$ ,

$$\lim \int (f_\infty - f_n)^+ \, dx = \int \lim (f_\infty - f_n)^+ \, dx.$$

□

**Example 1.16** (Counterexample: Weak convergence doesn't imply the convergence of PDF). Consider

$$f_n(x) = (1 - \cos 2\pi nx) \mathbf{1}_{\{0 \leq x \leq 1\}}.$$

Then for  $0 < x \leq 1$ ,

$$\begin{aligned} F_n(x) &= \int_0^x (1 - \cos 2\pi nx) \, dx \\ &= x - \frac{\sin 2\pi nx}{2\pi n} \\ &\rightarrow x \end{aligned}$$

It means the limit of weak convergence is the uniform distribution on  $[0, 1]$ . However,  $(f_n)$  doesn't converge to  $f(x) = \mathbf{1}_{\{0 \leq x \leq 1\}}$ .

**Proposition 1.17** (Connection between  $\xrightarrow{\mathbb{P}}$  and  $\xrightarrow{w}$ ).

- $X_n \xrightarrow{\mathbb{P}} X \implies X_n \xrightarrow{w} X$ .
- $X_n \xrightarrow{\mathbb{P}} c \iff X_n \xrightarrow{w} c$ .

**Theorem 1.18** (Skorohod's representation theorem). *If RVs  $\{X_n\}$  and  $X$  have DFs  $\{F_n\}$  and  $F$  such that  $X_n \xrightarrow{w} X$ , then there exists a probability space  $(\Omega', \mathcal{F}', \mathbb{P}')$  and real-value RVs  $\{Y_n\}$  and  $Y$  such that they have DFs  $\{F_n\}$  and  $\{F\}$ , and  $Y_n \xrightarrow{a.s.} Y$ .*

*Remark.*  $X_n \stackrel{D}{=} Y_n$ ; however, in general, the distribution of  $(X_n, X_m)$  doesn't equal to that of  $(Y_n, Y_m)$ !

The proof of CMT is an application of Skorohod's representation theorem.

**Theorem 1.19** (Continuous mapping theorem). *Let  $(X_n)$  be a sequence RVs and  $f : \mathbb{R} \rightarrow \mathbb{R}$  be continuous.*

- $X_n \xrightarrow{w} X$   
 $\implies f \circ X_n \xrightarrow{w} f \circ X$ .

*Proof.* By Skorohod's representation theorem,  $Y_n \xrightarrow{a.s.} Y$  in  $(\Omega', \mathcal{F}', \mathbb{P}')$ . Using CMT of a.s. convergence,

$$f \circ Y_n \xrightarrow{a.s.} f \circ Y.$$

Note that the weak convergence is implied by a.s. convergence. Thus,

$$f \circ X_n \stackrel{D}{=} f \circ Y_n \xrightarrow{w} f \circ Y \stackrel{D}{=} f \circ X.$$

□

**Theorem 1.20** (Portmanteau). *Omitted.*

In the rest of part, we will focus on the tightness and vague convergence.

**Question.** *Does every sequence of DFs converge?*

**Theorem 1.21** (Helly's selection Theorem). *Any seq. of DFs contains a convergent subseq.*

*Remark.* This theorem gives a partial answer for the preceding question.

**Question.** *If a sequence of DFs  $F_n \rightarrow F$ , is  $F$  a DF? When is the limit of a sequence of DFs a DF?*

**Example 1.22.** Let  $X_n = n$  almost surely. Then  $F_n(x) = \begin{cases} 0 & x < n \\ 1 & x \geq n \end{cases}$ .  
 $\implies F_n(x) = \mathbf{1}_{[n, \infty)} \rightarrow 0$  as  $n \rightarrow \infty$  for all  $x \in \mathbb{R}$ ; however, 0 is not a DF.

When  $\{F_n\}$  is tight, the limit must be a DF.

**Definition 1.23.**  $\{F_n\}$  is tight, if for all  $\epsilon > 0$ , there exists  $M = M_\epsilon > 0$  s.t.

$$F_n([-M_\epsilon, M_\epsilon]) > 1 - \epsilon, \forall n.$$

*Remark.* It is equivalent to require

$$\sup_n \mathbb{P}(|X_n| > M_\epsilon) \leq \epsilon.$$

The following theorems can be used to check tightness.

**Theorem 1.24** (Prohorov). *Every subsequential limit the DF of a prob. measure  $\iff \{F_n\}$  is tight.*

**Theorem 1.25.** *Let  $(X_n)$  be RVs with DFs  $\{F_n\}$ . If there is  $\varphi \geq 0$  s.t.  $\varphi \uparrow \infty$  as  $|x| \uparrow \infty$  and*

$$C = \sup_n \mathbb{E}\varphi(X_n) < \infty$$

$\implies \{F_n\}$  is tight.

*Remark.* The most commonly choice of  $\varphi$  is  $\varphi(x) = |x|^r$  for  $r > 0$ .

## Exercises

**Exercise 1.3.1.** Let  $X_n$  be symmetric and  $X_n \xrightarrow{w} X$ . Define  $Y_n = -X_n$ .

(a) Show that  $Y_n \xrightarrow{w} X$ . (b) Give an example where  $X_n + Y_n \not\xrightarrow{w} 2X$ .

*Proof.* content... □

**Exercise 1.3.2.** (a) Let  $X_n$  and  $X$  be positive integer-valued RVs. Prove

$$X_n \xrightarrow{w} X \iff \mathbb{P}(X_n = k) \rightarrow \mathbb{P}(X = k), \forall k \in \mathbb{N}.$$

(b) Let  $(X_n)$  be a sequence of RV with  $\mathbb{P}(X_n = 1 - 1/n) = 1/2$  and with  $\mathbb{P}(X_n = 1 + 1/n) = 1/2$ . Show that  $X_n \xrightarrow{w} 1$  but the pmf of  $X_n$  doesn't converge to that of 1.

*Proof.* (a)  $\Leftarrow$  : The CDF of  $X_n$  is

$$F_n(t) = \sum_{k=1}^{\lfloor t \rfloor} \mathbb{P}(X_n = k).$$

Because  $\mathbb{P}(X_n = k) \rightarrow \mathbb{P}(X = k)$  as  $n \rightarrow \infty$ , we have

$$F_n(t) \rightarrow F(t).$$

$\Rightarrow$  : Notice that

$$\begin{aligned} \mathbb{P}(X_n = k) &= \mathbb{P}(X_n \leq k) - \mathbb{P}(X_n \leq k-1) \\ \lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) &= \lim_{n \rightarrow \infty} (\mathbb{P}(X_n \leq k) - \mathbb{P}(X_n \leq k-1)) \\ \lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) &= \mathbb{P}(X = k) \end{aligned}$$

(b) Compute the CDF of  $X_n$ :

$$\begin{aligned} F_n(t) = \mathbb{P}(X_n \leq t) &= \begin{cases} 1 & t \geq 1 + \frac{1}{n} \\ \frac{1}{2} & 1 + \frac{1}{n} > t \geq 1 - \frac{1}{n} \\ 0 & t < 1 - \frac{1}{n} \end{cases} \\ &= \frac{1}{2} \mathbf{1}_{\{t \geq 1 + \frac{1}{n}\}} + \frac{1}{2} \mathbf{1}_{\{t \geq 1 - \frac{1}{n}\}} \end{aligned}$$

$$\Rightarrow \lim_{n \rightarrow \infty} F_n(t) = \frac{1}{2} \mathbf{1}_{\{t > 1\}} + \frac{1}{2} \mathbf{1}_{\{t \geq 1\}} = \mathbf{1}_{\{t > 1\}} + \frac{1}{2} \mathbf{1}_{\{t = 1\}}$$

And we notice that the distribution function of  $X$  is

$$F(t) = \mathbb{P}(X \leq t) = \mathbf{1}_{\{t \geq 1\}}.$$

Therefore,  $X_n \xrightarrow{w} X$ ; because  $t = 1$  is the unique discontinuous point of  $F$ .

Now we compute the pmf of  $X_n$ :

$$p_n(x) = \mathbb{P}(X_n = x) = \begin{cases} \frac{1}{2} & x = 1 \pm \frac{1}{n} \\ 0 & \text{o.w.} \end{cases}$$

$$\Rightarrow \lim_{n \rightarrow \infty} p_n(x) = 0.$$

However,  $p(x) = \mathbb{P}(X = x) = \begin{cases} 1 & x = 1 \\ 0 & \text{o.w.} \end{cases}$ . Thus, the pmf of  $X_n$  doesn't converge to that of  $X$ . □

**Exercise 1.3.3** (Discrete RVs with a continuous weak limit). Suppose  $F_n$  puts mass  $\frac{1}{n}$  at points  $\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$ . If  $F(x) = x$  on  $[0, 1]$ , then show that  $F_n \rightarrow F$ .

*Proof.* Compute  $F_n$ :

$$F_n(t) = \sum_{k=1}^n \frac{1}{n} \mathbf{1}_{\{t \geq \frac{k}{n}\}}.$$

If  $t \geq 1$  or  $t \leq 0$ , there is nothing to prove. Assume  $0 < t < 1$ .

Notice that for every  $\epsilon > 0$ , letting  $N_0 > \frac{1}{\epsilon}$ ,  $\forall n > N_0$ ,  $\exists k = k(n)$  s.t.  $\frac{k}{n} \leq t \leq \frac{k+1}{n}$ ,

$$|F_n(t) - t| \leq \frac{1}{n} < \epsilon.$$

$$\implies F_n(t) \rightarrow t.$$

□

**Exercise 1.3.4** (Convergence of maxima). Let

**-Vague convergence-**

**Definition.** A sequence...

**Exercise 1.3.5.** Define ...

**-Applications of Skorohod's representation theorem-**

**Exercise 1.3.6.** Let  $X_n \xrightarrow{w} X$ ,  $a_n \rightarrow a$ , and  $b_n \rightarrow b$ . Show that  $a_n X_n + b_n \xrightarrow{w} aX + b$ .

*Proof.* By Skorohod's, there exists  $\{Y_n\}$  and  $Y$  such that  $Y_n \xrightarrow{a.s.} Y$ . So  $a_n Y_n + b_n \xrightarrow{a.s.} aY + b$ .

$$\implies a_n Y_n + b_n \xrightarrow{w} aY + b$$

$$\implies a_n X_n + b_n \xrightarrow{w} aX + b, \text{ by noticing that } a_n X_n + b_n \stackrel{D}{=} a_n Y_n + b_n, \text{ and } aX + b \stackrel{D}{=} aY + b.$$

□

**Exercise 1.3.7** (Uniqueness). Let  $X_n \xrightarrow{w} X$  and  $X_n \xrightarrow{w} Y$ . Show that  $X \stackrel{D}{=} Y$ .

*Proof.* ???

□

**Exercise 1.3.8** (Fatou's lemma). Let  $g \geq 0$  be continuous, and  $X_n \xrightarrow{w} X$ . Show that

$$\liminf_n \mathbb{E}g(X_n) \geq \mathbb{E}g(X).$$

*Proof.* By Skorohod's,  $\exists Y_n \xrightarrow{a.s.} Y$  with  $Y_n \stackrel{D}{=} X_n$  and with  $Y \stackrel{D}{=} X$ . By Fatou's lemma,

$$\liminf \mathbb{E}g(Y_n) \geq \mathbb{E}g(Y).$$

Because  $Y_n \stackrel{D}{=} X_n$  and  $Y \stackrel{D}{=} X \implies \mathbb{E}g(Y_n) = \mathbb{E}g(X_n)$ ,  $\mathbb{E}g(Y) = \mathbb{E}g(X)$ . Thus,

$$\liminf \mathbb{E}g(X_n) \geq \mathbb{E}g(X).$$

□

**-Slutsky's theorem-**

**Exercise 1.3.9.** Let  $X_n \xrightarrow{w} X$ . Show the following results:

(a)  $|Y_n - X_n| \xrightarrow{w} 0$ , then  $Y_n \xrightarrow{w} X$ .

(b)  $Y_n \xrightarrow{w} c$ , then  $(X_n, Y_n) \xrightarrow{w} (X, c)$ .

(c)  $Y_n \xrightarrow{w} c$ , then  $X_n + Y_n \xrightarrow{w} X + c$ ,  $X_n Y_n \xrightarrow{w} cX$ , and  $X_n/Y_n \xrightarrow{a.s.} X/c$ .

## 1.4 Examples

Figure 1: Relations among modes of convergence

**Example 1.26** (Counterexample:  $X_n \xrightarrow{a.s.} X$  but not  $X_n \xrightarrow{L^p} X$ ).

Let  $\{X_n\}$  be a sequence of independent RVs such that  $\mathbb{P}(X_n = n^3) = \frac{1}{n^2}$  and  $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n^2}$ .

It converges to 0 almost surely by Borel-Canteli (a); however, it doesn't converge in  $L^1$  because we have  $\mathbb{E}|X_n| = n$ .

**Example 1.27.**

Let  $N_t$  be a Poisson process with  $\mathbb{E}N_t = \lambda t$ . Define  $X_t = N_t/t$ .

- We consider its limit when  $t \rightarrow \infty$ .

It is easy to see  $\mathbb{E}|X_t - 0| = \lambda \not\rightarrow 0$ ; moreover,  $X_t \rightarrow \lambda$  almost surely. See here for another proof.

First, for  $n \in \mathbb{N}$ , we re-write  $N_n = \sum_{i=1}^n (N_i - N_{i-1})$ . By SLLN,

$$N_n/n \xrightarrow[n.s.]{L^2} \mathbb{E}(N_1 - N_0) = \lambda.$$

Then, for  $t \in \mathbb{R}_+$ , we re-write  $N_t = N_{[t]} + (N_t - N_{[t]})$ ; so

$$N_t/t = N_{[t]}/[t] \cdot \frac{[t]}{t} + (N_t - N_{[t]})/t.$$

It suffices to prove  $\limsup_t (N_t - N_{[t]})/t = 0$ . We can use the tail probabilities of Poisson RV

$$\mathbb{P}(X \geq x) \leq e^{-\lambda} (e\lambda)^x / x^x.$$

Then  $\sum_{n=1}^{\infty} \mathbb{P}(N_{t_n} - N_{[t_n]} \geq t\epsilon)$  converges. Finally, use Borel-Cantelli (a).

- And we are also interested in the limit when  $t \rightarrow 0$ . Its  $L^1$  limit is same.

However, its almost sure limit is 0.

**Example 1.28** (Counterexample:  $X_n \xrightarrow{L^p} X$  but not  $X_n \xrightarrow{a.s.} X$ ).

Let  $\{X_n\}$  be a sequence of independent RVs such that  $\mathbb{P}(X_n = 1) = \frac{1}{n}$  and  $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$ . Then  $X_n \xrightarrow{L^1} 0$  but  $X_n$  doesn't converge to 0 almost surely.

See Exercise 1.2.6.

**Example 1.29** (Counterexample:  $X_n \xrightarrow{\mathbb{P}} X$  but not  $X_n \xrightarrow{L^p} X$ ).

Let  $\{X_n\}$  be a sequence of independent RVs such that  $\mathbb{P}(X_n = n^3) = \frac{1}{n^2}$  and  $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n^2}$ . Then  $X_n \xrightarrow{\mathbb{P}} 0$  but  $X_n$  doesn't converge to 0 in  $L^p$  for  $p \geq 1$ .

Because for every  $\epsilon > 0$ ,  $\mathbb{P}(|X_n| > \epsilon) = \mathbb{P}(X_n = n^3) = \frac{1}{n^2} \rightarrow 0$ . However,

$$\mathbb{E}X^p = \frac{1}{n^2} n^{3p} = n^{3p-2} \rightarrow \infty.$$

**Example 1.30** (Two Counterexamples:  $X_n \xrightarrow{\mathbb{P}} X$  but not  $X_n \xrightarrow{a.s.} X$ ).

- $\Omega = (0, 1]$ ;  $\mathcal{H}$  = the Borel  $\sigma$ -algebra on  $\Omega$ ;  $\mathbb{P}$  = the Lebesgue measure.

Define the following sequence of RVs

$$\begin{aligned} X_1 &= \mathbf{1}_{(0,1]}; \\ X_2 &= \mathbf{1}_{(0,\frac{1}{2}]}, \quad X_3 = \mathbf{1}_{(\frac{1}{2},1]}; \\ X_4 &= \mathbf{1}_{(0,\frac{1}{3}]}, \quad X_5 = \mathbf{1}_{(\frac{1}{3},\frac{2}{3}]}, \quad X_6 = \mathbf{1}_{(\frac{2}{3},1]}; \\ &\dots \end{aligned}$$

It converges in probability, because  $\mathbb{P}\{X_n > \epsilon\}$  form the sequence  $(1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \dots)$ .

But it doesn't converge a.s. because  $\liminf X_n(\omega) = 0$  and  $\limsup X_n(\omega) = 1$  for every  $\omega \in \Omega$ .

- Let  $(X_n)$  be a sequence of independent RVs with  $\mathbb{P}(X_n = 1) = p_n$  and  $\mathbb{P}(X_n = 0) = 1 - p_n$ , and  $\sum p_n = \infty$ . Then  $X_n \xrightarrow{\mathbb{P}} 0$  but  $X_n$  doesn't converge to 0 almost surely.

It is because  $\mathbb{P}(|X_n - 0| > \epsilon) = p_n \rightarrow 0$ , so  $X_n \xrightarrow{\mathbb{P}} 0$ ; however, letting  $A_n = \{|X_n| > \epsilon\}$ , we have

$$\sum_{n=1}^{\infty} \mathbb{P}A_n = \sum_{n=1}^{\infty} p_n = \infty.$$

By Borel-Canteli (b),  $X_n \not\xrightarrow{\text{a.s.}} 0$  almost surely.

**Example 1.31** (Counterexample:  $X_n \xrightarrow{\mathbb{P}} X$  but not  $X_n \xrightarrow{L^p} X$ ).

Let  $\Omega = (0, 1]$ ;  $\mathcal{H}$  = the Borel  $\sigma$ -algebra on  $\Omega$ ;  $\mathbb{P}$  = the Lebesgue measure.

Define  $X_n = 2^n \mathbf{1}_{(0, \frac{1}{n}]}$ . Then  $X_n \xrightarrow{\mathbb{P}} 0$  but  $X_n$  doesn't converge to 0 in  $L^p$  for all  $p \geq 1$ .

$\mathbb{P}(|X_n - 0| > \epsilon) = \frac{1}{n} \rightarrow 0$ , so  $X_n \xrightarrow{\mathbb{P}} 0$ ; however,

$$\mathbb{E}|X_n - 0| = \mathbb{E}X_n = 2^n \cdot \frac{1}{n} \rightarrow \infty.$$

Thus, it doesn't converge to 0 in  $L^1$ .

**Example 1.32** (Two Counterexamples:  $X_n \xrightarrow{w} X$  but not  $X_n \xrightarrow{\mathbb{P}} X$ ).

- Let  $X \sim \text{Bin}(1, \frac{1}{2})$ , and  $\{X_n\}$  be a sequence of RVs given by  $X_n = X$  for all  $n$ . Then  $X_n \xrightarrow{w} 1 - X$ ; however,  $X_n$  doesn't converge to  $1 - X$  in probability.

And the weak limit is unique (in distribution) because  $X \stackrel{D}{=} 1 - X$ .

- Let  $N \sim N(0, 1)$  and  $X_n = (-1)^n N$ .

It is easy to see  $X_n \xrightarrow{w} N$ , for  $X_n \sim N(0, 1)$ ,  $\forall n$ .

However,  $X_n$  doesn't converge to  $N$ , because when  $n$  is odd,  $\mathbb{P}(|X_n - N| > \epsilon) = \mathbb{P}(|N| > \frac{\epsilon}{2}) > 0$ .

## 1.5 The Moment Problem

**Theorem 1.33.** If  $X_n \xrightarrow{w} X$ , then  $\forall \beta > 0$ , the following are equivalent:

- $\mathbb{E}|X_n|^\beta < \infty$  for all  $n$ ,  $\mathbb{E}|X| < \infty$ , and  $\mathbb{E}|X_n|^\beta \rightarrow \mathbb{E}|X|^\beta$ .
- $\{|X_n|^\beta\}$  is u.i.

**Question.** Let  $X$  be a RV with DF  $F$ , and its all finite-order moments  $m_k = \mathbb{E}X^k$ . Is  $F$  the only DF with this moment sequence?

**Example 1.34** (Heyde). Assume  $X \sim N(0, 1)$ ,  $Y = e^X$ . Consider the following PDFs:

$$f_Y(x) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}(\log x)^2} \mathbf{1}_{\{x>0\}}$$

$$f_a(x) = f_Y(x)(1 + a \sin(2\pi \log x)) \quad |a| \leq 1$$

In fact,  $f_a$  is a density function with finite moments of all orders, none of which depend on the value of  $a$ . Thus, we construct a family of RVs that admit the same moment sequence.

**Theorem 1.35.** The moment problem has a unique solution, if one of the following conditions is satisfied:

a)

$$\limsup_k \frac{(m_{2k})^{1/2k}}{2k} = r < \infty$$

b) Carleman's.

$$\sum_{k=1}^{\infty} \frac{1}{(m_{2k})^{1/2k}} < \infty$$



## Exercises

**Exercise 1.5.1.** Show that the moments completely determine the probability distribution if it is concentrated on a finite interval (that is,  $\mathbb{P}(X_n \in [a, b]) = 1$ , for all  $n$ ).

*Proof.*

□

**Exercise 1.5.2.** Show that the moments  $m_{2k+1} = 0$ ,  $m_{2n} = \frac{(2n!)}{2^n n!} \sigma^{2n}$  uniquely determine the corresponding distribution.

*Proof.* Check Carleman's condition:

$$\sum_{k=1}^{\infty} \frac{1}{(m_{2k})^{1/2k}} =$$

□

**Exercise 1.5.3.** Let  $X$  and  $\{X_n\}$  be uniformly bounded RVs. Prove  $X_n \xrightarrow{w} X \iff \mathbb{E}X_n^k \rightarrow \mathbb{E}X^k$  for all  $k \geq 1$ .

*Proof.* content...

□

## 2 Law of Large Numbers and Central Limit Theorem

### 2.1 Strong Law of Large Numbers and Weak Law of Large Numbers

**Theorem 2.1** (SLLN).

- $(X_n)$  - i.i.d. RVs.  $\mathbb{E}X_i = \mu$ .  $\mathbb{E}X_i^2 < \infty$ . Then

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[L^2]{a.s.} \mu.$$

- $X_i$  are uncorrelated RVs.  $\mathbb{E}X_i = \mu$  and  $\text{Var}X_i \leq C < \infty$ . Then

$$S_n/n \xrightarrow[L^2]{a.s.} \mu.$$

- $(X_n)$  - i.i.d. RVs. Then  $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{a.s.} \mu$  for some  $\mu$  if and only if  $\mathbb{E}|X_i| < \infty$ .

*Remark.* SLLN holds  $\iff \mathbb{P}(\lim_{n \rightarrow \infty} S_n/n = \mu) = 1$ . Therefore, to show SLLN doesn't hold, we have two methods:

- When  $X_i$  are iid, show  $\mathbb{E}|X_i| = \infty$ .
- Use B-C lemma (b) to show  $\mathbb{P}(|S_n/n - \mu| > 0 \text{ i.o.}) = 1$ . Then  $\mathbb{P}(\lim_{n \rightarrow \infty} S_n/n = \mu) < 1$ .

**Theorem 2.2** (WLLN).  $\{X_n\}$  are independent RVs. Define  $S_n = \sum_{j=1}^n X_j$ , and

$$a_n = \sum_{j=1}^n \mathbb{E}(X_j \mathbf{1}_{\{|X_j| \leq n\}}).$$

Then

$$\frac{S_n - a_n}{n} \xrightarrow{\mathbb{P}} 0$$

if and only if the following conditions hold:

- $\sum_{j=1}^n \mathbb{P}(|X_j| > n) \rightarrow 0$  as  $n \rightarrow \infty$ .
- $\frac{1}{n^2} \sum_{j=1}^n \mathbb{E}(X_j^2 \mathbf{1}_{\{|X_j| \leq n\}}) \rightarrow 0$  as  $n \rightarrow \infty$ .

*Remark.* No assumptions of moments of  $X_j$ 's are made. Note that now we have two methods to show WLLN holds:

- Check these conditions.
- Show  $\mathbb{P}(|S_n/n - \mu| > \epsilon) \rightarrow 0$  for all  $\epsilon > 0$ . (i.e. check the def. of conv. in prob.)

The following corollary is an example of how to check the conditions of WLLN.

**Corollary 2.3** (Feller's WLLN). Let  $\{X_n\}$  be i.i.d. with  $\lim_{n \rightarrow \infty} n\mathbb{P}(|X| > n) = 0$ . Then

$$S_n/n - \mathbb{E}(X \mathbf{1}_{\{|X| \leq n\}}) \xrightarrow{\mathbb{P}} 0,$$

as  $n \rightarrow \infty$ .

*Proof.* Check (a) + (b).

To check (a), we notice

$$\begin{aligned} \sum_{j=1}^n \mathbb{P}(|X_j| \geq n) &= n\mathbb{P}(|X| \geq n) \quad (\text{by iid}) \\ &\longrightarrow 0 \quad (\text{it is given}) \end{aligned}$$

To check (b), we have

$$\begin{aligned}
\frac{1}{n^2} \sum_{j=1}^n \mathbb{E}(X_j^2 \mathbf{1}_{\{|X_j| \leq n\}}) &= \frac{1}{n} \mathbb{E}(X^2 \mathbf{1}_{|X| \leq n}) \quad (\text{by iid}) \\
&= \frac{1}{n} \int_{|X| \leq n} X^2 \, d\mathbb{P} \\
&= \frac{1}{n} \int_{|X| \leq n} \left( \int_0^{|X|} 2y \, dy \right) d\mathbb{P} \\
&= \frac{1}{n} \int_{|x| \leq n} \left( \int_0^{|x|} 2y \, dy \right) d\mathbb{P} \circ X^{-1} \\
&= \frac{1}{n} \int_0^n \left( \int_y^n 2y \, d\mathbb{P} \circ X^{-1} \right) dy \quad (\text{Fubini}) \\
&= \frac{1}{n} \int_0^n 2y \left( \mathbb{P}(|X| > y) - \mathbb{P}(|X| > n) \right) dy \\
&= \frac{2}{n} \int_0^n y \mathbb{P}(|X| > y) dy - n \mathbb{P}(|X| > n)
\end{aligned}$$

Let  $\tau(y) = y\mathbb{P}(|X| > y)$ . We just showed

$$\frac{1}{n^2} \sum_{j=1}^n \mathbb{E}(X_j^2 \mathbf{1}_{\{|X_j| \leq n\}}) = \frac{1}{n} \int_0^n \tau(y) dy - \tau(n).$$

Note that  $\tau(n) \rightarrow 0$  is given. So it remains to show that

$$\frac{1}{n} \int_0^n \tau(y) dy \rightarrow 0.$$

Let  $M = \sup_{y \geq 0} \tau(y)$ , and  $\epsilon_k = \sup\{\tau(y) : y > k\}$ .  
For  $0 < k < n$ ,

$$\begin{aligned}
\frac{1}{n} \int_0^n \tau(y) dy &= \frac{1}{n} \int_0^k \tau(y) dy + \frac{1}{n} \int_k^n \tau(y) dy \\
&\leq \frac{1}{n} \int_0^k M dy + \frac{1}{n} \int_k^n \epsilon_k dy \\
&= \frac{1}{n} kM + \frac{1}{n} (n - k)\epsilon_k.
\end{aligned}$$

$\implies \limsup \frac{1}{n} \int_0^n \tau(y) dy \leq \epsilon_k$ . Then, we let  $k \rightarrow \infty$ . □

**Corollary 2.4** (Khinchin's WLLN). *Let  $\{X_n\}$  be i.i.d. RVs such that  $\mathbb{E}|X| = \mu < \infty$ . Then*

$$S_n/n \xrightarrow{\mathbb{P}} \mu.$$

*Proof.* It suffices to check the condition of Feller's WLLN. We notice that

$$\begin{aligned}
x\mathbb{P}(|X| > x) &= x \cdot \mathbb{E}\mathbf{1}_{\{|X| > x\}} \\
&\leq \mathbb{E}|X| \mathbf{1}_{\{|X| > x\}} \\
&\longrightarrow 0
\end{aligned}$$

because  $X$  is integrable. Thus,  $S_n/n \xrightarrow{\mathbb{P}} \mu$ . □

## Exercises

**Exercise 2.1.1.** Let  $\{X_n\}$  be a sequence i.i.d RVs with  $\mathbb{E}X_n = 0$  and  $\mathbb{E}X_n^2 < \infty$ . Let  $(c_n)$  be a bounded sequence of real numbers and define  $Z_n = \frac{1}{n} \sum_{i=1}^n c_i X_i$ . Show that  $Z_n \xrightarrow{\mathbb{P}} 0$  as  $n \rightarrow \infty$ .

*Proof.* Directly use SLLN.  $\square$

**Exercise 2.1.2** (WLLN for a Triangular Array). Let  $\{X_{n,k}, 1 \leq k \leq n, n \geq 1\}$  be a family of RVs and  $S_n = X_{n,1} + \cdots + X_{n,n}$ . Let  $\mathbb{E}S_n = a_n$  and  $\sigma_n^2 = \text{Var}S_n$ . If  $\sigma_n^2/b_n^2 \rightarrow 0$  for some sequence  $\{b_n\}$ , then

$$\frac{S_n - a_n}{b_n} \xrightarrow{\mathbb{P}} 0.$$

*Proof.* Check the definition of  $\xrightarrow{\mathbb{P}}$ .

For all  $\epsilon > 0$ :

$$\begin{aligned} \mathbb{P}\left(\left|\frac{S_n - a_n}{b_n}\right| > \epsilon\right) &= \mathbb{P}(|S_n - a_n| > \epsilon|b_n|) \\ &\leq \frac{\text{Var}S_n}{\epsilon^2 b_n^2} \quad (\text{Chebyshev}) \\ &\longrightarrow 0 \quad \left(\frac{\sigma^2}{b_n^2} \rightarrow 0\right) \end{aligned}$$

$\square$

**Exercise 2.1.3** (Monte Carlo Integration). Let  $f$  be a measurable function on  $[0, 1]$  with  $\int_0^1 |f(x)| dx < \infty$ , and  $\{U_n\}_{n \in \mathbb{N}}$  be i.i.d. RVs uniform on  $[0, 1]$ . (a) Prove

$$I_n = \frac{f(U_1) + \cdots + f(U_n)}{n} \xrightarrow{\mathbb{P}} I = \int_0^1 f(x) dx.$$

(b) Estimate  $\mathbb{P}(|I - I_n| > a/n^{1/2}), a > 0$ .

*Proof.* (a) Check the conditions of Khinchin's WLLN.

First,  $f(U_1), \dots, f(U_n)$  are iid. And second,  $\mathbb{E}f(U) = \int_0^1 f(u) du \leq \int_0^1 |f(u)| du < \infty$ .

By Khinchin's WLLN,  $\frac{f(U_1) + \cdots + f(U_n)}{n} \xrightarrow{\mathbb{P}} \mathbb{E}f(U) = \int_0^1 f(u) du$ .

(b) Note  $\mathbb{E}I_n = I$ .

$$\begin{aligned} \mathbb{P}(|I - I_n| > \frac{a}{n^{1/2}}) &\leq \frac{\text{Var}I_n}{a^2/n} \\ &= \frac{\frac{1}{n^2} \cdot n \text{Var}f(U)}{a^2/n} \\ &= \frac{1}{a^2} \text{Var}f(U) \end{aligned}$$

$\square$

**Exercise 2.1.4.** Prove the following WLLN using the method of CFs:

Let  $(X_n)$  be i.i.d. with mean  $\mu < \infty$ , and  $S_n = \sum_{i=1}^n X_i$ . Then  $S_n/n \xrightarrow{\mathbb{P}} \mu$  as  $n \rightarrow \infty$ .

**Exercise 2.1.5** (WLLN fails). Let  $(Y_n)$  be i.i.d Cauchy RVs. Show that  $\bar{Y}_n = \sum_{i=1}^n Y_i/n$  also has the Cauchy density. Check the condition of Feller's WLLN, and explain why the WLLN does not hold.

**Exercise 2.1.6.** Let  $X_i$  be iid RVs such that  $\mathbb{E}|X_i| = \infty$ . Show that

$$\mathbb{P}(|X_n| > n \text{ i.o.}) = 1$$

and

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \bar{X} \in (-\infty, +\infty)\right) = 0.$$

**Exercise 2.1.7.** Let  $(X_n)$  be a sequence of independent RVs such that  $\mathbb{E}X_n = 0$  and  $\mathbb{E}X_n^2 = Cn^\alpha$  for some  $C > 0$ ,  $\alpha \geq 0$ , and  $n = 1, 2, \dots$ . Describe the set of values of  $\alpha$  for which the WLLN holds.

**-Examples: SLLN doesn't hold-**

**Exercise 2.1.8.** Let  $1 - F(x) = \frac{e}{2x \log x} \mathbf{1}_{\{x \geq e\}}$ . And  $X_n \stackrel{iid}{\sim} F$ . Show that WLLN holds but SLLN fails.

*Proof.* To show SLLN doesn't hold, it suffices to show  $\mathbb{E}|X_i| = \infty$  (Just compute its integral). Then use the result of Exercise 2.1.6.

And to prove WLLN holds, we use Feller's WLLN: we only need to check

$$x\mathbb{P}(|X| > x) = x \cdot 2(1 - F(x)) \rightarrow 0;$$

therefore,  $S_n/n \xrightarrow{\mathbb{P}} 0$ . □

**Exercise 2.1.9.** Let  $X_i$  be a sequence of indep. RVs.  $\mathbb{P}(X_n = n) = \mathbb{P}(X_n = -n) = \frac{1}{2}[n \log(n+2)]^{-1}$  and  $\mathbb{P}(X_n = 0) = 1 - [n \log(n+2)]^{-1}$ .

- (a) Show that  $\{X_n\}$  is tight.
- (b) Show that SLLN doesn't hold.
- (c) Show that WLLN holds.

*Proof.* Omitted. See Exercise 2.1.11. □

**-Trancation technique-**

**Exercise 2.1.10.** Let  $X_n, X \in L^p$  be two RVs with  $X_n \xrightarrow{a.s.} X$  and with  $\|X_n\|_p \rightarrow \|X\|_p$ . Show that

$$X_n \xrightarrow{L^p} X.$$

*Proof.* **Trancation technique.** Define

$$X_n^* = \begin{cases} X_n & |X_n| \leq |X|, \\ |X| \cdot \text{sgn} X_n & |X_n| > |X|. \end{cases}$$

Notice that

$$\|X_n - X\|_p \leq \|X_n - X_n^*\|_p + \|X - X_n^*\|_p.$$

To show  $X_n \xrightarrow{L^p} X$ , it suffices to prove that

$$\|X_n - X_n^*\|_p \rightarrow 0 \quad \text{and} \quad \|X - X_n^*\|_p \rightarrow 0$$

as  $n \rightarrow \infty$ .

- We have  $|X - X_n^*| \leq |X| + |X_n^*| \leq 2|X|$ . Thus,  $|X - X_n^*|$  is dominated by an integrable RV  $2|X|$ . By the dominated convergence theorem

$$\lim \mathbb{E}|X - X_n^*|^p = \mathbb{E} \lim |X - X_n^*|^p = 0;$$

that is  $\|X - X_n^*\|_p \rightarrow 0$ .

- Moreover, we can get  $\lim \mathbb{E}|X_n^*|^p = \mathbb{E} \lim |X_n^*|^p = \mathbb{E}|X|^p$ , by using the dominated convergence theorem and noticing that  $X_n^* \rightarrow X$ .
- We have  $|X_n - X_n^*| = |X_n| - |X_n^*|$ , by noticing the definition of  $X_n^*$  and using  $X_n = |X_n| \cdot \text{sgn} X_n$ . So we can get

$$|X_n - X_n^*|^p \leq |X_n|^p - |X_n^*|^p$$

by the convexity of  $x \mapsto x^p$ . Then

$$\mathbb{E}|X_n - X_n^*|^p \leq \mathbb{E}|X_n|^p - \mathbb{E}|X_n^*|^p \rightarrow \mathbb{E}|X|^p - \mathbb{E}|X|^p = 0;$$

that is  $\|X_n^* - X_n\| \rightarrow 0$ .

□

### -Tightness and SLLN-

**Exercise 2.1.11.** (a) If  $\{X_n\}$  is a tight sequence of independent RVs with its  $r$ -th moments ( $r > 1$ ) uniformly bounded, then  $\{X_n\}$  satisfies the SLLN.

(b) Let  $\{X_n\}$  be a sequence of independent RVs such that

$$X_n = \begin{cases} n & w.p. \frac{1}{2}[n \log(n+2)]^{-1} \\ 0 & w.p. 1 - [n \log(n+2)]^{-1} \\ -n & w.p. \frac{1}{2}[n \log(n+2)]^{-1} \end{cases}.$$

Show that  $\{X_n\}$  is tight.

(d) Show that the first moments of  $\{X_n\}$  are uniformly bounded.

(c) Show that SLLN doesn't hold for  $\{X_n\}$ .

## 2.2 Central Limit Theorem (CLT)

**Theorem 2.5** (iid sequence). Let  $X_i$  be iid RVs with  $\mathbb{E}X_i = \mu$  and  $\text{Var}X_i = \sigma^2 \in (0, \infty)$ . Then

$$n^{1/2}(\bar{X} - \mu)/\sigma \xrightarrow{w} Z$$

where  $Z \sim N(0, 1)$ .

**Theorem 2.6** (Lindeberg-Feller). For each  $n$ , let  $X_{n,m}$ ,  $1 \leq m \leq n$ , be independent RVs with  $\mathbb{E}X_{n,m} = 0$ . Suppose

(i)  $\sum_{m=1}^n \mathbb{E}X_{n,m}^2 \rightarrow \sigma^2 > 0$ ; and

(ii) For all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \sum_{m=1}^n \mathbb{E}(X_{n,m}^2 \mathbf{1}_{\{|X_{n,m}| > \epsilon\}}) = 0.$$

Then  $S_n = X_{n,1} + \cdots + X_{n,n} \xrightarrow{w} \sigma Z$ .

*Remark.* Lindeberg-Feller conditions are not necessary. For example, let  $X_k \sim N(0, 2^{k-2})$  for  $k \geq 2$  and let  $X_1 \sim N(0, 1)$ .

## Exercises

**Exercise 2.2.1.** Let  $\{X_n\}$  be a sequence i.i.d RVs with  $\mathbb{E}X_n = 0$  and  $\mathbb{E}X_n^2 < \infty$ . Let  $S_n = \sum_{i=1}^n X_i$ . Show that  $S_n^2$ , appropriately normalized, converges in distribution as  $n \rightarrow \infty$ .

*Proof.* First, use CLT to show  $S_n/n \xrightarrow{w} N(0, \sigma^2/n)$ . Then use CMT to get its limit, chi-square distribution. □

## 2.3 Delta's Method

**Theorem 2.7.** Let  $g$  be a smooth function which has non-zero derivative at  $\mu$ . If  $X_n \xrightarrow{w} N(\mu, \sigma_n^2)$  with  $\sigma_n^2 \rightarrow 0$ , then

$$g(X_n) \xrightarrow{w} N(g(\mu), [g'(\mu)]^2 \sigma_n^2).$$

### 3 Discrete-Time Martingales

#### 3.1 Conditional Expectation

**Definition 3.1.** Let  $X$  be a RV on  $(\Omega, \mathcal{F}, \mathbb{P})$  s.t.  $\mathbb{E}|X| < \infty$ ,  $\mathcal{F}_1 \subset \mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathcal{F}$ .

Then  $Y = \mathbb{E}(X|\mathcal{F}_1)$  (the conditional expectation of  $X$  given  $\mathcal{F}_1$ ) is a RV s.t.

- $Y \in \mathcal{F}_1$  and  $\mathbb{E}|Y| < \infty$
- $\forall A \in \mathcal{F}_1, \int_A X \, d\mathbb{P} = \int_A Y \, d\mathbb{P}$  (i.e.  $\mathbb{E}(X1_A) = \mathbb{E}(Y1_A)$ ).

*Remark.* In  $L^2$  space, the conditional expectation can be considered as the orthogonal projection in m.s. sense; that is

$$\|X - \mathbb{E}(X|\mathcal{F}_1)\| = \min_{Z \in \mathcal{F}_1} \|X - Z\|.$$

**Proposition 3.2.**

- **Linearity.**  $\mathbb{E}(\alpha X + \beta Y|\mathcal{F}_1) = \alpha \mathbb{E}(X|\mathcal{F}_1) + \beta \mathbb{E}(Y|\mathcal{F}_1)$
- $X$  is independent of  $\mathcal{F}_1 \implies \mathbb{E}(X|\mathcal{F}_1) = \mathbb{E}X$ .
- $X \in \mathcal{F}_1 \implies \mathbb{E}(XY|\mathcal{F}_1) = X \mathbb{E}(Y|\mathcal{F}_1)$ .
- $\mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)) = \mathbb{E}(X)$ .
- $\sigma$ -algebra  $\mathcal{F}_2 \subset \mathcal{F}_1 \subset \mathcal{F} \implies \mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2) = \mathbb{E}(X|\mathcal{F}_2)$ .
- **Monotone Convergence Theorem.** If  $X_n \geq 0$  and  $X_n \uparrow X$  almost surely, then

$$\mathbb{E}(X_n|\mathcal{F}_1) \uparrow \mathbb{E}(X|\mathcal{F}_1), \text{ a.s.}$$

- **Dominated Convergence Theorem.** If  $X_n \xrightarrow{\text{a.s.}} X$  and  $|X_n| \leq Y$  with  $\mathbb{E}Y < \infty$ , then

$$\mathbb{E}(X_n|\mathcal{F}_1) \xrightarrow{\text{a.s.}} \mathbb{E}(X|\mathcal{F}_1).$$

- **Jensen's Inequality.** Let  $g$  be convex. Then

$$\mathbb{E}(g(X)|\mathcal{F}_1) \geq g(\mathbb{E}(X|\mathcal{F}_1)).$$

#### Exercises

**Exercise 3.1.1.** Let  $X$  be an integrable RV on  $(\Omega, \mathcal{F}, \mathbb{P})$ , and  $H \in \mathcal{F}$ .

- Let  $\mathcal{F}_1 = \{\emptyset, \Omega\}$ . Find all RVs which are measurable w.r.t.  $\mathcal{F}_1$ , and find  $\mathbb{E}(X|\mathcal{F}_1)$ .
- Let  $\mathcal{F}_H = \{\emptyset, H, H^c, \Omega\}$ . Find  $\mathbb{E}(X|\mathcal{F}_H)$ .

*Proof.* (a) **First**, we prove all constant maps are measurable w.r.t  $\mathcal{F}_1$ .

For all  $c \in \mathbb{R}$ , we define  $c : \Omega \rightarrow \mathbb{R}$  by  $\omega \mapsto c$  (constant maps). Then

$$\{c \leq x\} = \begin{cases} \Omega & x \geq c; \\ \emptyset & x < c. \end{cases}$$

$$\implies c \in \mathcal{F}_1.$$

**Then**, we prove if  $f \in \mathcal{F}_1$ , then  $f$  is a constant map.

Assume  $f$  is not a constant map; that is  $\exists \omega_1, \omega_2$  s.t.  $f(\omega_1) \neq f(\omega_2)$ .

Let  $c_1 = f(\omega_1)$  and  $c_2 = f(\omega_2)$ . WLOG,  $c_1 < c_2$ . Consider the following set

$$A = \{f < c_2\}.$$

Notice that  $\omega_1 \in A$  and  $\omega_2 \notin A$ , which implies  $A \notin \mathcal{F}_1$ ; that is,  $f$  is not measurable. Contradiction.

**Finally**, to compute  $\mathbb{E}(X|\mathcal{F}_1)$ , it suffices to notice that  $\mathbb{E}(X|\mathcal{F}_1) \in \mathcal{F}_1$ . So it is a constant map.

Let  $c = \mathbb{E}(X|\mathcal{F}_1)$ . Take expectation on both sides. We get  $\mathbb{E}(X|\mathcal{F}_1) = \mathbb{E}(X)$ .

(b) Define  $\mathbb{E}(X | H) = \frac{1}{\mathbb{P}(H)} \int_H X \, d\mathbb{P}$ . We will show  $\mathbb{E}(X | \mathcal{F}_H) = \mathbb{E}(X | H)$ .

Directly check the definition: For all  $A \in \mathcal{F}_H$ ,

$$\mathbb{E}(\mathbb{E}(X | H) \cdot \mathbf{1}_A) = \mathbb{E}(X \mathbf{1}_A).$$

Therefore,  $\mathbb{E}(X | \mathcal{F}_H) = \mathbb{E}(X | H)$ . □

**Exercise 3.1.2.** Let  $X_1, X_2$  and  $Y$  be RVs with zero means and finite variances. Let  $\mathbb{E}X_1X_2 = 0$ . Find the best mean square linear approximation of  $Y$  in terms of  $X_1$  and  $X_2$ .

*Remark.* In this exercise, we will see that  $\text{span}\{X_1, X_2\}$  is different from  $L^2(\Omega, \mathcal{F}_1, \mathbb{P})$  where  $\mathcal{F}_1$  is the sigma algebra generated by  $X_1$  and  $X_2$ .

*Proof.* Let  $\hat{Y} = \alpha_1 X_1 + \alpha_2 X_2$ . We want to minimize:

$$\begin{aligned} \mathbb{E}(Y - \hat{Y})^2 &= \mathbb{E}Y^2 - 2\mathbb{E}Y\hat{Y} + \mathbb{E}\hat{Y}^2 \\ &= \mathbb{E}Y^2 - 2\alpha_1\mathbb{E}X_1Y - 2\alpha_2\mathbb{E}X_2Y + \alpha_1^2\mathbb{E}X_1^2 + \alpha_2^2\mathbb{E}X_2^2 \end{aligned}$$

Let  $\frac{\partial}{\partial \alpha_i} \mathbb{E}(Y - \hat{Y})^2 = 0$  for  $i = 1, 2$ . We get:

$$\begin{cases} \alpha_1^* = \frac{\mathbb{E}X_1Y}{\mathbb{E}X_1^2} \\ \alpha_2^* = \frac{\mathbb{E}X_2Y}{\mathbb{E}X_2^2} \end{cases}$$

Because  $J = \begin{pmatrix} 2\mathbb{E}X_1^2 & 0 \\ 0 & 2\mathbb{E}X_2^2 \end{pmatrix}$  is positive definite, the loss is a convex function of  $\alpha_1, \alpha_2$ . Therefore, it is minimized by  $(\alpha_1^*, \alpha_2^*)$ . □

### -Monotone class theorem-

#### Definition 3.3.

- The collection  $\mathcal{C}$  is called a  $\pi$ -system (or  $\pi$ -system) if

$$A, B \in \mathcal{C} \implies A \cap B \in \mathcal{C}.$$

- The collection  $\mathcal{D}$  is called a  $\lambda$ -system (or  $\lambda$ -system) if

- $\Omega \in \mathcal{D}$ .
- $A, B \in \mathcal{D}$  and  $B \subset A \implies A \setminus B \in \mathcal{D}$ .
- $(A_n) \subset \mathcal{D}$  and  $A_n \uparrow A \implies A \in \mathcal{D}$ .

- The collection  $\mathcal{M}$  is called a monotone class if

- $1 \in \mathcal{M}$ .
- $f, g \in \mathcal{M}_b$  and  $a, b \in \mathbb{R} \implies af + bg \in \mathcal{M}$ .
- $(f_n) \subset \mathcal{M}$  and  $f_n \uparrow f \implies f \in \mathcal{M}$ .

#### Theorem.

- If  $\mathcal{P}$  is a  $\pi$ -system contained in a  $\lambda$ -system  $\mathcal{L}$ , then  $\sigma(\mathcal{P}) \subset \mathcal{L}$ .
- Let  $\mathcal{E}$  be a  $\sigma$ -algebra generated by the  $\pi$ -system  $\mathcal{P}$  and  $\mathcal{M}$  is a monotone class. If  $\mathbf{1}_A \in \mathcal{M}$  for all  $A \in \mathcal{P}$ , then  $\mathcal{M}$  contains all positive  $\mathcal{E}$ -measurable functions and all bounded  $\mathcal{E}$ -measurable functions.

**Exercise 3.1.3.** Let  $X_i$  be a sequence of indep. RVs and  $M_0 = 1$ ,  $M_n = \prod_{i=1}^n X_i$ . Let  $\mathcal{D}_n = \sigma\{M_1, \dots, M_n\}$  and  $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$ . Show that in general  $\mathcal{D}_n \subset \mathcal{F}_n$ .

**Exercise 3.1.4.** Let  $Y$  be a RV with  $\mathbb{E}|Y| < \infty$ . Let  $X_1$  and  $X_2$  be RVs such that  $X_2$  is indep. of  $Y$  and indep. of  $X_1$ . Prove that

$$\mathbb{E}(Y|X_1, X_2) = \mathbb{E}(Y|X_1).$$



### 3.2 Basic Properties

**Definition 3.4.**  $X$  is  $L^1$ -bounded ( $\mathbb{E}|X_n| < \infty \forall n$ ) and adapted to  $\mathcal{F}$ .

$(X, \mathcal{F})$  is called a martingale if

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n.$$

... a submartingale if

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n.$$

... a supermartingale if

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n.$$

**Proposition 3.5.**  $(X, \mathcal{F})$  is a martingale.

- $\mathbb{E}X_n = \mathbb{E}X_m = \mathbb{E}X_0 \quad \forall m, n$
- $\mathbb{E}(X_n|\mathcal{F}_m) = X_m \quad \forall m \leq n$
- $\mathbb{E}((X_n - X_m)X_l) = 0 \quad \forall l \leq m \leq n$
- $\mathbb{E}((X_n - X_m)^2|\mathcal{F}_m) = \mathbb{E}(X_n^2|\mathcal{F}_m) - X_m^2 \quad \forall m \leq n$
- $\forall n \mathbb{E}X_n^2 \leq K \implies X \text{ has a m.s. limit.}$
- $\varphi \text{ is convex s.t. } \mathbb{E}|\varphi(X_n)| < \infty. \implies (\varphi(X_n), \mathcal{F}) \text{ is a submartingale.}$

**Proposition 3.6.**

a) Let  $(X, \mathcal{F})$  be a submartingale, and  $(H, \mathcal{F})$  be a bounded predictable process with  $H_n \geq 0$ . Define

$$(H \cdot X)_n = \sum_{m=1}^n H_m (X_m - X_{m-1}).$$

Then  $((H \cdot X), \mathcal{F})$  is a submartingale.

b) Let  $(X, \mathcal{F})$  be a submartingale, and  $\varphi$  be a non-decreasing convex function such that  $\mathbb{E}|\varphi(X_n)| < \infty$  for all  $n$ . Then  $(\varphi(X), \mathcal{F})$  is a submartingale.

c) Let  $(X, \mathcal{F})$  be a martingale, and  $\varphi$  be a convex function such that  $\mathbb{E}|\varphi(X_n)| < \infty$  for all  $n$ . Then  $(\varphi(X), \mathcal{F})$  is a submartingale. In particular,  $X_n^2, |X_n|, X_n^+, X_n \vee a$  all are submartingales.

#### Examples of Martingales

**Example 3.7** (Sums of independent RVs with mean 0). Let  $\{X_i\}$  be indep.  $\mathbb{E}X_i = 0 \forall i$ .  $\mathbb{E}|X_i| < \infty$ .

Let  $S_0 = 0, S_n = X_1 + X_2 + \dots + X_n$ .  $\mathcal{F}$  = the filtration generated by  $X$ .

$\implies (S, \mathcal{F})$  = martingale.

**Example 3.8** (Products of nonneg. indep. RVs with mean 1). Let  $\{X_i\}$  be indep.  $\mathbb{E}X_i \geq 0 \forall i$ .  $\mathbb{E}X_i = 1$ .

Let  $M_0 = 1, M_n = X_1 X_2 \dots X_n$ .  $\mathcal{F}$  = the filtration generated by  $X$ .

$\implies (M, \mathcal{F})$  = martingale.

**Example 3.9.** Let  $\xi$  be a RV with  $\mathbb{E}|\xi| < \infty$ ,  $\mathcal{F}$  be a filtration.

Let  $M_n = \mathbb{E}(\xi|\mathcal{F}_n)$ .

$\implies (M, \mathcal{F})$  = martingale.

## Exercises

**Exercise 3.2.1.** Let  $X$  and  $Y$  be martingales w.r.t.  $\mathcal{F}$ . And  $T$  is a stopping time w.r.t.  $\mathcal{F}$ . And assume  $X_T = Y_T$  on  $\{T < \infty\}$ . Define

$$Z_n = X_n \mathbf{1}_{\{n < T\}} + Y_n \mathbf{1}_{\{n \geq T\}}.$$

Show  $(Z, \mathcal{F})$  is a martingale.

*Proof.* Notice that

$$\begin{aligned} Z_n &= X_n \mathbf{1}_{\{n < T\}} + Y_n \mathbf{1}_{\{n \geq T\}} \\ (X, Y \text{ are mart.}) &= \mathbb{E}(X_{n+1} | \mathcal{F}_n) \mathbf{1}_{\{n < T\}} + \mathbb{E}(Y_{n+1} | \mathcal{F}_n) \mathbf{1}_{\{n \geq T\}} \\ (T \text{ is st. time.}) &= \mathbb{E}(\underbrace{X_{n+1} \mathbf{1}_{\{n < T\}} + Y_{n+1} \mathbf{1}_{\{n \geq T\}}}_{X_{n+1} \mathbf{1}_{\{n < T\}} + Y_{n+1} \mathbf{1}_{\{n \geq T\}}} | \mathcal{F}_n) \end{aligned}$$

And because

$$\begin{aligned} Z_{n+1} - (\underbrace{X_{n+1} \mathbf{1}_{\{n < T\}} + Y_{n+1} \mathbf{1}_{\{n \geq T\}}}_{X_{n+1} \mathbf{1}_{\{n < T\}} + Y_{n+1} \mathbf{1}_{\{n \geq T\}}}) &= X_{n+1} \mathbf{1}_{\{T = n+1\}} - Y_{n+1} \mathbf{1}_{\{T = n+1\}} \\ (\text{b/c } X_T = Y_T) &= 0 \end{aligned}$$

$$\implies Z_n = \mathbb{E}(Z_{n+1} | \mathcal{F}_n).$$

□

## 3.3 Doob's Decomposition

**Definition 3.10.**

- **Predictable Processes.**  $X_0 \in \mathcal{F}_0$ ,  $X_n \in \mathcal{F}_{n-1}$  for  $n \geq 1$ .
- **Increasing Processes.** predictable +  $X_0 = 0$ ,  $\mathbb{P}(X_n \leq X_{n+1}) = 1$ .

**Theorem 3.11. Doob's Decomposition**

- Let  $(X, \mathcal{F})$  be an adapted process with  $\mathbb{E}|X_n| < \infty \forall n$ .  
 $\implies (X, \mathcal{F})$  has a unique Doob's decomposition:

$$X_n = X_0 + M_n + A_n$$

where  $(M, \mathcal{F}) = \text{martingale}$ ;  $(A, \mathcal{F}) = \text{predictable}$ .

- Let  $(X, \mathcal{F})$  be a submartingale.  
 $\iff (A, \mathcal{F}) = \text{increasing}.$

*Proof.* The constructions of  $M$  and  $A$ :

$$\begin{aligned} A_0 &= 0; \quad A_n = A_{n-1} + \mathbb{E}(X_n - X_{n-1} | \mathcal{F}_{n-1}) \\ M_0 &= 0; \quad M_n = M_{n-1} + (X_n - \mathbb{E}(X_n | \mathcal{F}_{n-1})) \end{aligned}$$

□

## Exercises

**Exercise 3.3.1.** Let  $(X, \mathcal{F})$  be a  $L^1$ -bounded submartingale. Its Doob's decomposition is given by

$$X = X_0 + M + A.$$

Prove (a)  $M$  and  $A$  are both convergent, and (b) their limits  $M_\infty$  and  $A_\infty$  are both integrable.

*Proof.* First, notice that  $A_n = \sum_{i=1}^n \mathbb{E}(X_i - X_{i-1} | \mathcal{F}_{i-1}) \geq 0$ .  $\mathbb{E}|A_n| = \mathbb{E}A_n = \mathbb{E}X_n - \mathbb{E}X_0 < \infty$ . Define  $A_\infty(\omega) = \lim_{n \rightarrow \infty} A_n(\omega)$  (thus,  $A_n \uparrow A_\infty$  almost surely).

By the monotone convergence theorem,  $A_\infty$  is  $\mathcal{F}$ -measurable and  $\mathbb{E}A_\infty = \lim_{n \rightarrow \infty} \mathbb{E}A_n < \infty$ .

Then, notice that  $M_n = X_n - X_0 - A_n$ . We have  $M_\infty = X_\infty - X_0 - A_\infty$ . So

$$\mathbb{E}|M_\infty| \leq \mathbb{E}|X_\infty| + \mathbb{E}|X_0| + \mathbb{E}|A_\infty| < \infty.$$

□

**Exercise 3.3.2** (Another decomposition of submartingales). Let  $(X, \mathcal{F})$  be a  $L^1$ -bounded submartingale and  $X_\infty = \lim_{n \rightarrow \infty} X_n$ . Then

$$X = M + V$$

where  $M$  is a u.i. martingale and  $V$  is a submartingale with  $\lim_{n \rightarrow \infty} V = 0$ . Show this by the steps below:

- a) Define  $M_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ .  $M$  is a u.i. martingale.
- b) Define  $V_n = X_n - M_n$ .  $V$  is a submartingale with  $\lim_{n \rightarrow \infty} V = 0$  almost surely.
- c) The decomposition is unique.

*Proof.*

- a) Know:  $(X, \mathcal{F})$  is a  $L^1$ -bounded submartingale.  
 $\implies X_\infty$  is integrable. Theorem 3.12.  
 $\implies M$  is a u.i. martingale. Theorem 3.14.
- b)  $\mathbb{E}(V_{n+1} | \mathcal{F}_n) = \mathbb{E}(X_{n+1} | \mathcal{F}_n) - \mathbb{E}(\mathbb{E}(X_\infty | \mathcal{F}_{n+1}) | \mathcal{F}_n) \geq X_n - M_n = V_n$ .  
 $\implies$  submartingale.

Take  $\lim_{n \rightarrow \infty}$  on both sides. We get  $\lim V_n = \lim X_n - \lim M_n = X_\infty - \mathbb{E}(X_\infty | \mathcal{F}_\infty) = 0$ .

- c) Assume  $X = M + V = \tilde{M} + \tilde{V}$ . Letting  $n \rightarrow \infty$ , we get

$$M_\infty = \tilde{M}_\infty =: Z.$$

Because  $M$  and  $\tilde{M}$  are u.i. martingale, they can be written as  $M_n = \tilde{M}_n = \mathbb{E}(Z | \mathcal{F}_n)$ . Then

$$V = X - M = X - \tilde{M} = \tilde{V}$$

holds almost surely.

□

### -Riesz decomposition theorem for supermartingales-

**Exercise 3.3.3** (Potentials). Let  $X$  be a positive supermartingale.  $X$  is called a potential if  $\lim_{n \rightarrow \infty} X_n = 0$  almost surely. Show that  $X$  is a potential if  $\lim_{n \rightarrow \infty} \mathbb{E}X_n = 0$ .

*Remark.* For submartingales and supermartingales,  $X_n \xrightarrow{a.s.} 0$  is implied by  $X_n \xrightarrow{L^1} 0$ .

*Proof.* Know:  $X$  be a positive supermartingale.

$\implies X_\infty$  exists and is non-negative.  $\implies \mathbb{E}X_\infty = 0$ , since  $\lim_{n \rightarrow \infty} \mathbb{E}X_n = 0$ .

$\implies X_\infty = 0$  almost surely.

□

**Exercise 3.3.4** (Decomposition of supermartingales). Let  $X$  be a supermartingale with  $\sup \mathbb{E}X_n^- < \infty$ . Show that there is a unique decomposition

$$X = M + V$$

where  $M$  is a u.i. martingale and  $V$  is a potential.

*Proof.* By Exercise 3.3.2,

$$-X = \tilde{M} + \tilde{V},$$

where  $\tilde{M}$  is a u.i. martingale and  $\tilde{V}$  is a submartingale with  $\lim_{n \rightarrow \infty} \tilde{V}_n = 0$ .

Let  $M = -\tilde{M}$  and  $V = -\tilde{V}$ . We are done.  $\square$

**Exercise 3.3.5** (Riesz decomposition). Every positive supermartingale  $X$  has the following decomposition

$$X = Y + Z$$

where  $Y$  is a positive martingale and  $Z$  is a potential. Show this by the steps below:

- a) The limit  $Y_m = \lim_{n \rightarrow \infty} \mathbb{E}(X_{m+n} | \mathcal{F}_m)$  exists.
- b)  $Y = (Y_m)$  is a positive martingale.
- c)  $Z = X - Y$  is a positive supermartingale.
- d)  $Z$  is a potential.

*Remark.* Recall that for any supermartingale  $X$ , we have a natural monotone sequence  $\{\mathbb{E}X_n\}$ .

*Proof.*

- a) It suffices to prove for every  $m$ ,  $\{\mathbb{E}(X_{m+n} | \mathcal{F}_m)\}_{n \geq 0}$  is a positive supermartingale.

$X$  is positive  $\implies \{\mathbb{E}(X_{m+n} | \mathcal{F}_m)\}_{n \in \mathbb{N}}$  is positive.

Fix  $m$ . We only need to consider  $m < n$ :

$$\begin{aligned} \mathbb{E}(\mathbb{E}(X_{m+n+1} | \mathcal{F}_m) | \mathcal{F}_n) &= \mathbb{E}(X_{m+n+1} | \mathcal{F}_m) \\ (X \text{ is supermart.}) &\leq \mathbb{E}(X_{m+n} | \mathcal{F}_m) \end{aligned}$$

Thus,  $\{\mathbb{E}(X_{m+n} | \mathcal{F}_m)\}_{n \geq 0}$  is a positive supermartingale. Its a.s. limit is defined as  $Y_m$ .

- b)  $\{\mathbb{E}(X_{m+n} | \mathcal{F}_m)\}_{n \in \mathbb{N}}$  is positive  $\implies Y$  is positive.

Notice that

$$\begin{aligned} \mathbb{E}(Y_{m+1} | \mathcal{F}_m) &= \mathbb{E}\left(\lim_{n \rightarrow \infty} \mathbb{E}(X_{m+n+1} | \mathcal{F}_{m+1}) | \mathcal{F}_m\right) \\ (\text{Monotone conv. thm.}) &= \lim_{n \rightarrow \infty} \mathbb{E}(X_{m+n+1} | \mathcal{F}_m) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}(X_{m+n} | \mathcal{F}_m) \\ &= Y_m \end{aligned}$$

Therefore,  $Y$  is a positive martingale.

- c)  $\mathbb{E}(Z_{n+1} | \mathcal{F}_n) = \mathbb{E}(X_{n+1} | \mathcal{F}_n) - \mathbb{E}(Y_{n+1} | \mathcal{F}_n) \leq X_n - Y_n = Z_n$ .

- d) Because

$$\lim_{n \rightarrow \infty} \mathbb{E}Z_n = \lim_{n \rightarrow \infty} \mathbb{E}X_n - \lim_{n \rightarrow \infty} \mathbb{E}Y_n = 0,$$

$Z$  is a potential by Exercise 3.3.3.  $\square$

**Exercise 3.3.6.** The martingale  $Y$  in the preceding decomposition is the maximal submartingale majorized by  $X$ .

*Proof.* Let  $W$  be a submartingale s.t.  $W_n \leq X_n$  for every  $n$ . We will prove  $W_n \leq Y_n$ .

$$W_n \leq \mathbb{E}(W_\infty | \mathcal{F}_n) \leq \mathbb{E}(Y_\infty | \mathcal{F}_n) + \mathbb{E}(Z_\infty | \mathcal{F}_n) = Y_n.$$

$\square$

### -Krickeberg decomposition-

**Exercise 3.3.7** (Krickeberg decomposition). Let  $(X, \mathcal{F})$  be a  $L^1$ -bounded martingale. Then

$$X = Y - Z$$

where  $Y$  and  $Z$  are positive and  $L^1$ -bounded martingales. Show this by the steps below:

- a)  $Y_n = \lim_m \mathbb{E}(X_{n+m}^+ | \mathcal{F}_n)$  and  $Z_n = \lim_m \mathbb{E}(X_{n+m}^- | \mathcal{F}_n)$  exist.
- b)  $Y = (Y_n)$  and  $Z = (Z_n)$  are both positive and  $L^1$ -bounded martingales.

*Proof.* Omitted. See Exercise 3.3.5. It is similar.  $\square$

**Exercise 3.3.8.** A martingale is  $L^1$ -bounded  $\iff$  it is the difference of two positive  $L^1$ -bounded martingales.

*Proof.*

$\implies$  : By Krickeberg decomposition.

$\impliedby$  :  $\mathbb{E}|X_n| = \mathbb{E}|Y_n - Z_n| \leq \mathbb{E}|Y_n| + \mathbb{E}|Z_n| < \infty$ .  $\square$

**Exercise 3.3.9.** In the Krickeberg decomposition of an  $L^1$ -bounded martingale  $X$ , the process  $Y$  is the minimal positive martingale majorizing  $X$ , and the process  $Z$  is the minimal positive martingale majorizing  $-X$ .

## 3.4 Martingale Convergence

### a.s. convergence

**Theorem 3.12.**  $(X, \mathcal{F}) = \text{submartingale}$ .  $\sup_n \mathbb{E}X_n^+ < \infty$ .

- $\exists X_\infty$  s.t.  $X_n \xrightarrow{a.s.} X_\infty$
- $\mathbb{E}|X_\infty| < \infty$

*Remark.*

- For a submartingale,  $\mathbb{E}X_n^+$  is increasing in  $n$ .
- Quick check:  
negative submartingales, positive supermartingales, martingales bounded by an integrable RV.
- Because  $\mathbb{E}X_n^+ \leq \mathbb{E}|X_n| = 2\mathbb{E}X_n^+ - \mathbb{E}X_n \leq 2\mathbb{E}X_n^+ - \mathbb{E}X_0$ ,  $X^+$  is  $L^1$ -bounded is equivalent to that  $X$  is  $L^1$ -bounded.

**Example 3.13** (a.s. convergence  $\not\Rightarrow L^1$ -convergence). **Double or Nothing.**

Let  $X_0 = 1$ .  $\xi_1 = \mathbf{1}\{\text{wins on step } i\} - \mathbf{1}\{\text{lost at step } i\}$ , and  $S_n = \sum_{i=1}^n \xi_i$ . Define

$$H_n = \begin{cases} 2^{n-1} & S_{n-1} = n-1 \\ 0 & \text{o.w.} \end{cases}$$

and define

$$X_n = 1 + \sum_{m=1}^n H_m(S_m - S_{m-1}).$$

As we can see,  $(X_n)$  converges to 0 almost surely, but  $\mathbb{E}X_n = 1 \not\rightarrow 0$ .

## $L^1$ convergence and u.i. (sub)martingales

**Theorem 3.14.**  $(X, \mathcal{F}) = \text{submartingale}$ . The following are equivalent:

- a)  $X = \text{u.i.}$
- b)  $X$  converges in  $L^1$ .

If  $(X, \mathcal{F}) = \text{martingale}$ , then they are also equivalent to

- c)  $\exists X_\infty$  integrable RV s.t.  $X_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ .

*Remark.* For submartingales, we always have  $\xrightarrow{L^1} \implies \xrightarrow{a.s.}$ , because  $\xrightarrow{L^1}$  implies  $L^1$  bounded.

*Proof.*

- (a)  $\Rightarrow$  (b):

$$\text{u.i.} \implies \sup_n \mathbb{E}X_n^+ < \infty \implies X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{\mathbb{P}} X. \text{ And under } \xrightarrow{\mathbb{P}}:$$

$$\text{u.i.} \iff \xrightarrow{L^1}.$$

- (a)  $\Leftarrow$  (b):

$$X_n \xrightarrow{L^1} X \implies X_n \xrightarrow{\mathbb{P}} X. \text{ Use "u.i.} \iff \xrightarrow{L^1}" \text{ again.}$$

- (b)  $\Rightarrow$  (c): Note that for now  $X$  is a martingale.

Let  $X = \lim_{n \rightarrow \infty} X_n$ . Use the continuity of conditional expectation in  $L^1$  norm.

- (c)  $\Rightarrow$  (a):

Let  $X_{\mathcal{G}} = \mathbb{E}(X | \mathcal{G})$  where  $\mathcal{G}$  is a sub- $\sigma$ -field of  $\mathcal{F}$ . Then

$$\sup_{\mathcal{G}} \mathbb{E}(|X_{\mathcal{G}}| \mathbf{1}_{\{|X_{\mathcal{G}}| > b\}}) \rightarrow 0$$

as  $b \rightarrow \infty$ . It is because  $\{X_{\mathcal{G}}\}$  is  $L^1$ -bounded by  $\mathbb{E}|X|$ .

□

## Exercises

**Exercise 3.4.1** (Hunt's dominated convergence theorem). Let  $(X_n)$  be dominated by  $Z$ , an integrable RV, and  $X_\infty = \lim_n X_n$  exists.

$$\implies \mathbb{E}(X_n | \mathcal{F}_n) \xrightarrow[L^1]{a.s.} \mathbb{E}(X_\infty | \mathcal{F}_\infty)$$

*Remark.* No assumption of adaptedness for  $(X_n)$ .

*Proof.* Notice that

$$|\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_\infty)| \leq |\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| + |\mathbb{E}(X_\infty | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_\infty)|.$$

Thus, it suffices to prove

$$(a) \quad |\mathbb{E}(X_\infty | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_\infty)| \xrightarrow[L^1]{a.s.} 0$$

$$(b) \quad |\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| \xrightarrow[L^1]{a.s.} 0$$

$(X_n)$  is dominated by  $Z$ , and  $X_\infty = \lim_n X_n$  exists  $\implies X_\infty$  is integrable (by the dominated convergence theorem). Then we define  $Y_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ . By Theorem 3.14, we immediately have

(a)  $Y_n \xrightarrow[L^1]{a.s.} \mathbb{E}(X_\infty | \mathcal{F}_\infty)$ ; that is

$$|\mathbb{E}(X_\infty | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_\infty)| \xrightarrow[L^1]{a.s.} 0.$$

(b) To prove  $|\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| \xrightarrow[L^1]{a.s.} 0$ , we firstly consider the  $L^1$ -convergence.

(i) Notice that

$$\mathbb{E}|\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| \leq \mathbb{E}|X_n - X_\infty|.$$

Obviously,  $\mathbb{E}|X_n - X_\infty| \rightarrow 0$ , by Theorem 1.12 (u.i. + a.s. conv.). Thus,

$$|\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| \xrightarrow[L^1]{} 0.$$

(ii) Define  $Z_m = \sup_{n \geq m} |X_n - X_\infty|$ . Observe that

$$\begin{aligned} \limsup_{n \rightarrow \infty} |\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| &\leq \limsup_{n \rightarrow \infty} \mathbb{E}(|X_n - X_\infty| | \mathcal{F}_n) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{E}(Z_m | \mathcal{F}_n) \\ &= \mathbb{E}(Z_m | \mathcal{F}_\infty) \end{aligned}$$

And since  $|Z_m| \leq \sup_{n \geq m} (|X_n| + |X_\infty|) \leq 2|Z|$ , by the dominated convergence theorem in the Proposition 3.2, we have

$$\lim \mathbb{E}(Z_m | \mathcal{F}_\infty) = \mathbb{E}(\lim Z_m | \mathcal{F}_\infty) = 0.$$

Therefore,  $\limsup_{n \rightarrow \infty} |\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| \xrightarrow[L^1]{a.s.} 0$  as  $n \rightarrow \infty$ .

Combining (i) and (ii), we get  $|\mathbb{E}(X_n | \mathcal{F}_n) - \mathbb{E}(X_\infty | \mathcal{F}_n)| \xrightarrow[L^1]{a.s.} 0$ .

□

**Exercise 3.4.2.** Let  $X$  be a u.i. submartingale,  $X_\infty = \lim_{n \rightarrow \infty} X_n$ , and  $M_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ .

$\implies$  (a)  $M$  is a u.i. martingale, and (b)  $X_n \leq M_n$  for every  $n \in \mathbb{N}$ .

*Proof.* (a) Because  $X_\infty$  is integrable (Theorem 3.12),  $M$  is a u.i. martingale.

(b) By the definition of submartingales.

□

**Exercise 3.4.3.** Let  $(X, \mathcal{F})$  be a martingale and  $p > 1$ . If  $\sup_n \mathbb{E}|X_n|^p < \infty$ , then

$$X_n \xrightarrow[L^p]{} X$$

where  $X$  is the a.s. limit of  $X_n$ .

*Proof.*

- First, we notice that

$$\sup_n \mathbb{E}|X_n|^p < \infty \implies (X_n) \text{ is u.i.}$$

Thus, by Theorem 3.13,  $X_n \xrightarrow[L^1]{a.s.} X$  and  $X_n \xrightarrow[L^1]{} X$ .

- Because  $|X_n - X|^p \leq (2 \sup_n |X_n|)^p$ . If  $\mathbb{E}(2 \sup_n |X_n|)^p < \infty$ , the dominated convergence theorem will imply

$$\mathbb{E}|X_n - X|^p \rightarrow 0.$$

We prove it below.

- $|X_n|$  is a submartingale. Apply Doob's inequality for  $|X_n|$ :

$$\mathbb{E}(\sup_{0 \leq m \leq n} |X_m|)^p \leq \left(\frac{p}{1-p}\right)^p \mathbb{E}|X_n|^p,$$

and take  $n \rightarrow \infty$ .

□

### 3.5 Optimal Stopping Theorem

**Definition 3.15.**

- $T : \Omega \rightarrow \mathbb{N}$  is called a stopping time w.r.t.  $\mathcal{F} = (\mathcal{F}_k)_{k \in \mathbb{N}}$  if

$$\{T \leq k\} \in \mathcal{F}_k$$

for all  $k \in \mathbb{N}$ .

- Define  $\mathcal{F}_T = \{A : A \cap \{T = n\} \in \mathcal{F}_n, \forall n\}$ .

*Remark.* This definition is equivalent to  $\{T = k\} \in \mathcal{F}_k$  for all  $k$ .

**Example 3.16** (The first entrance/passage time). Let  $B$  is a Borel set, and  $T = \min\{n \in \mathbb{N} : X_n \in B\}$ , where  $X$  is an adapted process. Then  $T$  is a stopping time because

$$\begin{aligned} \{T = n\} &= \{X_1 \notin B, \dots, X_{n-1} \notin B, X_n \in B\} \\ &= \{X_1 \notin B\} \cap \dots \cap \{X_{n-1} \notin B\} \cap \{X_n \in B\} \\ &\in \mathcal{F}_n. \end{aligned}$$

Or we can use  $\{T \leq n\} = \bigcup_{i=1}^n \{X_i \in B\}$ .

**Example 3.17** (A map that is not a stopping time). Let  $B$  is a Borel set, and  $L = \sup\{n \in \mathbb{N} : n \leq 10, X_n \in B\}$ , where  $X$  is an adapted process. Then  $L$  is not a stopping time in general.

For example, let  $X$  be a simple symmetric random walk, and  $B = \{1\}$ . By Wald's identity,  $\mathbb{E}X_L = 0$ ; however,  $\mathbb{E}X_L = \mathbb{P}(X_L = 1) \neq 0$ .

**Proposition 3.18.** Let  $(X_t)_{t \in \mathbb{N}}$  be a submartingale w.r.t.  $\mathcal{F}$ , and  $T$  be a stopping time w.r.t.  $\mathcal{F}$ . Then  $(X_{t \wedge T})_{t \in \mathbb{N}}$  is also a submartingale w.r.t.  $\mathcal{F}$ .

**Theorem 3.19** (Optional stopping theorem for submartingales).

(I)  $0 \leq S \leq T \leq K$  almost surely, then

$$\mathbb{E}X_T \geq \mathbb{E}X_S \geq \mathbb{E}X_0$$

(II)  $X$  is u.i. submartingale, then

$$\mathbb{E}X_\infty \geq \mathbb{E}X_T \geq \mathbb{E}X_S \geq \mathbb{E}X_0$$

(III)  $X_{n \wedge T}$  is a u.i. martingale, then

$$\mathbb{E}X_T \geq \mathbb{E}X_S \geq \mathbb{E}X_0$$

*Remark.* For part (III), it is natural to ask when  $X_{n \wedge T}$  is a u.i. submartingale. There are several sufficient conditions:

- (i)  $X_n$  is u.i.  $\implies X_{n \wedge T}$  is u.i.
- (ii)  $\mathbb{E}|X_T| < \infty$  and  $\mathbb{E}(|X_n| \mathbf{1}_{\{T > n\}}) \rightarrow 0$  as  $n \rightarrow \infty \implies X_{n \wedge T}$  is u.i.
- (iii)  $\mathbb{E}(|X_{n+1} - X_n| \mid \mathcal{F}_n) \leq C$  and  $\mathbb{E}T < \infty \implies X_{n \wedge T}$  is u.i.
- (iv)  $X_{n \wedge T}$  is bounded by an integrable RV  $\implies X_{n \wedge T}$  is u.i.

*Proof.* We will only prove (iii).

- (iii) Because  $X_n = X_0 + (X_1 - X_0) + \dots + (X_n - X_{n-1})$ , for  $m \geq T$ ,

$$\begin{aligned} X_{n \wedge T} &= X_0 + (X_1 - X_0) + \dots + (X_{m \wedge T} - X_{m-1}) + (X_{(m+1) \wedge T} - X_{m \wedge T}) + \dots \\ &= X_0 + (X_1 - X_0) + \dots + (X_{m \wedge T} - X_{m-1}) \end{aligned}$$

$$|X_{n \wedge T}| \leq |X_0| + \sum_{m=0}^{\infty} |X_{m+1} - X_m| \mathbf{1}_{\{m < T\}}$$



Letting  $Y = |X_0| + \sum_{m=0}^{\infty} |X_{m+1} - X_m| \mathbf{1}_{\{m < T\}}$ , it suffices to prove  $\mathbb{E}Y < \infty$ . Notice that

$$\begin{aligned} \mathbb{E}|X_{m+1} - X_m| \mathbf{1}_{\{m < T\}} &= \mathbb{E}\left(\mathbb{E}(|X_{m+1} - X_m| \cdot \mathbf{1}_{\{m < T\}} \mid \mathcal{F}_m)\right) \\ &= \mathbb{E}\left(\mathbf{1}_{\{m < T\}} \cdot \mathbb{E}(|X_{m+1} - X_m| \mid \mathcal{F}_m)\right) \\ (\text{by the given cond.}) &\leq C\mathbb{P}(T > m) \end{aligned}$$

Thus,  $\mathbb{E}Y \leq \mathbb{E}|X_0| + C \sum_{m=0}^{\infty} \mathbb{P}(T > m) = \mathbb{E}|X_0| + C\mathbb{E}T < \infty$ .

By (v) we know  $X_{n \wedge T}$  is bounded by an integrable RV  $Y$ , so it is u.i.

□

**Theorem 3.20** (Optimal stopping theorem for martingales). *(Y, F) = a martingale. T = a stopping time w.r.t F such that T < ∞ almost surely. If one of the following holds*

- (I) a)  $\mathbb{E}(|Y_T|) < \infty$   
b)  $\mathbb{E}(Y_n \mathbf{1}_{T > n}) \rightarrow 0$
- (II) Y is u.i.
- (III) a)  $\mathbb{E}T < \infty$   
b)  $\exists c$  s.t.  $\mathbb{E}(|Y_{n+1} - Y_n| \mid \mathcal{F}_n) \leq c \quad \forall n < T$

$\implies \mathbb{E}Y_T = \mathbb{E}Y_0$ .

**Corollary 3.21** (Wald's identity). *Let  $S_n = \xi_1 + \dots + \xi_n$  be a RW. For any stopping time T with  $\mathbb{E}T < \infty$ :*

$$\mathbb{E}S_T = \mathbb{E}\xi \cdot \mathbb{E}T.$$

## Exercises

**Exercise 3.5.1.** Let  $Q = \min\{n \geq 1 : X_{n-1} \in B\}$  and  $R = \min\{n \geq 1 : X_{n+1} \in B\}$ . Are Q and R stopping times?

**Exercise 3.5.2.** Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be an increasing map such that  $f(n) \geq n$  for all  $n$ , and  $T$  be a stopping time w.r.t  $\mathcal{F}$ . Then  $f(T)$  is a stopping time w.r.t  $\mathcal{F}$ .

*Proof.* Because  $f$  is increasing,

$$\{f(T) \leq k\} \in \mathcal{F}_k \iff \{T \leq f^{-1}(k)\} \in \mathcal{F}_k.$$

And notice that  $f^{-1}(k)$  is an integer that is not larger than  $k$  (by  $f(n) \geq n$ ). So

$$\{T \leq f^{-1}(k)\} \in \mathcal{F}_{f^{-1}(k)} \subset \mathcal{F}_k.$$

□

**Exercise 3.5.3.** Let  $\tau$  be a stopping time w.r.t.  $\mathcal{F}$ . Then  $\mathcal{F}_{k \wedge \tau} \uparrow \mathcal{F}_\tau$  as  $k \rightarrow \infty$ .

### -Optional stopping theorem for continuous martingales-

**Theorem.** *Let  $(X_t)_{t \geq 0}$  be a uniformly integrable martingale with right-continuous paths. Let S and T be two stopping times with  $S < T$ . Then  $X_S$  and  $X_T$  are in  $L^1$  and*

$$X_S = \mathbb{E}(X_T \mid \mathcal{F}_S).$$

*In particular, for every stopping time S, we have*

$$X_S = \mathbb{E}(X_\infty \mid \mathcal{F}_S)$$

*and*

$$\mathbb{E}(X_S) = \mathbb{E}(X_\infty) = \mathbb{E}(X_0).$$

**Exercise 3.5.4.** Let  $B$  be a Brownian motion. Define

$$T_a = \inf\{t \geq 0 : B_t = a\}$$

and

$$U_a = \inf\{t \geq 0 : |B_t| = a\}.$$

- (a) Show  $T_a$  and  $U_a$  are stopping times for all  $a$ .
- (b) Find  $\mathbb{P}(T_a < T_b)$  for  $a < 0 < b$ .
- (c) Find  $\mathbb{E}U_a$ .
- (d) Compute  $\mathbb{E}(e^{-\lambda T_a})$  for  $\lambda > 0$  (The Laplace transform of  $T_a$ ).
- (e) Compute  $\mathbb{P}(T_a < \infty)$ .

*Remark.* In this exercise, we should avoid using the distribution of  $T_a$  directly. Try to apply the optional stopping theorem.

*Proof.* (a) Note that  $\{T_a < t\} = \bigcup_{s < t} \{B_s = a\}$ . And  $\{B_s = a\} \in \mathcal{F}_s$  and  $\mathcal{F}_t = \bigcup_{s < t} \mathcal{F}_s$ ; thus,  $T_a$  is a stopping time.

Note that  $\{U_a < t\} = \bigcup_{s < t} \{B_s = \pm a\}$ . And  $\{B_s = \pm a\} \in \mathcal{F}_s \subset \mathcal{F}_t$  for  $s < t$ ; thus,  $U_a$  is a stopping time.

- (b) **Tips 1:** In most cases, (stopped) martingale is bounded.

Let  $T = T_a \wedge T_b$  and  $M_t = B_{t \wedge T}$ . Because  $T_a$  and  $T_b$  are stopping times,  $T$  is stopping time. Then  $M_t$  is a u.i. martingale (its boundedness is given by the definition of  $T$ ). Apply the optional stopping theorem:

$$\mathbb{E}M_T = \mathbb{E}M_0 = 0$$

and notice that

$$\mathbb{E}M_T = a\mathbb{P}(T_a < T_b) + b\mathbb{P}(T_a > T_b).$$

Therefore, we have the following two equations:

$$\begin{cases} a\mathbb{P}(T_a < T_b) + b\mathbb{P}(T_a > T_b) &= 0 \\ \mathbb{P}(T_a < T_b) + \mathbb{P}(T_a > T_b) &= 1 \end{cases}$$

Solve it. We get  $\mathbb{P}(T_a < T_b) = \frac{b}{b-a}$ .

- (c) **Tips 2:** Construct an appropriate martingale.

Let  $M_t = B_t^2 - t$ . Then  $M_{t \wedge U_a}$  is uniform integrable.

It is because

$$M_{t \wedge U_a} \leq B_{t \wedge U_a}^2 \leq a^2.$$

Apply the optional stopping theorem:

$$\mathbb{E}M_{U_a} = M_0 = 0$$

so we have

$$\mathbb{E}B_{t \wedge U_a}^2 = \mathbb{E}(t \wedge U_a).$$

Notice that  $(t \wedge U_a) \uparrow U_a$ ; thus, by the monotone convergence theorem,

$$\mathbb{E}(t \wedge U_a) \uparrow \mathbb{E}U_a.$$

And because  $B_{t \wedge U_a}^2 \leq a^2$  and  $B_{U_a} = \pm a$ ; by the dominated convergence theorem,

$$\mathbb{E}B_{t \wedge U_a}^2 \rightarrow \mathbb{E}B_{U_a}^2 = a^2.$$

Therefore, taking  $t \rightarrow \infty$  on both sides, we have

$$a^2 = \mathbb{E}U_a.$$

(d) **Tips 3:** Always check the uniform integrability.

Consider the exponential martingale

$$N_t^\lambda = e^{\lambda B_t - \frac{\lambda^2}{2}t}.$$

- $\lambda > 0$ .

In this case,  $N_{t \wedge T_a}^\lambda$  is a uniformly integrable martingale, because

$$\exp(\lambda B_{t \wedge T_a} - \frac{\lambda}{2}(t \wedge T_a)) \leq \exp(\lambda B_{t \wedge T_a}) \leq e^{\lambda a}.$$

Apply the optional stopping theorem:

$$e^{\lambda B_{T_a} - \frac{\lambda^2}{2}T_a} = 1;$$

that is

$$e^{-\lambda a} = e^{-\frac{\lambda^2}{2}T_a}.$$

Then take  $\lambda \mapsto \sqrt{2\lambda}$ , we have  $\mathbb{E}(e^{-\lambda T_a}) = e^{-a\sqrt{2\lambda}}$ .

- $\lambda < 0$ .

Be careful! In this case,  $N_{t \wedge T_a}^\lambda$  is not uniformly integrable.

(e) Letting  $\lambda \rightarrow 0$ , we have

$$\mathbb{P}(T_a < \infty) = \lim_{\lambda \rightarrow 0} e^{-\lambda T} = 1.$$

□

**Exercise 3.5.5.** Let  $M$  be a martingale with continuous path such that  $M_0 = x \in \mathbb{R}_+$ . We assume that  $M_t \geq 0$  for every  $t \geq 0$ , and that  $M_t \rightarrow 0$  when  $t \rightarrow \infty$  almost surely. Show that for every  $y > x$ ,

$$\mathbb{P}(\sup_{t \geq 0} M_t \geq y) = \frac{x}{y}.$$

*Proof.* Define  $T = \inf\{t : M_t = y\}$ . It is easy to see that  $M_{t \wedge T}$  is bounded (so it is a u.i. martingale).

So  $\mathbb{E}M_T = \mathbb{E}M_0$ . Notice that actually  $T$  is not almost surely finite. Therefore, we have

$$\mathbb{E}M_T \mathbf{1}_{\{T < \infty\}} + \mathbb{E}M_\infty \mathbf{1}_{\{T = \infty\}} = x.$$

It is known that  $M_\infty$  is 0. So

$$y\mathbb{P}(T < \infty) = x.$$

Finally, notice that  $\{T < \infty\} \iff \{\sup_{t \geq 0} M_t \geq y\}$ .

□

**Exercise 3.5.6.** Recall that  $T_a = \inf\{t > 0 : B_t = a\}$ . Give the law of  $\sup_{t \leq T_0} B_t$  where  $B$  is a Brownian motion started from  $x > 0$ .

*Proof.* Notice that for  $y > x$

$$\{\sup_{t \leq T_0} B_t \leq y\} \iff \{T_y \geq T_0\}.$$

Consider  $B_{T_y \wedge T_0}$  for  $y > 0$ . It is a u.i. martingale. So we have

$$\mathbb{E}B_{T_y \wedge T_0} = \mathbb{E}B_0.$$

It implies that  $y \cdot \mathbb{P}(T_y < T_0) + 0 \cdot \mathbb{P}(T_y \geq T_0) = x$ . Therefore, we have

$$\mathbb{P}(\sup_{t \leq T_0} B_t \leq y) = \begin{cases} 1 - \frac{x}{y} & y > x; \\ 0 & y \leq x. \end{cases}$$

□

**Exercise 3.5.7.** Let  $B$  be a Brownian motion started from 0, and let  $\mu > 0$ . Show that

$$\sup_{t \geq 0} (B_t - \frac{1}{2}\mu t)$$

is exponentially distributed with parameter  $\mu$ .

*Proof.* Note that  $N_t^\lambda = e^{\lambda B_t - \frac{\lambda^2}{2}t}$  forms a martingale with  $\lim_{t \rightarrow \infty} N_t^\lambda = 0$  and with  $N_0 = 1$ . So by the preceding exercise, for all  $\lambda \in \mathbb{R}$ ,

$$\mathbb{P}(\sup_{t \geq 0} N_t^\lambda \geq y) = \frac{1}{y}.$$

We want to find the distribution of  $\sup_{t \geq 0} (B_t - \frac{1}{2}\mu t)$ ; that is

$$\begin{aligned} \mathbb{P}(\sup_{t \geq 0} (B_t - \frac{1}{2}\mu t) \leq s) &= \mathbb{P}(\sup_{t \geq 0} e^{\mu(B_t - \frac{1}{2}\mu t)} \leq e^{\mu s}) \\ &= \mathbb{P}(\sup_{t \geq 0} N_t^\mu \leq e^{\mu s}) \\ &= 1 - e^{-\mu s} \end{aligned}$$

Therefore, the distribution of  $\sup_{t \geq 0} (B_t - \frac{1}{2}\mu t)$  is  $\text{Exp}(\mu)$ . □

**Exercise 3.5.8.** Let  $B$  be a Brownian motion started from 0. Let  $a > 0$  and

$$\sigma_a = \inf\{t \geq 0 : B_t \leq t - a\}.$$

- (a) Show that  $\sigma_a$  is a stopping time and that  $\sigma_a < \infty$  almost surely.
- (b) Show that for every  $\lambda \geq 0$ ,

$$\mathbb{E}e^{-\lambda\sigma_a} = e^{-a(\sqrt{1+2\lambda}-1)}.$$

*Proof.* (a) First,  $\{\sigma_a \leq x\} = \bigcup_{t \leq x} \{B_t \leq t - a\}$ ; therefore, it is a stopping time.

Then notice that

$$\{\sigma_a < \infty\} \iff \{\inf_{t \geq 0} (B_t - t) \leq -a\}.$$

Finally, use  $(B_t - t)/t \xrightarrow{a.s.} -1$ .

- (b) For  $\mu \leq 0$ , define  $M_t = e^{\mu B_t - \frac{\mu^2}{2}t}$ . Then  $M_{t \wedge \sigma_a}$  is a u.i martingale.

Apply the optional stopping theorem:

$$\mathbb{E}e^{\mu B_{t \wedge \sigma_a} - \frac{\mu^2}{2}(t \wedge \sigma_a)} = 1.$$

It implies

$$\mathbb{E}e^{(\mu - \frac{\mu^2}{2})\sigma_a} = e^{a\mu}.$$

Letting  $\lambda = -(\mu - \frac{\mu^2}{2})$  (obviously, we have  $\lambda \leq 0$ ), we get the desired result. □

### 3.6 Inequalities

**Definition 3.22.**

- Let  $X$  be a sequence of RVs,  $T_0 = 0$ , and  $[a, b]$  be an interval. For  $k \in \mathbb{N}$ , define

$$T_{2k-1} = \min\{n > T_{2k-2} : X_n \leq a\};$$

$$T_{2k} = \min\{n > T_{2k-1} : X_n \geq b\}.$$

Then  $[T_{2k-1}, T_{2k}]$  is called an upcrossing of  $[a, b]$ .

- $U_n(a, b; X)$  = The number of upcrossing of  $[a, b]$  by  $X$  up to time  $n$ .
- $U(a, b; X) = \lim_{n \rightarrow \infty} U_n(a, b; X)$ .

**Theorem 3.23** (The upcrossing inequality). *Let  $(X, \mathcal{F})$  be a submartingale.*

$$\mathbb{E}(U_n(a, b; X)) \leq \frac{1}{b-a} \mathbb{E}((X_n - a)^+)$$

**Theorem 3.24** (Doob's inequality). *Let  $(X, \mathcal{F})$  be a submartingale, and*

$$A = \left\{ \max_{0 \leq m \leq n} X_m \geq \lambda \right\}.$$

a) For  $\lambda > 0$ ,

$$\lambda \mathbb{P}\left(\max_{1 \leq m \leq n} X_m \geq \lambda\right) \leq \mathbb{E}(X_n^+ \mathbf{1}_A) \leq \mathbb{E}X_n^+.$$

b) For  $p > 1$ ,

$$\mathbb{E}\left(\max_{1 \leq m \leq n} X_m^+\right)^p \leq \left(\frac{p}{1-p}\right)^p \mathbb{E}(X_n^+)^p.$$

*Proof.*

- (a) • Notice the following important relation:

$$\left\{ \max_{1 \leq m \leq n} X_m \geq \lambda \right\} = \{T \leq n\}$$

where  $T = \min\{n \geq 1 : X_n \geq \lambda\}$ . So we have

$$\begin{aligned} \lambda \mathbb{P}(\max X_m \geq \lambda) &= \lambda \mathbb{P}(T \leq n) \\ &= \lambda \mathbb{E} \mathbf{1}_{\{T \leq n\}} \\ &\leq \mathbb{E}(X_T^+ \mathbf{1}_{\{T \leq n\}}) \end{aligned} \tag{1}$$

- Apply the optional stopping theorem for  $(X_{n \wedge T}^+)$ :

$$\mathbb{E}X_0^+ \leq \mathbb{E}X_{n \wedge T}^+ \leq \mathbb{E}X_n^+ \tag{2}$$

where  $\mathbb{E}X_{n \wedge T}^+ = \mathbb{E}X_T^+ \mathbf{1}_{\{T \leq n\}} + \mathbb{E}X_n^+ \mathbf{1}_{\{T > n\}}$ .

- Finally, just put (1) and (2) together.

(b) Omitted. □

**Corollary 3.25** (Doob-Kolmogorov inequality for martingales). *Let  $(X, \mathcal{F})$  be a martingale such that  $\mathbb{E}X_n^2 < \infty$ . Then*

$$\mathbb{P}\left(\max_{1 \leq m \leq n} |X_m| \geq \lambda\right) \leq \frac{\mathbb{E}X_n^2}{\lambda^2}.$$

*Proof.* Use the previous theorem for  $(X_n^2)$ . □

**Theorem 3.26** (Hoeffding's inequality). *Let  $(Y, \mathcal{F})$  be a martingale. Suppose  $\exists \{k_n\}_{n \geq 1}$  of real numbers s.t.  $\mathbb{P}(|Y_n - Y_{n-1}| \leq k_n) = 1$  for all  $n$ . Then*

$$\mathbb{P}(|Y_n - Y_0| \geq x) \leq 2e^{-\frac{1}{2}x^2 / \sum_{i=1}^n k_i^2}.$$

*Proof.* • First, we show the following lemma:

Let  $D$  be a RV s.t.  $\mathbb{E}(D) = 0$  and  $|D| \leq 1$  almost surely, then

$$\mathbb{E}(e^{\psi D}) < e^{\frac{1}{2}\psi^2} \quad (*)$$

holds for all  $\psi > 0$ .

• **Proof.**

For  $\psi > 0$ , define

$$g : d \mapsto e^{\psi d},$$

which is a convex function. So for  $|d| \leq 1$ , we have

$$e^{\psi d} \leq \frac{1}{2}(1-d)e^{-\psi} + \frac{1}{2}(1+d)e^{\psi}.$$

Plug it in  $\mathbb{E}e^{\psi D}$ , we have

$$\begin{aligned} \mathbb{E}e^{\psi D} &\leq \mathbb{E}\left(\frac{1}{2}(1-D)e^{-\psi} + \frac{1}{2}(1+D)e^{\psi}\right) \\ &= \frac{1}{2}e^{-\psi} + \frac{1}{2}e^{\psi} \\ (\text{Taylor series}) \quad &< \exp\left(\frac{1}{2}\psi^2\right) \end{aligned}$$

- Now we consider  $\mathbb{P}(Y_n - Y_0 \geq x)$ . Let  $D_n = Y_n - Y_{n-1}$  be the martingale difference of  $Y$ . Then for all  $\theta > 0$ , we have

$$\begin{aligned} \mathbb{P}(Y_n - Y_0 \geq x) &= \mathbb{P}\left(e^{\theta(Y_n - Y_0)} \geq e^{\theta x}\right) \\ (\text{Markov}) \quad &\leq \mathbb{E}\left(e^{\theta(Y_n - Y_0)}\right) \cdot e^{-\theta x} \\ &= e^{-\theta x} \mathbb{E}\left(e^{\theta D_n} \cdot e^{\theta(Y_{n-1} - Y_0)}\right) \\ &= e^{-\theta x} \mathbb{E}\left(\mathbb{E}\left(e^{\theta D_n} \cdot e^{\theta(Y_{n-1} - Y_0)} \mid \mathcal{F}_{n-1}\right)\right) \\ &= e^{-\theta x} \mathbb{E}\left(e^{\theta(Y_{n-1} - Y_0)} \cdot \mathbb{E}(e^{\theta D_n} \mid \mathcal{F}_{n-1})\right) \end{aligned}$$

- Notice that  $|\frac{D_n}{k_n}| < 1$  and  $\mathbb{E}D_n = 0$ . Applying  $(*)$ , we have

$$\begin{aligned} \mathbb{E}(e^{\theta D_n} \mid \mathcal{F}_{n-1}) &= \mathbb{E}(e^{\theta k_n \frac{D_n}{k_n}} \mid \mathcal{F}_{n-1}) \\ &\leq e^{\frac{1}{2}\theta^2 k_n^2} \end{aligned}$$

Therefore, by induction, we have

$$\begin{aligned} \mathbb{P}(Y_n - Y_0 \geq x) &\leq e^{-\theta x} \cdot e^{\frac{1}{2}\theta^2 k_n^2} \cdot \mathbb{E}\left(e^{\theta(Y_{n-1} - Y_0)}\right) \\ &\leq \exp\left(-\theta x + \frac{1}{2}\theta^2 \sum k_i^2\right) \\ &\leq \exp\left(-\frac{1}{2}x / \sum k_i^2\right) \end{aligned}$$

- Finally, note that

$$\begin{aligned} \mathbb{P}(|Y_n - Y_0| \geq x) &= \mathbb{P}(Y_n - Y_0 \geq x) + \mathbb{P}(Y_0 - Y_n \geq x) \\ &\leq 2 \exp\left(-\frac{1}{2}x / \sum k_i^2\right) \end{aligned}$$

□

## Exercises

**Exercise 3.6.1.** Let  $X$  be a submartingale bounded in  $L^1$ . Let  $a < b$  and

$$U(a, b; X) = \lim_{N \rightarrow \infty} U_N(a, b; X).$$

Show that

$$(b - a)\mathbb{E}U(a, b; X) \leq |a| + \sup_n \mathbb{E}|X_n|.$$

*Proof.* Use the upcrossing inequality and the monotone convergence theorem together. □

## 4 Examples of Discrete Stochastic Processes

### 4.1 Finite Martingales

**Example 4.1.** Let  $\eta$  be an integrable RV and  $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_n$  be an increasing sequence of  $\sigma$ -fields.

- a) Define  $\xi_k = \mathbb{E}(\eta|\mathcal{F}_k)$ . Then  $(\xi, \mathcal{F})$  is a martingale.
- b) Conversely, if  $(\xi, \mathcal{F})$  is a finite martingale, there exists an integrable RV  $\eta$  such that

$$\xi_k = \mathbb{E}(\eta|\mathcal{F}_k).$$

- c) However, for infinite martingales, (b) doesn't hold in general.

**Example 4.2.** Let  $Y_1, \dots, Y_4$  be i.i.d. RVs with  $\mathbb{E}|Y_i| < \infty$  for all  $i$ . Define

$$X_1 = \frac{1}{4}(Y_1 + Y_2 + Y_3 + Y_4), \quad X_2 = \frac{1}{3}(Y_1 + Y_2 + Y_3), \quad X_3 = \frac{1}{2}(Y_1 + Y_2), \quad X_4 = Y_1.$$

Then  $X$  is a martingale.

To prove this, we notice the symmetry:

$$\begin{aligned} \mathbb{E}(X_2|X_1) &= \frac{1}{3}\mathbb{E}(Y_1 + Y_2 + Y_3|Y_1 + Y_2 + Y_3 + Y_4) \\ &= \frac{1}{3}\mathbb{E}(Y_2 + Y_3 + Y_4|Y_1 + Y_2 + Y_3 + Y_4) \\ &= \frac{1}{3}\mathbb{E}(Y_3 + Y_4 + Y_1|Y_1 + Y_2 + Y_3 + Y_4) \\ &= \frac{1}{3}\mathbb{E}(Y_4 + Y_1 + Y_2|Y_1 + Y_2 + Y_3 + Y_4) \end{aligned}$$

Thus,  $4\mathbb{E}(X_2|X_1) = \mathbb{E}(Y_1 + Y_2 + Y_3 + Y_4|Y_1 + Y_2 + Y_3 + Y_4) = Y_1 + Y_2 + Y_3 + Y_4 = 4X_1$ . Then it is easy to see that  $X$  is a martingale.

### 4.2 Markov Chains (MC)

**Definition.** The following definitions are equivalent:

- A process  $X$  is a Markov chain if it satisfies

$$\mathbb{P}(X_n = s \mid X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = s \mid X_{n-1} = x_{n-1}).$$

- Let  $\mathcal{F}$  be the natural filtration of  $X$ .  $X$  is a Markov chain if it satisfies

$$\mathbb{E}(f \circ X_{n+1} | \mathcal{F}_n) = (P_n f) \circ X_n$$

for every non-negative bounded measurable function  $f$  and  $n \in \mathbb{N}$ , where

$$P_n(s, B) := \mathbb{P}(X_{n+1} \in B | X_n = s)$$

is a Markov kernel on the state space of  $X$ .

*Remark.*

- Let  $f = \mathbf{1}_{\{s\}}$ , we have  $\mathbb{P}(X_{n+1} = s | \mathcal{F}_n) = P_n(X_n, s) = \mathbb{P}(X_{n+1} = s | X_n)$ . Recall that for any non-negative bounded measurable function  $g$ , we can find a sequence of simple functions  $\{f_n\}_{n \in \mathbb{N}}$  such that  $f_n \uparrow g$  almost surely as  $n \rightarrow \infty$ .
- Every discrete-time Markov chain has strong Markov property; that is,

$$\mathbb{P}(X_{T+k} = x | \mathcal{F}_T) = \mathbb{P}(X_{T+k} = x | X_T),$$

where  $T$  is a stopping time.



More properties of MCs will be given in Section 6. The following three examples will show the relation between MCs and martingales.

**Example 4.3** ( $X$  is both a MC and a mart.). Let  $X_i$  be i.i.d. RVs with  $\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1) = \frac{1}{2}$ . Define  $S_0 = 0$  and

$$S_n = S_0 + \sum_{i=1}^n X_i.$$

$S_n$  is a simple symmetric random walk. It is both a Markov chain and a martingale.

**Example 4.4** ( $X$  is a MC but is not a mart.). Let  $X_i \sim \text{Bin}(1, \frac{1}{2})$  be i.i.d.. Define  $S_0 = 0$  and

$$S_n = S_0 + \sum_{i=1}^n X_i.$$

$S_n$  is a Markov chain; however, it is not a martingale.

**Example 4.5** ( $X$  is not a MC but is a mart.). Let  $X_i$  be i.i.d. RVs with  $\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1) = \frac{1}{2}$ ; and  $S_0 \sim \text{Bin}(1, \frac{1}{2})$ , indep. of  $X$ . Define

$$S_n = S_{n-1} + X_n S_0.$$

It is a martingale; but it is not a Markov chain. Because the Markov property doesn't hold:

$$\begin{aligned}\mathbb{P}(S_2 = 1 \mid S_1 = 0, S_0 = 1) &= 1/2; \\ \mathbb{P}(S_2 = 1 \mid S_1 = 0, S_0 = 0) &= 0; \\ \mathbb{P}(S_2 = 1 \mid S_1 = 0) &= 1/4.\end{aligned}$$

The following example shows another difference between Markov chains and martingales. Let  $\tau$  be a stopping time w.r.t.  $\mathcal{F}$ . If  $(X_n)_{n \in \mathbb{N}}$  is a martingale, then  $(X_{n \wedge \tau})_{n \in \mathbb{N}}$  is also a martingale. What will happen if  $X$  is a Markov chain?

**Example 4.6** (Stopped Markov chains). Let  $X = (X_n)_{n \in \mathbb{N}}$  be a Markov chain with countable state space  $S$  and transition kernel  $P$ . Let  $A \subset S$  and  $\tau_A = \inf\{n \geq 1 : X_n \in A\}$ .

- Define  $Y_n = X_{n \wedge \tau_A}$ . We consider two cases: If  $n > \tau_A$ , then

$$Y_{\tau_A} = Y_{\tau_A+1} = \dots = Y_{n-1} = Y_n;$$

obviously, the value of  $Y_n$  totally depends on the value of  $Y_{n-1}$ . If  $n \leq \tau$ , then  $Y_k = X_k$  for all  $k \leq n$ . And notice that  $X$  is a Markov chain.

Therefore,  $Y$  is a Markov chain.

- Define  $\tau = \inf\{n > \tau_A : X_n \in A\}$  as the second entrance time of  $A$ . Let  $Y_n = X_{n \wedge \tau}$ . Then  $Y$  is not a Markov chain. We still consider two cases: If  $n \leq \tau$ , it is same. However, if  $n > \tau$ , it is possible that  $Y_n \neq Y_{n-1}$ .

It is well-known that the sum of two martingales are also a martingale, and that the sum of two Lévy processes are also a Lévy process. However, the following example shows that the sum of two Markov chains may not be Markov chains.

**Example 4.7** (Sums of two MCs may not be MC).

See here. Let  $X$  be a RV such that  $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2}$ . Set  $X_n = X$  for all  $n$ .

Let  $Y$  be another MC with state space  $S = \{-1, 0, 1\}$  that is independent of  $X$ . Let  $Y_0 = 0$  and its transition matrix be

$$P = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix}.$$

Now set  $Z_n = X_n + Y_n$ . Notice that

$$\begin{aligned}\mathbb{P}(Z_2 = 0 | Z_1 = 1) &= \frac{5}{12}, \\ \mathbb{P}(Z_2 = 0 | Z_1 = 1, Z_0 = 1) &= \frac{1}{4}.\end{aligned}$$

Therefore,  $X + Y$  is not a Markov chain.

The following proposition characterizes Markov chains in terms of martingales; and there is a more general result which holds for the continuous case. The proof of the following proposition is trivial by directly observing the martingale difference.

**Proposition.** *Let  $X$  be adapted to  $\mathcal{F}$ . Then  $X$  is a Markov chain with transition kernel  $P$  w.r.t.  $\mathcal{F}$  if and only if*

$$M_n = f \circ X_n - \sum_{m=1}^{n-1} (Pf - f) \circ X_m, \quad n \in \mathbb{N},$$

*is a martingale w.r.t  $\mathcal{F}$  for every bounded  $f \in \mathcal{E}_+$ .*

## Exercises

**Exercise 4.2.1.** Consider a colony of cells that evolves as follows. Initially, there is one cell. During each discrete time step, each cell either dies or splits into two new cells, each possibility having probability one half. Suppose cells die or split independently. Let  $X_k$  denote the number of cells alive at time  $k$ ;  $X_0 = 1$ . Determine which of the following properties are possessed by  $X$ :

- (a) Markov; (b) martingale; (c) independent increments; (d) uniform integrability.

## 4.3 Generating functions

**Definition 4.8.** content...

## 4.4 Branching Processes

**Definition.** ...

*Remark.*

- In general,  $(Z_n, \mathcal{F}_n)$  is not a martingale. Because

$$\mathbb{E}Z_{n+1} = \mathbb{E}\left(\mathbb{E}(Z_{n+1,1} + \dots + Z_{n+1,z_n} | \mathcal{F}_n)\right) = m \cdot \mathbb{E}Z_n = m^{n+1}$$

relies on  $n$ . Define  $M_n = Z_n/m^n$ ; then  $(M_n, \mathcal{F}_n)$  is a martingale.

- And  $(M_n, \mathcal{F}_n)$  gives an example of martingale which is not uniformly integrable (Example 4.7).

**Example 4.9.** Let  $Z_n$  denote the number of offsprings in the  $n$ -th generation of a branching process. Assume that  $Z_0 = 1$  and that the offspring distribution has mean  $\mu$  and variance  $\sigma^2 > 0$ . Define  $M_n = Z_n/\mu^n$ ,  $\mathcal{F} = \sigma\{Z_0, \dots, Z_n\}$ .

- a) For  $\mu \leq 1$ ,  $(M_n, \mathcal{F}_n)$  is a martingale but not a u.i. martingale.
- b) For  $\mu > 1$ ,  $(M_n, \mathcal{F}_n)$  is a u.i. martingale; and hence  $M_n \xrightarrow{a.s.} M_\infty$ . Show that  $M_\infty$  is not a constant RV.
- c) Suppose that  $\mu > 1$  and  $\mathbb{P}(Z_1 \geq 2) = 0$ . Show that

$$\mathbb{E}(\sup_n Z_n) \leq \eta/(\eta - 1),$$

where  $\eta$  is the largest root of the equation  $x = G(x)$  and  $G$  is the probability generating function of  $Z_1$ .

## 4.5 Birth-Death Processes

## 4.6 Random Walks

**Definition.** content...

**Example 4.10.** Let  $\xi_i$  be i.i.d. RVs with  $\mathbb{E}\xi_i = 0$ . Let  $S_0 = 0$ ,  $S_n = \sum_{i=1}^n \xi_i$ . Define  $T$  as the time until the first positive sum; that is

$$T = \min\{n \geq 1 : S_n > 0\}.$$

$\implies \mathbb{E}T = \infty$ .

Assume  $\mathbb{E}T < \infty$ .

**Example 4.11** (Symmetric simple random walk). Let  $X$  be a symmetric simple random walk on  $\{0, 1, \dots, N\}$  with absorbing boundaries. Suppose that  $X_0 = a$  ( $0 < a < N$ ). Let  $T = \min\{n : X_n = 0 \text{ or } N\}$ .

a)

## 5 General Stochastic Processes

### 5.1 Finite-Dimensional Distributions (fdd)

**Definition 5.1.** Let  $X = (X_t, t \in T)$  be a stochastic process. The finite-dimensional distributions of  $X$  are the distributions of  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  for all possible choices of  $t_1, t_2, \dots, t_n \in T$  and for all  $n \geq 1$ .

**Definition 5.2.** Let  $X = (X_t, t \in T)$ ,  $Y = (Y_t, t \in T)$  be two stochastic processes. If their fdds are same, they are called equivalent, or versions of one another.

*Remark.*

- Fdds can be considered as Borel probability measures on  $\mathbb{R}^n$  for all  $n \geq 1$ . And they satisfy Kolmogorov consistency conditions.

- **Kolmogorov's existence theorem**

Conversely, if there is a family of Borel probability measures satisfying Kolmogorov consistency conditions, we can find a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a RV  $X = (X_t)_{t \in T}$  defined on it such that its fdds are exactly the given Borel probability measures.

**Definition 5.3.** A stochastic process is called Gaussian if all fdds are (multivariate) normal.

**Proposition 5.4.** Fdds of a Gaussian process are completely determined by its covariance functions.

### Exercises

**Exercise 5.1.1.** Let  $U \sim U[0, 1]$ . And let  $f$  be a continuous function on  $[0, 1]$ . Define  $X_t = f(t)$ , for all  $t \in [0, 1]$ , and  $Y_t = X_t + \mathbf{1}_{\{U=t\}}$ .

- (a) Prove  $X$  and  $Y$  have the same fdds. (b) Prove  $\mathbb{P}\{Y_t \text{ is continuous on } T\} = 0$ .

*Proof.* (a) We divide  $\{Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n\}$  into two disjoint parts:

$$\begin{aligned} & \mathbb{P}(Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n) \\ &= \mathbb{P}\left(\left(\{Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n\} \cap \{\forall i, U(\omega) \neq t_i\}\right) \cup \left(\{Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n\} \cap \{\exists i, U(\omega) = t_i\}\right)\right) \\ &= \mathbb{P}\left(\{X_{t_1} \leq s_1, \dots, X_{t_n} \leq s_n\}\right) + \mathbb{P}\left(\{Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n\} \cap \{\exists i, U(\omega) = t_i\}\right) \end{aligned}$$

And notice

$$\begin{aligned} \mathbb{P}\left(\{Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n\} \cap \{\exists i, U(\omega) = t_i\}\right) &\leq \mathbb{P}(\exists i, U(\omega) = t_i) \\ &= 0 \end{aligned}$$

$$\implies \mathbb{P}(Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n) = \mathbb{P}(X_{t_1} \leq s_1, \dots, X_{t_n} \leq s_n)$$

- (b) Fix  $\omega = \omega^*$ . Let  $t_0 = U(\omega^*)$ . Then  $Y_{t_0} = f(t_0) + 1$ , while  $Y_{t_0-} = Y_{t_0+} = f(t_0)$ .

$\implies$  For each path of  $Y$ , there exists a discontinuous point.

$\implies \mathbb{P}\{Y_t \text{ is continuous on } T\} = 0$ .

□

**Exercise 5.1.2.** Let  $\Omega = [0, 1]$ ,  $\mathcal{F} = \mathcal{B}[0, 1]$ , and  $\mathbb{P} = \text{Leb}[0, 1]$ . Define  $X_t(\omega) = \mathbf{1}_{\{t=\omega\}}$  and  $Y_t(\omega) = 0$ .

- (a) Prove  $X$  and  $Y$  are equivalent processes.

- (b) Calculate  $\mathbb{P}(\sup_{0 \leq t \leq 1} X_t = 0)$  and  $\mathbb{P}(\sup_{0 \leq t \leq 1} Y_t = 0)$ .

*Proof.* (a) To prove  $X$  and  $Y$  are equivalent, we need to calculate their fdds:

$$\begin{aligned} \mathbb{P}(Y_{t_1} \leq s_1, \dots, Y_{t_n} \leq s_n) &= \mathbb{P}(0 \leq s_1, \dots, 0 \leq s_n) \\ \mathbb{P}(X_{t_1} \leq s_1, \dots, X_{t_n} \leq s_n) &= \mathbb{P}(\mathbf{1}_{\{t_1=\omega\}} \leq s_1, \dots, \mathbf{1}_{\{t_n=\omega\}} \leq s_n) \\ &= \mathbb{P}\left(\{0 \leq s_1, \dots, 0 \leq s_n\} \cup \{\omega : \exists i, 1 \leq s_i, \omega = t_i\}\right) \\ &= \mathbb{P}(0 \leq s_1, \dots, 0 \leq s_n) \end{aligned}$$

$\implies X$  and  $Y$  are equivalent.

(b) It is easy to notice that

$$\begin{aligned}\mathbb{P}\left(\sup_{0 \leq t \leq 1} Y_t = 0\right) &= 1 \\ \mathbb{P}\left(\sup_{0 \leq t \leq 1} X_t = 0\right) &= \mathbb{P}\left(\sup_{0 \leq t \leq 1} \mathbf{1}_{\{\omega=t\}=0}\right) \\ &= \mathbb{P}(\emptyset) = 0\end{aligned}$$

□

### -Gaussian processes-

**Exercise 5.1.3** (Brownian Bridge). Let  $(B_t)$  be a BM,  $X_t = B_t - tB_1$ ,  $0 \leq t \leq 1$ , and  $Y_t = X_{1-t}$ . Prove  $Y$  has the same fdds as  $X$ .

*Proof.* Notice that  $X$  and  $Y$  are both Gaussian. Calculate their covariance functions.

$$\begin{aligned}\text{Cov}(X_t, X_s) &= \text{Cov}(B_t - tB_1, B_s - sB_1) \\ &= \text{Cov}(B_t, B_s) - st \\ &= \min(s, t) - st \\ \text{Cov}(Y_t, Y_s) &= \text{Cov}(B_{1-t} - (1-t)B_1, B_{1-s} - (1-s)B_1) \\ &= \text{Cov}(B_{1-t}, B_{1-s}) - (1-t)(1-s) \\ &= \min(1-s, 1-t) - (1-t)(1-s)\end{aligned}$$

Obviously, they are same.

□

**Exercise 5.1.4.** Show the following statements:

- (a)  $X$  is a Gaussian process  $\iff$  every finite linear combination  $Z = \sum_{i=1}^n a_i X_{t_i}$  is a Gaussian RV.
- (b) Let  $X_t = Z \cos(2\pi t + \theta)$  where  $Z$  and  $\theta$  are independent,  $\theta \sim U[0, 2\pi]$ , and the density function of  $Z$  is  $p_Z(z) = ze^{-z^2/2} \cdot \mathbf{1}_{\{z \geq 0\}}$ .  
 $X$  is a Gaussian process.

*Proof.* (a) Trivial.

(b) Notice that

$$\begin{aligned}\sum_{i=1}^n a_i X_i &= \sum_{i=1}^n a_i \cdot Z \cos(2\pi t_i + \theta) \\ &= \left(\sum_{i=1}^n a_i \cos(2\pi t_i)\right) Z \cos \theta - \left(\sum_{i=1}^n a_i \sin(2\pi t_i)\right) Z \sin \theta.\end{aligned}$$

It suffices to prove  $\begin{pmatrix} Z \cos \theta \\ Z \sin \theta \end{pmatrix}$  is bi-normal.

In fact,  $\begin{pmatrix} Z \cos \theta \\ Z \sin \theta \end{pmatrix} \sim N_2(0, \mathbf{1}_{2 \times 2})$ . We're done.

□

### -Stationarity-

**Definition.**

- $X$  is called strictly stationary if for all  $n \geq 1$ ,  $t_1, \dots, t_n \in T$  and  $h > 0$ :

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{D}}{=} (X_{t_1+h}, \dots, X_{t_n+h})$$

- $X$  is called weakly stationary if for all  $t_1, t_2 \in T$  and  $h > 0$ :

$$\mathbb{E}X_{t_1} = \mathbb{E}X_{t_2} \quad \text{and} \quad \text{Cov}(X_{t_1}, X_{t_2}) = \text{Cov}(X_{t_1+h}, X_{t_2+h}).$$

**Exercise 5.1.5.** Let  $\{X_t\}_{t \in \mathbb{N}}$  be independent RVs such that  $X_t + 1 \sim \exp(1)$  when  $t$  is odd and  $X_t \sim N(0, 1)$  when  $t$  is even. Prove  $X$  is weakly stationary but not strictly stationary.

*Proof.* To prove the weak stationarity, we calculate its mean and covariance:

$$\begin{aligned} \mathbb{E}X_t &= 0 \\ \text{Cov}(X_{t_1}, X_{t_2}) &= \begin{cases} 0 & \forall t_1 \neq t_2 \\ 1 & o.w. \end{cases} \end{aligned}$$

Thus,  $X$  is weakly stationary.

And it is not strictly stationary because  $X_1 \stackrel{\mathcal{D}}{\neq} X_2$ . □

**Exercise 5.1.6.** Let  $S_n = \sum_{i=1}^n X_i$  where  $X_1, X_2, \dots$  are i.i.d. RVs with mean 0 and variance  $\sigma^2$ . Prove  $S$  is not stationary.

*Proof.* Notice  $\text{Cov}(S_{t_1}, S_{t_2}) = \min(t_1, t_2)\sigma^2 \implies$  not stationary. □

**Exercise 5.1.7.** Let  $X_t = Z_t + \theta Z_{t-1}$ ,  $t = 1, 2, \dots$ , where  $Z_0, Z_1, \dots$  are i.i.d RVs with mean 0 and variance  $\sigma^2$ . Prove  $X$  is both weakly stationary and strictly stationary.

*Proof.* It is easy to see the weak stationarity:

$$\mathbb{E}X_t = 0, \quad \text{Var}(X_t) = (1 + \theta^2)\sigma^2, \quad \text{and} \quad \text{Cov}(X_s, X_t) = 0 \text{ for } s \neq t.$$

And we can use the connection between MGFs and fdds to find the strict stationarity. Notice the moment generating function of  $(X_1, \dots, X_n)$  is:

$$\begin{aligned} \mathbb{E} \exp\left(\sum_{i=1}^n \lambda_i X_i\right) &= \mathbb{E} \exp\left(\lambda_n Z_n + (\theta \lambda_n + \lambda_{n-1})Z_{n-1} + \dots + (\theta \lambda_2 + \lambda_1)Z_1 + \theta \lambda_1 Z_0\right) \\ &= \mathbb{E} \exp(\lambda_n Z) \cdot \mathbb{E} \exp[(\theta \lambda_n + \lambda_{n-1})Z] \cdot \dots \cdot \mathbb{E} \exp[(\theta \lambda_2 + \lambda_1)Z] \cdot \mathbb{E} \exp(\theta \lambda_1 Z) \end{aligned}$$

Obviously, it is invariant under index-shift; that is,  $\mathbb{E} \exp(\sum_{i=1}^n \lambda_i X_i) = \mathbb{E} \exp(\sum_{i=1}^n \lambda_i X_{i+h})$ .

Now consider the random vector  $(X_{t_1}, \dots, X_{t_n})$ . It can be divided into several independent disjoint parts in which the index sequence is successive. Then apply for the invariance. □

## 5.2 Continuity of Sample Paths

**Definition 5.5.** Let  $X = (X_t, t \in T)$  be a stochastic process.  $\omega \mapsto X_t$  is called a sample path.

**Theorem 5.6** (Kolmogorov Continuity Criterion). *Let  $(X_t)$  be a stochastic process. If  $\forall T^* > 0$ ,  $\exists \alpha, \beta, C > 0$  such that*

$$\mathbb{E}|X_{t+h} - X_t|^\alpha \leq Ch^{1+\beta} \quad \forall h > 0, \quad 0 < t < T^* - h$$

*then there exists a continuous version of  $X$  on  $[0, T^*]$ .*

**Example 5.7** (Brownian Motion). Let  $(B_t)$  be a Brownian Motion. Take  $\alpha = 4$ .

$$\begin{aligned} \mathbb{E}|B_{t+h} - B_t|^4 &= 3(\mathbb{E}|B_{t+h} - B_t|^2)^2 \\ &= 3h^2 \end{aligned}$$

Note that we use the fact  $\mathbb{E}X^4 = 3(\mathbb{E}X^2)^2$  when  $X \sim N(0, h)$ . Therefore, we can take  $\alpha = 4$ ,  $C = 3$ ,  $\beta = 1$  in the criterion. There exists a continuous version of BM.

## Exercises

**Exercise 5.2.1.** Let  $\{X_t, 0 \leq t \leq 1\}$  be a family of i.i.d  $N(0, 1)$  RVs. Show that  $X$  cannot have a.s. continuous paths.

*Proof. Note:*  $X$  has a.s. continuous paths  $\implies \forall \epsilon > 0, \mathbb{P}\{\omega : \forall t, |X_t - X_{t+\frac{1}{n}}| > 2\epsilon\} \rightarrow 0$  as  $n \rightarrow \infty$ .

Because  $\{X_t > \epsilon, X_{t+\frac{1}{n}} < -\epsilon\} \subset \{|X_t - X_{t+\frac{1}{n}}| > 2\epsilon\}$ , it suffices to prove  $\mathbb{P}\{X_t > \epsilon, X_{t+\frac{1}{n}} < -\epsilon\}$  doesn't converge to 0. And by i.i.d.

$$\begin{aligned} \mathbb{P}\{X_t > \epsilon, X_{t+\frac{1}{n}} < -\epsilon\} &= \mathbb{P}\{X_t > \epsilon\} \mathbb{P}\{X_{t+\frac{1}{n}} < -\epsilon\} \\ &= [1 - \Phi(\epsilon)] \Phi(-\epsilon) \end{aligned}$$

where  $\Phi$  is the CDF of  $N(0, 1)$ . Thus,  $X$  cannot have a.s. continuous paths.  $\square$

**Exercise 5.2.2.** Let  $\{X_t, 0 \leq t \leq 1\}$  be a mean 0 Gaussian process with  $\mathbb{E}(X_{t+h} - X_t)^2 = h^\gamma, \gamma > 0$ . Show that  $X$  must be a.s. continuous.

*Proof. Note:*  $X$  has a.s. continuous paths  $\iff$  for all  $\epsilon > 0, \mathbb{P}\{\omega : \forall t, |X_t - X_{t+h}| > \epsilon\} \rightarrow 0$  as  $h \rightarrow 0$ .

$$\begin{aligned} \mathbb{P}\{\omega : |X_t - X_{t+h}| > \epsilon\} &= \mathbb{P}\left\{\omega : \frac{X_t - X_{t+h}}{h^{\gamma/2}} > \frac{\epsilon}{h^{\gamma/2}}\right\} + \mathbb{P}\left\{\omega : \frac{X_t - X_{t+h}}{h^{\gamma/2}} < -\frac{\epsilon}{h^{\gamma/2}}\right\} \\ &= 1 - \Phi\left(\frac{\epsilon}{h^{\gamma/2}}\right) + \Phi\left(-\frac{\epsilon}{h^{\gamma/2}}\right) \\ &\longrightarrow 0 \quad \text{as } h \rightarrow 0 \end{aligned}$$

Thus,  $X$  has a.s. continuous paths.  $\square$

## 5.3 Stopping Times

**Definition 5.8.** Let  $\mathbb{T}$  be an index set, and  $\mathcal{F}$  be a filtration on  $\mathbb{T}$ .

- A random time  $T : \Omega \rightarrow \mathbb{T} \cup \{+\infty\}$  is called a stopping time of  $\mathcal{F}$  if

$$\{T \leq t\} \in \mathcal{F}_t$$

for all  $t \in \mathbb{T}$ .

- The  $\sigma$ -field of the past before  $T$  is defined as

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \forall t \in \mathbb{T}, A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

*Remark.*

- $T$  is a stopping time of  $\mathcal{F}$  if and only if

$$\{T < t\} \in \mathcal{F}_t$$

for all  $t$ . This is also equivalent to saying that  $T \wedge t$  is  $\mathcal{F}_t$ -measurable for every  $t > 0$ .

- $T$  is a stopping time of  $\mathcal{F}$ , then

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \forall t \in \mathbb{T}, A \cap \{T < t\} \in \mathcal{F}_t\}.$$

## Exercises

**Exercise 5.3.1.** Let  $T$  and  $S$  be stopping times of  $\mathcal{F}$ . Show the following properties:

- $T$  is  $\mathcal{F}_T$ -measurable.
- Let  $A \in \mathcal{F}_\infty$ . Define

$$T^A(\omega) = \begin{cases} T(\omega) & \omega \in A, \\ +\infty & \omega \notin A. \end{cases}$$

Show that  $A \in \mathcal{F}_T$  if and only if  $T^A$  is a stopping time.

(c) If  $S \leq T$ , then  $\mathcal{F}_S \subset \mathcal{F}_T$ .

(d)  $S \vee T$  and  $S \wedge T$  are also stopping times.

Furthermore,  $\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$ ,  $\{S \leq T\} \in \mathcal{F}_{S \vee T}$ , and  $\{S = T\} \in \mathcal{F}_{S \vee T}$ .

(e) Let  $(X_t)$  be a stochastic process. Show that  $\omega \mapsto X_T(\omega)$  defined a  $\mathcal{F}_T$ -measurable map.

*Proof.* (a) Recall that  $\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \forall t \in \mathbb{T}, A \cap \{T \leq t\} \in \mathcal{F}_t\}$ . Because  $\{T \leq t\} \in \mathcal{F}_t$  is given in the definition of the stopping time,  $A = \{T \leq t\} \in \mathcal{F}_T$ .

(b) We just need to notice that

$$\begin{aligned} \{T^A \leq t\} &= (A \cap \{T^A \leq t\}) \cup (A^c \cap \{T^A \leq t\}) \\ &= (A \cap \{T \leq t\}) \cup (A^c \cap \emptyset) \\ &= A \cap \{T \leq t\}. \end{aligned}$$

(c)  $S \leq T$  implies  $\{T \leq t\} \subset \{S \leq t\}$ . So for  $A \in \mathcal{F}_S$ , we have

$$A \cap \{T \leq t\} = (A \cap \{S \leq t\}) \cap \{T \leq t\} \in \mathcal{F}_t.$$

(d) It is easy to see that  $S \vee T$  and  $S \wedge T$  are stopping times:

$$\begin{aligned} \{S \vee T \leq t\} &= \{S \leq t\} \cup \{T \leq t\} \in \mathcal{F}_t \\ \{S \wedge T \leq t\} &= \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t \end{aligned}$$

To show  $\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$ , we firstly notice that  $\mathcal{F}_{S \wedge T} \subset \mathcal{F}_S \cap \mathcal{F}_T$  by (c); thus, it suffices to prove that  $\mathcal{F}_S \cap \mathcal{F}_T \subset \mathcal{F}_{S \wedge T}$ . And notice that

$$A \cap \{S \wedge T \leq t\} = A \cap \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t.$$

To show  $\{S \leq T\} \in \mathcal{F}_{S \vee T}$ , we need to notice that **not done**

$$\begin{aligned} \{S \leq T\} \cap \{T \leq t\} &= \\ \{S \leq T\} \cap \{S \leq t\} &= \end{aligned}$$

And to show  $\{S = T\} \in \mathcal{F}_{S \vee T}$ , note that  $\{S = T\} = \{S \leq T\} \cap \{S \geq T\}$ .

(e)

□

### Exercise 5.3.2.

(a) Let  $(T_n)$  be a monotone sequence of stopping times of  $\mathcal{F}$ , then  $T = \lim T_n$  is also a stopping time.

(b) For any stopping time  $T$ , construct a sequence of stopping times that decreases to  $T$ .

(c) Let  $(T_n)$  be a sequence of stopping times of  $\mathcal{F}$ , then  $\sup T_n$  is also a stopping time; however,  $\inf T_n$  is not, in general.

**Exercise 5.3.3.** Let  $N_t = \sum_{n=1}^{\infty} \mathbf{1}_{[0,t]} \circ T_n$  be a counting process where  $0 = T_0 < T_1 < T_2 < \dots$  are some random times and  $\lim_{n \rightarrow \infty} T_n = +\infty$ . For fixed  $a, b \in \mathbb{R}_+$ , define

$$T = \inf\{t \geq a : N_t = N_{t-a}\}$$

and

$$L = \inf\{t \geq 0 : N_t = N_b\}.$$



- (a) Show that  $T_i$  and  $T$  are stopping times.
- (b) Show that  $L$  is not a stopping time.
- (c) Compute  $\mathbb{P}(T < \infty)$  if  $N$  is a Poisson process with intensity  $\lambda$ .

*Proof.* (a) First, we show that  $T_i$  is a stopping time for all  $i$ . Notice that

$$\{T_i \leq t\} = \{N_t \geq i\} \in \mathcal{F}_t.$$

Then, we show  $T$  is a stopping time.

**Option 1.**

Observing the definition of  $T$ , we notice that for every  $\omega \in \Omega$ ,

$$T = T_k + a$$

for some  $k$ . And  $T = T_k + a$  means the length of the first  $k$  intervals is less than  $a$  and the length of the  $(k+1)$ -th interval is larger than  $a$ ; that is,

$$\{T = T_k + a\} = \{T_1 - T_0 \leq a, \dots, T_k - T_{k-1} \leq a\} \cup \{T_{k+1} - T_k > a\}.$$

Moreover, we have

$$\{T = T_k + a\} \cap \{T_k + a \leq t\} \in \mathcal{F}_t,$$

by noting that

$$\begin{aligned} \{T_1 - T_0 \leq a, \dots, T_k - T_{k-1} \leq a\} \cap \{T_k + a \leq t\} &= \{T_1 \leq a, \dots, T_k \leq ka\} \cap \{T_k \leq t - a\} \\ &\in \mathcal{F}_{(t-a) \wedge a} \subset \mathcal{F}_t. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \{T \leq t\} &= \bigcup_{n \in \mathbb{N}_0} \left( \{T = T_n + a\} \cap \{T \leq t\} \right) \\ &\in \mathcal{F}_t. \end{aligned}$$

**Option 2.**

For each  $t$  we have  $T_k(\omega) \leq t < T_{k+1}(\omega)$  for some  $k$ . Thus, we can define

$$A_t(\omega) = t - T_k(\omega).$$

Then each path of  $A$  is strictly increasing and right-continuous. Moreover,  $A$  is adapted to  $\mathcal{F}$ .

Notice that  $T = \inf\{t \geq 0 : A_t \geq a\}$ . Therefore,  $T$  is a stopping time of  $\mathcal{F}$ . □

**-Galmarino's test-**

**Theorem.** Let  $(X_t)$  be a process,  $\mathcal{F}$  be a  $\sigma$ -algebra generated by  $(X_t)$ , and  $T$  be a RV. Then the following are equivalent

(i)  $T$  is a stopping time w.r.t.  $\mathcal{F}$ .

(ii)  $\forall t \geq 0 \forall \omega, \omega' \in \Omega$

$$T(\omega) \leq t \text{ and } \forall s \leq t \ X_s(\omega) = X_s(\omega') \implies T(\omega) = T(\omega'). \quad (3)$$

**Exercise 5.3.4.** Show the Galmarino's test as follows:

- (a) Define  $a_t : \omega(s) \mapsto \omega(s \wedge t)$ . Show  $\mathcal{F}_t = a_t^{-1}(\mathcal{F})$  for all  $t \geq 0$ .

(b) Let  $Y$  be a RV. Then  $Y$  is measurable w.r.t.  $\mathcal{F}_t$  if and only if

$$Y = Y \circ a_t.$$

(c) Show this theorem.

**Exercise 5.3.5.** Let  $N$  be a Poisson process with intensity  $\lambda$ . Is  $T = \sup\{n \in \mathbb{N} : N_n = 0\}$  a stopping time?

**Exercise 5.3.6.** Let  $(X_t)_{t \in \mathbb{R}_+}$  be a continuous stochastic process started at 0. For any  $b \geq 0$ , define

$$T = \inf\{t \geq 0 : X_t > b\}.$$

Is  $T$  a stopping time?

*Proof.* No, in general. Choose  $\omega$  and  $\omega'$  such that  $X_t(\omega) = X_t(\omega') = b$ ,  $X_{t+}(\omega) > b$  and  $X_{t+}(\omega') < b$ . Then  $T(\omega) = t$  while  $T(\omega') > t$ .

Figure 2: To be added

□

## 5.4 Total Variation and Quadratic Variation

**Definition 5.9.**

- Let  $[a, b]$  be an interval of  $\mathbb{R}$ . A countable family  $\mathcal{A} \subset 2^{[a, b]}$  is called a subdivision of  $[a, b]$  if
  - a) Every element in  $\mathcal{A}$  is of form  $(s, t]$ .
  - b)  $\cup \mathcal{A} = (a, b]$ .
  - c) If  $U_1 \neq U_2$  in  $\mathcal{A}$ , then  $U_1 \cap U_2 = \emptyset$ .
- Let  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  be right-continuous,  $p > 0$ , and  $\mathcal{A}$  be a subdivision of  $[a, b]$ . Define

$$\mathcal{V}_p = \sup_{\mathcal{A}} \sum_{(s, t] \in \mathcal{A}} |f(t) - f(s)|^p.$$

It is called the  $p$ -variation of  $f$  on  $[a, b]$ . For  $p = 1$ , it is called the total variation. For  $p = 2$ , it is called the quadratic variation.

- $f$  is said to have bounded  $p$ -variation if  $\mathcal{V}_p < \infty$ .

**Definition 5.10** (Quadratic variation for real-valued processes). Let  $Y$  be a real-valued process. Define

$$\langle Y, Y \rangle_t = \lim_{\delta^{(n)} \rightarrow 0} \sum_{i=1}^{K_n-1} |Y_{t_{i+1}^{(n)}} - Y_{t_i^{(n)}}|^2,$$

where the limit is defined using convergence in probability.

**Example 5.11** (Quadratic variation for Brownian motion). We will see that  $\langle B, B \rangle_t = t$ . Define

$$S_n = \sum_i |B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}|^2$$

and take  $\{\Delta^{(n)}\}$  as a sequence of partitions of interval of  $[a, b]$ .

- (i) Because  $B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}} \sim N(0, t_{i+1}^{(n)} - t_i^{(n)})$ , we have

$$\mathbb{E} S_n = \sum_i \mathbb{E} |B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}|^2 = \sum_i (t_{i+1}^{(n)} - t_i^{(n)}) = b - a.$$

(ii) If  $X \sim N(0, t)$ ,  $\mathbb{E}X^4 = 3t^2$ ; and  $(B_t)$  is a Lévy process. Thus,

$$\begin{aligned}
\mathbb{E}S_n^2 &= \mathbb{E}\left(\sum_i |B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}|^2\right)^2 \\
&= \mathbb{E}\sum_i |B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}|^4 + 2\sum_{i < j} \mathbb{E}\left\{(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2(B_{t_{j+1}^{(n)}} - B_{t_j^{(n)}})^2\right\} \\
&= \sum_i 3(t_{i+1}^{(n)} - t_i^{(n)})^2 + 2\sum_{i < j} (t_{i+1}^{(n)} - t_i^{(n)})(t_{j+1}^{(n)} - t_j^{(n)}) \\
&= 2\sum_i (t_{i+1}^{(n)} - t_i^{(n)})^2 + \sum_i \sum_j (t_{i+1}^{(n)} - t_i^{(n)})(t_{j+1}^{(n)} - t_j^{(n)}) \\
&= 2\sum_i (t_{i+1}^{(n)} - t_i^{(n)})^2 + (b - a)^2 \\
&\leq 2\delta^{(n)}(b - a) + (b - a)^2 \\
&\longrightarrow (b - a)^2
\end{aligned}$$

(iii) Then we have

$$\mathbb{E}(S_n - (b - a))^2 = \text{Var}S_n^2 = \mathbb{E}S_n^2 - (b - a)^2 \rightarrow 0,$$

which implies that  $S_n \rightarrow b - a$  in  $L^2$  and in probability. To prove  $\langle B, B \rangle_t = t$ , we need to choose  $[a, b] = [0, t]$ .

**Example 5.12** (Total variation for Brownian motion). For almost every  $\omega \in \Omega$ , we will see the path  $\omega := W(\omega)$  has infinite total variation, where  $W$  is a Brownian motion.

Let  $v^*$  be the total variation of  $\omega$ ; that is

$$v^* = \sup_{\mathcal{A}} \sum_{(s,t] \in \mathcal{A}} |\omega(s) - \omega(t)|.$$

Let  $\mathcal{A}_n$  be a sequence of subdivision. Then notice that

$$\begin{aligned}
\sum |\omega(t) - \omega(s)|^2 &\leq \sup |\omega(t) - \omega(s)| \sum |\omega(t) - \omega(s)| \\
&\leq \sup |\omega(t) - \omega(s)| \cdot v^*
\end{aligned}$$

Letting  $n \rightarrow \infty$ , by computing its quadratic variation, we have

$$\sum |\omega(t) - \omega(s)|^2 \rightarrow b - a,$$

and by the continuity of sample path, we have

$$\sup |\omega(t) - \omega(s)| \rightarrow 0.$$

Therefore,  $v^*$  cannot be finite.

## Exercises

**Exercise 5.4.1.** For each  $n \in \mathbb{N}$ , let  $\mathcal{A}$  be the subdivision of  $[a, b]$  that consists of  $2^n$  intervals of the same length. Then

$$V_n = \sum_{(s,t]} |B_t - B_s|^2$$

converges to  $b - a$  almost surely.

**Exercise 5.4.2.**

(a) Find the total variation and the quadratic variation of the Poisson process  $N$ .

- (b) Find the quadratic variation of  $N_t - \lambda t$ .
- (c) Find the quadratic variation of  $N_{\int_0^t f(s)ds}$ .

*Proof.* (a) First, we compute its **total variation**.

Let  $\{\mathcal{A}_n\}$  be a sequence subdivision of  $[a, b]$  such that  $\|\mathcal{A}_n\| \rightarrow 0$ . Define

$$S_n = \sum_{(s,t] \in \mathcal{A}_n} |N_t - N_s|.$$

Simplify it:

$$S_n = N_b - N_a.$$

Letting  $n \rightarrow \infty$  and  $[a, b] = [0, t]$ , the total variation of  $N$  is  $N$  itself.

Second, we compute its **quadratic variation**.

Notice that

$$[N]_t = \sum_{0 \leq s \leq t} (\Delta X_s)^2$$

where  $\Delta X_s := X_s - X_{s-}$  is the jump of  $N$  at  $s$ . And re-write  $N$  as

$$N_t = \sum_{i=1}^{\infty} \mathbf{1}_{[0,t]} \circ T_i.$$

Then  $\Delta N_t = \sum_{i=1}^{\infty} \mathbf{1}_{\{t\}} \circ T_i$ . Therefore,

$$\begin{aligned} [N]_t &= \sum_{0 \leq s \leq t} (\Delta N_s)^2 = \sum_{0 \leq s \leq t} \Delta N_s \\ &= \sum_{0 \leq s \leq t} \left( \sum_{i=1}^{\infty} \mathbf{1}_{\{s\}} \circ T_i \right) = \sum_{i=1}^{\infty} \left( \sum_{0 \leq s \leq t} \mathbf{1}_{\{s\}} \circ T_i \right) \\ &= \sum_{i=1}^{\infty} \mathbf{1}_{[0,t]} \circ T_i = N_t \end{aligned}$$

- (b) It is easy to see

$$[N_t - \lambda t]_t = \sum_{0 \leq s \leq t} (\Delta(N_s - \lambda s))^2 = \sum_{0 \leq s \leq t} (\Delta N_s)^2 = N_t.$$

- (c) ...

□

## 6 Discrete and Continuous Time Markov Chains

### 6.1 Characterization

**Definition 6.1.** Let  $(\Omega, \mathcal{H}, \mathbb{P})$  be a probability space, and  $\mathcal{F} = (\mathcal{F}_t)$  be a filtration on it. Let  $X = (X_t)$  be a stochastic process with some state space  $(E, \mathcal{E})$  and be adapted to  $\mathcal{F}$ .

- $X$  is Markovian if for every  $t, u > t$ , and  $f \in \mathcal{E}_+$ ,

$$\mathbb{E}(f \circ X_u | \mathcal{F}_t) = \mathbb{E}(f \circ X_u | X_t).$$

- For  $t \leq u$ , define the transition function on  $(E, \mathcal{E})$  as

$$P_{t,u}(x, A) = \mathbb{P}(X_u \in A | X_t = x).$$

- Chapman-Kolmogorov equation. As the product of transition kernels,

$$P_{s,t}P_{t,u} = P_{s,u}.$$

- Let  $X$  be Markovian and admit  $(P_{t,u})$  as its transition function, then it is called time-homogeneous if  $P_{t,u}$  is only dependent on  $u - t$ ; that is

$$P_{t,u} = P_{u-t}.$$

*Remark.* In the time-homogeneous case, we can re-write Chapman-Kolmogorov equation as

$$P_t P_u = P_{t+u}, \quad t, u \in \mathbb{R}_+;$$

and usually, we call  $X$  a Markov process with transition function  $(P_t)$ .

And suppose  $X$  is a Markov chain, then  $Q = P_{t,t+1}$  is free of  $t$ . Thus,

$$P_{t,u} = Q^n, \quad u - t = n \in \mathbb{N};$$

and usually, we call  $X$  a Markov chain with transition kernel  $Q$ .

**Example 6.2** (Chapman-Kolmogorov equations are not sufficient). content...

### Exercises

#### Exercise 6.1.1.

- Let  $Y$  be a Lévy process and  $\pi_t$  be the distribution of  $Y_t$ . Define  $X = X_0 + Y$  where  $X_0$  is a RV independent of  $Y$ . Show  $X$  is a Markov process and compute its transition function.
- Give an example of Markov process which doesn't have independent increments.

*Proof.* (a) Let  $A - x = \{a - x : a \in A\}$ . We have

$$\begin{aligned} & \mathbb{P}(X_{t_n} \in A | X_{t_{n-1}} = s_{n-1}, \dots, X_{t_1} = s_1) \\ &= \mathbb{P}(X_{t_n} - X_{t_{n-1}} \in A - s_{n-1} | X_{t_{n-1}} = s_{n-1}, X_{t_{n-1}} - X_{t_{n-2}} = s_{n-1} - s_{n-2}, \dots, X_{t_{n-1}} - X_{t_1} = s_{n-1} - s_1) \\ &= \mathbb{P}(Y_{t_n} - Y_{t_{n-1}} \in A - s_{n-1} | X_{t_{n-1}} = s_{n-1}, Y_{t_{n-1}} - Y_{t_{n-2}} = s_{n-1} - s_{n-2}, \dots, Y_{t_{n-1}} - Y_{t_1} = s_{n-1} - s_1) \\ &= \mathbb{P}(Y_{t_n} - Y_{t_{n-1}} \in A - s_{n-1} | X_{t_{n-1}} = s_{n-1}) \quad (\text{indep. increments}) \\ &= \mathbb{P}(X_{t_n} \in A | X_{t_{n-1}} = s_{n-1}) \end{aligned}$$

Therefore,  $X$  is a Markov process. And it is easy to see

$$P_t(A, x) = \mathbb{P}(X_{u+t} \in A | X_u = x) = \mathbb{P}(Y_{t+u} - Y_u \in A - x | X_u = x) = \pi_t(A, x).$$

is the transition function of  $X$ .

(b) Ornstein–Uhlenbeck process.

□

**Exercise 6.1.2.** Show the following processes are Markovian, compute the transition semigroups, and determine the time-homogeneity.

(a)

(b)

(c)

**Exercise 6.1.3.** Let  $(X_n)$  be a time-homogeneous Markov chain with a discrete state space  $(E, \mathcal{E})$  and with a transition kernel  $Q$ . If the distribution of  $X_n$  is  $\mu_n$ , find the distribution of  $X_{n+k}$ .

*Proof.* content...

□

### -Random dynamical systems-

**Exercise 6.1.4.** content...

## 6.2 Markov Chains: Classification of States

In this section, we are only interested in the time-homogeneous Markov chains with a discrete state space. In this case, we can re-write  $P_t$  as

$$P_t(x, A) = \mathbb{P}(X_t \in A \mid X_0 = x) = \sum_{y \in A} \mathbb{P}(X_t = y \mid X_0 = x).$$

**Definition 6.3.** Let  $(X_n)$  be a time-homogeneous Markov chain with a discrete state space.

- The transition matrix is defined as  $\mathbf{P} = (p_{ij})$ , where

$$p_{ij} = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

means the probability from the state  $i$  to  $j$  in one step.

- The  $k$ -step transition matrix is defined as  $\mathbf{P}(\mathbf{n}, \mathbf{n} + \mathbf{k}) = (p_{ij}(n, n + k))$ , where

$$p_{ij}(n, n + k) = \mathbb{P}(X_{n+k} = j \mid X_n = i)$$

means the probability from the state  $i$  to  $j$  in  $k$  steps.

*Remark.* With this notation, we can re-write the transition kernel as

$$Q(i, J) = \sum_{j \in J} p_{ij};$$

so obviously,  $\mathbf{P}^k = \mathbf{P}(\mathbf{n}, \mathbf{n} + \mathbf{k})$ , which is the matrix form of C-K equation.

**Definition 6.4.** Let  $(X_n)$  be a time-homogeneous Markov chain with a discrete state space.

- Let  $\mu_i^{(n)} = \mathbb{P}(X_n = i)$  be the mass function of  $X_n$ . And they form a row vector  $\boldsymbol{\mu}^{(n)}$ .
- The probability that the first visit to state  $j$ , starting from  $i$ , takes place at the  $n$ th step, is

$$f_{ij}(n) = \mathbb{P}(X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j \mid X_0 = i).$$

- The probability that the chain ever visits  $j$  is

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n).$$

*Remark.* The transition matrix  $\mathbf{P}$  and the initial mass function  $\boldsymbol{\mu}^{(0)}$  determine the chain. Many questions about the chain can be expressed in terms of these quantities.

**Definition 6.5.** Let  $(X_n)$  be a time-homogeneous Markov chain with a discrete state space.

a) For any state  $i$ :

- State  $i$  is called persistent (or recurrent) if

$$\mathbb{P}(X_n = i, \exists n \geq 1 \mid X_0 = i) = 1.$$

- State  $i$  is called transient if

$$\mathbb{P}(X_n = i, \exists n \geq 1 \mid X_0 = i) < 1.$$

b) Define  $T_j = \min\{n \geq 1 : X_n = j\}$  be the time of the first visit to  $j$ . For a **persistent state**  $i$ :

- State  $i$  is called null if

$$\mu_i := \mathbb{E}(T_i \mid X_0 = i) = \infty.$$

- State  $i$  is called non-null (or positive) if

$$\mu_i := \mathbb{E}(T_i \mid X_0 = i) < \infty.$$

c) Define  $d(i) = \gcd\{n : p_{ii}(n) > 0\}$ .

- State  $i$  is called periodic if

$$d(i) > 1.$$

- State  $i$  is called aperiodic if

$$d(i) = 1.$$

d) A state is called ergodic if it is persistent, non-null, and aperiodic.

**Theorem 6.6.** Let  $(X_n)$  be a time-homogeneous Markov chain with a discrete state space.

a) For any state  $j$ :

- State  $j$  is persistent if

$$\sum_n p_{jj}(n) = \infty.$$

- State  $j$  is transient if

$$\sum_n p_{jj}(n) < \infty.$$

b) A persistent state is null if and only if  $p_{ii}(n) \rightarrow 0$  as  $n \rightarrow \infty$ .

### 6.3 Markov Chains: Classification of Chains

In this section, we are only interested in the Markov chains with a discrete state space.

**Definition 6.7.**

- We say  $i$  communicates with  $j$ ,  $i \rightarrow j$ , if  $p_{ij}(m) > 0$  for some  $m \geq 0$ .
- We say  $i$  intercommunicates with  $j$ ,  $i \leftrightarrow j$ , if  $i \rightarrow j$  and  $j \rightarrow i$ .

*Remark.* If  $i \leftrightarrow j$ , then

- They have the same period.
- $i$  is transient  $\iff j$  is transient.
- $i$  is null persistent  $\iff j$  is null persistent.

**Definition 6.8.** A set  $C$  of states is called:

- closed if  $p_{ij} = 0$  for all  $i \in C$ ,  $j \notin C$ ,
- irreducible if  $i \leftrightarrow j$  for all  $i, j \in C$ .

## 6.4 Stationary distributions and the limit theorem

## 6.5 Markov Chains: Examples

In this section, we will mainly focus on the Markov chains with a discrete state space; however, some examples of Markov chains with a non-discrete state space are also given.

**Proposition 6.9.** *content...*

*Remark.* If the state space  $(E, \mathcal{E})$  is standard, we can construct every Markov chain in this way. And  $\varphi$  is called the structure function and  $(Z_n)$  is called the driving variables.

**Example 6.10** (Random walks). Let  $E = D = \mathbb{R}^d$ , and  $\varphi(x, z) = x + z$ . Then

$$\begin{aligned} X_{n+1} &= \varphi(X_n, Z_{n+1}) \\ &= X_n + Z_{n+1} \\ &= \sum_{i=1}^{n+1} Z_i \end{aligned}$$

is called a random walk on  $\mathbb{R}^d$ .

**Example 6.11** (Gauss-Markov chains). Let  $E = D = \mathbb{R}^d$ , and  $\varphi(x, z) = Ax + Bz$ , where  $A$  and  $B$  are some  $d \times d$  matrices. Assume  $Z_i \sim N_d(0, \mathbf{I}_{d \times d})$  is a sequence of i.i.d. RVs. Then the resulting chain is called a Gauss-Markov chain.

**Example 6.12** (Product of random matrices). *content...*

## 6.6 Markov Processes: Examples

**Example 6.13** (Markov chains subordinated to Poisson). Let  $(Y_n)_{n \in \mathbb{N}}$  be a Markov chain with state space  $(E, \mathcal{E})$  and transition kernel  $Q$ . Let  $(N_t)$  be a Poisson process with rate  $c$ , independent of  $(Y_n)$ . Suppose that

$$X_t = Y_{N_t}, \quad t \in \mathbb{R}_+.$$

Then,  $X$  is a Markov process, by the strong Markov property of  $Y$ .  
And the transition function is

$$\begin{aligned} P_t(x, A) &= \mathbb{P}(X_t \in A \mid X_0 = x) \\ &= \mathbb{P}(Y_{N_t} \in A \mid Y_0 = x) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(Y_n \in A \mid Y_0 = x, N_t = n) \cdot \mathbb{P}(N_t = n \mid Y_0 = 0) \\ &= \sum_{n=0}^{\infty} \frac{e^{-ct}(ct)^n}{n!} Q^n(x, A). \end{aligned}$$

**Example 6.14** (Delayed uniform motion). Let  $T$  be an exponentially distributed RV. Define  $X_t = (t - T)^+$ .  $X$  can be depicted as the motion of a particle. Firstly, this particle stays at the origin for  $T$  of time. Then, it moves upward at unit speed.

First, we show that  $X$  is a Markov process.

Then, we compute its transition kernel.

Finally, we can see  $X$  has no strong Markov property.



## 7 Poisson Processes

### 7.1 Characterization

**Definition 7.1.**

- **Counting processes**

$N = (N_t)_{t \in \mathbb{R}_+}$  is called a counting process, if for a.e.  $\omega$ , the path  $t \mapsto N_t(\omega)$  is an increasing right-continuous step function with  $N_0(\omega) = 0$  and whose every jump size is one.

- **Poisson processes**

$N = (N_t)_{t \in \mathbb{R}_+}$  is called a Poisson process with rate  $c$  if it is a counting process, and  $\forall s, t \in \mathbb{R}_+$ ,  $N_{s+t} - N_s \sim \text{Poisson}(ct)$  is independent of  $\mathcal{F}_s$ .

*Remark.*  $N$  can be written as

$$N_t(\omega) = \sum_{k=1}^{\infty} \mathbf{1}_{[0, t]} \circ T_k(\omega),$$

where  $T$  is an increasing sequence of RVs.

**Theorem 7.2.** For fixed  $c \in (0, \infty)$ , the following are equivalent:

a)  $N$  is a Poisson counting process with rate  $c$ .

b) **Martingale connection**

$N$  is a counting process and  $\tilde{N} = (N_t - ct)_{t \in \mathbb{R}_+}$  is a martingale.

c)  $(T_k)$  is an increasing sequence of stopping times, and the differences  $(T_k - T_{k-1})_{k \in \mathbb{N}}$  are i.i.d. with the distribution  $\text{Exp}(c)$ .

d) **Characterization as a Lévy process.**

$N$  is both a counting process and a Lévy process.

*Proof.*

- $a) \Rightarrow b)$ : Notice that for  $s \leq t$ ,

$$\begin{aligned} \mathbb{E}(\tilde{N}_t | \mathcal{F}_s) &= \mathbb{E}(N_t - ct | \mathcal{F}_s) \\ &= \mathbb{E}(N_t - N_s | \mathcal{F}_s) + N_s - ct \\ &= c(t - s) + N_s - ct = \tilde{N}_s. \end{aligned}$$

Therefore,  $\tilde{N}$  is a martingale.

$b) \Rightarrow c)$ : First, we notice that  $\{N_{T_k+t} - N_{T_k} = 0\} = \{T_{k+1} - T_k > t\}$ . (Recall that  $T_k$  means the arrival time of the  $k$ th visitor.)

By the strong Markov property,  $N_{T_k+t} - N_{T_k}$  is a Poisson distribution with rate  $ct$ , and is independent of  $\mathcal{F}_{T_k}$ . And

$$\mathbb{P}(N_{T_k+t} - N_{T_k} = 0) = e^{-ct};$$

that is,

$$\mathbb{P}(T_{k+1} - T_k \leq t) = 1 - e^{-ct},$$

which is the CDF of the exponential distribution with parameter  $c$ .

- $c) \Rightarrow a)$ :
- $a) \Rightarrow d)$ : Trivial.
- $d) \Rightarrow a)$ :

□

**Proposition 7.3.** Let  $N$  be a Poisson process with intensity  $\lambda$ . Then

$$a) (T_1, \dots, T_n) \stackrel{D}{=} (U_{(1)}, \dots, U_{(n)}).$$

b)

## Exercises

### Exercise 7.1.1.

- (a) Let  $N$  be a Poisson process with rate  $c$ . Compute the mean and the covariance function of  $N$ .
- (b) Let  $N^{(1)}$  and  $N^{(2)}$  be two independent Poisson processes with rate  $c_1$  and  $c_2$ . Prove  $N^{(1)} + N^{(2)}$  is a Poisson process with rate  $c_1 + c_2$ .
- (c)

**Exercise 7.1.2.** Let  $N^{(1)}$  and  $N^{(2)}$  be two independent Poisson processes with rate  $c_1$  and  $c_2$ . Define  $X_t = N_t^{(1)} - N_t^{(2)}$ .

- (a) Compute the mean and the covariance function of  $X$ .
- (b) Is  $X$  a Poisson process?
- (c) Is  $X$  a Lévy process?
- (d) Is  $X$  a Markov process? If so, compute its transition semigroup.
- (e) Is  $X$  a martingale?

*Proof.*

- (a)  $\mathbb{E}X_t = (c_1 - c_2)t$ ; and  $\mathbb{E}X_t X_s = (c_1 + c_2) \cdot t \wedge s$ .
- (b) No, obviously.
- (c) Yes. If  $X$  is a Lévy process, then  $cX$  is a Lévy process. If  $X$  and  $Y$  are two independent Lévy processes, then  $X + Y$  is a Lévy process.
- (d) Yes, even better. Because  $X$  is a Lévy process, it has strong Markov property.
- (e) Let  $\mathcal{F}$  be the natural filtration generated by  $X$ .

$$\begin{aligned} \mathbb{E}(X_t | \mathcal{F}_s) &= \mathbb{E}(X_t - X_s | \mathcal{F}_s) + X_s \\ &= \mathbb{E}((N_t^{(1)} - N_s^{(1)}) - (N_t^{(2)} - N_s^{(2)}) | \mathcal{F}_s) + X_s \\ &= (c_1 - c_2)(t - s) + X_s \end{aligned}$$

Therefore,  $X$  is a martingale if and only if  $c_1 = c_2$ .

□

**Exercise 7.1.3.** Let  $N$  be a Poisson process with intensity  $c$  and  $\mathcal{F}$  be the filtration generated by  $N$ . Show that (a)  $(N_t - ct)^2 - ct$  is a martingale w.r.t  $\mathcal{F}$ . (b)  $\exp(\ln(1 - u)N_t + uct)$ , for any  $u \in (0, 1)$ , is a martingale w.r.t  $\mathcal{F}$ .

*Proof.* Use  $N_t = (N_t - N_s) + N_s$ ; then  $N_t - N_s$  is independent of  $\mathcal{F}_s$ , and  $N_s$  is measurable w.r.t  $\mathcal{F}_s$ .

(a) First, notice that

$$\mathbb{E}(N_t^2 | \mathcal{F}_s) = c(t-s) + N_s^2 - 2N_s \cdot c(t-s).$$

Therefore,

$$\begin{aligned} \mathbb{E}((N_t - ct)^2 - ct | \mathcal{F}_s) &= \mathbb{E}(N_t^2 | \mathcal{F}_s) - 2ct \cdot \mathbb{E}(N_t | \mathcal{F}_s) + c^2 t^2 - ct \\ &= (N_s - cs)^2 - cs \end{aligned}$$

(b) We use  $N_t = (N_t - N_s) + N_s$  again.

$$\begin{aligned} \mathbb{E}\left(\exp(\ln(1-u)N_t + uct) \middle| \mathcal{F}_s\right) &= \mathbb{E}\left(\exp(\ln(1-u)(N_t - N_s) + \ln(1-u)(N_s)) \cdot \exp(uct) \middle| \mathcal{F}_s\right) \\ &= e^{uct} \cdot e^{\ln(1-u)N_s} \cdot \mathbb{E}\left(e^{\ln(1-u)(N_t - N_s)} \middle| \mathcal{F}_s\right) \\ (\text{Use the MGF of Poisson dist.}) &= e^{uct} \cdot e^{\ln(1-u)N_s} \cdot e^{c(t-s)(-u)} \\ &= \exp(\ln(1-u)N_s + ucs) \end{aligned}$$

□

## 7.2 Strong Markov Properties

**Theorem 7.4.** Let  $N$  be a Poisson process with intensity  $c$ , and  $S$  be a finite stopping time. Then

$$\mathbb{E}\left(e^{-r \cdot (N_{S+t} - N_S)} \middle| \mathcal{F}_S\right) = \sum_{k=0}^{\infty} \frac{e^{-ct}(ct)^k}{k!} e^{-rk},$$

i.e.  $N_{S+t} - N_S$  is independent of  $\mathcal{F}_S$  and  $N_{S+t} - N_S \stackrel{D}{=} N_t$ .

### Exercises

**Exercise 7.2.1** (Total unpredictability of jumps). Let  $S$  be a stopping time such that  $0 \leq S < T$  almost surely. Then  $S = 0$  almost surely.

*Proof.* content...

□

## 7.3 Compound Poisson

**Definition 7.5.** Let  $N$  be a Poisson process with intensity  $\lambda$ , and  $Y$  be a sequence of i.i.d. RVs. Then

$$X_t = \sum_{k=1}^{N_t} Y_k$$

is called a compound Poisson process.

**Proposition 7.6.** content...

### Exercises

**Exercise 7.3.1.** .

## 8 Brownian Motion

### 8.1 Basic Properties

**Definition 8.1.** The following two definitions are equivalent:

- A stochastic process  $X$  is called a Brownian motion if
  - a)  $X$  is continuous;
  - b)  $X$  has stationary independent increments;
  - c)  $B_0 = 0$ ,  $\mathbb{E}B_t = 0$ , and  $\mathbb{E}B_t^2 = t$ .
- A stochastic process  $X$  is called a Brownian motion if
  - a)  $X$  is continuous;
  - b)  $X$  is Gaussian;
  - c)  $B_0 = 0$ ,  $\mathbb{E}B_t = 0$ , and  $\text{Cov}(B_t, B_s) = \min(s, t)$ .

**Proposition 8.2.** Let  $(B_t)_{t \in \mathbb{R}^+}$  be a BM. Then the following processes are also BMs:

- *Symmetry.*  $(-B_t)_{t \in \mathbb{R}^+}$
- *Scaling.*  $(c^{-1/2}B_{ct})_{t \in \mathbb{R}^+}$
- *Time inversion.*  $(tB_{1/t})_{t \in \mathbb{R}^+}$

### Exercises

**Exercise 8.1.1.** Show  $Z_t = B_{T-t} - B_T$  is a BM, where  $T$  is a constant and  $0 \leq t \leq T$ .

*Proof.*

- $Z_0 = 0$ ,  $\mathbb{E}Z_t = 0$ , and  $\mathbb{E}Z_t^2 = \mathbb{E}B_{T-t}^2 - 2\mathbb{E}B_{T-t}B_T + \mathbb{E}B_T^2 = (T-t) - 2\min(T-t, T) + T = t$ .
- Continuous paths:  $B_{T-t} - B_T$  is continuous in  $t$ .
- Independent increments: Assume  $t_0 < t_1 < \dots < t_n$ ,

$$\begin{aligned} Z_{t_n} - Z_{t_{n-1}} &= B_{T-t_n} - B_{T-t_{n-1}} \\ Z_{t_{n-1}} - Z_{t_{n-2}} &= B_{T-t_{n-1}} - B_{T-t_{n-2}} \\ &\dots \\ Z_{t_1} - Z_{t_0} &= B_{T-t_1} - B_{T-t_0} \end{aligned}$$

Because  $B_{T-t_n} - B_{T-t_{n-1}}, \dots, B_{T-t_1} - B_{T-t_0}$  are independent,

$\implies Z_{t_n} - Z_{t_{n-1}}, \dots, Z_{t_1} - Z_{t_0}$  are independent.

- Stationary increments:  $Z_{t+u} - Z_t = B_{T-t-u} - B_{T-t} \sim N(0, u)$ .

□

**Exercise 8.1.2.** Show that  $2B_t - B_s$  for  $t > s$  is not independent of  $\mathcal{F}_s$ .

*Proof.* We can check the independence by compute its conditional expectation given  $\mathcal{F}_s$ .

$$\begin{aligned} \mathbb{E}(2B_t - B_s \mid \mathcal{F}_s) &= \mathbb{E}(B_t \mid \mathcal{F}_s) + \mathbb{E}(B_t - B_s \mid \mathcal{F}_s) \\ &= B_s + \mathbb{E}(B_t - B_s) \\ &= B_s \neq 2B_t - B_s \end{aligned}$$

□

**Exercise 8.1.3.** Calculate  $\mathbb{P}(B_t \leq 0, t = 0, 1, 2)$ .

*Proof.* We notice that  $\mathbb{P}(B_t \leq 0, t = 0, 1, 2) = \mathbb{P}(B_1 \leq 0, B_2 \leq 0)$ . Thus, it suffices to compute the joint distribution of  $B_1$  and  $B_2$ .

$$\begin{aligned} \text{Because } \begin{pmatrix} B_2 - B_1 \\ B_1 \end{pmatrix} &\sim N_2(0, \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}), \\ \implies \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} &= \begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} B_2 - B_1 \\ B_1 \end{pmatrix} \sim N_2(0, \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}). \end{aligned}$$

Therefore,

$$\mathbb{P}(B_1 \leq 0, B_2 \leq 0) = \int_0^\infty \int_0^\infty \frac{1}{2\pi} e^{-(x^2 + \frac{y^2}{2} - xy)} dx dy = \frac{3}{8}.$$

□

**Exercise 8.1.4.** Let  $T$  be an exponential RV with parameter  $\lambda$  independent of  $B$ . Compute the characteristic function of  $B_T$ .

*Proof.*

$$\begin{aligned} \mathbb{E}e^{isB_T} &= \mathbb{E}(\mathbb{E}(e^{isB_T} | T)) \\ &= \mathbb{E}(e^{-Ts^2/2}) \\ &= \int_0^\infty \lambda e^{-\lambda t} \cdot e^{-ts^2/2} dt \\ &= \frac{1}{1 + \frac{s^2}{2\lambda}} \end{aligned}$$

It is easy to see it is the CF of Laplace distribution with mean 0 and variance  $\frac{1}{\lambda}$ .

□

**-Time inversion-**

**Exercise 8.1.5.** Show  $(tB_{1/t})_{t \in \mathbb{R}^+}$  is a BM.

**Exercise 8.1.6.** For  $0 < s < t$ , find the distribution of  $W_s | W_t = x$ .

*Proof.*

• **Option 1**

First, we should notice that  $(W_s | W_t = x) \stackrel{D}{=} (\bar{W}_s | \bar{W}_t = x)$ , where  $\bar{W}_t = tW_{1/t}$ . It is easy to check by considering their fdds. Then, it suffices to find the distribution of  $\bar{W}_s | \bar{W}_t = x$

$$\begin{aligned} \mathbb{P}(\bar{W}_s \leq y | \bar{W}_t = x) &= \mathbb{P}(\bar{W}_s - sW_{1/t} + sW_{1/t} \leq y | \bar{W}_t = x) \\ &= \mathbb{P}(sW_{1/s} - sW_{1/t} + sW_{1/t} \leq y | W_{1/t} = x/t) \\ &= \mathbb{P}(s(W_{1/s} - W_{1/t}) + sx/t \leq y | W_{1/t} = x/t) \\ (\text{indep. incre.}) \quad &= \mathbb{P}(s(W_{1/s} - W_{1/t}) + sx/t \leq y) \end{aligned}$$

$$\text{Because } W_{1/s} - W_{1/t} \sim N(0, \frac{1}{s} - \frac{1}{t}), \implies \bar{W}_s | \bar{W}_t = x \sim N(\frac{sx}{t}, s^2(\frac{1}{s} - \frac{1}{t}))$$

• **Option 2**

We know for  $0 < s < t$ ,  $(W_s, W_t) \sim N_2(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} s & s \\ s & t \end{pmatrix})$ . Then we directly know

$$W_s | W_t \sim N(\frac{s}{t}W_t, s^2(\frac{1}{s} - \frac{1}{t})).$$

□

## 8.2 Martingale Connection

**Theorem 8.3.** Let  $B$  be a continuous process starting at  $B_0 = 0$ . The following are equivalent:

- a)  $B = (B_t)_{t \in \mathbb{R}_+}$  is a BM w.r.t  $\mathcal{F}$ .
- b) For all  $r \in \mathbb{R}$ ,  $\{\exp(rB_t - \frac{1}{2}r^2t)\}_{t \in \mathbb{R}_+}$  is a  $\mathcal{F}$ -martingale.
- c)  $\{B_t^2 - t\}_{t \in \mathbb{R}_+}$  is a  $\mathcal{F}$ -martingale.
- d) For every twice-differentiable function  $f : \mathbb{R} \rightarrow \mathbb{R}$  that is bounded along with its first derivative  $f'$  and second derivative  $f''$ , the process

$$M_t = f \circ B_t - \frac{1}{2} \int_0^t f'' \circ B_s ds \quad t \in \mathbb{R}_+$$

is an  $\mathcal{F}$ -martingale.

*Proof.*

- a)  $\Rightarrow$  b). Trivial.
- b)  $\Rightarrow$  a). To show  $B$  is a BM, it suffices to find the distribution of  $B_{s+t} - B_s$  where  $t > 0$ .

$$\mathbb{E}(\exp(r(B_{s+t} - B_s)) \mid \mathcal{F}_s) = \exp(\frac{1}{2}r^2t)$$

holds for all  $r \in \mathbb{R}_+$ . It is the MGF of  $N(0, t)$ .

- a)  $\Rightarrow$  c). Trivial.
- c)  $\Rightarrow$  a). Omitted. It is well-known as the Lévy characterization.
- a)  $\Leftrightarrow$  d). Omitted.

□

## 8.3 Strong Markov Properties (SMP)

**Definition 8.4.** Given a stochastic process  $(X_t)$ , For all  $A \in \mathcal{B}(\mathbb{R})$  and for all  $t > 0$ , define

$$\mathcal{P}_t(x, A) = \mathbb{P}(X_{t+s} \in A \mid X_s = x).$$

It is called the transition function of  $(X_t)$ .

*Remark.* When  $(X_t)$  is a time-homogeneous Markov process,  $(\mathcal{P}_t)$  forms a semigroup. The relation

$$\mathcal{P}_s \mathcal{P}_t = \mathcal{P}_{s+t}$$

is called Chapman-Kolmogorov equations.

**Proposition 8.5.**

- a)  $(B_t)$  is a time-homogeneous Markov process with the transition function

$$\int_A \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/2t} dy.$$

- b)  $\mathbb{E}(f(B_t) \mid \mathcal{F}_s) = \mathbb{E}(f(B_t) \mid B_s)$  for all  $f =$  bounded and continuous functions.

*Remark.*  $\frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/2t}$  is called the transition pdf.

*Proof.* We only compute its transition function here:

$$\begin{aligned}
\mathcal{P}_t(x, A) &= \mathbb{P}(B_{t+s} \in A \mid B_s = x) \\
&= \mathbb{P}(B_{t+s} - B_s \in A - x \mid B_s = x) \\
(\text{indep. incre.}) &= \mathbb{P}(B_{t+s} - B_s \in A - x) \\
(\text{Gaussian incre.}) &= \int_{A-x} \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} dy \\
&= \int_A \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/2t} dy
\end{aligned}$$

□

**Proposition 8.6.**

- a) *MP says that  $\{Y_t = B_{t+s} - B_s\}_{t \geq 0}$  is a standard BM indep. of  $\mathcal{F}_s$  for all fixed  $s$ .*
- b) *SMP means  $\{Y_t = B_{t+S} - B_S\}_{t \geq 0}$  is a standard BM indep. of  $\mathcal{F}_S$  for all stopping times  $S$ .*

**Example 8.7** (Reflected BM). Define

$$B_t^{(x)} = \begin{cases} B_t & 0 < t \leq T_x \\ 2x - B_t & t > T_x \end{cases}$$

We can show that  $B^{(x)}$  is a standard BM.

- $B_0^{(x)} = 0$ .
- For  $t > T_x$ ,

$$\begin{aligned}
B_t^{(x)} &= 2x - B_t = 2B_{T_x} - B_t \\
&= B_{T_x} - (B_t - B_{T_x}) \\
(\text{symmetry of normal dist.}) &= B_{T_x} + (B_t - B_{T_x})
\end{aligned}$$

**Exercises**

**Exercise 8.3.1.** Compute the transition semigroups:

(a) **Reflected BM**

Let  $X = x + B$ . Define  $R = |X|$ . Compute the transition semigroup of  $R$ .

(b) **Geometric BM**

Let  $X_t = e^{at+bB_t}$ . Prove  $X$  is a Markov process and compute its transition semigroup.

(c) **BM with a drift**

Let  $Y_t = B_t + t\mu$ . Compute its transition density function.

*Proof.* (a) Directly compute it:

$$\begin{aligned}
\mathcal{P}_t(x, A) &= \mathbb{P}(R_{t+s} \in A \mid R_s = y) \\
&= \mathbb{P}(B_{s+t} \in (A - x) \cup (-x - A) \mid B_s = y - x) \\
&= \int_{(A-x) \cup (-x-A)} \frac{1}{\sqrt{2\pi t}} e^{-(z-y+x)^2/2t} dz
\end{aligned}$$

If we assume  $(A - x) \cap (-x - A) = \emptyset$  (e.g.  $A = (x, +\infty)$ ),  $\mathcal{P}_t(x, A)$  can be simplified as

$$\mathcal{P}_t(x, A) = \int_A \frac{1}{\sqrt{2\pi t}} e^{-(u-y)^2/2t} du$$

(b) To prove  $X$  is a Markov process, we have two options:

• **Option 1.**

Note that  $B_{s+t} - B_t$  is independent of  $\mathcal{F}_t$ .

$$\begin{aligned}
& \mathbb{P}(X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) \\
&= \mathbb{P}(e^{at_n + bB_{t_n}} = x_n \mid X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) \\
&= \mathbb{P}(e^{a(t_n - t_{n-1}) + b(B_{t_n} - B_{t_{n-1}})} = x_n / X_{t_{n-1}} \mid X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) \\
&= \mathbb{P}(e^{a(t_n - t_{n-1}) + b(B_{t_n} - B_{t_{n-1}})} = x_n / X_{t_{n-1}} \mid X_{t_{n-1}} = x_{n-1}) \\
&= \mathbb{P}(X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1})
\end{aligned}$$

• **Option 2.**

If  $X$  is a Markov process and  $f$  is injective, then  $f(X)$  is also a Markov process.

Now we compute its transition semigroup:

$$\mathbb{P}(X_{t+s} \in A \mid X_s = x) = \mathbb{P}(B_t \in \frac{\ln A - a(t+s)}{b} \mid B_s = \frac{\ln x - as}{b})$$

(c)

$$\begin{aligned}
\mathbb{P}(X_{t+s} \in A \mid X_s = x) &= \mathbb{P}(B_{t+s} + (t+s)\mu \in A \mid B_s + s\mu = x) \\
&= \mathbb{P}(B_{t+s} \in A - (t+s)\mu \mid B_s = x - s\mu)
\end{aligned}$$

□

**-Ornstein-Uhlenbeck process-**

**Exercise 8.3.2.** Define

$$X_t = X_0 e^{-at} + b e^{-at} B_{e^{2at}-1},$$

where  $B$  is independent of  $X_0$  and  $a, b$  are strictly positive real number.

(a) Show that  $(X)$  is a Markov process and compute its transition semigroup.

(b) Prove it Gaussian if  $X_0 = x$  or if  $X_0$  is Gaussian.

(c) Show that, as  $t \rightarrow \infty$ ,  $X_t \xrightarrow{w} X$  where  $X \sim N(0, b^2)$ .

(d) If  $X_0 \sim N(0, b^2)$ ,  $X_t$  has the same distribution as  $X_0$  for all  $t$ .

*Proof.* (a)

$$\begin{aligned}
& \mathbb{P}(X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) \\
&= \mathbb{P}(X_{t_n} - e^{-a(t_n - t_{n-1})} X_{t_{n-1}} = x_n - e^{-a(t_n - t_{n-1})} x_{n-1} \mid X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) \\
&= \mathbb{P}(b e^{-at_n} \cdot (B_{e^{2at_n}-1} - B_{e^{2at_{n-1}}-1}) = x_n - e^{-a(t_n - t_{n-1})} x_{n-1} \mid X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) \\
&= \mathbb{P}(b e^{-at_n} \cdot (B_{e^{2at_n}-1} - B_{e^{2at_{n-1}}-1}) = x_n - e^{-a(t_n - t_{n-1})} x_{n-1} \mid X_{t_{n-1}} = x_{n-1}) \\
&= \mathbb{P}(X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1})
\end{aligned}$$

(b) Trivial. Just compute its fdds.

(c) Obviously,  $X_0 e^{-at} \xrightarrow{w} 0$  and  $b e^{-at} B_{e^{2at}-1} \sim N(0, b^2 - e^{-2at})$  and  $b e^{-at} B_{e^{2at}-1} \xrightarrow{w} X$ . So,

$$X_t \xrightarrow{w} X$$

by Slutsky's theorem.



(d) Again, we use  $be^{-at}B_{e^{2at}-1} \sim N(0, b^2 - e^{-2at})$ . And it is independent of  $X_0$ . □

**Exercise 8.3.3.** Define  $Y_t = e^{-\alpha t/2}B_{e^{\alpha t}}$  where  $\alpha > 0$ . (a) Prove  $Y$  is stationary. (b) Show that  $Y$  has a.s. continuous path.

*Proof.* Because  $Y$  is a Gaussian process, we only need to check its 2nd moment. Obviously, it is stationary. Now, we prove that  $Y$  has a.s. continuous path. For almost all  $\omega$ ,  $t \mapsto B_t$  is continuous. And  $t \mapsto e^{\alpha t}$  is continuous. So

$$t \mapsto e^{-\alpha t/2}B_{e^{\alpha t}}$$

is continuous. □

## 8.4 Hitting Time and Running Maximum

**Definition 8.8.**

- Hitting time.  $T_x = \inf\{t > 0 : B_t(\omega) \geq x\}$ .
- Running maximum.  $M_t(\omega) = \max_{0 \leq s \leq t} B_s(\omega)$ .

**Proposition 8.9** (Properties of  $T$ ).

- $\mathbb{P}(T_a \leq t) = \mathbb{P}(M_t > a) = \mathbb{P}(|B_t| > a)$ .
- The distribution of  $T_x$  is the Inverse Gamma distribution:

$$f_{T_x}(t) = \frac{x}{\sqrt{2\pi t^3}} e^{-x^2/2t} \mathbf{1}_{\{t>0\}}.$$

- $\mathbb{E}T_x = \infty$  and  $\mathbb{P}(T_x < \infty) = 1$ .

*Proof.*

- $\mathbb{P}(T_a \leq t) = \mathbb{P}(M_t > a)$  is obvious.
- It suffices to find its Laplace transform. Consider the exponential martingale  $N_t^\lambda = e^{\lambda B_t - \frac{\lambda^2}{2}t}$  and use the optional stopping theorem. We can get:

$$\mathbb{E}e^{-\lambda T_a} = e^{-a\sqrt{2\lambda}}$$

- Its expectation is

$$\lim_{\lambda \rightarrow 0} e^{-a\sqrt{2\lambda}} \cdot \frac{1}{2}(2\lambda)^{-\frac{1}{2}} = \infty.$$

And

$$\mathbb{P}(T < \infty) = \lim_{\lambda \rightarrow 0} e^{-a\sqrt{2\lambda}} = 1.$$

□

**Proposition 8.10.**

- $M - B$  and  $|B|$  have the same law.
- $2M - B$  is a Bessel process with index 3.

## Exercises

**Exercise 8.4.1.**  $T_a \stackrel{D}{=} a^2 T_1$

*Proof.* Note that  $f_{a^2 T_1}(t) = \frac{1}{a^2} f_{T_1}(\frac{t}{a^2}) = f_{T_a}(t)$ . □

**Exercise 8.4.2.**

- Compute

$$\mathbb{P}(T_a \leq t, B_t < a) = \mathbb{P}(B_t > a).$$

- Show that

$$\mathbb{P}(M_t > a, B_t \leq x) = \int_{2a-x}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} dy.$$

And compute its derivative w.r.t.  $a$  and  $x$ .

**Exercise 8.4.3.** Define  $\tilde{T}_a = \inf\{t > 0 : B_t(\omega) > a\}$ . (a) Show that  $T_a = \tilde{T}_a$  almost surely. (b) Both of  $T_a$  and  $\tilde{T}_a$  are stopping times.

**Exercise 8.4.4.** Let  $X$  and  $Y$  be two independent BMs, and  $T$  be the hitting time processes of  $Y$ . Show that  $(X_{T_a})$  is a Cauchy process.

*Proof.* • First, we prove that  $(X_{T_a})$  is a Lévy process.

- For almost every  $\omega$ ,  $a \mapsto X_{T_a}$  is right-continuous and left-limited starting from  $X_0(\omega) = 0$ . This is trivial.
- $X_{T_{a+u}} - X_{T_a}$  is independent of  $\mathcal{F}_{T_a}$  and has the same distribution as  $X_{T_u}$ .

- Second, we compute the distribution of  $X_{T_a}$ .

$$\begin{aligned} \mathbb{P}(X_{T_a} \in A) &= \mathbb{P}\left\{\bigcup_t [(X_t \in A) \cap (T_a = t)]\right\} \\ &= \int_{t \in \mathbb{R}_+} \mathbb{P}(X_t \in A) dF_{T_a} \\ f_{X_{T_a}}(x) &= \int_0^\infty \frac{a}{\sqrt{2\pi t^3}} e^{-a^2/2t} \cdot \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} dt \\ &= \frac{a}{2\pi} \int_0^\infty e^{-\frac{x^2+a^2}{2t}} d\left(-\frac{1}{t}\right) \\ &= \frac{a}{\pi} \cdot \frac{1}{x^2 + a^2} \end{aligned}$$

□

**Exercise 8.4.5.** Let  $(X, Y)$  be a BM in  $\mathbb{R}^2$  with initial state  $(X_0, Y_0) = (0, y)$  for some  $y < 0$ . Let  $S$  be the first time that  $(X, Y)$  touches the x-axis. Find the distribution of  $X_S$ .

*Proof.* Let  $S = T_{-y}$ , the hitting time of  $Y$ . As we just computed,  $X_S$  has the Cauchy distribution with the pdf  $f_{X_S}(x) = -\frac{y}{\pi} \cdot \frac{1}{y^2 + x^2}$ . □

### -Arcsin law-

**Exercise 8.4.6.** Let  $W$  be a standard Brownian motion. Define

$$G_t = \sup s \in [0, t] : W_s = 0$$

and

$$D_t = \inf\{u \in (t, \infty) : W_u = 0\}.$$

(a) Show that  $D_t$  is a stopping time but  $G_t$  is not.

(b) Let  $A$  be a RV with  $\mathbb{P}(A \leq u) = \frac{2}{\pi} \arcsin \sqrt{u}$  where  $0 \leq u \leq 1$ . Show that

$$G_t \stackrel{D}{=} tA$$

and

$$D_t \stackrel{D}{=} t/A.$$

(c) Show that

$$\mathbb{P}(W_t \in \mathbb{R} \setminus \{0\}, \forall t \in [s, u]) = \frac{2}{\pi} \arcsin \sqrt{\frac{s}{u}}.$$

## 8.5 Path Properties

**Definition 8.11.**  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  is said to be Hölder continuous of order  $\alpha$  on  $A \subset \mathbb{R}_+$ , if there exists a constant  $k$  such that

$$|f(t) - f(s)| \leq k \cdot |t - s|^\alpha \quad \forall s, t \in A.$$

**Proposition 8.12.** For almost every  $\omega \in \Omega$ , the path  $t \mapsto B_t$  has the following properties:

- a) Infinite total variation over every interval;
- b) Not Hölder continuous on every interval for  $\alpha \geq 1/2$ ;
- c) Nowhere differentiable;
- d) Locally Hölder continuous of order  $\alpha$  for every  $\alpha < 1/2$ .

*Proof.*

a) See Example 6.12.

b) This is a partial proof. Suppose it is Hölder continuous on  $[a, b]$  for  $\alpha > 1/2$ ; i.e.

$$|B_t(\omega) - B_s(\omega)| \leq k|t - s|^\alpha$$

for all  $s, t \in [a, b]$ . Then we compute its total variation on  $[a, b]$ :

$$\sum |B_t(\omega) - B_s(\omega)|^2 \leq k^2 \cdot |t - s|^{2\alpha} \leq k^2 \cdot (b - a) \cdot \sup |t - s|^{2\alpha-1}.$$

If  $2\alpha - 1 > 0$ , it implies the total variation of  $B$  on  $[a, b]$  is 0. It is impossible. However, when  $\alpha = 1/2$ , this method doesn't work (the proposition is still true).

c) Note that if  $f$  is differentiable on  $[a, b]$ , then  $f$  is Hölder continuous on  $[a, b]$  for  $\alpha = 1$ .

d) Check Kolmogorov's moment condition.

□

## Exercises

**Exercise 8.5.1.** Show that the path of Brownian motion is monotone in no interval.

*Proof.* First, we show the following fact: Let  $f = g - h$ . If  $g$  and  $h$  are increasing on  $[a, b]$ , then  $f$  has bounded variation on  $[a, b]$ .

Then suppose  $B(\omega)$  is monotone on  $[a, b]$ , then it has bounded variation. Contradiction. □

**-Law of the iterated logarithm-**

**Exercise 8.5.2.** Let  $(W_t)$  be a Brownian motion. The law of the iterated logarithm describes the oscillatory behavior of Brownian motion near the time 0 and for very large time. Define

$$h(t) = \sqrt{2t \log \log \left(\frac{1}{t}\right)}$$

for  $t \in [0, 1]$ . And it can be shown that for almost every  $\omega$ ,

$$\limsup_{t \rightarrow 0} \frac{1}{h(t)} W_t(\omega) = 1 \quad \text{and} \quad \liminf_{t \rightarrow 0} \frac{1}{h(t)} W_t(\omega) = -1.$$

Show it by following these steps:

(a) For all  $p, q > 0$ ,

$$\mathbb{P}\left\{\sup_{t \leq 1} (W_t - \frac{1}{2}pt) > q\right\} \leq e^{-pq}.$$

(b) Define

$$\alpha(\omega) = \limsup_{t \rightarrow 0} \frac{1}{h(t)} W_t(\omega).$$

Show that

$$\alpha(\omega) \leq 1.$$

(c) By considering the process  $(-W_n)$ , show that

$$\alpha(\omega) \geq 1.$$

(d) Let  $Z \sim N(0, 1)$ . Then for  $b > 0$ ,

$$\frac{1}{4} \cdot \frac{b^2}{1 + b^2} e^{-b^2/2} < \mathbb{P}(Z > b) < \frac{1}{2b} e^{-b^2/2}.$$

(e) Show

$$\liminf_{t \rightarrow 0} \frac{1}{h(t)} W_t(\omega) = -1.$$

## 8.6 Zeros of Brownian motion

**Definition 8.13.** Fix  $\omega$ , define

$$C_\omega = \{t \in \mathbb{R}_+ : B_t(\omega) = 0\}.$$

**Proposition 8.14.** For almost every  $\omega$ , we have:

- $C_\omega$  is perfect and unbounded.
- $C_\omega^\circ = \emptyset$ .
- $\text{Leb}(C_\omega) = 0$ .
- $C_\omega$  is uncountable.

## 9 Basic Stochastic Calculate

### 9.1 Definition of $\int_0^t X_s \, ds$

**Definition 9.1.** Let  $g(t)$  be a real-valued function on  $[0, T]$ ,  $X_t$  be a process, and  $\mathcal{A}_n$  be a sequence of subdivision of  $[0, T]$  with  $\|\mathcal{A}_n\| \rightarrow 0$ . Define

$$I_n := \sum_{(s,t]} g(s) X_s \cdot (t - s);$$

then its m.s. limit is written as

$$\int_0^T g(s) X_s \, ds := \text{m.s.} \lim_{n \rightarrow \infty} I_n.$$

**Proposition 9.2** (Integration by parts). *If  $f(t)$  is differentiable, then*

$$\int_0^T f(t) \, dB_t = B(T)f(T) - \int_0^T B_t f'(t) \, dt.$$

*Remark.* We write  $d(f(t)B_t) = f'(t)B_t dt + f(t)dB_t$ . Then take the integral on both sides.

### 9.2 Itô Integral: Definitions of $\int_0^t X_s \, dB_s$

**Definition 9.3.** Let  $X$  be a step function; that is

$$X_t = \sum_{i=1}^{n-1} X_{t_i} \mathbf{1}_{[t_i, t_{i+1}]}(t),$$

where  $0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T$  is a partition of the interval  $[0, T]$ .

For  $t \in [0, T]$ , it must be in  $[t_k, t_{k+1}]$  for some  $k$ . Define

$$\begin{aligned} \int_0^t X_s \, dB_s &= X_0(B_{t_1} - B_0) + X_{t_1}(B_{t_2} - B_{t_1}) + \dots + X_{t_k}(B_t - B_{t_k}) \\ &= \sum_{i=0}^{k-1} X_{t_i}(B_{t_{i+1}} - B_{t_i}) + X_{t_k}(B_t - B_{t_k}) \end{aligned}$$

**Definition 9.4.**  $X \in \mathcal{L}_C^2([0, T])$  if

- (a)  $X_t \in \mathcal{F}_t$ .
- (b)  $\mathbb{E}X_t^2 \leq K < \infty \, \forall t \in [0, T]$ .
- (c)  $X$  is continuous in q.m. on  $[0, T]$ .

*Remark.*  $\mathcal{L}_C^2 = \mathcal{L}^2 + \text{continuous}$ .

**Definition 9.5** (Stochastic integral for  $X \in \mathcal{L}_C^2$ ).

Approach: Simple-function approximation to  $X$ .

- (i) Divide  $[0, t]$  to  $2^n$  subintervals of equal length. Let  $\lfloor s \rfloor^n$  be the partition point which is closest to  $s$  from below.
- (ii) Define  $X(s) = X(\lfloor s \rfloor^n)$ . It is a simple function.
- (iii) Notice that

$$\mathbb{E}(X_{\lfloor s \rfloor^n} - X_s)^2 \rightarrow 0$$

as  $n \rightarrow \infty$ . By bounded convergence theorem, we have

$$\int_0^t \mathbb{E}(X_{\lfloor s \rfloor^n} - X_s)^2 \, ds \rightarrow 0$$

as  $n \rightarrow \infty$ .

(iv) The limit  $\int_0^t X_s \, dB_s$  is unique.

**Example 9.6.** Calculate  $\int_0^t B_s \, dB_s$  using definition.  
Note that  $X_t = B_t \in \mathcal{L}_C^2$ .

(i) Recall the definition:

$$\int_0^t B_s \, dB_s = \lim_{n \rightarrow \infty} \sum_{\Delta_n} B_{t_i^{(n)}} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})$$

(ii) Notice the following fact:

$$\sum (B_{t_{i+1}^{(n)}}^2 - B_{t_i^{(n)}}^2) = B_t^2$$

and

$$\begin{aligned} \sum (B_{t_{i+1}^{(n)}}^2 - B_{t_i^{(n)}}^2) &= \sum (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}} + 2B_{t_i^{(n)}}) \\ &= \sum (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2 + 2 \sum (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})B_{t_i^{(n)}} \end{aligned}$$

(iii) Combine two facts:

$$B_t^2 = \sum (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2 + 2 \sum (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})B_{t_i^{(n)}}$$

Let  $n \rightarrow \infty$ , we have

$$B_t^2 = \langle B, B \rangle_t - 2 \int_0^t B_s \, dB_s.$$

(iv) Therefore,  $\int_0^t B_s \, dB_s = \frac{1}{2}(B_t^2 - t)$ .

**Definition 9.7.**  $X \in \mathcal{L}^2([0, T])$  if

(a)  $X_t \in \mathcal{F}_t$ .

(b)  $\int_0^T \mathbb{E} X_s^2 \, ds < \infty$ .

*Remark.* Notice that  $\mathcal{L}_C^2$  is dense in  $\mathcal{L}^2$ ; then follow the plan of Definition 10.7.

### 9.3 Properties of Stochastic Integral

**Proposition 9.8.**

- $\int_0^t (\alpha X_s + \beta Y_s) \, dB_s = \alpha \int_0^t X_s \, dB_s + \beta \int_0^t Y_s \, dB_s$ .
- $\mathbb{E}(\int_0^t X_s \, dB_s) = 0$ .
- $\mathbb{E}(\int_0^t X_s \, dB_s)^2 = \mathbb{E}(\int_0^t X_s^2 \, ds) = \int_0^t \mathbb{E} X_s^2 \, ds$ .
- $\mathbb{E}(\int_0^t X_s \, dB_s \cdot \int_0^t Y_s \, dB_s) = \int_0^t \mathbb{E}(X_s Y_s) \, ds$ .

*Proof.*

- Omitted.
- Omitted.

- It suffices to prove it for the simple function; because we can use the  $L^2$ -continuity of the expectation operator to get the similar statement for the sgeneral case. Let

$$X_t = \sum_{i=1}^{n-1} X_{t_i} \mathbf{1}_{[t_i, t_{i+1}]}(t).$$

Note that  $X$  is independent of  $B$ . Then

$$\begin{aligned} \mathbb{E}\left(\int_0^t X_s \, dB_s\right) &= \mathbb{E}\left\{\sum_i X_{t_i}^2 (B_{t_{i+1}} - B_{t_i})^2 + 2 \sum_{i < j} X_{t_i} X_{t_j} (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right\} \\ (\text{indep. incre.}) \quad &= \sum_i \mathbb{E}(X_{t_i}^2)(t_{i+1} - t_i) + 0 \\ (\text{def. of integral}) \quad &= \int_0^t \mathbb{E}X_s^2 \, ds \end{aligned}$$

- Use  $4ab = (a+b)^2 - (a-b)^2$ .

□

**Proposition 9.9.** Let  $Z_t = \int_0^t X_s \, dB_s$ .

- $Z$  has a.s. continuous path.
- $Z$  is a martingale.
- $\langle Z, Z \rangle_t = \int_0^t X_s^2 \, dB_s$ .
- If  $X_t = \varphi(t)$  (deterministic), then  $Z$  is a Gaussian, mean zero process with covariance

$$\mathbb{E}(Z_t, Z_s) = \int_0^{t \wedge s} \varphi^2(u) \, du.$$

*Proof.*

- Let  $X$  be a step function; that is

$$X_t = \sum_{i=1}^{n-1} X_{t_i} \mathbf{1}_{[t_i, t_{i+1}]}(t).$$

Then  $Z_t = \int_0^{t_k} X_s \, dB_s + X_{t_k}(B_t - B_{t_k})$ . By induction,  $Z$  has a.s. continuous path.

- Omitted.
- Omitted.
- Use Proposition 10.8.

□

**Definition 9.10.**  $X \in \mathcal{L}_{loc}^2$  if

- $X_t \in \mathcal{F}_t$ .
- $\int_0^T X_s^2 \, ds < \infty$  almost surely.

*Remark.* Some properties fail in this case.

## Exercises

**Exercise 9.3.1.** Find the distribution of the following processes.

(a)  $X_t = \int_0^t \ln(s) \, dB_s$ .

(b)  $Z_t = \int_0^t \frac{B_s}{s} \, ds$ .

*Proof.* (a) By Proposition 10.5,  $X_t$  is a Gaussian RV with mean zero and covariance

$$\mathbb{E}(X_t X_s) = \int_0^{t \wedge s} (\ln(u))^2 \, du.$$

Letting  $a = s \wedge t$ ,  $\mathbb{E}(X_t X_s) = a[2 + \ln^2(a) - 2\ln(a)]$ .

(b) By the definition of integral,  $Z_t$  is a linear combination of Gaussian variables; so  $Z_t$  is also Gaussian. And obviously, its mean is 0. Then we need to compute its covariance.

Notice that  $d(\ln(t)B_t) = \ln(t)dB_t + \frac{B_t}{t}dt$ . Therefore,

$$\ln(s)B_s|_0^t = \int_0^t \ln(s) \, dB_s + \int_0^t \frac{B_s}{s} \, ds.$$

Recall that for all  $\alpha > 1/2$ ,  $\lim_{t \rightarrow \infty} t^{-\alpha} B_t = 0$ ; that is,  $\lim_{\frac{1}{t} \rightarrow 0} t^{-\alpha} \cdot t B_{\frac{1}{t}} = 0$ . Letting  $u = \frac{1}{t}$ , we have

$$\lim_{u \rightarrow 0} \frac{B_u}{u^\beta} = 0$$

holds for all  $\beta < \frac{1}{2}$ . Therefore,  $\ln(s)B_s|_0^t = \ln(t)B_t$ . Finally, we get

$$Z_t = \int_0^t \frac{B_s}{s} \, ds = \ln(t)B_t - \int_0^t \ln(s) \, dB_s.$$

By Proposition 10.5, the covariance is

$$\begin{aligned} \mathbb{E}(Z_t Z_s) &= \ln(t) \ln(s) s \wedge t - \ln(t) \mathbb{E}(B_t \int_0^s \ln(u) \, dB_u) \\ &\quad + \int_0^{t \wedge s} \ln^2(u) \, du - \ln(s) \mathbb{E}(B_s \int_0^t \ln(u) \, dB_u) \\ &= \ln(t) \ln(s) s \wedge t + \int_0^{t \wedge s} \ln^2(u) \, du - \ln(t) \int_0^{t \wedge s} \ln(u) \, du - \ln(s) \int_0^{t \wedge s} \ln(u) \, du \\ &= s \wedge t \cdot [\ln(s \vee t) - \ln(s \wedge t) + 2] \end{aligned}$$

□

## -Itô isometry-

**Theorem.**

$$\mathbb{E} \left[ \int_0^T X_t \, dW_t \right]^2 = \mathbb{E} \left[ \int_0^T X_t^2 \, dt \right].$$

*Remark.* Usually, we use the Itô isometry with the polarization identity  $4ab = (a+b)^2 - (a-b)^2$ .

**Exercise 9.3.2.** Compute  $\mathbb{E}[W_t \int_0^t s dW_s]$ .



*Proof.* Note that

$$\begin{aligned}
4\mathbb{E}[W_t \int_0^t s dW_s] &= 4\mathbb{E}[\int_0^t dW_s \int_0^t s dW_s] \\
&= \mathbb{E}[\int_0^t dW_s + \int_0^t s dW_s]^2 - \mathbb{E}[\int_0^t dW_s - \int_0^t s dW_s]^2 \\
&= \int_0^t (s+1)^2 ds - \int_0^t (s-1)^2 ds \\
&= 4 \cdot \frac{1}{2}t^2
\end{aligned}$$

Therefore,  $\mathbb{E}[W_t \int_0^t s dW_s] = \frac{t^2}{2}$ . □

**Exercise 9.3.3.** Compute  $\mathbb{E}\left[\exp(\int_0^t W_s ds) \mid W_t\right]$ .

*Proof.* Let  $Y = \int_0^T W_s dt$  and  $X = W_T$ .

- First, we compute the joint distribution of  $(X, Y)$ . Obviously, they are Gaussian with mean 0; so it suffices to find their covariance matrix.

- It is easy to see that  $\mathbb{E}X^2 = T$ .
- Using the integration by parts,

$$Y = TW_T - \int_0^T s dW_s.$$

So we have

$$\mathbb{E}Y^2 = \frac{1}{3}T^3.$$

- And

$$\mathbb{E}XY = \mathbb{E}\left[W_T \int_0^T W_s ds\right] = \frac{1}{2}T^2.$$

- Recall that if  $(X, Y) \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & r \\ r & \sigma_Y^2 \end{pmatrix}\right)$ , then

$$X|Y = y \sim N\left(\mu_1 + \frac{r}{\sigma_Y^2}(y - \mu_2), \sigma_X^2 - \frac{r^2}{\sigma_Y^2}\right).$$

Therefore, in our case,

$$Y|X \sim N\left(\frac{T}{2}X, \frac{1}{12}T^3\right).$$

- Finally, we have

$$\exp(Y)|X \sim \text{Lognormal}\left(\frac{T}{2}X, \frac{1}{12}T^3\right).$$

Its mean is

$$\mathbb{E}(\exp(Y)|X) = \exp\left(\frac{T}{2}W_T + \frac{T^3}{24}\right).$$

□

## 9.4 Itô Formula

**Definition 9.11.** A continuous adapted process  $X$  is called an Itô process if can be expressed as

$$X_t = X_0 + \int_0^t U_s \, ds + \int_0^t V_s \, dB_s$$

where  $U, V \in \mathcal{L}_{loc}^2$ . Or we write

$$dX_t = U_t dt + V_t dB_t.$$

**Theorem 9.12.** Suppose that  $f(x) \in C^2$ . For all  $x \geq 0$ ,

$$f(B_t) - f(B_0) = \int_0^t f'(B_s) \, dB_s + \frac{1}{2} \int_0^t f''(B_s) \, ds.$$

**Theorem 9.13.** Suppose that  $f(t, x) \in C^2$ . For all  $t \geq 0$ ,

$$\begin{aligned} f(t, B_t) - f(0, B_0) &= \int_0^t \frac{\partial f}{\partial t}(s, B_s) \, ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s) \, dB_s \\ &\quad + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s) \, ds. \end{aligned}$$

**Example 9.14.** Let  $f(x) = x^n$ . By Itô formula, we have

$$B_t^n = n \int_0^t B_s^{n-1} \, dB_s + \frac{n(n-1)}{2} \int_0^t B_s^{n-2} \, ds.$$

If  $n = 2$ , it is easy to notice that

$$B_t^2 = 2 \int_0^t B_s \, dB_s + t$$

and  $B_t^2$  is the sum of a martingale (Proposition 10.5) and an increasing function. So  $B_t^2$  is a submartingale (Doob's decomposition theorem).

**Example 9.15.** Let  $g(x)$  be convex ( $g'' > 0$ ) and  $X_t = \int_0^t V_s \, dB_s$ . We want to show  $Y_t = g(X_t)$  is a submartingale.

Notice that  $X$  is a Itô process with  $dX_t = 0 \cdot dt + V_t dB_t$ . Using Itô formula, we get

$$\begin{aligned} g(X_t) - g(X_0) &= \int_0^t g'(X_s) \, dX_s + \frac{1}{2} \int_0^t g''(X_s) \, d\langle X, X \rangle_s \\ &= \int_0^t g'(X_s) V_s \, dB_s + \frac{1}{2} \int_0^t g''(X_s) V_s^2 \, ds \end{aligned}$$

If  $g'(X_s) V_s \in \mathcal{L}^2$ , by Doob's decomposition theorem,  $g(X_t)$  is a submartingale.

**Example 9.16.** Let  $f(x) = e^x$  and  $X_t = \int_0^t V_s \, dB_s - \frac{1}{2} \int_0^t V_s^2 \, ds$ . Then

$$\begin{aligned} e^{X_t} - 1 &= \int_0^t e^{X_s} \, dX_s + \frac{1}{2} \int_0^t e^{X_s} \, d\langle X, X \rangle_s \\ &= \left( \int_0^t e^{X_s} V_s \, dB_s - \frac{1}{2} \int_0^t e^{X_s} V_s^2 \, ds \right) + \frac{1}{2} \int_0^t e^{X_s} V_s^2 \, ds \\ &= \int_0^t e^{X_s} V_s \, dB_s \end{aligned}$$

Let  $Y_t = e^{X_t}$ . We differentiate on both sides:

$$\begin{cases} Y_0 &= 1 \\ dY_t &= Y_t V_t \, dB_t \end{cases}$$

If  $V_t = a$ , then  $Y_t = e^{aB_t - \frac{1}{2}at^2}$ .

## 9.5 Applications of Itô formula