Stochastic Process

LU Shan

Preface

These notes originate from the graduate course MAT7092 Stochastic Processes, taught by Dr. Wangjun Yuan in the Fall semester of 2025 at the Southern University of Science and Technology (SUSTech). Presented entirely on the blackboard with remarkable clarity and depth, the course wove together rigorous theory and profound insight.

Coming from an undergraduate background in physics and without prior formal training in measure theory, I initially found its pace challenging. The elegant abstraction of martingale, dynamic structure of Markov process, and rich behavior of Brownian motion often seemed daunting amidst the swift flow of ideas. It was in striving to keep up, to translate each lecture into something tangible and retainable, that this notebook began, first as hurried annotations on an iPad, later as a structured and reflective compilation.

As the course unfolded, so did my appreciation. What once felt like a rapid succession of definitions and theorems gradually revealed itself as a coherent and beautifully constructed edifice. It is designed for those who have encountered elementary stochastic processes and now wish to delve deeper. This text is therefore more than a transcription. It is a digested and reorganized account, meant to capture both the logical skeleton and the intuitive spirit of the material.

In these pages, I have sought to preserve the lecture's rigor while supplementing explanations, examples, and occasional commentary that helped me cross conceptual bridges. If this notebook offers others a clearer path through the same rewarding terrain, or simply conveys something of the course's intellectual elegance, it will have achieved its purpose.

To the fellow student who also once felt lost in the landscape of random evolution: may you find here not only clarity, but also inspiration. The beauty of stochastic processes lies not in their complexity alone, but in their power to model, to reveal, and to predict the uncertain rhythms of the world.

Contents

1 Conditional Expectation

Intuitively speaking, **stochastic processes**: A mathematical model for occurrence of a random phenomenon at each moment after intial time. And if $X = \{X_t : t \geq 0\}$ is a stochastic process, then for every fixed t, X_t is a random variables. The goal of this subject is to study the relationship between X_t and X_s.
Roughly speaking the course content is:

1. **Tools:** Conditional Expectation, Martingales, Markov Property.
2. **Processes:** Markov process, Brownian motions, Poisson process, Lévy process, Brownian sheet.

1.1 Brief Probability Theory

Firstly, let's state out some important probability theory concepts for further formalisation. In order to describe probability space, we need sample space and sample points:

Definition 1.1. Sample space denoted as Ω, and sample points: $\omega \in \Omega$.

Then on this sample space, we will restrict to only study certain subsets:

Definition 1.2. Let $\mathcal{F}$ be a collection of subsets of Ω, such that:

(a) $\Omega \in \mathcal{F}$.

(b) $\forall\ A \in \mathcal{F} \implies A^c \in \mathcal{F}$.

(c) $\forall\ A_i \in \mathcal{F} \implies \cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

, then $\mathcal{F}$ is called σ**-algebra** on Ω.

For better understanding, we provide some trivial examples:

Example 1.1. The most trivial one would be: $\mathcal{F} = \{\emptyset, \Omega\}$. And we also can add in one element, then $\mathcal{F} = \{\emptyset, \Omega, A, A^c\}$ is also a σ-algebra.

After two major elements, we need to assign value, then comes with measure:

Definition 1.3. If a set function: $\mathsf{P} : \mathcal{F} \mapsto [0, 1]$ satisfies:

(a) $\mathsf{P}(\emptyset) = 0, \mathsf{P}(\Omega) = 1$.

(b) $\forall$ disjoint $A_i \in \mathcal{F}$, $\mathsf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathsf{P}(A_i)$.

, then P is **probability measure**, and we further call $(\Omega, \mathcal{F}, \mathsf{P})$ a probability space.

Since measure taking on element in abstract space to $[0, 1]$, and we are more into studying real value, therefore we need a function to bring us back, there comes random variable:

Definition 1.4. A measurable function $X : \Omega \mapsto \mathbb{R}$ is called a random variables.

Then the correspongding "measure" also need to be pushed forward:

Definition 1.5. Let X be a r.v. on $(\Omega, \mathcal{F}, \mathsf{P})$, on $(\mathbb{R}, \text{Borel } \sigma\text{-algebra on } \mathbb{R}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we define set function: $\mu_X : \mathcal{B}(\mathbb{R}) \mapsto \mathbb{R}([0,1])$, such that:

$$\mu_X(B) = \mathsf{P}(X^{-1}(B)) = \mathsf{P}(X \in B)$$

, then μ_X is probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, called probability distribution of X.

Remark 1.1. Note that: $\mu_X((-\infty, a]) = \mathsf{P}(X \in (-\infty, a]) = \mathsf{P}(X \leq a)$.

Once we have the distribution, we could study the relationship of r.v.s:

Definition 1.6. Let X, Y be two r.v.s on $(\Omega, \mathcal{F}, \mathsf{P})$. If $\mu_X = \mu_Y$ on $\mathbb{R}^d$, then we say X and Y are **identically distributed**.

Done with major definitions, we could delve into r.v's properties, like expectation:

Definition 1.7. Let X be a r.v. on $(\Omega, \mathcal{F}, \mathsf{P})$. The **expectation** of X is:

$$\mathsf{E}X = \int_\Omega X(\omega)\mathsf{P}(\mathrm{d}\omega)$$

, if it is finite, otherwise it is undefined.

This is well defined, since generally, from probability theory class, you should know that, we could seperate this into four progressive stages:

1. If $X(\omega) = \mathbf{1}_A(\omega), A \in \mathcal{F}$, then $\mathsf{E}X = \mathsf{E}\mathbf{1}_A = \mathsf{P}(A)$.
2. If $X(\omega) = \sum_{i=1}^m a_i \mathbf{1}_{A_i}, A_i \in \mathcal{F}, a_i \in \mathbb{R}$, then $\mathsf{E}X = \sum_{i=1}^m a_i \mathsf{P}(A_i)$.
3. If $X(\omega) \geq 0$, take $\varphi_n = \sum a_i \mathbf{1}_{A_i}, s.t.\ 0 \leq \varphi \leq X, \varphi_n \uparrow X$, then, $\mathsf{E}X = \lim_{n\to\infty} \mathsf{E}\varphi_n$.
4. For general case, $X(\omega) = X^+(\omega) + X^-(\omega)$, here $X^+ = \max(X, 0), X^- = \min(X, 0)$.

Remark 1.2.
1. Note here: $\mathsf{E}f(X) = \int_\Omega f(X(\omega))\mathsf{P}(\mathrm{d}\omega) = \int_\mathbb{R} f(x)\mu_X(\mathrm{d}x)$.
2. If X be a r.v. on $(\Omega, \mathcal{F}, \mathsf{P})$ satisfies $\mathsf{E}|X| < +\infty$ then X is integrable. We write: $X \in L^1(\Omega, \mathcal{F}, \mathsf{P})$. More generally, if $p \in \mathbb{R}, X^p \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, then $X \in L^p(\Omega, \mathcal{F}, \mathsf{P})$.

In order to further proceed, we need an important concepts, independence:

Definition 1.8. On $(\Omega, \mathcal{F}, \mathsf{P})$, $\mathcal{A}_i \subseteq \mathcal{F}$ is a σ-algebra for $1 \leq i \leq n$. If $\forall\ A_{i_k} \in \mathcal{A}_{i_k}, i_k \in \{1, \cdots, n\}$ it holds that:

$$\mathsf{P}\left(\bigcap_k^l A_{i_k}\right) = \prod_k^l \mathsf{P}(A_{i_k})$$

, then we say $\mathcal{A}_1, \cdots, \mathcal{A}_n$ are independent.

After the definition of independence in σ-algebra, we move to r.v.s:

Definition 1.9. Let $\{X_i : 1 \leq i \leq n\}$ be r.v.s If $\forall$ Borel sets $\{\mathcal{B}_j : 1 \leq j \leq n\}$:

$$\mathsf{P}\left(X_1 \in \mathcal{B}_1, \cdots, X_n \in \mathcal{B}_n\right) = \prod_{i=1}^{n} \mathsf{P}(X_i \in \mathcal{B}_j)$$

, then $\{X_i : 1 \leq i \leq n\}$ are **independent**.

Remark 1.3. $\{X_i : 1 \leq i \leq n\}$ are independent $\Leftrightarrow \sigma(X_1), \cdots, \sigma(X_n)$ are independent.

Lasly, we extend the standard definition into the quantity we are interested in:

Proposition 1.1. *If $X_1, \cdots, X_n$ are independent r.v.s then:*

$$\mathsf{E}X_1 \cdots X_n = \mathsf{E}X_1 \cdots \mathsf{E}X_n$$

Proof. The proof is simple and we leave for reader to verify. □

1.2 Conditional Expectation

Finally, we are about to study the main basic tool of this subjects:

Definition 1.10. On $(\Omega, \mathcal{F}, \mathsf{P})$, X be a r.v., and $\mathcal{A} \subseteq \mathcal{F}$ be a σ-algebra, then a r.v. Y is the conditional expectation of X relative to $\mathcal{A}$, if Y satisfies:

(a) Y is measurable with respect to $\mathcal{A}$.

(b) $\forall\ A \in \mathcal{A}$, the following integration holds:

$$\int_A Y \mathrm{d}\mathsf{P} = \int_A X \mathrm{d}\mathsf{P} \text{ or } \mathsf{E}X\mathbf{1}_A = \mathsf{E}Y\mathbf{1}_A$$

Further more, using indicate function, we have conditional probability:

Definition 1.11. For $\mathcal{B} \in \mathcal{F}$, we call $\mathsf{P}(\mathcal{B} \mid \mathcal{A}) = \mathsf{E}\left[\mathbf{1}_{\mathcal{B}} \mid \mathcal{A}\right]$ the conditional probability of $\mathcal{B}$ relative to $\mathcal{A}$.

After defining conditional expectation, reader may wonder the existence and uniqueness:

Remark 1.4. If $X \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, the conditional expectation, $\mathsf{E}\left[X \mid \mathcal{A}\right]$ exists and is unique in the sense of a.s.

Proof. Consider the set function ν on $\mathcal{A}$ satisfying $\nu(A) = \int_A X \mathrm{d}\mathsf{P}, \forall\ A \in \mathcal{A}$. Then ν is well-defined, finite-valued and countably additive. Thus, ν is a (sign) measurable. If $\mathsf{P}(A) = 0$ then $\nu(A) = 0$, so $\nu \ll \mathsf{P}$ (absolutely continous). By **Radon-Nikodym Theorem**. $Y = \nu/\mathrm{d}\mathsf{P}$ exists and is unique except a P-null set. □

The Radon-Nikodym Theorem is not the main focus, then we just ignore it, and to better understand how the conditional expectation means, we consider:

Remark 1.5. By 1.10'(b), we could rewrite into:

$$\forall\ A \in \mathcal{A}, \int_A (X - \mathsf{E}\left[X \mid \mathcal{A}\right]) \mathrm{d}\mathsf{P} = 0\ a.s.$$

Then we may plug into indicator function, $\forall A_i \in \mathcal{A}$

$$\int_\Omega (X - \mathsf{E}[X \mid \mathcal{A}]) \left(\sum_{i=1}^{k} \mathbf{1}_{A_i} \right) \mathrm{d}\mathsf{P} = 0$$

Then by DCT, $\forall Z \in \mathcal{A}$:

$$\mathsf{E}\{[X - \mathsf{E}(X \mid \mathcal{A})] Z\} = \int_\Omega [X - \mathsf{E}(X \mid \mathcal{A})] Z \mathrm{d}\mathsf{P} = 0$$

Thus, we have decomposition $X = X_1 + X_2$ with:

$$X_1 = \mathsf{E}[X \mid \mathcal{A}] \in \mathcal{A},\ X_2 = X - \mathsf{E}[X \mid \mathcal{A}] \perp \mathcal{A}$$

In particular, $X_1 \perp X_2$.

The above remark is to say our expectation is indeed the best "guess" we can make under the information we have, somewhat like a projection since perpendicular property.

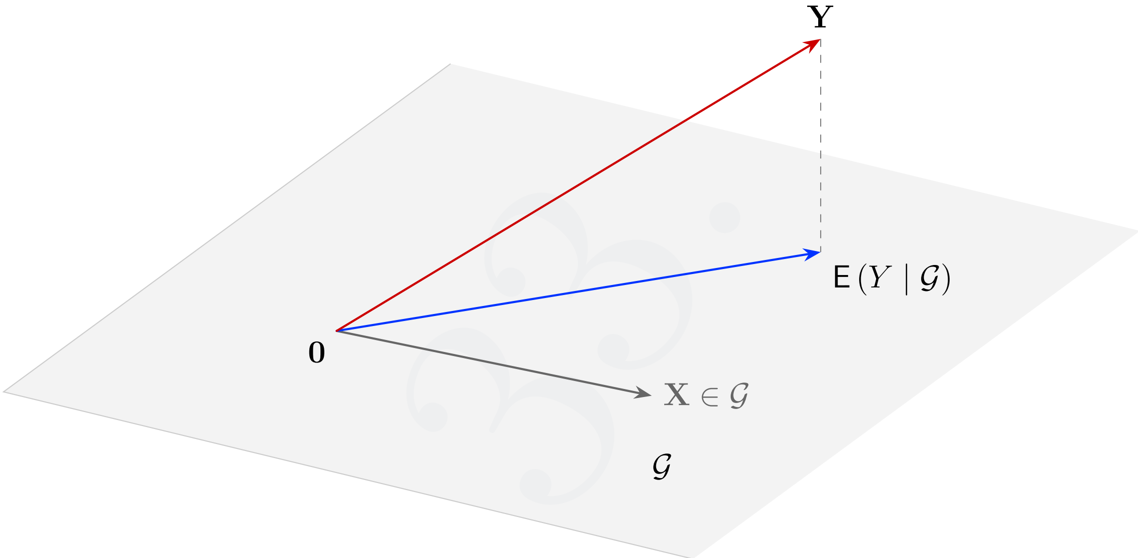

Figure 1: Geometric Interpretation of Conditional Expectation

Then next, for r.v.s $X_1, \cdots, X_n \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, let $\sigma(X_1, \cdots, X_n)$ be the σ-algebra generated by $X_1, \cdots, X_n$, then we write:

$$\mathsf{E}[X \mid X_1, \cdots, X_n] := \mathsf{E}[X \mid \sigma(X_1, \cdots, X_n)]$$

Lastly, to conclude, we provide some properties of conditional expectations:

Proposition 1.2. *Let a r.v. $X \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, and $\mathcal{A} \in \mathcal{F}$ be a σ-algebra. If $X \in \mathcal{A}$, then $\mathsf{E}[X \mid \mathcal{A}] = X$.*

Proof. Clearly, by definition, for any $A \in \mathcal{A}$:

$$\int_A X \mathrm{d}\mathsf{P} = \int_A X \mathrm{d}\mathsf{P} \implies \mathsf{E}[X \mid \mathcal{A}] = X$$

And since $X \in \mathcal{A}$, then we completed the proof. $\square$

Proposition 1.3. *Let a r.v. $X \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, and $\mathcal{A} \in \mathcal{F}$ be a σ-algebra. If X is independent of $\mathcal{A}$, then $\mathsf{E}[X \mid \mathcal{A}] = \mathsf{E}X$.*

Proof. Just check definition, write: $Y = \mathsf{E}[X \mid \mathcal{A}]$, then $\forall\, A \in \mathcal{A}$:

$$\mathsf{E}Y\mathbf{1}_A = \mathsf{E}X\mathbf{1}_A = \mathsf{E}X\mathsf{E}\mathbf{1}_A = \mathsf{E}\left[(\mathsf{E}X)\mathbf{1}_A\right]$$

As $\sigma(\mathsf{E}X) = \{\emptyset, \Omega\} \in \mathcal{A}$, we have $Y = \mathsf{E}X$ by uniqueness. $\square$

Also more concrete properties with inequalities:

Proposition 1.4. *Let a r.v. $X \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, and $\mathcal{A} \in \mathcal{F}$ be a σ-algebra, then:*

(a) $\mathsf{E}[X \mid \mathcal{F}] = X$.

(b) $\mathsf{E}[X \mid \{\emptyset, \Omega\}] = \mathsf{E}X$

(c) If $X \geq 0$, a.s., then $\mathsf{E}[X \mid \mathcal{A}] \geq 0$, a.s.

Proof. (a) Clearly, $X \in \mathcal{F}$, then by 1.2, $\mathsf{E}[X \mid \mathcal{F}] = X$.

(b) If $X \in \{\emptyset, \Omega\}$, then X is constant, so it holds. Otherwise ,use independence.

(c) Let $Y = \mathsf{E}[X \mid \mathcal{A}] \in \mathcal{A}$, then set: $G = \{Y < 0\}$, and by definition:

$$\mathsf{E}Y\mathbf{1}_G = \mathsf{E}X\mathbf{1}_G \geq 0$$

If $\lambda(G) \neq 0$, then this is a contradiction.

Then we completed the proof. $\square$

Corollary 1.5. *If X, Y be two r.v.s, and $X \geq Y$, a.s., then $\mathsf{E}[X \mid \mathcal{A}] \geq \mathsf{E}[Y \mid \mathcal{A}]$*

Proof. This is simply an application of 1.4's(c). $\square$

Followings are more advanced tools we are going to use:

Proposition 1.6. *For r.v.s $X, Y \in L^1(\Omega, \mathcal{F}, \mathsf{P}), \mathcal{A} \subseteq \mathcal{F}$ is a σ-algebra, then:*

(a) ***Linearity:*** $\alpha\mathsf{E}[X \mid \mathcal{A}] + \beta\mathsf{E}[Y \mid \mathcal{A}] = \mathsf{E}[\alpha X + \beta Y \mid \mathcal{A}]$.

(b) ***Jensen's Inequality:*** *for convex integrable function φ:*

$$\varphi\left(\mathsf{E}[X \mid \mathcal{A}]\right) \leq \mathsf{E}\left[\varphi(X) \mid \mathcal{A}\right] \ a.s.$$

(c) ***Hölder's Inequality:*** *for $1 < p, q < +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$:*

$$\left|\mathsf{E}[XY \mid \mathcal{A}]\right| \leq \mathsf{E}\left[|XY| \mid \mathcal{A}\right] \leq \left(\mathsf{E}\left[|X|^p \mid \mathcal{A}\right]\right)^{1/p} + \left(\mathsf{E}\left[|X|^q \mid \mathcal{A}\right]\right)^{1/q} \ a.s.$$

Proof. The proof is left as homework. $\square$

Proposition 1.7. *For r.v.s X, Y on $(\Omega, \mathcal{F}, \mathsf{P}), \mathcal{A} \subseteq \mathcal{F}$ is a σ-algebra, if $X, XY \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ and $Y \in \mathcal{A}$, then:*

$$\mathsf{E}[XY \mid \mathcal{A}] = Y\mathsf{E}[X \mid \mathcal{A}], \ a.s.$$

Proof. By linearity and approximation, it suffices to only show that it works for indicator function, write: $Z = \mathsf{E}[X \mid \mathcal{A}]$, then $\forall\ B \in \mathcal{A}, \mathsf{E}Z\mathbf{1}_B = \mathsf{E}X\mathbf{1}_B$, now for $A, B \in \mathcal{A}$:

$$\mathsf{E}(Z\mathbf{1}_A \cdot \mathbf{1}_B) = \mathsf{E}(X\mathbf{1}_A \cdot \mathbf{1}_B) \implies \mathbf{1}_A\mathsf{E}[X \mid \mathcal{A}] = Z\mathbf{1}_A = \mathsf{E}[X\mathbf{1}_A \mid \mathcal{A}]$$

Therefore, as $Z\mathbf{1}_A \in \mathcal{A}$, we showed that work for $Y = \mathbf{1}_A$ □

Proposition 1.8. *For a r.v. $X \in L^1(\Omega, \mathcal{F}, \mathsf{P}), \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{F}$ are σ-algebras, then:*

$$\mathsf{E}[\mathsf{E}(X \mid \mathcal{A}_1) \mid \mathcal{A}_2] = \mathsf{E}[X \mid \mathcal{A}_1] = \mathsf{E}[\mathsf{E}(X \mid \mathcal{A}_2) \mid \mathcal{A}_1]$$

Proof. (1) For first equality, since $\mathsf{E}[X \mid \mathcal{A}_1] \in \mathcal{A}_1 \subseteq \mathcal{A}_2$, then by 1.2, it holds:

$$\mathsf{E}[\mathsf{E}(X \mid \mathcal{A}_1) \mid \mathcal{A}_2] = \mathsf{E}[X \mid \mathcal{A}_1]$$

(2) For second equality, write: $Y = \mathsf{E}[X \mid \mathcal{A}_2], Z = \mathsf{E}[Y \mid \mathcal{A}_1]$, then $\forall A \in \mathcal{A}_1 \subseteq \mathcal{A}_2$:

$$\mathsf{E}Z\mathbf{1}_A = \mathsf{E}[\mathsf{E}(X \mid \mathcal{A}_2)\mathbf{1}_A] = \mathsf{E}X\mathbf{1}_A = \mathsf{E}Y\mathbf{1}_A$$

This gives us: $\mathsf{E}[X \mid \mathcal{A}_1] = Z = \mathsf{E}[\mathsf{E}(X \mid \mathcal{A}_2) \mid \mathcal{A}_1]$.
Therefore, we completed the proof. □

1.3 Homeworks

The followings are exercises for this section:

Problem 1.1. Let $X, Y \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ be random variables and $\mathcal{A} \subseteq \mathcal{F}$ be a σ-algebra.

(1) Prove that if $X \geq Y$ a.s., then $\mathsf{E}[X \mid \mathcal{A}] \geq \mathsf{E}[Y \mid \mathcal{A}]$.

(2) For any $\alpha, \beta \in \mathbb{R}$, prove that $\alpha\mathsf{E}[X \mid \mathcal{A}] + \beta\mathsf{E}[Y \mid \mathcal{A}] = \mathsf{E}[\alpha X + \beta Y \mid \mathcal{A}]$.

(3) For convex integrable function φ, prove that

$$\varphi(\mathsf{E}[X|\mathcal{A}]) \leq \mathsf{E}[\varphi(X) \mid \mathcal{A}].$$

(4) For $1 < p, q < +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, prove that

$$|\mathsf{E}[XY \mid \mathcal{A}]| \leq (\mathsf{E}[|X|^p \mid \mathcal{A}])^{1/p} (\mathsf{E}[|Y|^q \mid \mathcal{A}])^{1/q}.$$

2 Discrete-time Stochastic Process

After establishing the most thrilling tools, we are ready to start the journey of stochastic process, and firstly, we focus on general properties of discrete time stochastic process, where we first define some general concepts, then filtration, stopping time, a certain type of r.v., then basic disrete martingale theory.

2.1 Formalise Process

Firstly, we need to make sure our focus, what kind of stochastic process, and how:

Definition 2.1. On $(\Omega, \mathcal{F}, \mathsf{P})$, a sequence $\{X_n : n \in \mathbb{N}\}$, where each X_n is a r.v., is called a **discrete time stochastic process**. Moreover:

(a) If the r.v.s, $X_1, X_2, \cdots$ are independent, then the process, $\{X_n : n \in \mathbb{N}\}$ is an **independent process**.

(b) If the r.v.s, $X_1, X_2, \cdots$ are i.i.d., then the process $\{X_n : n \in \mathbb{N}\}$ is a **stationary independent process**.

Above sounds abstract, but we could see some familiarities from below:

Definition 2.2. Let $\{X_n : n \in \mathbb{N}\}$ be a stationary independent process, we write $S_0 = 0$ and $S_n = \sum_{i=1}^n X_i$, then the process $\{S_n : n \in \mathbb{N}\}$ is a **random walk**.

From this definition, we could provide following example:

Example 2.1. Let $X_1, X_2, \cdots$ be i.i.d. Bernoulli r.v.s, i.e. $\mathsf{P}(X_1 = 1) = p$ and $\mathsf{P}(X_1 = -1) = 1 - p$. Then $S_n = \sum_{i=1}^n X_i$ can be viewed as a n-step walk from 0.

Correspondingly, we state a property of random walk:

Definition 2.3. Let $S = \{S_n : n \in \mathbb{N}\}$ be a random walk and $x \in \mathbb{R}$ if:

$$\forall\, \varepsilon > 0, \mathsf{P}\left(\limsup_{n\to\infty} \{|S_n - x| < \varepsilon\}\right) = 1$$

, then x is called a **recurrent value** of S. If $\forall\, \varepsilon > 0, \exists\, n \in \mathbb{N}$, s.t:

$$\mathsf{P}(|S_n - x| < \varepsilon) > 0$$

, then x is a **possible value** of S.

Remark 2.1. Recall: $\limsup_{n\to\infty}\{|S_n - x| < \varepsilon\} = \cap_{N=1}^{\infty} \cup_{n=N}^{\infty} \{|S_n - x| < \varepsilon\}$.

We point out here but won't delve into, the tools to study random walk is Borel-Cantelli Lemma, Law of Large Numbers, Characteristic function, etc.

2.2 Filtration

A process evolves as time, so as the information, then how to depict this phenomenon:

Definition 2.4. On $(\Omega, \mathcal{F}, \mathsf{P})$, a sequence $\{\mathcal{F}_n : n \in \mathbb{N}\}$ is called a **filtration** if $\mathcal{F}_n \subseteq \mathcal{F}$ is a σ-algebra satisfying, $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}, \forall\, n$. If $X = \{X_n : n \in \mathbb{N}\}$ is a process satisfying $X_n \in \mathcal{F}_n\ \forall\, n$, then we say X is **adapted** to $\{\mathcal{F}_n : n \in \mathbb{N}\}$.

Remark 2.2.

1. $\forall\, n \in \mathbb{N}, X_1, \cdots, X_n \in \mathcal{F}_n$, that is $\mathcal{F}_n$ contains information of $X_1, \cdots, X_n$.
2. For a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$, we write $\mathcal{F}_\infty$ for the σ-algebra generated by $\cup_{n=1}^{\infty} \mathcal{F}_n$.
3. For a process $X = \{X_n : n \in \mathbb{N}\}$, we set $\mathcal{F}_n^X = \sigma(X_1, \cdots, X_n)$, the $\{\mathcal{F}_n^X : n \in \mathbb{N}\}$ is a filtration and X is adapted to $\{\mathcal{F}_n^X : n \in \mathbb{N}\}$. Then we say $\{\mathcal{F}_n^X : n \in \mathbb{N}\}$ is the filtration generated by X, and indeed the **smallest** filtration s.t. X is adapted to.

After introducing filtration, we could define a type of r.v.s called optional r.v.:

Definition 2.5. For a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$, if a r.v. α satisfies:

(a) $\alpha \in \mathbb{N} \cup \{\infty\}$.

(b) $\forall\, n \in \mathbb{N} \cup \{\infty\},\ \{\omega : \alpha(\omega) = n\} \in \mathcal{F}_n$.

, then we say α is **optional** relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$.

Strictly equal is too demanding, and using the knowledge in *Remark* 2.2:

Remark 2.3.

1. $\{\alpha = n\}$ can be replaced by $\{\alpha \le n\}$.
2. For a stochastic process $X = \{X_n : n \in \mathbb{N}\}$, if α is optional relative to $\{\mathcal{F}_n^X : n \in \mathbb{N} \cup \{\infty\}\}$, then we say α is optional relative to X.

To avoid abstractness, we provide an example:

Example 2.2. For $X = \{X_n : n \in \mathbb{N}\}$, $c \in \mathbb{R}$, we set $\alpha = \min\{n : X_n < c\}$, then α is optional relative to X.

Proof. Simply, check:

$$\{\alpha = n\} = \{X_1 \ge c\} \cap \cdots \cap \{X_{n-1} \ge c\} \cap \{X_n < c\} \in \mathcal{F}_n^X$$

Then we smoothly completed the proof. □

A natural question is that optional r.v. tends to represent time related event, then does it has σ-algebra. (Actually, stopping time is a new way of seperation of time)

Definition 2.6. For a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$ and an optional r.v., α, define:

$$\mathcal{F}_\alpha = \{A \in \mathcal{F}_\infty : A \cap \{\alpha \le n\} \in \mathcal{F}_n, \forall\, n \in \mathbb{N} \cup \{\infty\}\}$$

, is called **pre-α** σ-algebra.

Remark 2.4.

1. $\mathcal{F}_\alpha$ is indeed a σ-algebra. (the verification leave as homework)
2. Useful tool when we dealing with optional r.v.:

$$\forall\, A \in \mathcal{F}_\infty,\ A = \bigcup_{n \in \mathbb{N} \cup \{\infty\}} (A \cap \{\alpha = n\})$$

Instead of stating out more definitions, we first see some properties of optional r.v.s:

Proposition 2.1. *Let α, β be optional r.v.s relative to $\{\mathcal{F}_n : n \in \mathbb{N} \cup \{\infty\}\}$, then:*

(a) If $\alpha \leq \beta$ a.s, then $\mathcal{F}_\alpha \subseteq \mathcal{F}_\beta$.

(b) $\mathcal{F}_{\alpha \wedge \beta} = \mathcal{F}_\alpha \cap \mathcal{F}_\beta$.

Proof. (a) Since $\alpha \leq \beta$, then $\{\beta \leq n\} \subseteq \{\alpha \leq n\}, \forall\, n \in \mathbb{N}$, then, $\forall\, A \in \mathcal{F}_\alpha$:

$$A \cap \{\beta \leq n\} = (A \cap \{\alpha \leq n\}) \cap \{\beta \leq n\} \in \mathcal{F}_n$$

, for $n \in \mathbb{N} \cup \{\infty\}$, then $A \in \mathcal{F}_\beta$.

(b) As $\alpha \wedge \beta \leq \alpha, \alpha \wedge \beta \leq \beta$, then $\mathcal{F}_{\alpha\wedge\beta} \subseteq \mathcal{F}_\alpha \cap \mathcal{F}_\beta$. Reversely, if $A \in \mathcal{F}_\alpha \cap \mathcal{F}_\beta$:

$$A \cap \{\alpha \wedge \beta \leq n\} = A \bigcap (\{\alpha \leq n\} \cup \{\beta \leq n\}) = (A \cap \{\alpha \leq n\}) \bigcup (A \cap \{\beta \leq n\}) \in \mathcal{F}_n$$

, therefore $\mathcal{F}_\alpha \cap \mathcal{F}_\beta \subseteq \mathcal{F}_{\alpha\wedge\beta}$.

Hence, we completed the proof. □

2.3 Basic Martingale Theory

For discrete time optional r.v.s, we don't want to waste too much time in providing more propositions and properties, let's directly jump into the core of this course:

Definition 2.7. On $(\Omega, \mathcal{F}, \mathsf{P})$, $X = \{X_n : n \in \mathbb{N}\}$ is an adapted process relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$, and $\forall\, n,\ X_n \in L^1(\Omega, \mathcal{F}, \mathsf{P})$:

(a) If $X_n \geq \mathsf{E}\left[X_{n+1} \mid \mathcal{F}_n\right]$, then X_n is called $\{\mathcal{F}_n : n \in \mathbb{N}\}$-supermartingale.

(b) If $X_n \leq \mathsf{E}\left[X_{n+1} \mid \mathcal{F}_n\right]$, then X_n is called $\{\mathcal{F}_n : n \in \mathbb{N}\}$-submartingale.

(c) If $X_n = \mathsf{E}\left[X_{n+1} \mid \mathcal{F}_n\right]$, then X_n is called $\{\mathcal{F}_n : n \in \mathbb{N}\}$-martingale.

Remark 2.5. 1. X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-supermartingale, then $-X = \{-X_n : n \in \mathbb{N}\}$ is $\{\mathcal{F}_n : n \in \mathbb{N}\}$-submartingale.

2. X is $\{\mathcal{F}_n : n \in \mathbb{N}\}$-martingale iff X is $\{\mathcal{F}_n : n \in \mathbb{N}\}$-super- and sub-martingale.

It is easy to see that in general submartingale are "increasing" in some sense, formally:

Example 2.3. Choose $X_n = a_n, \forall\, n \in \mathbb{N}$, and adapted to any filtration,$\{\mathcal{F}_n : n \in \mathbb{N}\}$. If $\{a_n : n \in \mathbb{N}\}$ is a decreasing/increasing /constant sequence, then X_n is $\{\mathcal{F}_n : n \in \mathbb{N}\}$-super-/sub-martingale.

Following, we provide a more concrete example:

Example 2.4. Say $X = \{X_n : n \in \mathbb{N}\}$ be an independent process relative to $\mathcal{F}_n^X = \sigma(X_1, \cdots, X_n)$, also $\forall\, n,\ X_n \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, denote: $S_n = \sum_{k=1}^n X_k$, then:

$$\mathsf{E}\left[S_{n+1} \mid \mathcal{F}_n^X\right] = \mathsf{E}\left[S_n + X_{n+1} \mid \mathcal{F}_n^X\right] = S_n + \mathsf{E}\left[X_{n+1} \mid \mathcal{F}_n^X\right] = S_n + \mathsf{E}X_{n+1}$$

Therefore, we consider following situations:

1. If $\forall\, n,\ \mathsf{E}X_n \geq 0$, then S_n is submartingale.
2. If $\forall\, n,\ \mathsf{E}X_n \leq 0$, then S_n is supermartingale.
3. If $\forall\, n,\ \mathsf{E}X_n = 0$, then S_n is martingale.

Lastly, we provide a interesting lemma of **Doob martingale**:

Example 2.5. Let $Y \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, and a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$, then for a process $X = \{X_n : n \in \mathbb{N}\}$ is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-martingale by defining:

$$X_n = \mathsf{E}\left[Y_n \mid \mathcal{F}_n\right],\ \forall\, n \in \mathbb{N} \cup \{\infty\}$$

Remark 2.6. Normally: martingale + $\dfrac{\text{increasing}}{\text{decreasing}} = \dfrac{\text{submartingale}}{\text{supermartingale}}$.

Then a natural motivation is to ask: what the martingale looks like in general?

Definition 2.8. For a process $X = \{X_n : n \in \mathbb{N}\}$, $\forall\, n,\ X_n \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ with:

$$X_1 = 0 \text{ and } X_n \leq X_{n+1}, \forall\, n \in \mathbb{N}$$

, then we call X is an increasing process.

Besides that in order to introduce decomposition, we need one more friend:

Definition 2.9. A process $X = \{X_n : n \in \mathbb{N}\}$ and a filtration, $\{\mathcal{F}_n : n \in \mathbb{N}\}$,if:

$$\forall\, n \in \mathbb{N},\ X_{n+1} \in \mathcal{F}_n$$

, then we call X is a predictable prcess.

Based on the newly introduced concepts, we land on astounding result:

Theorem 2.2. ***(Doob's Decomposition Theorem)***
Any submartingale, $X = \{X_n : n \in \mathbb{N}\}$ relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$, there exists a martingale, $Y = \{Y_n : n \in \mathbb{N}\}$ relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$, and an increasing process, $Z = \{Z_n : n \in \mathbb{N}\}$, such that: $X_n = Y_n + Z_n, \forall\, n$.

Proof. To prove this, we need to construct the corresponding process, firstly $n \geq 2$:

$$Z_n = \sum_{k=2}^{n} \left[\mathsf{E}\left(X_k \mid \mathcal{F}_{k-1}\right) - X_{k-1}\right]$$

with convention $Z_1 = 0$, and X is submartingale, then Z is increasing process since:

$$\mathsf{E}\left(X_k \mid \mathcal{F}_{k-1}\right) - X_{k-1} \geq 0 \implies Z_n \uparrow$$

Then it is easily check following is a martingale by $Y_1 = X_1$, and for $n \geq 2$:

$$Y_n = X_n - Z_n = \sum_{k=2}^{n} \left[X_k - \mathsf{E}\left(X_k \mid \mathcal{F}_{k-1}\right)\right] + X_1$$

Therefore, we completed the proof by finding two processes out. $\square$

Remark 2.7. Notice that $Z_n \in \mathcal{F}_{n-1}$, that is to say Z is a predictable process. If in addition, we require Z to be predictable, then the Doob's decomposition is unique.

From above really intuitive (?) theorem, we have following useful corollary:

Corollary 2.3. *Let $X = \{X_n : n \in \mathbb{N}\}$ is a submartingale relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$, and for the Doob's decomposition, i.e. $X_n = Y_n + Z_n$:*

(a) If X is L^1 bounded, then Y and Z are also L^1 bounded.

(b) If X is uniformly integrable, then Y amd Z are also uniformly integrable.

Proof. (a) Firstly, since Y is martingale, then $\mathsf{E}|Y_1| < +\infty$, therefore:

$$0 \leq \mathsf{E}|Z_n| \leq \mathsf{E}|X_n| + \mathsf{E}|Y_n| \leq \sup_n \mathsf{E}|X_n| + \mathsf{E}|Y_1| < +\infty$$

, next, similarly, using above one can also show:

$$\mathsf{E}|Y_n| \leq \mathsf{E}|X_n| + \mathsf{E}|Z_n| \leq 2 \sup_n \mathsf{E}|X_n| + \mathsf{E}|Y_1| < +\infty$$

(b) Firstly, let's see what uniformly integrability of X gives, $\forall\, n \in \mathbb{N}$:

$\exists\, A$, s.t. $\mathsf{E}|X_n|\mathbf{1}_{\{|X_n|>A\}} < 1 \implies \mathsf{E}|X_n| = \mathsf{E}|X_n|\mathbf{1}_{\{|X_n|>A\}} + \mathsf{E}|X_n|\mathbf{1}_{\{|X_n|\leq A\}} < 1 + A$

Above gives X is L^1 bounded, so as Z, since its non-negative and increasing:

$$Z_n \geq 0, Z_n \uparrow \implies Z' = \lim_{n\to\infty} Z_n \;\exists \overset{\text{Fatou}}{\implies} \mathsf{E}Z' \leq \liminf_{n\to\infty} \mathsf{E}Z_n \leq \sup_n \mathsf{E}Z_n < +\infty$$

Hence $\forall\, n \in \mathbb{N}, \forall\, A \in \mathbb{R}$, by increasing $Z_n \leq Z'$, then:

$$0 \leq \mathsf{E}Z_n\mathbf{1}_{\{Z_n>A\}} \leq \mathsf{E}Z'\mathbf{1}_{\{Z'>A\}} \xrightarrow{A\to+\infty} 0$$

Therefore, Z is also u.i., for L^1 bounded Y, decompose into, $\forall\, E \in \mathcal{F}$:

$$\mathsf{E}|Y_n|\mathbf{1}_E \leq \mathsf{E}|X_n|\mathbf{1}_E + \mathsf{E}Z_n\mathbf{1}_E$$

Then we can control Y by X and Z, so we completed the proof. $\square$

Remark 2.8. Quick recall for the convergence concepts:

1. L^1 bounded: $\exists\, C \in \mathbb{R}, \mathsf{E}|X_n| < C < +\infty \Leftrightarrow \sup_{n\in\mathbb{N}} \mathsf{E}|X_n| < +\infty$.
2. Uniformly integrable: $\lim_{M\to\infty} \mathsf{E}|X_n|\mathbf{1}_{\{|X_n|>M\}} = 0$, another equivlance condition: X is L^1 bounded **plus** $\forall\, \varepsilon > 0, \exists\, \delta = \delta(\varepsilon) > 0, s.t.$:

$$\forall\, E \in \mathcal{F}, \mathsf{P}(E) < \delta \implies \forall\, n \in \mathbb{N}, \mathsf{E}|X_n|\mathbf{1}_E < \varepsilon$$

Then we restrict on submartingale, and introduce some properties:

Proposition 2.4. *If X is a $\{\mathcal{F}_n\}$-submartingale, then*

(a) $\mathsf{E}X_n \geq \mathsf{E}X_1$.

(b) $\forall\, A_n \in \mathcal{F}_n, \mathsf{E}X_{n+1}\mathbf{1}_{A_n} \geq \mathsf{E}X_n\mathbf{1}_{A_n}$.

Proof. Simply iteratively use tower rule:

(a) $\mathsf{E}X_n = \mathsf{E}\left[\mathsf{E}\left(X_n \mid \mathcal{F}_{n-1}\right)\right] \geq \mathsf{E}X_{n-1} = \mathsf{E}\left[\mathsf{E}\left(X_{n-1} \mid \mathcal{F}_{n-2}\right)\right] \geq \mathsf{E}X_{n-2} \geq \cdots \geq \mathsf{E}X_1$.

(b) $\mathsf{E}X_{n+1}\mathbf{1}_{A_n} = \mathsf{E}\left[\mathsf{E}\left(X_{n+1}\mathbf{1}_{A_n} \mid \mathcal{F}_n\right)\right] = \mathsf{E}\left[\mathsf{E}\left(X_{n+1} \mid \mathcal{F}_n\right)\mathbf{1}_{A_n}\right] \geq \mathsf{E}X_n\mathbf{1}_{A_n}$.

Then we completed the proof. □

Proposition 2.5. *If X, Y are $\{\mathcal{F}_n\}$-submartingale, then $\forall\, a, b \in \mathbb{R}^+$:*

$$aX + bY = \{aX_n + bY_n : n \in \mathbb{N}\}$$

is a $\{\mathcal{F}_n\}$-submartingale.

Proof. It is same for supermartingale, and can be extended to martingale even to whole real line, and the proof is omitted since it just checks the definitions. □

Proposition 2.6. *Let $X = \{X_n : n \in \mathbb{N}\}$ be $\{\mathcal{F}_n\}$-submartingale, and φ be an increasing convex function on $\mathbb{R}$. If $\varphi(X_n) \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, then $\{\varphi(X_n) : n \in \mathbb{N}\}$ is a $\{\mathcal{F}_n\}$-submartingale.*

Proof. It simply follows below:

$$\mathsf{E}\left(\varphi(X_{n+1}) \mid \mathcal{F}_n\right) \geq \varphi\left(\mathsf{E}\left[X_{n+1} \mid \mathcal{F}_n\right]\right) \geq \varphi(X_n)$$

And for integrability, it is obvious. □

Now we are about to preceed to reveal **Optional Sampling Theorem**:

Theorem 2.7. *Let $Y \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ and a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$, then set $X_n = \mathsf{E}\left(Y \mid \mathcal{F}_n\right)$, let α be a optional r.v. with $\{\mathcal{F}_n : n \in \mathbb{N}\}$, then, $X_\alpha = \mathsf{E}\left[Y \mid \mathcal{F}_\alpha\right]$.*

Proof. Firstly, we check integrability, by Jensen's inequality:

$$\begin{aligned}
\mathsf{E}|X_\alpha| &= \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathsf{E}|X_\alpha|\mathbf{1}_{\{\alpha=n\}} = \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathsf{E}|X_n|\mathbf{1}_{\{\alpha=n\}} \\
&\leq \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathsf{E}\left[\mathsf{E}\left(|Y| \mid \mathcal{F}_n\right)\mathbf{1}_{\{\alpha=n\}}\right] \\
&= \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathsf{E}\left[\mathsf{E}\left(|Y| \cdot \mathbf{1}_{\{\alpha=n\}} \mid \mathcal{F}_n\right)\right] \\
&= \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathsf{E}|Y|\mathbf{1}_{\{\alpha=n\}} = \mathsf{E}|Y| < +\infty
\end{aligned}$$

Next, we need to show that $X_\alpha \in \mathcal{F}_\alpha$, indeed, $\forall\, x \in \mathbb{R}$:

$$\{X_\alpha \le x\} \cap \{\alpha \le n\} = \bigcup_{k=1}^{n} (\{X_\alpha \le x\} \cap \{\alpha = k\}) = \bigcup_{k=1}^{n} (\{X_k \le x\} \cap \{\alpha = k\}) \in \mathcal{F}_n$$

, so $\{X_\alpha \le x\} \in \mathcal{F}_\alpha, \forall\, x \implies X_\alpha \in \mathcal{F}_\alpha$. Finally, only need to check, $\forall\, A \in \mathcal{F}_\alpha$:

$$\begin{aligned} \mathsf{E} X_\alpha \mathbf{1}_A &= \sum_n \mathsf{E} X_\alpha \mathbf{1}_A \mathbf{1}_{\{\alpha=n\}} = \sum_n \mathsf{E} X_n \mathbf{1}_A \mathbf{1}_{\{\alpha=n\}} \\ &= \sum_n \mathsf{E} X_n \mathbf{1}_{A\cap\{\alpha=n\}} = \sum_n \mathsf{E} Y \mathbf{1}_{A\cap\{\alpha=n\}} = \mathsf{E} Y \mathbf{1}_A \end{aligned}$$

Then we finished the proof, following the definition of conditional expectation. □

Corollary 2.8. *Let $Y \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ and a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$, then set $X_n = \mathsf{E}\,(Y \mid \mathcal{F}_n)$. Let α, β be two optional r.v.s with $\{\mathcal{F}_n : n \in \mathbb{N}\}$ and $\alpha \le \beta$, then, $\{X_\alpha, X_\beta\}$ is a martingale relative to $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$.*

Proof. Leave as homework. □

Remark 2.9. To generalise it, for $\{\alpha_n : n \in \mathbb{N}\}$ optional r.v.s, and $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_n \le \cdots$, then $\{X_{\alpha_n}\}$ is a martingale relative to $\{\mathcal{F}_{\alpha_n} : n \in \mathbb{N}\}$.

Theorem 2.9. *Let $\{X_n : n \in \mathbb{N}\}$ is $\{\mathcal{F}_n\}$-martingale, and α, β are bounded optional r.v.s satisfying $\alpha \le \beta$, then $\{X_\alpha, X_\beta\}$ is a $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$-martingale.*

Proof. As definition, first to check integrability:

$$\mathsf{E}|X_\alpha| = \sum_{j=1}^{N} \mathsf{E}|X_\alpha| \mathbf{1}_{\{\alpha=j\}} = \sum_{j=1}^{N} \mathsf{E}|X_j| \mathbf{1}_{\{\alpha=j\}} = \sum_{j=1}^{N} \mathsf{E}|X_1| \mathbf{1}_{\{\alpha=j\}} = \mathsf{E}|X_1| < +\infty$$

Then we prove it is martingale by show it is both super- and sub- martingale, $\forall\, A \in \mathcal{F}_\alpha$, $\forall\, k \ge j$, we have $A \cap \{\alpha \le j\} \in \mathcal{F}_j \subseteq \mathcal{F}_k$ and $\{\beta \le k\} \in \mathcal{F}_k$, then:

$$\mathsf{E} X_k \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta>k\}\}} \ge \mathsf{E} X_{k+1} \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta>k\}\}}$$

, one can rewrite it into:

$$\mathsf{E} X_k \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta\ge k\}\}} - \mathsf{E} X_{k+1} \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta\ge k+1\}\}} \ge \mathsf{E} X_k \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta=k\}\}}$$

Summing over k from j to N noticing that on $\{\alpha = j\}$, $j \le \beta \le N$:

$$\mathsf{E} X_\alpha \mathbf{1}_{\{A\cap\{\alpha=j\}\}} \ge \mathsf{E} X_\beta \mathbf{1}_{\{A\cap\{\alpha=j\}\}}$$

Then summing over j from 1 to N to get:

$$\mathsf{E} X_\alpha \mathbf{1}_A \ge \mathsf{E} X_\beta \mathbf{1}_A$$

, now choose $A = \{\omega : \mathsf{E}\,[X_\beta \mid \mathcal{F}_\alpha] > X_\alpha\} \in \mathcal{F}_\alpha$, then:

$$\mathsf{E} X_\alpha \mathbf{1}_A \ge \mathsf{E}\,[\mathsf{E}\,(X_\beta \mid \mathcal{F}_\alpha)\, \mathbf{1}_A] \implies \mathsf{P}(A) = 0$$

Therefore, it is same for submartingale, then finished the proof. □

Remark 2.10. It is true for sub/super-martingale, even is generalisation of corollary 2.8.

Now before we proceed, let's look at a simple example:

Example 2.6. The bounded condition is useful, but we can construct for a general optional r,.v. α, we define a truncation: $\alpha_k = \alpha \wedge k$, which is bounded optioanl r.v.

Then a natural question is that what if optional r.v. is not bounded, then comes following:

Theorem 2.10. *Let $\{X_n : n \in \mathbb{N} \cup \{\infty\}\}$ be a martingale relative to $\{\mathcal{F}_n : n \in \mathbb{N} \cup \{\infty\}\}$, and α, β are two optional r.v.s, with $\alpha \leq \beta$, then $\{X_\alpha, X_\beta\}$ is martingale relative to $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$.*

Proof. (1) We start with supermartingale with $X_\infty = 0$. Then $\forall\, n \in \mathbb{N}$,

$$X_n \geq \mathsf{E}[X_\infty \mid \mathcal{F}_n] = 0 \text{ a.s.} \implies X_n \geq 0, \forall\, n \in \mathbb{N} \text{ a.s.}$$

Next, $\forall j \in \mathbb{N}$, $X_\alpha \mathbf{1}_{\{\alpha=j\}} = X_{\alpha\wedge n}\mathbf{1}_{\{\alpha=j\}}$. Fix j and set $n \to \infty$, we have:

$$X_\alpha \mathbf{1}_{\{\alpha=j\}} = \liminf_{n\to\infty} X_{\alpha\wedge n}\mathbf{1}_{\{\alpha=j\}}$$

Sum over j, we get:

$$X_\alpha \mathbf{1}_{\{\alpha<\infty\}} = \liminf_{n\to\infty} X_{\alpha\wedge n}\mathbf{1}_{\{\alpha<\infty\}}$$

Besides, $X_\alpha \mathbf{1}_{\{\alpha=\infty\}} = 0 \leq \liminf_{n\to\infty} X_{\alpha\wedge n}\mathbf{1}_{\{\alpha=\infty\}}$. As a consequence, we have:

$$X_\alpha \leq \liminf_{n\to\infty} X_{\alpha\wedge n}$$

By Fatou's lemma,

$$\mathsf{E}X_\alpha \leq \mathsf{E}\liminf_{n\to\infty} X_{\alpha\wedge n} \leq \liminf_{n\to\infty} \mathsf{E}X_{\alpha\wedge n} \leq \mathsf{E}X_1$$

Since $\alpha \wedge n$ is bounded optional r.v. Thus, $X_\alpha, X_\beta \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, $\forall A \in \mathcal{F}_\alpha$:

$$\mathsf{E}X_k \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta\geq k\}\}} - \mathsf{E}X_{k+1}\mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta\geq k+1\}\}} \geq \mathsf{E}X_k \mathbf{1}_{\{A\cap\{\alpha=j,\beta=k\}\}}$$

, $\forall\, M \geq j$, sum over k from j to M to obtain:

$$\mathsf{E}X_\alpha \mathbf{1}_{\{A\cap\{\alpha=j\}\}} - \mathsf{E}X_{M+1}\mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta\geq M+1\}\}} \geq \mathsf{E}X_\beta \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta\leq M\}\}}$$

Thus, $\mathsf{E}X_\alpha \mathbf{1}_{\{A\cap\{\alpha=j\}\}} \geq \mathsf{E}X_\beta \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta\leq M\}\}}$. As $X_\beta \geq 0$, by MCT, we get:

$$\mathsf{E}X_\alpha \mathbf{1}_{\{A\cap\{\alpha=j\}\}} \geq \mathsf{E}X_\beta \mathbf{1}_{\{A\cap\{\alpha=j\}\cap\{\beta<\infty\}\}}$$

Sum over j, we get:

$$\mathsf{E}X_\alpha \mathbf{1}_{\{A\cap\{\alpha<\infty\}\}} \geq \mathsf{E}X_\beta \mathbf{1}_{\{A\cap\{\beta<\infty\}\}}$$

Besides, noting that $X_\infty = 0$,

$$\mathsf{E}X_\alpha \mathbf{1}_{\{A\cap\{\alpha=\infty\}\}} = \mathsf{E}X_\infty \mathbf{1}_{\{A\cap\{\alpha=\infty\}\}} = 0 = \mathsf{E}X_\infty \mathbf{1}_{\{A\cap\{\beta=\infty\}\}} = \mathsf{E}X_\beta \mathbf{1}_{\{A\cap\{\beta=\infty\}\}}$$

Combining two formula, we get:

$$\mathsf{E}X_\alpha \mathbf{1}_A \geq \mathsf{E}X_\beta \mathbf{1}_A \implies \mathsf{E}X_\alpha \mathbf{1}_A \geq \mathsf{E}[\mathsf{E}[X_\beta \mid \mathcal{F}_\alpha]\cdot \mathbf{1}_A]$$

Thus, $X_\alpha \geq \mathsf{E}[X_\beta \mid \mathcal{F}_\alpha]$.

(2) In general case of supermartingale, we set $X'_n = \mathsf{E}[X_\infty \mid \mathcal{F}_n]$, then $\{X'_n : n \in \mathbb{N} \cup \{\infty\}\}$ is a $\{\mathcal{F}_n : n \in \mathbb{N} \cup \{\infty\}\}$-martingale and $\{X_n - X'_n : n \in \mathbb{N}\}$ is a supermartingale relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$. On one hand, $\{X'_\alpha, X'_\beta\}$ is a martingale relative to $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$. On the other hand, from $X'_\infty = X_\infty$, we can conclude that $\{X_\alpha - X'_\alpha, X_\beta - X'_\beta\}$ is a supermartingale relative to $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$.

(3) In the case of submartingale, $\{-X_n : n \in \mathbb{N} \cup \{\infty\}\}$ is a supermartingale.

(4) In the case of martingale, by (2) and (3).

Finally, we completed the proof. □

After providing these thrilling theorems, we introduce **martingale transformation**.

Definition 2.10. Let $Y = \{Y_n : n \in \mathbb{N}\}$, on $(\Omega, \mathcal{F}, \mathsf{P})$ be a martingale relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$ with $Y_0 = 0$, and another $Z = \{Z_n : n \in \mathbb{N}\}$ is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-predictable process. Then if we set, $X_0 = 0$ and:

$$X_n = \sum_{k=1}^{n} Z_k (Y_k - Y_{k-1})$$

, we say $X = \{X_n : n \in \mathbb{N}\}$ is a martingale transfomation of Y through Z.

Remark 2.11. If you learnt stochastic calculus, then this is some how discrete version of:

$$X = \int Z \mathrm{d}Y$$

Next, we introduce two propositions of martingale transformation:

Proposition 2.11. *Let Y be a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-martingale, and Z is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-predictable process (non-negative), so if X, is a martingale transformation of Y through Z, is integrable, then X is $\{\mathcal{F}_n : n \in \mathbb{N}\}$-martingale.*

Proof. By definition, X is $\{\mathcal{F}_n : n \in \mathbb{N}\}$-adapted and integrable, only need to check:

$$\mathsf{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = \mathsf{E}\left[Z_{n+1}(Y_{n+1} - Y_n) + X_n \mid \mathcal{F}_n\right] = X_n + Z_{n+1}\left(\mathsf{E}\left[Y_{n+1} \mid \mathcal{F}_n\right] - Y_n\right) = X_n$$

, $\forall\, n \in \mathbb{N}$, then we completed the proof. □

Following is to extend the result to sub-/super-martingale:

Proposition 2.12. *Let Y be a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-sub/super-martingale, and Z is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-predictable process (non-negative), so if X, is a martingale transformation of Y through Z, is integrable, then X is $\{\mathcal{F}_n : n \in \mathbb{N}\}$-sub/super-martingale.*

Proof. Leave as homework. □

2.4 Backward Martingale

Since for continuous stochastic process, when dealing with optional sampling theorem, we need to use some properties in backward martingale, then we state some:

Definition 2.11. Let $X = \{X_n : n \in \mathbb{N}\}$ be an integrable process on $(\Omega, \mathcal{F}, \mathsf{P})$, let $\{\mathcal{F}_n : n \in \mathbb{N}\}$ be a sequence of **decreasing** σ-algebra with $X_n \in \mathcal{F}_n, \forall\, n$ a.s. then:

(a) If $X_{n+1} \geq \mathsf{E}\left(X_n \mid \mathcal{F}_{n+1}\right)$, then X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-backward supermartingale.

(b) If $X_{n+1} \leq \mathsf{E}\left(X_n \mid \mathcal{F}_{n+1}\right)$, then X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-backward submartingale.

(c) If $X_{n+1} = \mathsf{E}\left(X_n \mid \mathcal{F}_{n+1}\right)$, then X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-backward martingale.

Remark 2.12. $\forall\, n \leq 0, Y_n = X_{-n}, \mathcal{G}_n = \mathcal{F}_{-n}$, then $\{Y_n : n \in \mathbb{N}\}$ is martingale.
Next, listing some properties with backward martingale, which are similar than previous:

Proposition 2.13. *Let $\{\mathcal{F}_n : n \in \mathbb{N}\}$ be a sequence of decreasing σ-algebra, $X = \{X_n : n \in \mathbb{N}\}$ be a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-backward submartingale. Also φ be increasing convex function, with $\varphi(X_n) \in L^1(\Omega, \mathcal{F}, \mathsf{P}), \forall\, n$. Then $\{\varphi(X_n) : n \in \mathbb{N}\}$ is also a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-backward submartingale.*

Proof. By Jensen's inequality, we have:

$$\varphi(X_{n+1}) \leq \varphi\left(\mathsf{E}\left[X_n \mid \mathcal{F}_{n+1}\right]\right) \leq \mathsf{E}\left[\varphi(X_n) \mid \mathcal{F}_{n+1}\right]$$

,$\forall\, n \in \mathbb{N}$, therefore we completed the proof. □

Remark 2.13. For backward martingale, φ need not to be increasing.
And it also has really good properties:

Corollary 2.14. *If in addition, $\varphi \geq 0$, then $\{\varphi(X_n) : n \in \mathbb{N}\}$ is u.i.*

Proof. Since its submartingale, thus $\forall\, A \in \mathcal{F}_n$:

$$\mathsf{E}\varphi(X_n)\mathbf{1}_A \leq \mathsf{E}\left[\mathsf{E}\left(\varphi(X_1) \mid \mathcal{F}_n\right)\mathbf{1}_A\right] = \mathsf{E}\varphi(X_1)\mathbf{1}_A$$

If we choose $A = \{\omega : \varphi(X_n) > M\} = \{X_n > \varphi^{-1}(M)\} \in \mathcal{F}_n$, then:

$$\sup_{n\in\mathbb{N}} \mathsf{E}\varphi(X_n)\mathbf{1}_{\{\varphi(X_n)>M\}} \leq \sup_{n\in\mathbb{N}} \mathsf{E}\varphi(X_1)\mathbf{1}_{\{\varphi(X_n)>M\}}$$

Lastly, since as $M \to +\infty$, we have:

$$\sup_{n\in\mathbb{N}} \mathsf{P}(\varphi(X_n) > M) \leq \sup_{n\in\mathbb{N}} \frac{1}{M}\mathsf{E}\varphi(X_n) = \frac{1}{M}\mathsf{E}\varphi(X_1) \longrightarrow 0$$

, then the integrability of $\varphi(X_1)$ implies u.i. of $\{\varphi(X_n) : n \in \mathbb{N}\}$. □

Lastly, we extend this to backward martingale:

Proposition 2.15. *Let $\{\mathcal{F}_n : n \in \mathbb{N}\}$ be a decreasing σ-algebra, and $X = \{X_n : n \in \mathbb{N}\}$ be a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-backward martingale. Then X is uniformly integrable.*

Proof. Leave as homework. □

2.5 Markov Property

We stop the procedure of backward martingale, then into a more interesting topic:

Definition 2.12. A stochastic process $X = \{X_n : n \in \mathbb{N}\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$ satisfying:

$$\forall\, B \in \mathcal{B}(\mathbb{R}), \forall\, n,\ \mathsf{P}(X_{n+1} \in B \mid X_1, \cdots, X_n) = \mathsf{P}(X_{n+1} \in B \mid X_n)$$

, then X is said to have **Markov Property**.

Any process having Markov property is called markov process/chain, another definition:

Definition 2.13. A stochastic process $X = \{X_n : n \in \mathbb{N}\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$ is a markov process if and only if: $\forall$ integrable $Y \in \sigma(X_{n+1})$,

$$\mathsf{E}\left(Y \mid X_1, \cdots, X_n\right) = \mathsf{E}\left(Y \mid X_n\right)$$

Proof. (1) Forward direction, we only need to choose $Y = \mathbf{1}_{\{X_{n+1} \in B\}}$.

(2) For backward direction, by linearity it holds for $Y = \sum_{i=1}^{M} a_i \mathbf{1}_{A_i}$ for some $A_i \in \sigma(X_{n+1})$. By approximation from below, it holds for non-negative $Y \in \sigma(X_{n+1})$.

Then general cases can be splited into positive and negative parts. $\square$

The following corollary extends above to further:

Corollary 2.16. *A stochastic process $X = \{X_n : n \in \mathbb{N}\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$ is a markov process, then $\forall\, k \in \mathbb{N}, Y \in \sigma(X_{n+k})$:*

$$\mathsf{E}\left(Y \mid X_1, \cdots, X_n\right) = \mathsf{E}\left(Y \mid X_n\right)$$

Proof. By markov property, $\mathsf{E}\left[Y \mid X_1, \cdots, X_{n+k-1}\right] = \mathsf{E}\left[Y \mid X_{n+k-1}\right] \in \sigma(X_{n+k-1})$, so

$$\forall\, A \in \sigma(X_n, \cdots, X_{n+k-1}), \mathsf{E}\left[\mathsf{E}\left(Y \mid X_1, \cdots, X_{n+k-1}\right) \mathbf{1}_A\right] = \mathsf{E} Y \mathbf{1}_A$$

Hence, we have starting point that $\forall\, k \in \mathbb{N}$:

$$\mathsf{E}\left[Y \mid X_{n+k-1}\right] = \mathsf{E}\left[Y \mid X_1, \cdots, X_{n+k-1}\right] = \mathsf{E}\left[Y \mid X_n, \cdots, X_{n+k-1}\right]$$

Next, we prove by induction, for $k = 1$ is above, assume it holds for $k - 1$, then:

$$\begin{aligned} &\mathsf{E}\left[Y \mid X_1, \cdots, X_n\right] = \mathsf{E}\left[\mathsf{E}\left(Y \mid X_1, \cdots, X_{n+k-1}\right) \mid X_1, \cdots, X_n\right] \\ &= \mathsf{E}\left[\mathsf{E}\left(Y \mid X_{n+k-1}\right) \mid X_1, \cdots, X_n\right] = \mathsf{E}\left[\mathsf{E}\left(Y \mid X_{n+k-1}\right) \mid X_n\right] \\ &= \mathsf{E}\left[\mathsf{E}\left(Y \mid X_n, \cdots, X_{n+k-1}\right) \mid X_n\right] = \mathsf{E}\left[Y \mid X_n\right] \end{aligned}$$

Then we finished the proof. $\square$

To conclude the markov property, we present following more general theorem:

Theorem 2.17. *Let $X = \{X_n : n \in \mathbb{N}\}$ be a process on $(\Omega, \mathcal{F}, \mathsf{P})$, then following are equivalent:*

(a) X is a Markov process.

(b) $\forall\, n$, $\forall\, M \in \sigma(X_{n+1}, X_{n+2}, \dots), \mathsf{P}(M \mid X_1, \cdots, X_n) = \mathsf{P}(M \mid X_n)$.

(c) $\forall\, n$, $\forall M_1 \in \sigma(X_1, \dots, X_n), M_2 \in \sigma(X_{n+1}, X_{n+2}, \cdots)$, *then:*

$$\mathsf{P}(M_1 M_2 \mid X_n) = \mathsf{P}(M_1 \mid X_n)\mathsf{P}(M_2 \mid X_n)$$

Proof. (1) First to show (b) $\implies$ (c):

$$\begin{aligned}\text{RHS} &= \mathsf{E}\left[\mathbf{1}_{M_1} \mid X_n\right]\mathsf{E}\left[\mathbf{1}_{M_2} \mid X_n\right] = \mathsf{E}\left[\mathbf{1}_{M_1}\mathsf{E}\left[\mathbf{1}_{M_2} \mid X_n\right] \mid X_n\right] \\ &\overset{(b)}{=} \mathsf{E}\left[\mathbf{1}_{M_1}\mathsf{E}\left[\mathbf{1}_{M_2} \mid X_1, \cdots, X_n\right] \mid X_n\right] = \mathsf{E}\left[\mathsf{E}\left[\mathbf{1}_{M_1}\mathbf{1}_{M_2} \mid X_1, \cdots, X_n\right] \mid X_n\right] \\ &= \mathsf{E}\left[\mathbf{1}_{M_1}\mathbf{1}_{M_2} \mid X_n\right] = \text{LHS}\end{aligned}$$

(2) Then we check (c) $\implies$ (b), $\forall\, A \in \sigma(X_n), M' \in \sigma(X_1, \cdots, X_n)$, we have:

$$\begin{aligned}&\mathsf{E}\left[\mathsf{E}\left[\mathbf{1}_M \mid X_n\right]\mathbf{1}_{AM'}\right] = \mathsf{E}\left[\mathsf{E}\left[\mathbf{1}_M \mid X_n\right]\mathbf{1}_{M'} \cdot \mathbf{1}_A\right] \\ &= \mathsf{E}\left\{\mathsf{E}\left[\mathsf{E}\left(\mathbf{1}_M \mid X_n\right)\mathbf{1}_{M'} \mid X_n\right]\mathbf{1}_A\right\} = \mathsf{E}\left[\mathsf{E}\left[\mathbf{1}_M \mid X_n\right]\mathsf{E}\left[\mathbf{1}_{M'} \mid X_n\right] \cdot \mathbf{1}_A\right] \\ &\overset{(c)}{=} \mathsf{E}\left[\mathsf{E}\left[\mathbf{1}_{MM'} \mid X_n\right]\mathbf{1}_A\right] = \mathsf{E}\left[\mathbf{1}_{MM'}\mathbf{1}_A\right] = \mathsf{E}\left[\mathbf{1}_M\mathbf{1}_{AM'}\right]\end{aligned}$$

(3) Easily, for (b) $\implies$ (a), choose $M = \mathbf{1}_{\{X_{n+1} \in B\}}$.

(4) For (a) $\implies$ (b), similar argument, $\forall\, k \in \mathbb{N}$, $\mathsf{E}\left[Y \mid X_1, \cdots, X_n\right] = \mathsf{E}\left[Y \mid X_n\right]$ for all bounded $Y \in \sigma(X_{n+1}, \cdots, X_{n+k})$. Thus, for $M \in \bigcup_{k=1}^{\infty} \sigma(X_{n+1}, \cdots, X_{n+k})$, (b) holds. Hence, (b) holds for $M \in \sigma(\bigcup_{k=1}^{\infty} \sigma(X_{n+1}, \cdots, X_{n+k})) = \sigma(X_{n+1}, \cdots)$.

Therefore, we completed the proof. □

To see the existence of markov process, we provide an example:

Example 2.7. Let $X = \{X_n : n \in \mathbb{N}\}$ be an independent process, then X is markov process,besides, corresponding random walk (sum process), $S_n = X_1 + \cdots + X_n$, then $S = \{S_n : n \in N\}$, is also markov process.

Now, before proceeding, we define post-α process:

Definition 2.14. Let $\{X_n : n \in \mathbb{N}\}$ be a stochastic process, and α be an optional r.v. with $\alpha < \infty$ a.s. Then $\{X_{\alpha+n} : n \in \mathbb{N}\}$ is called the post-α process.

Correspondingly, $\sigma(\{X_{\alpha+n} : n \in \mathbb{N}\})$ is called the post-α σ-algebra, $\mathcal{F}'_\alpha$.

Theorem 2.18. *Let $X = \{X_n : n \in \mathbb{N}\}$ be a stationary independent process, α is an optional r.v. relative to $\{\mathcal{F}_n^X : n \in \mathbb{N}\}$ with $\alpha < \infty$ a.s. Then:*

(a) The pre-α and post-α are independent, i.e.:

$$\{A : A \cap \{\alpha \le k\} \in \mathcal{F}_k^X, \forall\, k\} = \mathcal{F}_\alpha^X \perp \mathcal{F}_\alpha^{X'} = \sigma(\{X_{\alpha+n} : n \in \mathbb{N}\})$$

(b) The post-α process, i.e. $\{X_{\alpha+n} : n \in \mathbb{N}\}$, is also a stationary independent process with the same distribution of X.

Proof. Firstly, $\forall\ A \in \mathcal{F}_\alpha, k \in \mathbb{N}, \forall\ B_j \in \mathcal{B}(\mathbb{R})$ with $1 \le j \le k$, we have:

$$\begin{aligned}
&\mathsf{P}\left(A \bigcap \{X_{\alpha+j} \in B_j, 1 \le j \le k\}\right) = \sum_{n=1}^{\infty} \mathsf{P}\left(A \cap \{\alpha = n\} \cap \{X_{\alpha+j} \in B_j, 1 \le j \le k\}\right) \\
&= \sum_{n=1}^{\infty} \mathsf{P}\left(A \cap \{\alpha = n\} \cap \{X_{n+j} \in B_j, 1 \le j \le k\}\right) \ (A \cap \{\alpha = n\} \in \mathcal{F}_n^X \perp \sigma(X_{n+j})) \\
&= \sum_{n=1}^{\infty} \mathsf{P}(A \cap \{\alpha = n\}) \cdot \prod_{j=1}^{k} \mathsf{P}(X_{n+j} \in B_j) = \mathsf{P}(A) \cdot \prod_{j=1}^{k} \mathsf{P}(X_{n+j} \in B_j)
\end{aligned}$$

Choose $A = \Omega, B_1 = \cdots = B_{k-1} = \mathbb{R}$ to obtain the indentically distributed of $X_{\alpha+k}$ and X_{n+k}. Choose $A = \Omega$ to obtain the independence of post-α process. Indepdence of σ-algebras follows from distribution. Then finish the proof. □

Next, we just present a complicated theorem but worthy to know:

Theorem 2.19. *Let $\{X_n : n \in \mathbb{N}\}$ be markov process, and a finite α optional r.v. relative to $\{\mathcal{F}_n^X : n \in \mathbb{N}\}$, then, $\forall\ M \in \mathcal{F}_\alpha^X$:*

$$\mathsf{P}(M \mid \mathcal{F}_\alpha^X) = \mathsf{P}(M \mid \alpha, X_\alpha)$$

Proof. Read the textbook at Theorem 9.2.5. □

Lastly, we end this chapter by providing some notification about markov process:

Remark 2.14. (1) The sum of two markov process may not be a markov process:

> *Example* 2.8. Let X be a r.v. and set $X_n = X, Y_n = (-1)^n X$, then $\{X_n : n \in \mathbb{N}\}$ and $\{Y_n : n \in \mathbb{N}\}$ are markov chains for sure. As $X_n + Y_n = 2X_n \mathbf{1}_{2|n}$, then $\{X_n + Y_n : n \in \mathbb{N}\}$ is not a markov chain.

(2) A markov chain may not be a martingale:

> *Example* 2.9. Let $X = \{X_n : n \in \mathbb{N}\}$ be an independent process and set $S = \{S_n : n \in \mathbb{N}\}$ where $S_n = X_1 + \cdots + X_n$, then both X and S are markov chain. However, if X_{n+1} is not deterministic, then by independence:
>
> $$\mathsf{E}\left[X_{n+1} \mid \mathcal{F}_n^X\right] = \mathsf{E}X_{n+1} \neq X_n$$
>
> , thus X is not a martingale. Same for S if $\mathsf{E}X_{n+1} \neq 0$.

(3) A martingale may not be a markov chain:

> *Example* 2.10. Let $X = \{X_n : n \in \mathbb{N}\}$ be an independent square integrable process with $\mathsf{E}X_n = 0, \forall\ n$. Let W be a bounded r.v. independent of X. For $n \in \mathbb{N}$, set $Y_n = W(X_{n-1} + \cdots + X_1 + 1)$ and $\mathcal{F}_n = \sigma(W, X_1, \cdots, X_{n-1})$. Then $Y = \{Y_n : n \in \mathbb{N}\}$ is martingale w.r.t $\{\mathcal{F}_n : n \in \mathbb{N}\}$. But, not a markov chain.

Proof. Indeed, $Y_{n+1} - Y_n = WX_n$ and $Y_1 = W$, then on one hand:

$$\begin{aligned}
&\mathsf{E}\left[Y_n^2 \mid Y_1, \cdots, Y_n\right] = \mathsf{E}\left[(Y_n + WX_n)^2 \mid Y_1, \cdots, Y_n\right] \\
&= Y_n^2 + 2Y_n \mathsf{E}\left[WX_n \mid Y_1, \cdots, Y_n\right] + \mathsf{E}\left[W^2 X_n^2 \mid Y_1, \cdots, Y_n\right] \\
&= Y_n^2 + 2Y_n W \mathsf{E} X_n + W^2 \mathsf{E} X_n^2 = Y_n^2 + W^2 \mathsf{E} X_n^2
\end{aligned}$$

On another hand,

$$\begin{aligned}
&\mathsf{E}\left[Y_{n+1}^2 \mid Y_n\right] = \mathsf{E}\left[(Y_n + WX_n)^2 \mid Y_n\right] \\
&= Y_n^2 + 2\mathsf{E}\left[Y_n W X_n \mid Y_n\right] + \mathsf{E}\left[W^2 X_n^2 \mid Y_n\right] \\
&= Y_n^2 + 2\mathsf{E}\left[\mathsf{E}\left(Y_n W X_n \mid Y_n, W\right) \mid Y_n\right] + \mathsf{E}\left[\mathsf{E}\left(W^2 X_n^2 \mid Y_n, W\right) \mid Y_n\right] \\
&= Y_n^2 + 2\mathsf{E}\left[Y_n W \mathsf{E} X_n \mid Y_n\right] + \mathsf{E}\left[W^2 \mathsf{E} X_n^2 \mid Y_n\right] \\
&= Y_n^2 + \mathsf{E}\left[W^2 \mid Y_n\right] \cdot \mathsf{E} X_n^2
\end{aligned}$$

And clearly, these two are not equal, then we completed the proof. $\square$

2.6 Homeworks

The followings are exercises for this section:

Problem 2.1. Let $\{\mathcal{F}_n : n \in \mathbb{N}\}$ be a filtration, $X = \{X_n : n \in \mathbb{N}\}$ be an adapted process relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$ such that $X_n \in L^1(\Omega, \mathcal{F}, \mathsf{P})$. Prove that

(1) X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-supermartingale if and only if $X_n \geq \mathsf{E}\left[X_{n+k} \mid \mathcal{F}_n\right]$ for any fixed $k \in \mathbb{N}$.

(2) X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-submartingale if and only if $X_n \leq \mathsf{E}\left[X_{n+k} \mid \mathcal{F}_n\right]$ for any fixed $k \in \mathbb{N}$.

(3) X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-martingale if and only if $X_n = \mathsf{E}\left[X_{n+k} \mid \mathcal{F}_n\right]$ for any fixed $k \in \mathbb{N}$.

(4) What is the case of backward supermartingale/submartingale/martingale if $\{\mathcal{F}_n : n \in \mathbb{N}\}$ is a family of decreasing σ-algebras? State the results and prove them.

Problem 2.2. If X_n is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-martingale and $X_{n+1} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$, prove that $X_n = X_1$ for all $n \in \mathbb{N}$ a.s. Use this result to show the uniqueness of Doob's decomposition $X = Y + Z$ for martingale Y and increasing predictable process Z.

Problem 2.3. For a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$ and an optional random variable α, prove that $\mathcal{F}_\alpha$ is a σ-algebra and $\alpha \in \mathcal{F}_\alpha$.

Problem 2.4. Let $Y \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ be a random variable and $\{\mathcal{F}_n : n \in \mathbb{N}\}$ be a filtration. Set $X_n = \mathsf{E}[Y \mid \mathcal{F}_n]$ for $n \in \mathbb{N} \cup \{\infty\}$.

(1) Prove that $\{X_n : n \in \mathbb{N} \cup \{\infty\}\}$ is a $\{\mathcal{F}_n : n \in \mathbb{N} \cup \{\infty\}\}$-martingale.

(2) If $\alpha \leq \beta$ are optional random variables relative to $\{\mathcal{F}_n, n \in \mathbb{N}\}$, then $\{X_\alpha, X_\beta\}$ is a martingale relative to $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$.

Problem 2.5. Let Y be a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-submartingale/$\{\mathcal{F}_n : n \in \mathbb{N}\}$-supermartingale, Z be a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-predictable non-negative process, and X be a martingale transformation of Y through Z. If X is integrable, then prove that X is a $\{\mathcal{F}_n : n \in \mathbb{N}\}$-submartingale/$\{\mathcal{F}_n : n \in \mathbb{N}\}$-supermartingale.

Problem 2.6. Let $\{\mathcal{F}_n : n \in \mathbb{N}\}$ be a sequence of decreasing σ-algebras, $X = \{X_n : n \in \mathbb{N}\}$ be a backward martingale relative to $\{\mathcal{F}_n : n \in \mathbb{N}\}$. Prove that X is uniformly integrable.

Problem 2.7. Let X_n be an integrable random variable for all $n \in \mathbb{N}$ satisfying:

$$\mathsf{E}[X_{n+1} | X_1, \ldots, X_n] = \frac{X_1 + \ldots + X_n}{n}$$

Prove that $\left\{\frac{X_1+\ldots+X_n}{n} : n \in \mathbb{N}\right\}$ is a martingale relative to $\{\mathcal{F}_n^X : n \in \mathbb{N}\}$.

3 Continuous-time Stochastic Process

The shift from discrete to continuous time marks a profound expansion in our modeling horizon. Here, randomness is no longer observed at isolated ticks of a clock, but evolves fluidly, moment by moment. This chapter formalizes that vision, introducing the core language of continuous-time processes: sample paths, stochastic continuity, modifications, and filtrations. We confront the subtle distinctions between different notions of equality for processes and meet the powerful Kolmogorov Continuity Theorem, a result that often grants us the gift of continuous paths. These concepts are the essential scaffolding for everything that follows, from martingales and Markov processes to the Brownian motion.

3.1 Formalisation of Process

As usual, we first define what is continuous stochastic process:

Definition 3.1. $X = \{X_t : t \geq 0\}$ of r.v.s on $(\Omega, \mathcal{F}, \mathsf{P})$ is a continuous-time stochastic process.

Remark 3.1.

1. If we want continuous stochastic process degenerate, by setting $X_t = X_n, t \in [n, n+1)$.
2. There are two ways of interpreting the r.v.s, $X_t(\omega) : \Omega \times [0, +\infty) \mapsto \mathbb{R}^d/\mathbb{C}^d$:
   (a) From function perspective: $(\omega, t) \mapsto X_t(\omega)$.
   (b) From r.v. perspective:
   $$\begin{aligned} X : &\Omega \mapsto \{\text{function on } [0, +\infty)\} \\ &\omega \mapsto X_t(\omega) \end{aligned}$$
   This is called **sample path/trajectory**, $t \mapsto X_t(\omega), \forall\, \omega \in \Omega$.

Then we could start to add contraint on continuous process:

Definition 3.2. Let X be a process on $(\Omega, \mathcal{F}, \mathsf{P})$,

(a) For a.s. (ω), $t \mapsto X_t(\omega)$ is continuous, then we say X is continuous. (sample path is continuous)

(b) For a.s. (ω), $t \mapsto X_t(\omega)$ is right continuous on $[0, +\infty)$ with finite left limits on $[0, +\infty)$, we say the smaple path of X is RCLL. (Right Continuous Left Limit)

(c) For a.s. (ω), $t \mapsto X_t(\omega)$ is left continuous on $[0, +\infty)$ with finite right limits on $[0, +\infty)$, we say the smaple path of X is LCRL. (Left Continuous Right Limit)

Beside, we propose a new continuous concept that is widely used later:

Definition 3.3. Let X be a stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$. If $\forall\, \varepsilon > 0, \forall\, t \geq 0$:

$$\lim_{s \to t} \mathsf{P}\left(|X_s - X_t| > \varepsilon\right) = 0$$

, then we say that X is **stochastic continuous**.

For continuous time, since we add more points in it, then new concepts come out:

Definition 3.4. Let X, Y be a stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$. If $\forall\, t \geq 0, \forall\, \omega \in \Omega$:

$$X_t(\omega) = Y_t(\omega)$$

, then we say they are same.

However, above is too demanding for a process:

Definition 3.5. Let X, Y be two stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$. If:

$$\mathsf{P}\left(X_t = Y_t, \forall\, t \geq 0\right) = 1$$

, then we say X, Y are **indistinguishable**.

Also we could loose our constraint, below is the most useful one:

Definition 3.6. Let X, Y be two stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$. If:

$$\mathsf{P}\left(X_t = Y_t\right) = 1, \forall\, t \geq 0$$

, then we say X is a **modification** of Y.

Remark 3.2. From the definition, we see: **Indistinguishable** $\Longrightarrow$ **Modification**.

Lastly, since r.v.s' most interesting property is just distribution:

Definition 3.7. Let X, Y be two stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$.
If $\forall\, n \geq 1, \forall\, 0 \leq t_1 < t_2 < \cdots < t_n < \infty, \forall\, \mathcal{B}(\mathbb{R}^d)$:

$$\mathsf{P}\left((X_{t_1}, X_{t_2}, \cdots, X_{t_n}) \in A\right) = \mathsf{P}\left((Y_{t_1}, Y_{t_2}, \cdots, Y_{t_n}) \in A\right)$$

, then we say that X and Y have the same **finite-dimensional distributions**.

Remark 3.3. From the definition, we see: **Modification** $\Longrightarrow$ **finite-dimensional distribution**.

As proposed above, it is still difficult to verify, then Kolmogorov gives following:

Theorem 3.1. ***(Kolmogorov Continuity Theorem)***
Let $X = \{X_t : t \in [0, T_0]\}$ be a process, $T_0 > 0$. If $\exists\, \alpha, \beta, C > 0, s.t\, \forall\, t, s \in [0, T_0]$:

$$\mathsf{E}\left|X_t - X_s\right|^{\alpha} \leq C \cdot |t - s|^{1+\beta}$$

, then there exists a process Y satisfying:

(a) Y is a modification of X.

(b) Y has continuous trajectory.

(c) $\forall\, r \in (0, \beta/\alpha), \exists\, r.v.\ \delta(\omega), C' > 0, s.t:$

$$\mathsf{P}\left\{\omega \in \Omega : \sup_{\substack{s,t\in[0,T_0]\\ 0<|t-s|<\delta(\omega)}} \frac{|Y_t(\omega) - Y_t(\omega)|}{|t-s|^r} < C'\right\} = 1$$

Proof. Firstly, let's check if the process satisfy the criterion, what properties it will have:

1. The process itself is stochastically continuous, $\forall\, \varepsilon > 0, \forall\, t, s \in [0, T_0]$:

$$\lim_{s\to t} \mathsf{P}(|X_t - X_s| > \varepsilon) \leq \lim_{s\to t} \varepsilon^{-\alpha}\mathsf{E}|X_t - X_s|^\alpha \leq C\varepsilon^{-\alpha} \lim_{s\to t} |t-s|^{1+\beta} \longrightarrow 0$$

2. The process itself have "good" properties on a "large" set, to see this, we set:

$$A_n = \left\{\omega \in \Omega : \max_{1\leq k\leq 2^n T_0} \left|X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)\right| \leq C^{1/\alpha}2^{-\gamma n}\right\}$$

, for some γ we will point out later. Now, its complement follows:

$$\begin{aligned}
\mathsf{P}(A_n^c) &= \mathsf{P}\left(\max_{1\leq k\leq 2^n T_0} \left|X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)\right| > C^{1/\alpha}2^{-\gamma n}\right) \\
&= \mathsf{P}\left(\bigcup_{k=1}^{2^n T_0} \left\{\left|X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)\right| > C^{1/\alpha}2^{-\gamma n}\right\}\right) \\
&\leq \sum_{k=1}^{2^n T_0} \mathsf{P}\left(\left|X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)\right| > C^{1/\alpha}2^{-\gamma n}\right) \\
&\leq \sum_{k=1}^{2^n T_0} C \cdot C^{-1} \cdot 2^{\alpha\gamma n} 2^{-n(1+\beta)} = T_0 2^{n(\alpha\gamma-\beta)}
\end{aligned}$$

Here, if we set our γ such that: $\alpha\gamma - \beta < 0$, then by Borel-Cantelli's Lemma:

$$\sum_{n=1}^{\infty} \mathsf{P}(A_n^c) < +\infty \implies \mathsf{P}\left(\liminf_{n\to\infty} A_n\right) = 1$$

Then by construction, we set: $\Omega^* = \liminf_{n\to\infty} A_n \in \mathcal{F}$, clearly it holds:

$$\forall\, \omega^* \in \Omega, \forall\, n \geq n^*(\omega),\ \max_{1\leq k\leq 2^n T_0} \left|X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)\right| \leq C^{1/\alpha}2^{-\gamma n}$$

Since on this set only countable but fortunately dense set has such property, define:

$$D = \bigcup_{n=0}^{\infty} \left\{\frac{k}{2^n} : 1 \leq k \leq 2^n T_0\right\} := \bigcup_{n=0}^{\infty} D_n$$

Clealy, it is countable and dense on $[0, T_0]$, then fix $\omega \in \Omega^*$, we claim that (leave for reader to verify): $\forall\, n \geq n^*(\omega), \forall\, m > n, \forall\, s, t \in D_m, 0 < t - s < 2^{-n}$:

$$|X_t(\omega) - X_s(\omega)| \leq 2C^{1/\alpha} \sum_{j=n+1}^{m} 2^{-\gamma j}$$

Lastly, set $\delta(\omega) = 2^{-n^*(\omega)}, \forall\, s, t \in D, 0 < t - s < \delta(\omega)$, we can choose $m, n \geq n^*(\omega)$ s.t: $s, t \in D_m, 2^{-n-1} \leq t - s < 2^{-n}$:

$$|X_t - X_s| \leq 2C^{1/\alpha} \sum_{j=n+1}^{m} 2^{-\gamma j} \leq 2C^{1/\alpha} \cdot 2^{-r(n+1)} \cdot \frac{1}{1 - 2^{-\gamma}} \leq \frac{2C^{1/\alpha}}{1 - 2^{-\gamma}} |t - s|^{\gamma}$$

After all the preparation, we reconstruct our Y, by defining:

$$Y_t(\omega) = \begin{cases} 0, & \forall\, \omega \notin \Omega^*, t \in [0, T_0] \\ X_t(\omega), & \forall\, \omega \in \Omega^*, t \in D \\ \lim\limits_{\substack{t_n \in D \\ t_n \to t}} X_{t_n}(\omega), & \forall\, \omega \in \Omega^*, t \notin D \end{cases}$$

Finally, reader can easily checked such defined process satisfy (a) - (c). □

The above proving trick is useful when we want to contruct some RCLL process, just to use its left limit, provided exists.

3.2 Filtration

Next, another problem is how to fill in the gap between discrete points:

Definition 3.8. On $(\Omega, \mathcal{F}, \mathsf{P})$, $\{\mathcal{F}_t : t \geq 0\}$ is a family of σ-algebra satisfying:

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}, \forall\, 0 \leq s < t < +\infty$$

Then $\{\mathcal{F}_t : t \geq 0\}$ is called **filtration**.

Remark 3.4. We use following notataion:

$$\mathcal{F}_\infty = \sigma\left(\bigcup_{t \geq 0} \mathcal{F}_t\right)$$

Also for a stochastic process, given process X, set $\mathcal{F}_t^X = \sigma(\{X_s : 0 \leq s \leq t\})$, i.e. the smallest σ-algebra such that X_s is measurable.

As usual to define our adapted process:

Definition 3.9. Let X be a process, $\{\mathcal{F}_t : t \geq 0\}$ be a filtration, if $X_t \in \mathcal{F}_t, \forall\, t \geq 0$, then X is adapted $\{\mathcal{F}_t : t \geq 0\}$.

Remark 3.5. 1. Every X is adapted relative to $\{\mathcal{F}_t^X, t \geq 0\}$.

2. We call $\{\mathcal{F}_t : t \geq 0\}$ is **right-continuous** if $\mathcal{F}_t = \mathcal{F}_{t+}, \forall\, t \geq 0$, and **left-continuous** if $\mathcal{F}_t = \mathcal{F}_{t-}, \forall\, t \geq 0$, here we use following notations:

$$\mathcal{F}_{t-} = \sigma\left(\bigcup_{s < t} \mathcal{F}_s\right), \mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s$$

3. $\forall\, A, \mathsf{P}(A) = 0 \implies A \in \mathcal{F}_0$.

4. Clearly, if Y is a modification of X, X is adapted to $\{\mathcal{F}_t : t \geq 0\}$, then Y is adapted to $\{\mathcal{F}_t : t \geq 0\}$.

Now, we are going to restrict ourself into more specific situation:

Definition 3.10. On $(\Omega, \mathcal{F}, \mathsf{P})$, X is measurable if $\forall\, A \in \mathcal{B}(\mathbb{R}^d)$:

$$\{(\omega, t) \in \Omega \times [0, +\infty) : X_t(\omega) \in A\} \subseteq \mathcal{F} \otimes \mathcal{B}([0, +\infty))$$

And another is "gradually" measurable as time goes:

Definition 3.11. On $(\Omega, \mathcal{F}, \mathsf{P})$, X is progressively measurable with respect to filtration $\{\mathcal{F}_t : t \geq 0\}$, if $\forall\, t \geq 0, \forall\, A \in \mathcal{B}(\mathbb{R}^d)$:

$$\{(\omega, s) \in \Omega \times [0, t] : X_t(\omega) \in A\} \subseteq \mathcal{F}_t \otimes \mathcal{B}([0, t])$$

Following proposition presents relation between measurable and progressively measurable:

Proposition 3.2. *(a) Any progressively measurable process is measurable process.*

(b) If a process is measurable and adapted to a filtration, then the process has a modification that is progressively measurable.

Proof. The proof is omitted since it just verifies the condition. $\square$

A natural question is how it relates to normal adapted process:

Proposition 3.3. *If the process X is adapted to the filtration $\{\mathcal{F}_t : t \geq 0\}$ and every sample path of X is right-continuous, then X is progressively measurable relative to $\{\mathcal{F}_t : t \geq 0\}$.*

Proof. Since the process is right-continuous, we define, $\forall\, t \geq 0, n \in \mathbb{N}, 0 \leq s \leq t$:

$$X_s^{(n)}(\omega) = X_{\frac{(k+1)t}{2^n}}(\omega), \text{ for } s \in \left(\frac{kt}{2^n}, \frac{(k+1)t}{2^n}\right], k \in \mathbb{N}$$

Then for this new process, or more precisely:

$$(\omega, s) \mapsto X_s^{(n)}(\omega) \text{ is } \mathcal{F}_t \otimes \mathcal{B}([0, t])$$

Lastly, using the right continuity: $\lim_{n\to\infty} X_s^{(n)}(\omega) = X_s(\omega)$, which holds naturally. $\square$

3.3 Properties of Stopping Times

Then we present the most important concepts, stopping time, before that:

Definition 3.12. On $(\Omega, \mathcal{F}, \mathsf{P})$, a random time, T, is a r.v. with values in $[0, +\infty]$. Let X be a process, X_T is a r.v. such that $X_T(\omega) = X_{T(\omega)}(\omega)$. Define the σ-algebra generated by X_T as:

$$\sigma(X_T) = \{\{X_T \in A\} \cup \{T = \infty\} : A \in \mathcal{B}(\mathbb{R}^d)\}$$

Remark 3.6. The verification of $\sigma(X_T)$ is indeed a σ-algebra left as homework.
Then a stopping and optional time is based on random time:

Definition 3.13. $(\Omega, \mathcal{F}, \mathsf{P})$ equipped with a filtration $\{\mathcal{F}_t : t \geq 0\}$. A random time T satisfies: $\{T \leq t\} \in \mathcal{F}_t, \forall\, t \geq 0$, is called a **stopping time** of $\{\mathcal{F}_t : t \geq 0\}$. And T is called **optional time** of $\{\mathcal{F}_t : t \geq 0\}$ if $\{T < t\} \in \mathcal{F}_t$.

Remark 3.7. Recall the same definition in discrete-time setting:

$$\{T \leq t\} \xrightarrow{\text{discrete}} \{T \leq n\}$$
$$\{T < t\} \xrightarrow{\text{discrete}} \{T < n\} = \{T \leq n-1\}$$

A natural eager is to see the connection between optional and stopping times:

Proposition 3.4. *(a) A stopping time of $\{\mathcal{F}_t : t \geq 0\}$ is always an optional time of $\{\mathcal{F}_t : t \geq 0\}$.*

(b) If $\{\mathcal{F}_t : t \geq 0\}$ is right-continuous, then any optional time of $\{\mathcal{F}_t : t \geq 0\}$ is a stopping time of $\{\mathcal{F}_t : t \geq 0\}$.

Proof. (a) Notice that for each $\{T \leq t - \frac{1}{n}\} \in \mathcal{F}_{t-\frac{1}{n}} \subseteq \mathcal{F}_t$, then it holds since:

$$\{T < t\} = \bigcup_{n=1}^{\infty} \left\{T < t - \frac{1}{n}\right\} \in \mathcal{F}_t,\ \forall\, t \geq 0$$

(b) To better use the right continuity, for any fixed $k \in \mathbb{N}$:

$$\{T \leq t\} = \bigcap_{n=k}^{\infty} \left\{T < t + \frac{1}{n}\right\} \in \mathcal{F}_{t+\frac{1}{k}}, \forall\, n \geq k$$

And above holds for any $k \in \mathbb{N}$, then $\{T \leq t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$, by right continuity. □

If we use the "right process", then it is obvious right-continuous:

Corollary 3.5. *If T is an optional time of $\{\mathcal{F}_t : t \geq 0\}$, and set $\mathcal{G}_t = \mathcal{F}_{t+}$, then T is a stopping time of $\{\mathcal{G}_t : t \geq 0\}$.*

Proof. It is easy to check $\mathcal{G}_t$ is right-continuous. □

To avoid abstractness, we present following two examples:

Example 3.1. Clearly, if T is deterministic, then T is a stopping time of corresponding filtration.

More concretely, hitting time is widely used for Brownian motion:

Example 3.2. Let X be an adapted process relative to filtration $\{\mathcal{F}_t : t \geq 0\}$. Choose $A \in \mathcal{B}(\mathbb{R}^d)$, then T is known as **hitting time** if we set:

$$T = \inf\{t \geq 0 : X_t(\omega) \in A\}$$

Consequently, we could verify followings properties: (left as homework)

1. If A is open, the sample path of X are right-continuous, then T is an optional time of $\{\mathcal{F}_t : t \geq 0\}$.
2. If A is closed, the sample path of X are continuous, then T is a stopping time of $\{\mathcal{F}_t : t \geq 0\}$.

Then we are about to show some properties of stopping times:

Proposition 3.6. *If T, S are stopping time of $\{\mathcal{F}_t : t \geq 0\}$, then $T \wedge S, T \vee S, T + S$ are all stopping time of $\{\mathcal{F}_t : t \geq 0\}$.*

Proof. Simply, we rewrite into:

1. $T \wedge S : \forall\, t, \{T \wedge S \leq t\} = \{T \leq t\} \cup \{S \leq s\} \in \mathcal{F}_t$.
2. $T \vee S : \forall\, t, \{T \vee S \leq t\} = \{T \leq s\} \cap \{S \leq t\} \in \mathcal{F}_t$.
3. $T + S$: This is non-trivial, but notice following seperarion:

$$\{T+S \leq t\} = \{T = 0, S \leq t\} \cup \{S = 0, T \leq t\} \cup \{0 < T < t, 0 < S < t, T+S \leq t\} \in \mathcal{F}_t$$

For third terms, we could use rational points to argue, and leave for reader. □

Similar but weaker result for optional times:

Proposition 3.7. *If T, S are optional times of $\{\mathcal{F}_t : t \geq 0\}$, then $T \wedge S, T \vee S$ are optional times of $\{\mathcal{F}_t : t \geq 0\}$.*

Proof. We skipped this proof, since it just follows 3.6. □

If we want to recover the third one in 3.6 for optional times, we need following condition:

Proposition 3.8. *If T, S are optional times of $\{\mathcal{F}_t : t \geq 0\}$, then $T + S$ is an optional time of $\{\mathcal{F}_t : t \geq 0\}$. If either $T, S > 0$ or $T > 0$ is a stopping time, then $T + S$ is a stopping time of $\{\mathcal{F}_t : t \geq 0\}$.*

Proof. For the first case $(T, S > 0)$, it is simple, since from 3.6, we know:

$$\{T + S \leq t\} = \{0 < T < t, 0 < S < t, S + T \leq t\} \in \mathcal{F}_t$$

For second case, we omit the proof, and encourage reader to try. □

Then we could also study its limiting behavior:

Proposition 3.9. *If $\{T_n : n \in \mathbb{N}\}$ is a sequence of optional times relative to $\{\mathcal{F}_t : t \geq 0\}$, then:*

$$\sup\{T_n : n \in \mathbb{N}\}, \quad \inf\{T_n : n \in \mathbb{N}\}, \quad \limsup_{n\in\mathbb{N}} T_n, \quad \liminf_{n\in\mathbb{N}} T_n,$$

are all optional times of $\{\mathcal{F}_t : t \geq 0\}$.

Proof. Simply, we could rewrite them into:

1. $\{\sup_n T_n < t\} = \cap_{n\in\mathbb{N}}\{T_n < t\}$.
2. $\{\inf_n T_n < t\} = \cup_{n\in\mathbb{N}}\{T_n < t\}$.
3. $\limsup_{n\to\infty} = \lim_{M\to\infty}\sup_{n\geq M} = \cap_{M\in\mathbb{N}}\sup_{n\geq M}$.
4. $\liminf_{n\to\infty} = \lim_{M\to\infty}\inf_{n\geq M} = \cup_{M\in\mathbb{N}}\sup_{n\geq M}$.

Then using the properties of σ-algebra, we completed the proof. □

Remark 3.8. If $\{T_n : n \in \mathbb{N}\}$ is stopping times, then $\sup T_n$ is also a stopping time.

Since we finally want to change time with stopping time, then need filtration:

Definition 3.14. Let T be stopping time of $\{\mathcal{F}_t : t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$. The σ-algebra:

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t, \forall\, t \geq 0\}$$

is called **pre-T σ-algebra.**

Interestingly, in this container, all random time moves to stopping time:

Proposition 3.10. *Let T be a stopping time of $\{\mathcal{F}_t : t \geq 0\}$. S be a random time such that $S \geq T$. If $S \in \mathcal{F}_T$, then S is a stopping time.*

Proof. Simply from definintion:

$$\forall\, t \geq 0, \{S \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t$$

Then we could finish the proof. □

For preparation of OST, we need to study the relation for multiple stopping time:

Proposition 3.11. *Let S, T be stopping times of $\{\mathcal{F}_t : t \geq 0\}$, then:*

(a) $\forall\, A \in \mathcal{F}_S, A \cap \{S \leq T\} \in \mathcal{F}_T$.

(b) $\{T < S\}, \{T \leq S\}, \{T = S\} \in \mathcal{F}_T \cap \mathcal{F}_S$.

(c) $\mathcal{F}_{T\wedge S} = \mathcal{F}_T \cap \mathcal{F}_S$.

Proof. Just iteratively use the defintion, and we omit the proof. □

Then we give a tease of changing time with stopping times:

Proposition 3.12. *Let $X = \{X_t : t \geq 0\}$ be progressively measurable process relative to $\{\mathcal{F}_t : t \geq 0\}$. Let T be finite stopping time of $\{\mathcal{F}_t : t \geq 0\}$. Then X_T is $\mathcal{F}_T$-measurable.*

Proof. Leave as homework $\square$

Remark 3.9.

1. $\{X_{T\wedge t, t\geq 0}\}$, called "stopping process", is progressively measurable relative to $\{\mathcal{F}_t : t \geq 0\}$.
2. If T be an optional time of $\{\mathcal{F}_t : t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$, $\mathcal{F}_{T+} = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_{t+}, \forall\, t \geq 0\}$, then $\mathcal{F}_{T+}$ is a $\sigma-$algebra.

For further process, we need to introduce more conditions:

Definition 3.15. A filtration $\{\mathcal{F}_t : t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$ is said to satisfy the **usual conditions** if it is right-continuous and $\mathcal{F}_0$ contains all $\mathsf{P}-$negligible-events in $\mathcal{F}$, i.e. $A \in \mathcal{F}, s.t.\ \mathsf{P}(A) = 0$.

Lastly, under this conditions, we show its useness of stopping times:

Proposition 3.13. *Let $\{\mathcal{F}_t : t \geq 0\}$ filtrations satisfies the usual conditions. Let X be an adapted process relative to $\{\mathcal{F}_t : t \geq 0\}$ s.t: the sample path of X is RCLL. Then there exists a sequence $\{T_n : n \in \mathbb{N}\}$ of stopping time of $\{\mathcal{F}_t : t \geq 0\}$ s.t:*

$$\{(\omega, t) \in \Omega \times \mathbb{R}_+ : X_t(\omega) \neq X_{t-}(\omega)\} \subseteq \bigcup_{n=1}^{\infty} \{(\omega, t) \in \Omega \times (0, +\infty) : T_n(\omega) = t\}$$

, i.e. $\{T_n : n \in \mathbb{N}\}$ exhausts the jumps of X.

Proof. The proof is omitted since its complexity. $\square$

3.4 Fundamental Theorem of Submartingale

Next, we move to most important tools of this course:

Definition 3.16. Let $X = \{X_t : t \geq 0\}$ be an adapted process of $\{\mathcal{F}_t : t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$ such that: $X_t \in L^1(\Omega, \mathcal{F}, \mathsf{P}), \forall\, t \geq 0$:

(a) If $0 \leq s < t < +\infty, \mathsf{E}[X_t \mid \mathcal{F}_s] \geq X_s$ *a.s.* then X is a **submartingale**. If in addition, $X_\infty \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ and $X_\infty \in \mathcal{F}_\infty$ satisfying: $\mathsf{E}[X_\infty \mid \mathcal{F}_s] \geq X_s$ *a.s.* $\forall\, s \geq 0$. Then $\{X_t : 0 \leq t \leq +\infty\}$ is a **submartingale with last element:** X_∞.

(b) If $0 \leq s < t < +\infty, \mathsf{E}[X_t \mid \mathcal{F}_s] \leq X_s$ *a.s.* then X is a **supermartingale**. If in addition, $X_\infty \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ and $X_\infty \in \mathcal{F}_\infty$ satisfying: $\mathsf{E}[X_\infty \mid \mathcal{F}_s] \leq X_s$ *a.s.* $\forall\, s \geq 0$. Then $\{X_t : 0 \leq t \leq +\infty\}$ is a **supermartingale with last element:** X_∞.

(c) If $0 \leq s < t < +\infty, \mathsf{E}[X_t \mid \mathcal{F}_s] = X_s$ *a.s.* then X is a **martingale**. If in addition, $X_\infty \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ and $X_\infty \in \mathcal{F}_\infty$ satisfying: $\mathsf{E}[X_\infty \mid \mathcal{F}_s] =$

X_s *a.s.* $\forall\ s \geq 0$. Then $\{X_t : 0 \leq t \leq +\infty\}$ is a **martingale with last element:** X_∞.

Remark 3.10. If $\{X_t, t \geq 0\}$ is a submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$ and $\{t_n : n \in \mathbb{N}\}$ be decreasing and non-negative numbers. Then $\{X_{t_n}, n \in \mathbb{N}\}$ is a backward submartingale. Similarly, we could put a convex increasing function to produce submartingale:

Proposition 3.14. *Let $X = \{X_t : t \geq 0\}$ be a submartingale of $\{\mathcal{F}_t : t \geq 0\}$. Let φ be a convex increasing function, such that $\varphi(X_t) \in L^1(\Omega, \mathcal{F}, \mathsf{P}), \forall\ t \geq 0$. Then $\varphi(X) = \{\varphi(X_t) : t \geq 0\}$ is a $\{\mathcal{F}_t : t \geq 0\}$-submartingale.*

Proof. Simply proved using Jensen's Inequality, then we omitted here. □

Remark 3.11. If X is martingale, φ be a convex function, then $\varphi(X)$ submartingale.

Then, we are going to introduce useful inequality for proving convergence:

Definition 3.17. Let $X = \{X_t : t \geq 0\}$ be a real-valued process, for $a < b$ and $I \subseteq [0, +\infty)$, we define:

$$\begin{aligned} \tau_1(\omega) &= \inf\{t \in I : X_t(\omega) \leq a\} \\ \sigma_1(\omega) &= \inf\{t \in I : t \geq \tau_1(\omega), X_t(\omega) \geq b\} \\ \tau_2(\omega) &= \inf\{t \in I : t \geq \sigma_1(\omega), X_t(\omega) \leq a\} \\ &\cdots\cdots\cdots \end{aligned}$$

Since it may be empty set, we set $\inf\{\emptyset\} = +\infty$. Next, we define the number of up-crossings of interval $[a, b]$ by sample path $\{X_t, t \geq 0\}$:

$$U_I\left([a, b]; X(\omega)\right) = \max\{j \in \mathbb{N}, \sigma_j(\omega) < +\infty\}$$

Similarly, the number of down-crossings of interval $[a, b]$ by sample path $\{X_t, t \geq 0\}$:

$$D_I\left([a, b]; X(\omega)\right) = \max\{j \in \mathbb{N}, \tau_j(\omega) < +\infty\}$$

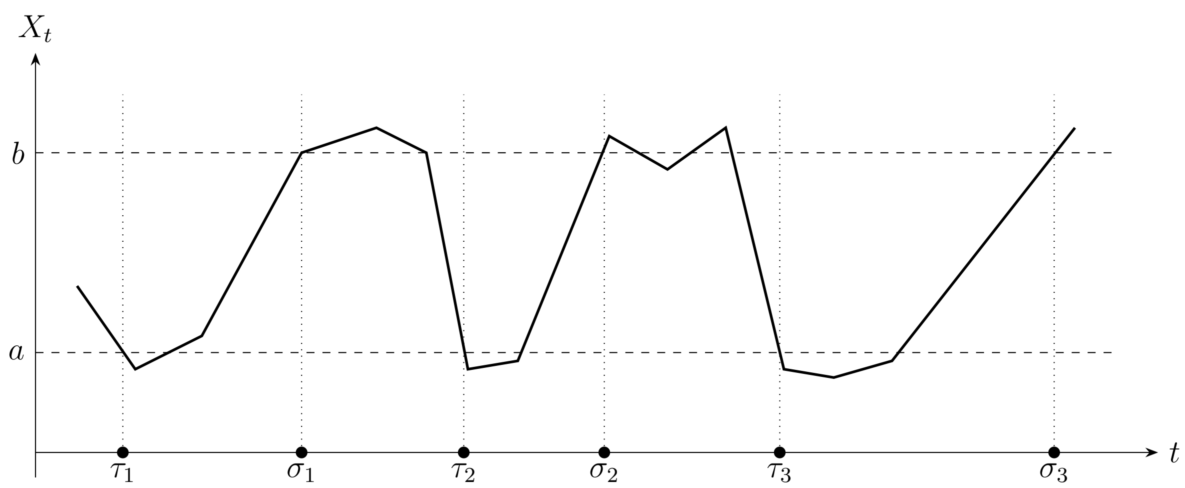

Figure 2: Illustration of upcrossings of the interval $[a, b]$ by a sample path $\{X_t : t \geq 0\}$.

Based on up/down-crossing, we could use to show convergence if we prepared toolkit:

Theorem 3.15. *Let $X = \{X_t : t \geq 0\}$ be a supermartingale of $\{\mathcal{F}_t : t \geq 0\}$ whose sample path are right-continuous, let $I \subseteq [0, +\infty)$ be a compact interval and $a < b, \lambda > 0$ be all real numbers:*

(a) *(First submartingale inequality):* $\lambda \mathsf{P}\left(\sup_{t \in I} X_t \geq \lambda\right) \leq \mathsf{E}X^+_{\sup I}$.

(b) *(Second submartingale inequality):* $\lambda \mathsf{P}\left(\inf_{t \in I} X_t \leq -\lambda\right) \leq \mathsf{E}X^+_{\sup I} - \mathsf{E}X_{\inf I}$.

(c) *(Up-crossing inequality):* $(b-a)\mathsf{E}U_I\left([a,b]; X(\omega)\right) \leq \mathsf{E}X^+_{\sup I} + |a|$.

(d) *(Doob's maximal inequality): If $X_t \geq 0, \forall\, t \geq 0$, then $\forall\, p > 1$,:*

$$\mathsf{E}\left(\sup_{t \in I} X_t\right)^p \leq \left(\frac{p}{p-1}\right)^p \mathsf{E}X^p_{\sup I}$$

(e) *Regularity of sample path of X:*

  (i) *Almost every sample path of $\{X_t : t \geq 0\}$ is bounded on compact intervals and admits left-hand limits everyone on $(0, +\infty)$.*

  (ii) *If $\{\mathcal{F}_t : t \geq 0\}$ satisfies the usual conditions, then jumps of X are exhausted by a sequence of stopping times.*

We postpone proving this theorem, and first introduce following two lemmas:

Lemma 3.16. *If $\{X_n : 1 \leq n \leq N\}$ is submartingale of $\{\mathcal{F}_n : 1 \leq n \leq N\}$. Then $\forall\, \lambda > 0$:*

1. $\lambda \mathsf{P}\left(\max_{1 \leq n \leq N} X_n \geq \lambda\right) \leq \mathsf{E}X_N \mathbf{1}_{\{\max_{1 \leq n \leq N} \geq \lambda\}} \leq \mathsf{E}X^+_N$.
2. $\lambda \mathsf{P}\left(\min_{1 \leq n \leq N} X_n \leq -\lambda\right) \leq \mathsf{E}(X_N - X_1) - \mathsf{E}X_N \mathbf{1}_{\{\min_{1 \leq n \leq N} \leq \lambda\}} \leq \mathsf{E}(X^+_N - X_1)$.

Proof. For this discrete version, we could treat with more confident:

1. Set $\alpha = \min\{nX_n \geq \lambda\}$ with convention: $\min\{\emptyset\} = N$, is an optional r.v. since:

$$\{\alpha < k\} = \bigcup_{j=1}^{k} \{X_j \geq \lambda\} \in \mathcal{F}_k$$

Then by discrete version of OST, $\{X_\alpha, X_N\}$ is a $\{\mathcal{F}_\alpha, \mathcal{F}_N\}$-submartingale. Notice:

$$\forall\, k < N, \{\alpha < k\} \bigcap \left\{\max_{1 \leq n \leq N} X_n \geq \lambda\right\} = \{\alpha < k\} \in \mathcal{F}_k \implies \left\{\max_{1 \leq n \leq N} X_n \geq \lambda\right\} \in \mathcal{F}_\alpha$$

Therefore, we could complete the proof since:

$$\begin{aligned}\lambda \mathsf{P}\left(\max_{1 \leq n \leq N} X_n \geq \lambda\right) &= \mathsf{E}\lambda \mathbf{1}_{\{\max_{1 \leq n \leq N} X_n \geq \lambda\}} \leq \mathsf{E}X_\alpha \mathbf{1}_{\{\max_{1 \leq n \leq N} X_n \geq \lambda\}} \\ &\leq \mathsf{E}\left[\mathsf{E}\left(X_N \mid \mathcal{F}_\alpha\right) \mathbf{1}_{\{\max_{1 \leq n \leq N} X_n \geq \lambda\}}\right] = \mathsf{E}X_N \mathbf{1}_{\{\max_{1 \leq n \leq N} X_n \geq \lambda\}}\end{aligned}$$

For second inequlity, notice $X_N \leq X^+_N, \mathbf{1}_A \leq 1$.

2. Simlarly, we define $\beta = \min\{n : X_n \le -\lambda\}$ with convention: $\min\{\emptyset\} = N$, then:

$$\begin{aligned}
\mathsf{E}X_1 &\le \mathsf{E}\left[\mathsf{E}\left(X_\beta \mid \mathcal{F}_1\right)\right] = \mathsf{E}X_\beta \\
&= \mathsf{E}X_\beta \mathbf{1}_{\{\beta \le N-1\}} + \mathsf{E}X_\beta \mathbf{1}_{\{\beta = N\} \cap \{\min\limits_{1\le n\le N} X_n \le -\lambda\}} + \mathsf{E}X_\beta \mathbf{1}_{\{\beta = N\} \cap \{\min\limits_{1\le n\le N} X_n > -\lambda\}} \\
&= \mathsf{E}X_\beta \mathbf{1}_{\{\beta \le N-1\}} + \mathsf{E}X_N \mathbf{1}_{\{\beta = N\} \cap \{\min\limits_{1\le n\le N} X_n \le -\lambda\}} + \mathsf{E}X_N \mathbf{1}_{\{\min\limits_{1\le n\le N} X_n > -\lambda\}} \\
&\le -\lambda \mathsf{P}(\beta \le N-1) - \lambda \mathsf{P}\left(\{\beta = n\} \cap \{\min_{1\le n\le N} X_n \le -\lambda\}\right) + \mathsf{E}X_N \mathbf{1}_{\{\min\limits_{1\le n\le N} X_n > -\lambda\}} \\
&= -\mathsf{P}\left(\min_{1\le n\le N} X_n \le -\lambda\right) + \mathsf{E}X_N \mathbf{1}_{\{\min_{1\le n\le N} X_n > -\lambda\}}
\end{aligned}$$

For this long derivation, we use three properties:

(a) On $\{\beta \le N-1\}$, $X_\beta \le -\lambda$.

(b) On $\{\beta = N\} \cap \{\min_{1\le n\le N} X_n \le -\lambda\}$, $X_N \le -\lambda$.

(c) $\{\beta \le N-1\} \cup (\{\beta = N\} \cap \{\min_{1\le n\le N} X_n \le -\lambda\}) = \{\min_{1\le n\le N} X_n \le -\lambda\}$.

Finally, we rearrange the terms to see:

$$\begin{aligned}
\mathsf{P}\left(\min_{1\le n\le N} X_n \le -\lambda\right) &\le \mathsf{E}X_N \mathbf{1}_{\{\min_{1\le n\le N} X_n > -\lambda\}} - \mathsf{E}X_1 \\
&= \mathsf{E}X_N - \mathsf{E}X_N \mathbf{1}_{\{\min_{1\le n\le N} X_n \le -\lambda\}} - \mathsf{E}X_1
\end{aligned}$$

□

And then another lemma for proving continuous version of up-crossing inequality:

Lemma 3.17. *If $X = \{X_n : 1 \le n \le N\}$ is a submartingale relative to $\{\mathcal{F}_n : 1 \le n \le N\}$ and $[a, b]$ is a compact interval, then:*

$$(b-a)\mathsf{E}U_N\left([a,b]; X(\omega)\right) \le \mathsf{E}(X_N - a)^+ \le \mathsf{E}X_N^+ + |a|$$

Proof. Firstly, we define: $\tau_1 = \min\{1 \le n \le N : X_n \le a\}$, $\forall\, j \in \mathbb{N}$:

$$\sigma_j = \min\{1 \le n \le N : n \ge \tau_j, X_n \ge b\}, \tau_j = \min\{1 \le n \le N : n \ge \sigma_{j-1}, X \le a\}$$

Then it is easy to see: $1 \le \tau_1 \le \sigma_1 \le \cdots \le N$ are all optional r.v.s. Therefore, by discrete version of OST, $\{X_1, X_{\tau_1}, \cdots, X_N\}$ is submartingale relative to $\{\mathcal{F}_1, \mathcal{F}_{\tau_1}, \cdots\}$, notice:

$$\forall\, 1 \le j \le M = U_N([a,b]; X(\omega)),\ X_{\sigma_j}(\omega) - X_{\tau_j}(\omega) \ge b - a$$

Then we could finish our proof by:

$$\begin{aligned}
(b-a)\mathsf{E}U_N([a,b]; X) &\le \mathsf{E}\sum_{j=1}^{M}(X_{\sigma_j} - X_{\tau_j}) = \mathsf{E}X_{\sigma_M} - \mathsf{E}X_{\tau_1} + \mathsf{E}\sum_{j=1}^{M-1}(X_{\sigma_j} - X_{\tau_{j+1}}) \\
&\le \mathsf{E}X_{\sigma_n} - \mathsf{E}X_{\tau_1} \le \mathsf{E}(X_N - a)^+ \le \mathsf{E}X_N^+ + |a|
\end{aligned}$$

Lastly, we use the property of X is submartingale so does $(X-a)^+$. □

Then we start to prove 3.15 using above lemma:

Proof. Finally, we are ready to prove the big theorem 3.15. But only for (a) - (c):

(a) Choose $\{\mathsf{F}_n\}$ an increasing sequence of finite set, s.t: $\inf I, \sup I \in \mathsf{F}_N, \forall\ N$, and:

$$\mathsf{F} = \bigcup_{N=1}^{\infty} = \{\inf I, \sup I\} \bigcup (I \cap \mathbb{Q})$$

Then considering $\{X_t, t \in \mathsf{F}_N\}$, by lemma 3.16:

$$\lambda \mathsf{P}\left(\max_{t \in \mathsf{F}_N} X_t > \lambda\right) \leq \mathsf{E}X^+_{\sup I}$$

Now, using the inclusion relation, i.e. $\mathsf{F}_N \subseteq \mathsf{F}_{N+1}$, of our chosing "field":

$$\left\{\max_{t \in \mathsf{F}_N} X_t > \lambda\right\} \subseteq \left\{\max_{t \in \mathsf{F}_{N+1}} X_t > \lambda\right\} \implies \bigcup_{N=1}^{\infty} \left\{\max_{t \in \mathsf{F}_N} X_t > \lambda\right\} = \left\{\max_{t \in \mathsf{F}} X_t > \lambda\right\}$$

Therefore, we extend to F, since it's dense and the process is right-continuous:

$$\lambda \mathsf{P}\left(\max_{t \in \mathsf{F}} X_t > \lambda\right) \leq \mathsf{E}X^+_{\sup I} \implies \lambda \mathsf{P}\left(\sup_{t \in I} X_t > \lambda\right) \leq \mathsf{E}X^+_{\sup I}$$

(b), (c) Same strategy as (a).

(d), (e) Leave for reader to verify.

Then we completed the proof. □

Then firstly using above tools, we could "fix" our process:

Proposition 3.18. *Let $X = \{X_t : t \geq 0\}$ be a submartingale of $\{\mathcal{F}_t : t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$, then:*

(a) $\exists\ \Omega^* \in \mathcal{F}$ with $\mathsf{P}(\Omega^*) = 1$ satifying, $\forall\ \omega \in \Omega^*$:

$$X_{t+}(\omega) = \lim_{\substack{s \to t+ \\ s \in \mathbb{Q}}} X_s(\omega) \ \exists\ \forall\ t \geq 0,\ X_{t-}(\omega) = \lim_{\substack{s \to t- \\ s \in \mathbb{Q}}} X_s(\omega) \ \exists\ \forall\ t > 0$$

(b) The followings holds:

$$\mathsf{E}(X_{t+} \mid \mathcal{F}_t) \geq X_t,\ a.s., \forall t > 0$$
$$\mathsf{E}(X_t \mid \mathcal{F}_{t-}) \geq X_{t-},\ a.s., \forall t > 0$$

(c) $\{X_t, t \geq 0\}$ is a submartingale with $\mathsf{P}-$almost all sample path being RCLL.

Proof. We only present a rough idea of proofs here:

(a) Using up-crossings inequality on $[0, n] \cap \mathbb{Q}$.

(b) Using $t_n \in \mathbb{Q}$ to approximate, then exchange the integral and limits.

(c) By (a) + (b).

□

Also, even though the process is not good enough, we could fix with modification:

Theorem 3.19. *Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration satisfying usual conditions, let $X = \{X_t : t \geq 0\}$ be a submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$. Then X has a modification that is right-continuous iff $t \mapsto \mathsf{E}X_t$ is right-continuous on $(0, +\infty)$.*

Proof. The proof is intentionally omitted here. □

The power of up-crossing inequality is to show the convergence:

Theorem 3.20. ***(Submartingale Convergence Theorem)*** *Let $X = \{X_t : t \geq 0\}$ be a right-continuous submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$ such that: $\sup_{t\geq 0} \mathsf{E}X_t^+ < +\infty$, then:*

(a) $\lim_{t\to\infty} X_t$ *exists a.s.*

(b) $\lim_{t\to\infty} X_t \in L^1(\Omega, \mathcal{F}, \mathsf{P})$.

Proof. (a) But we have a powerful tool of 3.15's (c), $\forall\, n \in \mathbb{N}$, $\forall\, a < b$:

$$\mathsf{E}U_{[0,n]}([a,b]; X) \leq \frac{\mathsf{E}X_n^+ + |a|}{b-a} \leq \frac{\sup_{t\geq 0} \mathsf{E}X_t^+ + |a|}{b-a}$$

Here, set $n \to +\infty$, and notice $U_{[0,n]} \uparrow$, then by MCT, following holds:

$$\mathsf{E}U_{[0,+\infty)}([a,b]; X) \leq \frac{\sup_{t\geq 0} \mathsf{E}X_t^+ + |a|}{b-a}, \ \forall\, a < b$$

Then observing below relations:

$$\begin{aligned}\left\{\omega : \lim_{t\to\infty} X_t(\omega) \nexists\right\} &= \left\{\omega : \limsup_{t\to\infty} X_t(\omega) > \liminf_{t\to\infty} X_t(\omega)\right\} \\ &\subseteq \bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \left\{\omega : U_{[0,+\infty)}([a,b]; X(\omega)) = \infty\right\} = \emptyset\end{aligned}$$

Therefore, we complete the proof of existence since:

$$\mathsf{P}\left\{\omega : \lim_{t\to\infty} X_t(\omega) \exists\right\} = \mathsf{P}\left\{\omega : \lim_{t\to\infty} X_t(\omega) \nexists\right\}^c = 1$$

(b) For integrability, notice the relation: $|X_t| = 2X_t^+ - X_t, \forall\, t$,

$$\mathsf{E}|X_t| = \mathsf{E}(2X_t^+ - X_t) \leq 2\sup_{t\geq 0} \mathsf{E}X_t^+ - \mathsf{E}X_0 < +\infty$$

Then we simply set both sides limit:

$$\mathsf{E}\left|\lim_{t\to\infty} X_t\right| = \mathsf{E}\liminf_{t\to\infty} |X_t| \leq \liminf_{t\to\infty} \mathsf{E}|X_t| < +\infty$$

Finally, we completed the proof. □

Easily, we could see following from above theorem:

Corollary 3.21. *Let $X = \{X_t : t \geq 0\}$ be a right-continuous, non-negative supermartingale relative to $\{\mathcal{F}_t : t \geq 0\}$, then $X_\infty = \lim_{t\to\infty} X_t$ $\exists$, and $\{X_t : 0 \leq t \leq +\infty\}$ is a supermartingale.*

Proof. Leave as homework. □

And naturally, we want to restrict the last element, to be finite:

Definition 3.18. A right-continuous, non-negative supermartingale $X = \{X_t : t \geq 0\}$ relative to $\{\mathcal{F}_t : t \geq 0\}$ with $\lim_{t\to\infty} \mathsf{E}X_t = 0$, then X is called a **potential**.

Remark 3.12. A potential is a supermartingale with last element 0 a.s.

Under this, we could present equivalent conditions for convergence:

Proposition 3.22. *Let $X = \{X_t : t \geq 0\}$ be a non-negative, right-continuous submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$, then the followings are equivalent:*

(a) The family $\{X_t, t \geq 0\}$ of r.v.s are uniformly integrable.

(b) X_t converges in L^1 as $t \longrightarrow \infty$.

(c) X_t converges to $X_\infty \in L^1(\Omega, \mathcal{F}, \mathsf{P})$, and $\{X_t : 0 \leq t \leq +\infty\}$ is a submartingale of $\{\mathcal{F}_t : t \geq 0\}$.

Proof. Leave as homework. □

Remark 3.13. If $X_n \geq Y$, *a.s.* $\forall\, t \geq 0$, then $X_t - Y$ is non-negative.

Lastly, we come to the final boss, to replace time with stopping times:

Theorem 3.23. ***(Optional Sampling Theorem)***
Let $X = \{X_t : 0 \leq t \leq +\infty\}$ be a right-continuous submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$ with last element X_∞. Let $S \leq T$ be two optinal times of $\{\mathcal{F}_t : t \geq 0\}$. Then:

$$\mathsf{E}\left[X_T \mid \mathcal{F}_{S+}\right] \geq X_S, \text{ a.s.}$$

Here if S is stopping time, then $\mathsf{E}\left[X_T \mid \mathcal{F}_S\right] \geq X_S$, a.s.

Proof. In order to use discrete OST, we discretize the time into, $\forall\, n \in \mathbb{N}$:

$$S_n = \begin{cases} +\infty & , S = +\infty \\ \frac{k}{2^n} & , \frac{k-1}{2^n} \leq S < \frac{k}{2^n} \end{cases}, T_n = \begin{cases} +\infty & , T = +\infty \\ \frac{k}{2^n} & , \frac{k-1}{2^n} \leq T < \frac{k}{2^n} \end{cases}$$

Dominance relation still holds, i.e. $S_n \leq T_n$, are stopping times of $\{\mathcal{F}_t : t \geq 0\}$, since:

$$\{S_n \leq t\} = \begin{cases} \{S < t\} & , t = \frac{k}{2^n}, \exists\, k \\ \left\{S < \frac{k}{2^n}\right\} & , \frac{k}{2^n} < t < \frac{k+1}{2^n}, \exists\, k \end{cases} \in \mathcal{F}_t$$

Besides, by right continuity, we can easily recover the process, because:

$$\lim_{t\to\infty} X_{S_n}(\omega) = X_S(\omega) \text{ and } \lim_{t\to\infty} X_{T_n}(\omega) = X_T(\omega)$$

After discretization, $\{X_{S_n}, X_{T_n}\}$ is submartingale of $\{\mathcal{F}_{S_n}, \mathcal{F}_{T_n}\}$, then

$$\forall\ A \in \mathcal{F}_{S+} = \bigcap_{n=1}^{\infty} \mathcal{F}_{S_n},\ \mathsf{E} X_{T_n} \mathbf{1}_A = \mathsf{E}\left[\mathsf{E}\left(X_{T_n} \mid \mathcal{F}_{S_n}\right) \mathbf{1}_A\right] \geq \mathsf{E} X_{S_n} \mathbf{1}_A$$

For taking limit, we first need to check its integrability, notice:

$$\forall\ n,\ S_n \geq S_{n+1}, T_n \geq T_{n+1}$$

It tells that $\{X_{S_n} : n \in \mathbb{N}\}, \{X_{T_n} : n \in \mathbb{N}\}$ are backward submartingale of $\{\mathcal{F}_{S_n}\}, \{\mathcal{F}_{T_n}\}$:

$$\mathsf{E} X_{S_n} \geq \mathsf{E} X_0, \mathsf{E} X_{T_n} \geq X_0 \implies \{X_{S_n}\}, \{X_{T_n}\} \text{ are u.i.}$$

Then since $\{X_{S_n}\}, \{X_{T_n}\}$ are u.i., it follows:

$$X_{S_n} \xrightarrow{L_1} X_S, X_{T_n} \xrightarrow{L_1} X_T \implies X_S, X_T \in L^1(\Omega, \mathcal{F}, \mathsf{P})$$

Moreover, $\forall\ A \in \mathcal{F}_{S+}, \{X_{S_n} \mathbf{1}_A\}, \{X_{T_n} \mathbf{1}_A\}$ are both u.i, then:

$$\mathsf{E}\left[\mathsf{E}\left(X_T \mid \mathcal{F}_{S+}\right) \mathbf{1}_A\right] = \mathsf{E} X_T \mathbf{1}_A = \lim_{n \to \infty} \mathsf{E} X_{T_n} \mathbf{1}_A \geq \lim_{n \to \infty} \mathsf{E} X_{S_n} \mathbf{1}_A = \mathsf{E} X_S \mathbf{1}_A$$

Just choose $A = \{\mathsf{E}\left(X_T \mid \mathcal{F}_{S+}\right) < X_S\}$, and state $\mathsf{P}(A) = 0$ to complete the proof. □

Presenting lots of properties of submartingale, then one want to how it looks like:

Definition 3.19. Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration on $(\Omega, \mathcal{F}, \mathsf{P})$. An adapted process, $\mathcal{A}$, is called increasing if:

(a) $\mathcal{A}_0 = 0,\ a.s.$

(b) For $a.s - \omega : t \mapsto \mathcal{A}_t$ is increasing and right-continuous.

(c) $\forall\ t \geq 0, \mathsf{E} \mathcal{A}_t < +\infty$.

An increasing process $\mathcal{A}$ is integrable, i.e. $\mathsf{E} \lim_{t \to \infty} \mathcal{A}_t < +\infty$.

Therefore, a submartingale can be decomposed into:

Theorem 3.24. ***(Doob-Meyer Decomposition)***
Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration satisfying the usual conditions. Let $X = \{X_t : t \geq 0\}$ be a right-continuous submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$, $\forall\ a \in \mathbb{R}^+$,

$$\varphi_a = \{T : T \textit{ is a stopping time s.t. } \mathsf{P}(T \leq a) = 1\}$$

Assume that $\{X_T : T \in \varphi_a\}$ is u.i. $\forall\ a \in \mathbb{R}^+$. Then $\forall\ t \geq 0$:

$$X_t = M_t + \mathcal{A}_t$$

, where $\{M_t : t \geq 0\}$ is a right-continuous martingale relative to $\{\mathcal{F}_t : t \geq 0\}$, and $\{\mathcal{A}_t\}$ is an increasing process.

Proof. The proof is omitted, we just present the result here. □

Remark 3.14.
1. Set $\varphi = \{T : T$ is stopping time of $\{\mathcal{F}_t : t \geq 0\}$ s.t. $\mathsf{P}(T < +\infty) = 1\}$ and $\{X_T : T \in \varphi\}$ is u.i., then $\{M_t, t \geq 0\}$ is u.i. and $\{\mathcal{A}_t, t \geq 0\}$ is integrable.
2. The uniqueness of theorem holds if $\mathcal{A}$ is nature. It needs Lebesgue-stieltjes integral.

3.5 Homeworks

The followings are exercises for this section:

Problem 3.1. On $(\Omega, \mathcal{F}, \mathsf{P})$, let $X = \{X_t : t \geq 0\}$ be a stochastic process that is measurable and T is a random time. Prove that

(1) X_T is a random variable if T is finite.

(2) All sets of the form $\{X_T \in A\}$ and $\{X_T \in A\} \cup \{T = \infty\}$ with $A \in \mathcal{B}(\mathbb{R}^d)$ forms a σ-algebra.

Problem 3.2. Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration and X be an adapted process relative to $\{\mathcal{F}_t : t \geq 0\}$. Set $T = \inf\{t \geq 0 : X_t \in A\}$. Prove that

(1) If A is open and the sample paths of X are right-continuous, then T is an optional time of $\{\mathcal{F}_t : t \geq 0\}$.

(2) If A is closed and the sample paths of X are continuous, then T is a stopping time of $\{\mathcal{F}_t : t \geq 0\}$.

Problem 3.3. Let $\{X_t : t \geq 0\}$ be a progressively measurable process relative to $\{\mathcal{F}_t : t \geq 0\}$ and let T be a finite stopping time of $\{\mathcal{F}_t : t \geq 0\}$. Prove that

(1) X_T is $\mathcal{F}_T$-measurable.

(2) Process $\{X_{T\wedge t} : t \geq 0\}$ is progressively measurable relative to $\{\mathcal{F}_t : t \geq 0\}$.

Problem 3.4. Let T be an optional time of $\{\mathcal{F}_t : t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathsf{P})$.

(1) Prove that $\mathcal{F}_{T+}$ is a σ-algebra and $\mathcal{F}_{T+} = \{A \in \mathcal{F} : A \cap \{T < t\} \in \mathcal{F}_t, \forall t \geq 0\}$.

(2) Prove that if T is a stopping time, then $\mathcal{F}_T \subseteq \mathcal{F}_{T+}$.

Problem 3.5. Let $\{X_t : t \geq 0\}$ be a right-continuous, nonnegative supermartingale relative to $\{\mathcal{F}_t : t \geq 0\}$. Prove that $X_\infty = \lim_{t\to\infty} X_t$ exists almost surely and $\{X_t : 0 \leq t \leq \infty\}$ is a supermartingale relative to $\{\mathcal{F}_t : 0 \leq t \leq \infty\}$.

Problem 3.6. Let $\{X_t, t \geq 0\}$ be a right-continuous, nonnegative submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$.

(1) The family $\{X_t : t \geq 0\}$ of r.v.s are uniformly integrable.

(2) X_t converges in L^1 as $t \to \infty$.

(3) X_t converges almost surely and the limit X_∞ is integrable. Moreover, $\{X_t : 0 \leq t \leq \infty\}$ is a submartingale relative to $\{\mathcal{F}_t : 0 \leq t \leq \infty\}$.

Prove that (1), (2) and (3) are equivalent.

Problem 3.7. Let $\{X_t : t \geq 0\}$ be a right-continuous supermartingale relative to $\{\mathcal{F}_t : t \geq 0\}$ and $S \leq T$ are stopping times of $\{\mathcal{F}_t : t \geq 0\}$. Prove that

(1) $\{X_{T\wedge t} : t \geq 0\}$ is supermartingale of $\{\mathcal{F}_t : t \geq 0\}$.

(2) $\mathsf{E}[X_{T\wedge t} \mid \mathcal{F}_S] \leq X_{S\wedge t}$.

Problem 3.8. Let $\{X_t : t \geq 0\}$ be a right-continuous process such that $X_t \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ for $\forall t \geq 0$. Prove that X is a submartingale relative to $\{\mathcal{F}_t : t \geq 0\}$ if and only if $\mathsf{E}X_T \geq \mathsf{E}X_S$ for all bounded stopping times $S \leq T$ of $\{\mathcal{F}_t : t \geq 0\}$.

4 Markov Process

At the heart of many stochastic models lies a powerful simplifying principle: the future, given the present, is independent of the past. This Markov property transforms complex dependence into manageable, local evolution. In this chapter, we give this idea precise measure-theoretic form, exploring its consequences through transition probabilities, evolution operators, and the foundational Chapman-Kolmogorov equations. We distinguish between the Markov property and its stronger counterpart adapted to stopping times, and see how processes like Brownian motion naturally fit within this framework. Markov theory provides a unifying lens, revealing a common structure behind many seemingly disparate random phenomena.

4.1 Basic Definitions

Firstly, we present what is a Markov process:

Definition 4.1. Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration on $(\Omega, \mathcal{F}, \mathsf{P})$, and $X = \{X_t : t \geq 0\}$ be an adapted process relative to $\{\mathcal{F}_t : t \geq 0\}$. Then X is a **Markov Process** if, $\forall\, 0 \leq s \leq t, \forall$ bounded Borel-measurable function f, it follows:

$$\mathsf{E}\left[f(X_t) \mid \mathcal{F}_s\right] = \mathsf{E}\left[f(X_t) \mid X_s\right] \; a.s.$$

The distribution of X_0 is called the initial distribution of X.

Remark 4.1. 1. Markov property can be expressed as:

$$\mathsf{P}(X_t \in A \mid \mathcal{F}_s) = \mathsf{P}(X_t \in A \mid X_s) \; a.s. \; \forall\, 0 \leq s \leq t, \forall\, A \in \mathcal{B}(\mathbb{R}^d)$$

2. Choose an increasing sequence $\{t_n\}$, $\{X_{t_n}\}$ is a discrete markov process/chain relative to $\{\mathcal{F}_{t_n}\}$.

3. Initial distribution μ of X, $\mathsf{P}(X_0 \in A) = \mu(A), \forall\, A \in \mathcal{B}(\mathbb{R}^d)$. Here $\mu = \mathsf{P} \circ X^{-1}$

Besides, we also have strong markov process, it allows optinal times:

Definition 4.2. Let $\{\mathcal{F}_t\}$ be a filtration on $(\Omega, \mathcal{F}, \mathsf{P})$ and $X = \{X_t : t \geq 0\}$ be an adapted and progressively measurable process relative to $\{\mathcal{F}_t : t \geq 0\}$. Then X is a **strong markov process** if:
$\forall$ optional T of $\{\mathcal{F}_t : t \geq 0\}, \forall\, t \geq 0, \forall\, A \in \mathcal{B}(\mathbb{R}^d)$, it holds that:

$$\mathsf{P}(X_{T+t} \in A \mid \mathcal{F}_{T_+}) = \mathsf{P}(X_{T+t} \mid X_T) \; a.s. \text{ on } \{T < +\infty\}$$

Remark 4.2. 1. We use $B_b(\mathbb{R}^d)$ = bounded, Borel-measurable function on $\mathbb{R}^d$ = Banach space (completed vector space with sup norm: $\|f\| = \sup_x |f(x)|$)

2. For any markov process X, we associate a family of operator $\{T_{s,t} : 0 \leq s \leq t\}$:

$$\forall\, s \leq t, T_{s,t} : B_b(\mathbb{R}^d) \mapsto \text{the space of bounded functions on } \mathbb{R}^d$$

, where typically we will know it is: $T_{s,t} : f \mapsto \mathsf{E}\left[f(X_t) \mid X_s = x\right]$

3. A markov process is normal if $\forall\, 0 \leq s \leq t, \forall\, f \in B_b(\mathbb{R}^d), T_{s,t} \circ f \in B_b(\mathbb{R}^d)$.

4.2 Markov Evolution

Following from that, we can generalise to richer properties of operator:

Theorem 4.1. *If X is normal markov process, then:*

(a) $T_{s,t}$ *is a linear operator on* $B_b(\mathbb{R}^d), \forall\, 0 \le s \le t$.

(b) $T_{s,s}$ *is* $\mathbf{I}_d, \forall\, s \ge 0$. *(Identity,* $T_{s,s}f = f$*)*

(c) $T_{r,s}T_{s,t} = T_{r,t}, \forall\, 0 \le r \le s \le t$.

(d) $\forall\, 0 \le s \le t, \forall\, f \ge 0, f \in B_b(\mathbb{R}^d), T_{s,t}f \ge 0$.

(e) $\forall\, 0 \le s \le t,\ \|T_{s,t}\| = \sup_{f \in B_b(\mathbb{R}^d), f \ne 0} \dfrac{\|T_{s,t}f\|}{\|f\|} \le 1$. *(Contraction)*

(f) If $f(x) = 1, \forall\, x \in \mathbb{R}^d, T_{s,t}f = f$.

Proof. First two of the properties are easy to check, we just left for reader.

(c) It suffices to prove: $T_{r,s}T_{s,t} \circ f = T_{r,t} \circ f, \forall\, f \in B_b(\mathbb{R}^d)$, starting from right we have:

$$\begin{aligned} T_{r,t}f(x) &= \mathsf{E}\left(f(X_t) \mid X_r = x\right) = \mathsf{E}\left[\mathsf{E}\left(f(X_t) \mid \mathcal{F}_s\right) \mid X_r = x\right] \\ &= \mathsf{E}\left[\mathsf{E}\left(f(X_t) \mid X_s\right) \mid X_r = x\right] = \mathsf{E}\left(T_{s,t} \mid X_r = x\right) = T_{r,s}T_{s,t}f(x) \end{aligned}$$

(d) By the definition of conditional expectation, it is easy to check.

(e) Directly applying the definition, it suffices to show:

$$\begin{aligned} \|T_{s,t}f\| &= \sup_{x \in \mathbb{R}^d} |T_{s,t}f(x)| = \sup_{x \in \mathbb{R}^d} |\mathsf{E}\left(f(X_t) \mid X_s = x\right)| \\ &\le \sup_{x \in \mathbb{R}^d} \mathsf{E}\left(|f(X_t)| \mid X_s = x\right) \le \sup_{x \in \mathbb{R}^d} \mathsf{E}\left(\|f\| \mid X_s = x\right) = \|f\| \end{aligned}$$

(f) Straightforwardly, we verify by applying the definition:

$$T_{s,t}f = \mathsf{E}\left(f(X_t) \mid X_s\right) = \mathsf{E}\left(1 \mid X_s\right) = 1 = f$$

□

Remark 4.3. Any family operator satisfy these (a) - (f) is called **markov evolution**.

Then for continous case, we define the transition probability through evolution:

Definition 4.3. Let X be a markov process on $(\Omega, \mathcal{F}, \mathsf{P})$ relative to $\{\mathcal{F}_t : t \ge 0\}$ with markov evolution $\{T_{s,t} : 0 \le s \le t\}$. For $0 \le s \le t$, we define the **transition probability** $\mathsf{P}_{s,t}$ on $\mathbb{R}^d \otimes \mathcal{B}(\mathbb{R}^d)$ by:

$$\mathsf{P}_{s,t}(x, A) = T_{s,t}\mathbf{1}_A(x) = \mathsf{E}\left[X_t \in A \mid X_s = x\right]$$

Remark 4.4. 1. If $\forall\, x \in \mathbb{R}^d, \mathsf{P}_{s,t}(x,\cdot)$ admits a density function $p_{s,t}(x,\cdot)$, i.e.

$$\mathsf{P}_{s,t}(x,A) = \int_A p_{s,t}(x,y)\mathrm{d}y, \forall A \in \mathcal{B}(\mathbb{R}^d)$$

then the family of density function $\{p_{s,t} : 0 \le s \le t\}$ is called **transition probability density function**.

2. From above, we know: $\forall\, f \in \mathcal{B}(\mathbb{R}^d), \forall\, x \in \mathbb{R}^d, \forall\, 0 \le s \le t$,

$$T_{s,t}f(x) = \int_{\mathbb{R}^d} f(y)\mathsf{P}_{s,t}(x,\mathrm{d}y) = \int_{\mathbb{R}^d} f(y)\rho_{s,t}(x,y)\mathrm{d}y$$

Now comes with most important and fundamental theorem:

Theorem 4.2. ***(Chapman-Kolmogorov equations)***
If X be a normal markov process with transition probability density function $\{\mathsf{P}_{s,t} : 0 \le s \le t\}$. Then $\forall\, 0 \le r \le s \le t, \forall\, x \in \mathbb{R}^d, \forall\, A \in \mathcal{B}(\mathbb{R}^d)$:

$$\mathsf{P}_{r,t}(x,A) = \int_{\mathbb{R}^d} \mathsf{P}_{s,t}(y,A)\mathsf{P}_{r,s}(x,\mathrm{d}y)$$

Proof. Since X is normal, then $\forall A \in \mathcal{B}(\mathbb{R}^d), \mathsf{P}_{s,t}(\cdot,A)$ is measurabl, therefore,

$$\begin{aligned}
\mathsf{P}_{r,t}(x,A) = T_{r,t}\mathbf{1}_A(x) &= T_{r,s}T_{s,t}\mathbf{1}_A(x) \\
&= \int_{\mathbb{R}^d} T_{s,t}\mathbf{1}_A(y)\mathsf{P}_{r,s}(x,\mathrm{d}y) \\
&= \int_{\mathbb{R}^d} \mathsf{P}_{s,t}(y,A)\mathsf{P}_{r,s}(x,\mathrm{d}y)
\end{aligned}$$

Then we completed the proof. □

If the density exists, we could extend this into:

Corollary 4.3. *Let X be a normal markov process with transition probability density function $\{\rho_{s,t} : 0 \le s \le t\}$, then $\forall\, 0 \le r \le s \le t, \forall\, x, z \in \mathbb{R}^d$:*

$$\rho_{r,t}(x,z) = \int_{\mathbb{R}^d} \rho_{r,s}(x,y)\rho_{s,t}(y,z)\mathrm{d}y$$

Proof. Leave as homework. □

4.3 Markov & Feller Semigroup

Lastly, we consider more general group that may be helpful when you read literature:

Theorem 4.4. *Let $\{\mathsf{P}_{s,t} : 0 \le s \le t\}$ be a family of mappings from $\mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d) \mapsto [0,1]$, satisfying: $\forall\, 0 \le s \le t$:*

(a) $\mathsf{P}_{s,t}(\cdot,A)$ *is measurable for* $\forall\, A \in \mathcal{B}(\mathbb{R}^d)$.

(b) $\mathsf{P}_{s,t}(x, \cdot)$ *is a probability measure on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)), \forall\, x \in \mathbb{R}^d$.

(c) Chapman-Kolmogorov equations are satisfied.

Let μ *be fixed probability measure on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, *then there exists a probability space* $(\Omega, \mathcal{F}, \mathsf{P}^\mu)$ *with filtration* $\{\mathcal{F}_t : t \geq 0\}$ *and a normal markov process* $X = \{X_t : t \geq 0\}$ *on* $(\Omega, \mathcal{F}, \mathsf{P}^\mu)$ *relative to* $\{\mathcal{F}_t : t \geq 0\}$ *such that:*

(1) $\mathsf{P}(X_t \in A \mid X_s = x) = \mathsf{P}_{s,t}(x, A), \forall\, 0 \leq s \leq t, \forall\, x \in \mathbb{R}^d, \forall A \in \mathcal{B}(\mathbb{R}^d)$.

(2) The initial distribution of X_0 *is* μ.

Proof. The proof is omitted since its complexity, see details textbook, theorem 3.17. □

To connect stationary properties, we define:

Definition 4.4. Let X be a markov process with markov evolution, $\{T_{s,t} : 0 \leq s \leq t\}$. Then X is said to be **time-homogenous** if:

$$T_{s,t} = T_{0,t-s}, \forall\, 0 \leq s \leq t.$$

If X is not time-homogenous, X is called time-inhomogenous.

Remark 4.5. 1. Time-homogenous is equivalent to $\mathsf{P}_{s,t} = \mathsf{P}_{0,t-s}$.

2. For time-homogenous markov process, we write T_t for $T_{0,t}$, P_t for $\mathsf{P}_{0,t}$, ρ_t for $\rho_{0,t}$.

3. For normal time-homogenous markov process, we have $T_s T_t = T_{s+t} \forall\, t, s \geq 0, T_0 = \mathbf{I}_d$. Then we call this T, the **markov transition semigroup**.

Finally, we define what Feller process is:

Definition 4.5. Let $C_0(\mathbb{R}^d)$ be the Banach space of continous function that vanish at ∞, i.e. $\lim_{|x|\to\infty} f(x) = 0$, equipped with the sup/max norm, i.e. $\|f\| = \sup_{x\in\mathbb{R}^d} |f(x)|$. If X is a time-homogenous markov process with markov evolution operator $\{T_t : t \geq 0\}$ satifying:

(a) $T_t C_0(\mathbb{R}^d) \subseteq C_0(\mathbb{R}^d)$.

(b) $\lim_{t\to 0^+} \|T_t f - f\| = 0, \forall\, f \in C_0(\mathbb{R}^d)$.

Then X is called a **Feller process**, $\{T_t : t \geq 0\}$ is called a **Feller semigroup**.

Remark 4.6. 1. Some literature use: $C_0(\mathbb{R}^d) \longleftrightarrow C_b(\mathbb{R}^d)$.

2. If 4.5's (a) replaced by: $T_t B_b(\mathbb{R}^d) \subseteq B_b(\mathbb{R}^d), \forall\, t \geq 0$, we call this process, **strong Feller process**, and $\{T_t : t \geq 0\}$ the **strong Feller semigroup**.

3. For a family $\{T_t, t \geq 0\}$ of semigroup. If $\lim_{t\to 0^+} \frac{T_t f - f}{t}$ exists, then $Af = \lim_{t\to 0^+} \frac{T_t f - f}{t}$ is called **infinitistimal generator** of $\{T_t, t \geq 0\}$.

4.4 Homework

Followings are homeworks of this chapter:

Problem 4.1. Let X be a normal Markov Process with transition density functions $\{\rho_{s,t} : 0 \leq s \leq t\}$. Prove that:

$$\rho_{r,t}(x, z) = \int_{\mathbb{R}^d} \rho_{r,s}(x, y)\rho_{s,t}(y, z)\mathrm{d}y$$

for any $0 \leq r \leq s \leq t$ and $x, z \in \mathbb{R}^d$.

5 Brownian Motion

Brownian motion stands as a cornerstone of stochastic processes, a singular object of immense beauty and utility. It is the unique continuous process with stationary, independent increments, a Gaussian martingale, and a fertile source of mathematical paradox. In this chapter, we study its defining properties and then delve into the fascinating, often pathological, nature of its paths: nowhere differentiable, yet Hölder continuous; of infinite variation, yet with quadratic variation that unfolds deterministically. We will see how Brownian motion connects scaling, time inversion, and a deep recurrence to the origin.

5.1 Define Brownian Motion

Firstly, we present the standard definition of brownian motion:

Definition 5.1. Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration on $(\Omega, \mathcal{F}, \mathsf{P})$. Then the continous stochastic process $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ adapted to $\{\mathcal{F}_t : t \geq 0\}$ is called $d-$dimensional **Brownian Motion** if:

(1) $\mathsf{B}_0 = \mathbf{0}_d, a.s.$

(2) $\forall\, 0 \leq s \leq t, \mathsf{B}_t - \mathsf{B}_s \sim \mathcal{N}(\mathbf{0}_d, (t-s)\mathbf{I}_d)$ and independent of $\mathcal{F}_s$

Remark 5.1. (a) If $d = 1$, we call B as a **standard brownian motion**.

(b) $\mathsf{B} = (\mathsf{B}^{(1)}, \cdots, \mathsf{B}^{(d)})$ is a standard $d-$dimensional brownian motion iff $\mathsf{B}^{(1)}, \cdots, \mathsf{B}^{(d)}$ are independent and standard brownian motion.

A more simple way to transform independent to σ field:

Definition 5.2. $X = \{X_t : t \geq 0\}$ be a stochastic process such that:

$$\forall\, n \in \mathbb{N}, \forall\, 0 = t_0 < t_1 < \cdots, t_n, \text{ the r.v.s } X_{t_0} \perp X_{t_1} - X_{t_0} \perp \cdots \perp X_{t_n} - X_{t_{n-1}}$$

then we say that X has **independent increments**.

Remark 5.2. The Standard Brownian Motion always have independent increments.

Then naturally, they are connected through:

Proposition 5.1. *Let X be a stochastic process with independent increments, then $\forall\, 0 \leq s \leq t, X_t - X_s$ is independent of $\mathcal{F}_s^X$.*

Proof. Leave as a homework, and use Dynkin system for σ-algebra. ☐

The following are equivalent definition, and we could verify they are equal with 5.1.

Definition 5.3. A continous process $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ adapted to filtration $\{\mathcal{F}_t : t \geq 0\}$ is a standard brownian motion if:

(a) B is a zero-mean/centered Gaussian process.

(b) $\mathsf{E}\mathsf{B}_t\mathsf{B}_s = \min\{s, t\}$. $\forall\, t, s \geq 0$

Proof. Fill in a proof for equivalence. □

This is easier to check at higher dimensions (in "time").

5.2 Properties of Sample Path

We leave the session of showing the existence of brownian motion, it can be constructed via distribution or sample path.

Proposition 5.2. *The sample paths of a standard brownian motion are locally Hölder continous of order* $\gamma \in (0, \frac{1}{2})$.

Proof. $\forall\, n \in \mathbb{N}$, by Kolmogorov's continuity theorem, the order γ of Hölder-continous of B satisfied $\gamma \in (0, \frac{n-1}{2n})$, then we take its limits: $\lim_{n\to\infty} \frac{n-1}{2n} = \frac{1}{2}$, we completed proof. □

Remark 5.3. This gives us a hint that how should we interpret the stochastic integral, like $\int \cdot\, \mathrm{d}\mathsf{B}_t$. Let's say for **Young integral**, $\int f \mathrm{d}g$ is well defined only if $\alpha + \beta > 1$, where the α, β are the Hölder continity of f, g respectively. Then this explains why we cannot use **Young integral** to study integral like: $\int \mathsf{B}_t \mathrm{d}\mathsf{B}_t$.

Brownian motion is well studied since it has lots of interesting properties:

Proposition 5.3. *A standard brownian motion* $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ *relative to* $\{\mathcal{F}_t : t \geq 0\}$ *is a martingale relative to* $\{\mathcal{F}_t : t \geq 0\}$.

Proof. The adaptation and integrability is easy to show, we omitted here, then check:

$$\mathsf{B}_t \sim \mathcal{N}(0, t\mathbf{I}_d) \implies \mathsf{E}\left(\mathsf{B}_t - \mathsf{B}_s \mid \mathcal{F}_s\right) = \mathsf{E}\left(\mathsf{B}_t - \mathsf{B}_s\right) = 0 \implies \mathsf{E}\left(\mathsf{B}_t \mid \mathcal{F}_s\right) = \mathsf{B}_s$$

, for $\forall\, 0 \leq s < t$. □

And it also is a (strong) Markov process:

Proposition 5.4. *A standard brownian motion relative to* $\{\mathcal{F}_t : t \geq 0\}$ *is time-homogeneous Markov Process. (Also a strong Markov Process)*

Proof. Leave as a homework □

Interestingly, the brownian motion can be constructed by optional times:

Theorem 5.5. *Let* S *be a finite optional time of a filtration* $\{\mathcal{F}_t : t \geq 0\}$, *and* $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ *be a* $d-$*dimensional brownian motion relative to* $\{\mathcal{F}_t : t \geq 0\}$. *Set* $W_t = \mathsf{B}_{S+t} - \mathsf{B}_S$, $W = \{W_t : t \geq 0\}$, *then* W *is a* $d-$*dimensional brownian motion relative to* $\{\mathcal{F}_t^W : t \geq 0\}$ *and is independent of* $\mathcal{F}_{S+}$.

Proof. We only sketch the idea of the proof here.
By 5.3, it suffices to show it is jointly centered normal, then using characteristic function:

1. $\forall\, n \in \mathbb{N}, 0 = t_0 < t_1 < \cdots < t_n, u_1, \cdots, u_n \in \mathbb{R}^d$, we can show by induction:

$$\mathsf{E}\left[\exp\left(i \sum_{k=1}^{n} u_k (W_{t_k} - W_{t_{k-1}})\right) \mid \mathcal{F}_s\right] = \prod_{k=1}^{n} \exp\left[-\frac{1}{2}\left(t_k - t_{k-1}\right)^2 \|u_k\|^2\right]$$

2. From above, we use tower rules:

$$\mathsf{E}\left\{\mathsf{E}\left[\exp\left(i\sum_{k=1}^{n}u_k(W_{t_k}-W_{t_{k-1}})\right)\mid \mathcal{F}_s\right]\right\}$$
$$=\mathsf{E}\exp\left(i\sum_{k=1}^{n}u_k(W_{t_k}-W_{t_{k-1}})\right)=\prod_{k=1}^{n}\exp\left[-\frac{1}{2}(t_k-t_{k-1})^2\|u_k\|^2\right]$$

Then by linear transformation, we can show:

$$(W_{t_1},W_{t_2},\cdots,W_{t_n})\sim\mathcal{N}\left(\mathbf{0}_d,\begin{bmatrix}\ddots & & \\ & t_k-t_{k-1} & \\ & & \ddots\end{bmatrix}\right)$$

3. Lastly, using independence of increments to show that $W_{t_k}-W_{t_{k-1}}\perp\mathcal{F}_{S+}$

The detailed proof, please refer to textbook. □

Remark 5.4. If S is deterministic, then $W_t\sim\mathcal{N}(0,t)$.

Following ways are useful when proving some theorem:

Proposition 5.6. *Let $W=\{W_t : t\geq 0\}$ be a $d-$dimensional brownian motion relative to $\{\mathcal{F}_t : t\geq 0\}$:*

(a) ***Scaling Property****: For $c>0$, let $X_t=\frac{1}{\sqrt{c}}W_{ct}$, then $\{X_t : t\geq 0\}$ is a standard brownian motion relative to $\{\mathcal{F}_{ct} : t\geq 0\}$.*

(b) ***Time-Inversion Property****: Let $Y_t=tW_{\frac{1}{t}},\forall\, t>0$ with $Y_0=0, a.s.$ Then $\{Y_t : t\geq 0\}$is a standard brownian motion relative to $\{\mathcal{F}_t^Y : t\geq 0\}$.*

(c) ***Time-Reversal Property****: For $T>0$, let $Z_t=W_T-W_{T-t},\forall\, 0\leq t\leq T$. Then $\{Z_t : 0\leq t\leq T\}$ is a standard brownian motion relative to $\{\mathcal{F}_t^Z : t\geq 0\}$.*

(d) ***Symmetry Property****: $-W=\{-W_t : t\geq 0\}$ is a standard brownian motion relative to $\{\mathcal{F}_t : t\geq 0\}$.*

Proof. We omitted this proof since it is only tedious checking conditions. □

Then we have below strange properties:

Proposition 5.7. *With probability 1, a standard brownian motion changes sign infinitely many times in any time interval $[0,\varepsilon]$ with any $\varepsilon>0$.*

Proof. See textbook theorem 7.18 □

Consequently, we can move to origin:

Proposition 5.8. *With probability 1, a standard brownian motion returns to origin infinitely often.*

Proof. See textbook theorem 9.15, hint: using last proposition + Time-Inversion. □

According to above, it seems like zeros of brownian motion should be large, but:

Theorem 5.9. *Let* $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ *be a standard 1-dimensional brownian motion on* $(\Omega, \mathcal{F}, \mathsf{P})$ *relative to* $\{\mathcal{F}_t : t \geq 0\}$*. For* $\omega \in \Omega$*, we define (zeros of* $\mathsf{B}(\omega)$*):*

$$\mathcal{L}(\omega) = \{t \geq 0 : \mathsf{B}_t(\omega) = 0\}$$

Then for $\mathsf{P} - a.s, \omega \in \Omega$*:*

(a) $\mathcal{L}(\omega)$ *has zero* ***Lebesgue measure****.*

(b) $\mathcal{L}(\omega)$ *is closed and unbounded.*

(c) $\mathcal{L}(\omega)$ *has an accumulation point at* $t = 0$*.*

(d) $\mathcal{L}(\omega)$ *has no isolated point in* $[0, +\infty)$*.*

Proof. (a) Since B is adapted to $\{\mathcal{F}_t : t \geq 0\}$ and has continous sample paths, then B is progressively measurable, consquently, measurable. This tells us:

$$\mathcal{L} = \{(t, \omega) \in [0, +\infty) \otimes \Omega : \mathsf{B}_t(\omega) = 0\} \in \mathcal{B}([0, +\infty)) \otimes \mathcal{F}$$

Then it is legal to consider:

$$\mathsf{E}\lambda(\mathcal{L}(\omega)) = \mathsf{E}\int_0^\infty \mathbf{1}_{\mathcal{L}(\omega)} \mathrm{d}t = \int_0^\infty \mathsf{E}\mathbf{1}_{\mathcal{L}(\omega)} \mathrm{d}t = \int_0^\infty \mathsf{P}(\mathsf{B}_t(\omega) = 0) \mathrm{d}t = 0$$

Therefore, combined with $\lambda(\mathcal{L}(\omega)) \geq 0$, we have $\lambda(\mathcal{L}(\omega)) = 0, a.s.$

(b) Notice that $\mathcal{L}(\omega)$ is pre-image of $\{0\}$, which is closed, then by continuity of B, we have $\mathcal{L}(\omega)$ is closed. And for unbounded property, it is easy to see from 5.8.

(c) We know from 5.7 that: $\exists$ decreasing $\{r_n : n \in \mathbb{N}\}, \{s_n : n \in \mathbb{N}\}, s.t.$

$$0 < r_{n+1} < s_{n+1} < r_n < s_n, \text{and } \mathsf{B}_{r_n}(\omega) > 0, \mathsf{B}_{s_n}(\omega) < 0, \lim_{n\to\infty} r_n = \lim_{n\to\infty} s_n = 0$$

Then by continuity, $\exists\, t_n \in (r_n, s_n), s.t. \mathsf{B}_{t_n}(\omega) = 0 \implies t_n \in \mathcal{L}(\omega)$. It is easy to see the accumulation point, since $\lim_{n\to\infty} t_n = 0$.

(d) Firstly, we want to split the event, and show the latter's measure is zero:

$$\{\omega \in \Omega : \mathcal{L}(\omega) \text{ has isolated points on } \mathbb{R}_+\} \subseteq \bigcup_{\substack{a,b \in \mathbb{Q} \\ 0<a<b}} \{\omega \in \Omega : \exists!\, s \in (a, b), s.t. \mathsf{B}_s(\omega) = 0\}$$

Now, set $\beta_t = \inf\{s > t : \mathsf{B}_s(\omega) = 0\}, t \geq 0$, by (c), $\beta_0 = 0, \beta_t < +\infty, a.s$, notice:

$$\{\beta_t < r\} = \{\exists\, s \in (t, r), s.t.\ \mathsf{B}_s = 0\} \in \mathcal{F}_r$$

Then β_t is optional time of $\{\mathcal{F}_t : t \geq 0\}$, and $\mathsf{B}_{\beta_t}(\omega) = 0\ a.s.$ More importantly:

$$\begin{aligned}\beta_{\beta_t} &= \inf\{s > \beta_t : \mathsf{B}_s(\omega) = 0\} = \beta_t + \inf\{s > 0 : \mathsf{B}_{\beta_t + s}(\omega) = 0\} \\ &= \beta_t + \inf\{s > 0 : \mathsf{B}_{\beta_t+s} - \mathsf{B}_{\beta_t} = 0\} = \beta_t\end{aligned}$$

Lastly, we are ready to check, $\forall\, a < b$:

$$\{\omega \in \Omega : \exists !\ s \in (a,b), s.t.\mathsf{B}_s(\omega) = 0\} \subseteq \{\omega \in \Omega : \beta_a < b, \beta_{\beta_a} \geq b\} = \emptyset$$

□

No suprisingly, monotonicity is also too demanding for brownian motion:

Theorem 5.10. *Let $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ ve a standard 1-dimensional brownian motion on $(\Omega, \mathcal{F}, \mathsf{P})$ relative to filtration $\{\mathcal{F}_t : t \geq 0\}$. Then the sample path $\mathsf{B}.(\omega)$ is monotone in no interval a.s.*

Proof. Firstly, we define following partition:

$$\mathsf{F}_{s,t} = \{\omega \in \Omega : \mathsf{B}.(\omega) \text{ is monotone on } [s,t]\}$$

Then it is clear that:

$$\mathsf{F} = \{\omega \in \Omega : \mathsf{B}.(\omega) \text{ is monotone in some interval}\} = \bigcup_{\substack{0 \leq s < t \\ s,t \in \mathbb{Q}}} \mathsf{F}_{s,t}$$

Since we could split monotonicity into increasing and decreasing, and they are techinically same, then we denote $\mathsf{F}^+_{s,t}$ the increasing part, and only consider:

$$A^n_{s,t} = \bigcap_{i=0}^{n-1} \{\omega \in \Omega : \mathsf{B}_{s+\frac{i+1}{n}(t-s)}(\omega) - \mathsf{B}_{s+\frac{i}{n}(t-s)}(\omega) \geq 0\} \in \mathcal{F}$$

Clearly, $\mathsf{F}^+_{s,t} \subseteq A^n_{s,t}, \forall\, n \in \mathbb{N}$, then we completed the proof, since:

$$\mathsf{P}(A^n_{s,t}) = \prod_{i=0}^{n-1} \left(\frac{1}{2}\right) = \frac{1}{2^n} \implies \mathsf{P}\left(\bigcap_{n=1}^{\infty} A^n_{s,t}\right) \leq \frac{1}{2^N}, \forall\, N \in \mathbb{N} \implies \mathsf{P}\left(\bigcap_{n=1}^{\infty} A^n_{s,t}\right) = 0$$

□

Remark 5.5. This is also the reason why $\int f \mathrm{d}g$ fails in **lebesgue-stieltjes integral** for defining stochastic integral, since it asks g to be increase, but B_t is not monotonic at all.

Recall that: let $f : [0, +\infty) \mapsto \mathbb{R}, t \geq 0$ is a point of increase of f, if $\exists\, \delta > 0$, s.t. $\forall\, t - \delta < r < t < s < t + \delta, f(r) \leq f(t) \leq f(s)$. And clearly this one is loose than monotonicity. Except above, the critical points is also "non-exist".

Theorem 5.11. *With probability 1, the sample paths of* B *have no point of increase or point of decrease.*

Proof. We intentionally omitted the proof. □

Based on above theorem, naturally we have:

Proposition 5.12. *For standard 1-dimensional brownian motion, for* $\mathsf{P}-$*almost all sample paths, all local maximum/minimum is strict.*

Proof. As usual, we do a split of events:

$$\begin{aligned}
&\{\omega \in \Omega : \text{all local maximum of } \mathsf{B}.(\omega) \text{ are strict}\} \\
&\supseteq \bigcap_{\substack{0 \le t_1 < t_2 < t_3 < t_4 \\ t_1, t_2, t_3, t_4 \in \mathbb{Q}}} \{\omega \in \Omega : \max_{t \in [t_1, t_2]} \mathsf{B}_t(\omega) \neq \max_{t \in [t_3, t_4]} \mathsf{B}_t(\omega)\}
\end{aligned}$$

Lastly, noticing following:

$$\begin{aligned}
&\max_{t \in [t_1, t_2]} \mathsf{B}_t(\omega) - \max_{t \in [t_3, t_4]} \mathsf{B}_t(\omega) = \\
&\max_{t \in [t_1, t_2]} (\mathsf{B}_t(\omega) - \mathsf{B}_{t_2}(\omega)) - \max_{t \in [t_3, t_4]} (\mathsf{B}_t(\omega) - \mathsf{B}_{t_3}(\omega)) + (\mathsf{B}_{t_2}(\omega) - \mathsf{B}_{t_3}(\omega))
\end{aligned}$$

And since each parts of them are independent, then we could easily check their probability using density function. Therefore, we completed the proof. □

Without suprising, since no monoticity, the set of local maximum should be large:

Corollary 5.13. *Let* B *be a standard 1-dimensional brownian motion on* $(\Omega, \mathcal{F}, \mathsf{P})$ *relative to* $\{\mathcal{F}_t : t \ge 0\}$*. Then the set of points of local maximum for sample path* $\mathsf{B}.(\omega)$ *is* ***countable*** *and* ***dense*** *in* $[0, +\infty)$*.*

Proof. Since $\mathsf{B}.(\omega)$ is continous and monetone on no interval. □

Then last properties is nowhere differentiable even though it is smooth:

Theorem 5.14. *Let* B *be a standard 1-dimensional brownian motion, then almost surely* $\omega, \mathsf{B}.(\omega)$ *is nowhere differentiable.*

Recall: f be a continous function:

$$D^+ f(t) = \limsup_{h \to 0^+} \frac{f(t+h) - f(t)}{h} \qquad D^- f(t) = \limsup_{h \to 0^-} \frac{f(t+h) - f(t)}{h}$$
$$D_+ f(t) = \liminf_{h \to 0^+} \frac{f(t+h) - f(t)}{h} \qquad D_- f(t) = \liminf_{h \to 0^-} \frac{f(t+h) - f(t)}{h}$$

These are **Dini derivatives** of f at t. And the typical derivatives are:

$$\lim_{h \to 0^+} \frac{f(t+h) - f(t)}{h} = \lim_{h \to 0^-} \frac{f(t+h) - f(t)}{h}$$

The function f is said to be *differentiable* at $t > 0$ if it is differentiable from both the right and the left and the four Dini derivatives agrees.

Proof. We will proof this by showing their upper-right/left derivatives are infinite, more precisely, what we want is to prove: $\exists\, F \in \mathcal{F}, \mathsf{P}(F) = 1, s.t:$

$$F \subseteq \{\omega \in \Omega : \forall\, t \ge 0, \text{ either } D^+\mathsf{B}_t(\omega) = +\infty \text{ or } D_+\mathsf{B}_t(\omega) = -\infty\}$$

It is enough to consider the interval $[0, 1]$. For fixed $j, k \ge 1$, we define:

$$A_{j,k} = \bigcup_{t \in [0,1]} \bigcap_{h \in [0, \frac{1}{k}]} \{\omega \in \Omega : |\mathsf{B}_{t+h}(\omega) - \mathsf{B}_t(\omega)| \le jh\}$$

And it is clear that, we need to find $C \in \mathcal{F}, s.t.\ A_{j,k} \subseteq C, \mathsf{P}(C) = 0$, since:

$$\{\omega \in \Omega : \exists\, t \in [0,1], -\infty < D_+\mathsf{B}_t(\omega) < D^+\mathsf{B}_t(\omega) < \infty\} = \bigcup_{j=1}^{\infty}\bigcup_{k=1}^{\infty} A_{j,k}$$

A crucial observation is that, for a fixed $\omega \in A_{j,k}$:

$$\exists\, t \in [0,1], \forall\, h \in \left[0, \frac{1}{k}\right], |\mathsf{B}_{t+h}(\omega) - \mathsf{B}_t(\omega)| \leq jh$$

Then by triangle inequality, following works:

$$\forall\, h_1, h_2 \in \left[0, \frac{1}{k}\right], |\mathsf{B}_{t+h_1}(\omega) - \mathsf{B}_{t+h_2}(\omega)| \leq j(h_1 + h_2)$$

Or more generally, $\forall\, n \geq 4k, \exists\, i, s.t : \frac{i-1}{n} \leq t \leq \frac{i}{n}$, then for $l = 1, 2, 3$:

$$\left|\mathsf{B}_{\frac{i+l}{n}}(\omega) - \mathsf{B}_{\frac{i+l-1}{n}}\right| \leq j\left(\left(\frac{i+l}{n} - t\right) + \left(\frac{i+l-1}{n} - t\right)\right) \leq j\left(\frac{2l+1}{n}\right)$$

Inspired by this control, notice that:

$$A_{j,k} \subseteq \bigcup_{j=1}^{n}\bigcap_{l=1}^{3}\left\{\omega \in \Omega : \left|\mathsf{B}_{\frac{i+l}{n}}(\omega) - \mathsf{B}_{\frac{i+l-1}{n}}\right| \leq j\left(\frac{2l+1}{n}\right)\right\}$$

Using the distribution information, we easily show the latter's probability is zero. □

Simiar to above technicals, one can also show:

Theorem 5.15. *The sample path $\mathsf{B}_{.}(\omega)$ of standard 1-dimensional brownian motion B are nowhere Hölder-continous with order $\gamma > \frac{1}{2}$, a.s.*

Proof. Leave as a homework □

Remark 5.6. $\gamma > \frac{1}{2}$ is the threshold of the order of Hölder continity.

Finally, we can also somehow measure the "derivatives":

Theorem 5.16. ***(Lévy modulus)***
Let B be a 1-dimensional brownian motion on $(\Omega, \mathcal{F}, \mathsf{P})$ relative to $\{\mathcal{F}_t : t \geq 0\}$:

$$\mathsf{P}\left(\liminf_{\delta \to 0^+} \frac{1}{\sqrt{2\delta \ln\left(\frac{1}{\delta}\right)}} \max_{\substack{s,t \in [0,1] \\ |t-s| < \delta}} |\mathsf{B}_t - \mathsf{B}_s| = 1\right) = 1$$

Proof. See the textbook Theorem 9.25 □

5.3 Homeworks

Followings are homeworks of this chapter:

Problem 5.1. Let X be a stochastic process with independent increments. Then, for $\forall\, 0 \leq s < t$, the increment $X_t - X_s$ is independent of $\mathcal{F}_s^X$.

Problem 5.2. (1) Prove that a standard Brownian motion is a time-homogeneous Markov process.

(2) Determine the family of Markov evolution operators.

(3) Determine whether a standard brownian motion is a (strong) Feller process of not.

(4) Determine the infinitestimal generator of standard brownian motion.

Problem 5.3. Let B be a brownian motion on $(\Omega, \mathcal{F}, \mathsf{P})$ relative to some filtration $\{\mathcal{F}_t : t \geq 0\}$. Prove that there exists $F \in \mathcal{F}$ with $\mathsf{P}(F) = 1$, such that:

$$F \subseteq \{\omega \in \Omega : \forall\, t \geq 0, \text{ either } D^{-}\mathsf{B}_t(\omega) = +\infty \text{ or } D_{-}\mathsf{B}_t(\omega) = -\infty\}$$

Problem 5.4. Prove that the sample paths of a standard 1-dimensional Brownian motion are nowhere Hölder-continous with Hölder order $\gamma > 1/2$ a.s.

6 Poisson & Lévy Process

The primary motivation for this chapter is to understand the behavior of Lévy processes under a stochastic clock. By replacing deterministic time t with a random process T_t (a subordinator), we can construct a wider class of processes that maintain the Lévy structure while allowing for more complex path behaviors. Furthermore, we aim to extend the simple Markov property to stopping times. Since Lévy processes are often used to model first-passage times and ruin probabilities, the **Strong Markov Property** is not merely a theoretical curiosity but a vital tool for analyzing the process at random horizons.

6.1 Infinitely Divisible Distribution

Firstly, we define what is a poisson r.v. by given the distribution function:

Definition 6.1. A r.v. X taking value in $\mathbb{N}\cup\{0\}$ is a **poisson** r.v. with parameter $\lambda > 0$, denoted $X \sim \mathrm{Poi}(\lambda)$, if:

$$\mathsf{P}(X = n) = \frac{\lambda^n}{n!}e^{-\lambda}$$

The unity (sum to one) can be easily verified by tayor expansion of e^λ at zero.

Remark 6.1. Simply calculation gives us:

1. $\mathsf{E}X = \lambda = \mathrm{Var}(X)$
2. $\phi_X(u) = \mathsf{E}e^{iuX} = \exp(\lambda(e^{iu} - 1))$

Now, we introduce the most important concept of studying poisson r.v.:

Definition 6.2. A r.v. X is **infinitely divisible** if:

$$\forall\, n \in \mathbb{N}, \exists\, i.i.d \text{ r.v.s } Y_1^{(n)}, Y_2^{(n)}, \cdots, Y_n^{(n)} \text{ s.t. } Y_1^{(n)} + \cdots + Y_n^{(n)} \overset{d.}{=} X$$

Remark 6.2. A probability measure μ on $\mathbb{R}^d$, is infinitely divisible, if the r.v. X is infinitely divisible with $\mu_X = \mu$, i.e. $\mathsf{P}(X \in A) = \mu(A)$.

We will come back to this later, and present another concept about poisson r.v.:

Definition 6.3. Let $\{Z_n : n \in \mathbb{N}\}$ be $i.i.d$ r.v.s with values in $\mathbb{R}^d$, let $Y \sim \mathrm{Poi}(\lambda)$ be independent of $\{Z_n : n \in \mathbb{N}\}$, then:

$$X = \sum_{i=1}^{Y} Z_i = Z_1 + \cdots + Z_Y$$

is called the **compounded Poisson** r.v.s.

Remark 6.3. If the distribution of Z_1 is μ, then we denoted as $X \sim \mathrm{Poi}(\lambda, \mu)$ and:

$$\phi_X(u) = \mathsf{E}e^{iuX} = \exp\left(\int_{\mathbb{R}^d} \lambda\left(e^{iuy} - 1\right)\mu(\mathrm{d}y)\right)$$

In order to smoothly introduce Lévy-Khintchine theorem, we first look at:

Definition 6.4. Let ν be a Borel measure on $\mathbb{R}^d/\{0\}$.

(a) If $\int_{\mathbb{R}^d/\{0\}} \min(y^2,1)\nu(\mathrm{d}y) < \infty$, then we say ν is a **Lévy measure**.

(b) If ν is absolutely continous w.r.t Lebesgue measure, then the **Radon-Nikodym derivative**, $\mathsf{P}_\nu = \frac{\mathrm{d}\nu}{\mathrm{d}x}$ is called **Lévy density**.

Remark 6.4. 1. $\int_{\mathbb{R}^d/\{0\}} \min(y^2,1)\nu(\mathrm{d}y) < \infty \Leftrightarrow \int_{\mathbb{R}^d/\{0\}} \frac{y^2}{1+y^2}\nu(\mathrm{d}y) < \infty$

2. A **finite measure** is always **Lévy measure**. (The reverse is generally not true)

Then from the definition, it should be clear that:

Proposition 6.1. *Every Lévy measure on $\mathbb{R}^d/\{0\}$ is $\sigma-$finite.*

Proof. Leave as a homework □

And one of the most important theorem in this chapter, which depicts what an infinitely divisible measure (hence r.v.) looks like, i.e. its general form of characteristic function:

Theorem 6.2. ***(Lévy-Khintchine)***
A probability measure, μ, on $\mathbb{R}^d$ is infinitely divisible if: $\exists\ a \in \mathbb{R}^d, A = \mathcal{M}_d(\mathbb{R})$ satisfying $A = A^\top, A \succeq 0$, and a Lévy measure ν on $\mathbb{R}^d/\{0\}$, such that:

$$\phi_\mu(x) = \exp\left(ia^\top x - \frac{1}{2}x^\top Ax + \int_{\mathbb{R}^d/\{0\}} \left(e^{ix^\top y} - 1 - ix^\top y\mathbf{1}_{B_1(0)}(y)\right)\nu(\mathrm{d}y)\right)$$

Conversely, $\forall\ \phi(x)$ satisfy the Lévy-Khintchine formula is the characteristic function of an infinitely divisible probability measure.

Proof. Omitted. □

Until now, one can see why we introduce compounded poisson process (compare).

Remark 6.5. 1. If $Y \sim \mathcal{N}(a,A)$, then $\phi_Y(x) = \exp(ia^\top x - \frac{1}{2}x^\top Ax)$.

2. We call (a,A,ν) the characteristic of μ. (Since μ is uniquely determined by (a,A,ν))

Then since that exponent is too complex, we define following symbol:

Definition 6.5. We call the map η:

$$\eta(x) = ia^\top x - \frac{1}{2}x^\top Ax + \int_{\mathbb{R}^d/\{0\}} \left(e^{ix^\top y} - 1 - ix^\top y\mathbf{1}_{B_1(0)}(y)\right)\nu(\mathrm{d}y)$$

a Lévy$-$symbol. (or Lévy exponent, charateristic exponent)

Remark 6.6. For inifinitely divisible probability measure μ, $\phi_\mu(x) = e^{\eta(x)}$.

Now, let's see two basic proposition to get familiar with these notions:

Proposition 6.3. *The Lévy−symbol satisfies followings:*

(a) $\Re(\eta(x)) \leq 0$.

(b) η *is continous on* $\mathbb{R}^d$. *It is uniformly continous on bounded sets.*

(c) $\exists\, c > 0$, *s.t.* $|\eta(x)| \leq c(1 + |x|^2)\ \forall\, x \in \mathbb{R}^d$.

Proof. (a) Simply, we pick out the real part of Lévy symbol:

$$\Re(\eta(x)) = -\frac{1}{2}x^\top A x + \int_{\mathbb{R}^d/\{0\}} \left[\cos(x^\top y) - 1\right] \nu(\mathrm{d}x) \leq 0$$

, first terms is negative since $A \succeq 0$, second is for $\cos(x) \leq 1\ \forall\, x \in \mathbb{R}$.

(b) To see this, we need to investigate into difference, $\forall\, x_1 \neq x_2 \in \mathbb{R}^d$:

$$\begin{aligned}\eta(x_1) - \eta(x_2) = {} & ia^\top(x_1 - x_2) - \frac{1}{2}(x_1^\top A x_1 - x_2^\top A x_2) \\ & + \int_{\mathbb{R}^d/\{0\}} \left[e^{ix_1^\top y} - e^{ix_2^\top y} - i(x_1^\top - x_2^\top) y \mathbf{1}_{B_1(0)}(y)\right] \nu(\mathrm{d}y)\end{aligned}$$

Then for the qudratic form, we could rewrite:

$$x_1^\top A x_1 - x_2^\top A x_2 = (x_1^\top - x_2^\top)Ax_1 + x_1^\top A(x_1 - x_2) + (x_1^\top - x_2^\top)A(x_1 - x_2)$$

And also using bound for imaginary number on exponential to get:

$$|\eta(x_1) - \eta(x_2)| \leq C_{x_1}|x_1 - x_2| + O(|x_1 - x_2|^2) \implies \lim_{x_1 \to x_2} \eta(x_2) = \eta(x_1)$$

, where $C_{x_1} = |a| + 2\|A\| + \int_{\mathbb{R}^d/\{0\}} \left[e^{ix_1 y} iy - iy\mathbf{1}_{B_1(0)}\right] \nu(\mathrm{d}y)$.

(c) Similar to (b), one can show:

$$|\eta(x)| \leq |a| \cdot |x| + \|A\| \cdot |x|^2 + C + C'|x| = O(1 + x^2)$$

Therefore, we completed the proof. □

The following theorem explains the connection between compounded poisson r.v. and infinitely divisible r.v., also serves as conclusion of this subchapter.

Theorem 6.4. *Any infinitely divisible probability measure* μ *on* $\mathbb{R}^d$ *is the* ***weak limit*** *of a sqeuence of compounded poisson distributions.*

Proof. Let ϕ be characteristic function of μ, then since it is infinitely divisible, there exists μ_n such that: $\mu = \mu_n * \cdots * \mu_n$, and its ch.f is $\phi^{1/n}(x)$, now write:

$$\phi_n(x) = \exp\left[n\left(\phi^{1/n}(x) - 1\right)\right] = \exp\left(n\int_{\mathbb{R}^d}(e^{ix^\top y} - 1)\mu_n(\mathrm{d}y)\right)$$

, which is exactly the compound poisson r.v.'s ch.f, moreover, one can chcek:

$$\phi_n(x) = \exp\left[n\left(e^{\frac{1}{n}\log(\phi(x))} - 1\right)\right] = \exp\left\{n\left[\frac{1}{n}\log(\phi(x)) + o\left(\frac{1}{n}\right)\right]\right\} \longrightarrow \phi(x)$$

as $n \to \infty$, then the result follows by Lévy continuity theorem. □

6.2 Lévy Process

Now, we come to a new subchapter, here we will firstly introduce Lévy process.

Definition 6.6. A stochastic process $X = \{X_t : t \geq 0\}$ is **Lévy process** if:

(a) $X_0 = 0$ a.s.

(b) X has independent and stationary increments.

(c) X is stochastic continous. (continous in probability)

Note under (2) the stochasticall continuous is equivalent to stochastic continuous at $t = 0$, and another property is that the modification of Lévy process is still a Lévy process.

Remark 6.7. 1. **Continous:** $\forall\, t > 0, \lim_{s\to t} X_s = X_t$.

2. **Stochastic Continous:** $\forall\, \varepsilon > 0, \lim_{s\to t} \mathsf{P}\left(|X_s - X_t| > \varepsilon\right) = 0, \forall\, t$

Clearly, from the definition, we have below trivial example:

Example 6.1. A standard brownian motion is a Lévy process.

And then comes with poisson process:

Definition 6.7. The poisson process with intensity $\lambda > 0$, i.e.

$$N = \{N_t : t \geq 0\} \text{ with values in natural number satisfy } N_t \sim \text{Poi}(\lambda t)$$

, is a Lévy process.

Remark 6.8. 1. The sample paths of poisson process are discontinous.

2. Set $\tilde{N}_t = N_t - \lambda t$, $\tilde{N} = \{\tilde{N}_t, t \geq 0\}$ is called compensated poisson process, $\mathsf{E}\tilde{N}_t = 0$.

Besides, the interesting compounded poisson process:

Definition 6.8. $\{Z_n : n \in \mathbb{N}\}$ is a sequence of $i.i.d$ r.v.s with values in $\mathbb{R}^d$, distribution μ. Let $N = \{N_t, t \geq 0\}$ be a poisson process with intensity $\lambda > 0$, such that N is independent of $\{Z_n : n \in \mathbb{N}\}$. Let $Y_t = \sum_{i=1}^{N_t} Z_i = Z_1 + \cdots + Z_{N_t}, \forall\, t \geq 0$ with convention if $N_t = 0, Y_t = 0$. Then Y_t is called compounded poisson process.

Remark 6.9. With same convention, we denoted as $\forall\, t \geq 0, Y_t \sim \text{Poi}(\lambda t, \mu)$.

This chapter mainly aim to study Lévy process, sicne lots of other processes are just special cases of Lévy process. Then following proposition presents the relationship between Lévy and compounded poisson process:

Proposition 6.5. *The compounded poisson process, Y, is Lévy process.*

Proof. First two conditions are easy to check, we mainly focus on stochastic continuous:

$$\begin{aligned}\mathsf{P}\left(|Y_t| > \varepsilon\right) &= \sum_{n=0}^{\infty} \mathsf{P}\left(|Y_t| > \varepsilon, N_t = n\right) = \sum_{n=0}^{\infty} \mathsf{P}\left(|Z_1 + \cdots + Z_n| > \varepsilon\right) \mathsf{P}(N_t = n) \\ &= \sum_{n=0}^{\infty} \mathsf{P}(|Z_1 + \cdots + Z_n| > \varepsilon) \frac{(\lambda t)^n}{n!} e^{-\lambda t} \xrightarrow{\text{DCT}} 0\end{aligned}$$

, as $t \to 0$, then we completed the proof. ☐

Remark 6.10. 1. For compound poisson process's sample path are discontinous.

2. A compound poisson proces is a poisson process iff $d = 1, Z_n = 1$ $a.s. \forall\, n$.

And by first subchapter's preparation, we could state following proposition:

Proposition 6.6. *Let $X = \{X_t : t \geq 0\}$ be Lévy process, then $\forall\ t \geq 0, X_t$ is infinitely divisible.*

Proof. $\forall\ n \in \mathbb{N}$, we could split into: $Y_k^{(n)} = X_{\frac{kt}{n}} - X_{\frac{(k-1)t}{n}}$, then by independent and stationary increment, we know that, $Y_1^{(n)}, \cdots, Y_n^{(n)}$ are i.i.d. and $X_t = \sum_{k=1}^{n} Y_k^{(n)}$. ☐

Lemma 6.7. *If $X = \{X_t : t \geq 0\}$ is stochastically continous, then $\forall\ u \in \mathbb{R}^d, t \mapsto \phi_{X_t}(u)$ is uniformly continous.*

Proof. The general idea is to control them by parts, firstly using the fact, $u \mapsto e^{iu}$ is continuous, then $\forall\, \varepsilon > 0, \exists\, \delta' > 0$, s.t:

$$\forall\ |y - 0| < \delta',\ |e^{iu^\top y} - e^{iu^\top 0}| = |e^{iu^\top y} - 1| < \varepsilon$$

Next, using the stochastic continuity, we could control, $\forall\, \varepsilon > 0, \exists\, \delta > 0$, s.t:

$$\forall\, 0 < |t - s| < \delta,\ \mathsf{P}(|X_t - X_s| \geq \delta') < \varepsilon$$

Therefore, we could use these two by carefully our event, $\forall\, 0 < |t - s| < \delta$:

$$\begin{aligned}|\phi_{X_t}(u) - \phi_{X_s}(u)| &\leq \mathsf{E}\left|e^{iu^\top X_t} - e^{iu^\top X_s}\right| = \mathsf{E}\left|e^{iu^\top (X_t - X_s)} - 1\right| \\ &= \mathsf{E}\left|e^{iu^\top (X_t - X_s)} - 1\right| \mathbf{1}_{\{|X_t - X_s| \geq \delta'\}} + \mathsf{E}\left|e^{iu^\top (X_t - X_s)} - 1\right| \mathbf{1}_{\{|X_t - X_s| < \delta'\}} \\ &\leq 2\mathsf{P}(|X_t - X_s| \geq \delta') + \sup_{|y| < \delta'} \left|e^{iu^\top y} - 1\right| < 3\varepsilon\end{aligned}$$

This completed the proof, since it is independent with interval we chose. ☐

Remark 6.11. For any bounded set A, the continuity is also uniformly continuous on A.

Theorem 6.8. *If X is a Lévy process, $\eta(a)$ is the Lévy symbol of X_1, then $\forall\ u \in \mathbb{R}^d, t \geq 0, \phi_{X_t}(u) = e^{t\eta(u)}$.*

Proof. Since there is no distribution information, then considering to establish equation:

$$\phi_{X_{t+s}}(u) = \mathsf{E}e^{iu^\top X_{t+s}} = \mathsf{E}e^{iu^\top (X_{t+s}-X_t)}\mathsf{E}e^{iu^\top X_t} = \mathsf{E}e^{iu^\top X_t}\mathsf{E}e^{iu^\top X_s} = \phi_{X_t}(u)\phi_{X_s}(u)$$

, $\forall\, t, s \geq 0$. Therefore, set $f(t) = \phi_{X_t}(u)$, above gives: $f(t+s) = f(t)f(s)$. Notice:

$$\forall\, n \in \mathbb{N},\; f(n) = f(n-1)f(1) = f(n-2)f^2(1) = \cdots = f^n(1)$$

This also can be generalised to rational number since:

$$\forall\, n \in \mathbb{N}, f(nt) = f^n(t) \implies \forall\, p, q \in \mathbb{Z}/\{0\}, f(p/q) = f^{1/q}(p) = f^{p/q}(1)$$

Then we land in $\forall\, x \in \mathbb{Q}, f(x) = f^x(1)$. Now, $\forall\, x \notin \mathbb{Q}$, we use rational number to approximate, then it could be generalise to real number, so: $\phi_{X_t}(u) = \phi^t_{X_1}(u) = e^{t\eta(u)}$. □

Corollary 6.9. *(Lévy-Khintchine formula for Lévy process)*
Let X be a Lévy process, and (a, A, ν) be the characteristic of X_1, then:

$$\phi_{X_t}(u) = \exp\left\{t\left[ia^\top x - \frac{1}{2}x^\top Ax + \int_{\mathbb{R}^d/\{0\}}\left(e^{ix^\top y} - 1 - ix^\top y\mathbf{1}_{B_1(0)}(y)\right)\nu(\mathrm{d}y)\right]\right\}$$

Proof. The proof is directly followed by 6.2. □

Remark 6.12. We define Lévy symbol and characteristic of X as the Lévy symbol and characteristic of X_1.

Then following we present some properties about Lévy process.

Proposition 6.10. *Let X be a Lévy process with characteristic (a, A, ν), then $-X = \{-X_t, t \geq 0\}$ is also a Lévy process with characteristic $(-a, A, \tilde{\nu})$, where $\tilde{\nu}$ is the Lévy measure satisfy $\tilde{\nu}(A) = \nu(A), \forall\, A \in \mathcal{B}(\mathbb{R}^d)$.*

Proof. Leave as a homework □

A natural question is that the limit of a sequence of Lévy processes is Lévy or not.

Theorem 6.11. *Let $X^{(n)}$ be Lévy process, $\forall\, n \in \mathbb{N}$. Let X be a process. If:*

(a) $\forall\, t \geq 0, X_t^{(n)}$ *converges in probability to* X_t.

(b) $\forall\, \varepsilon > 0, \lim_{n\to\infty}\limsup_{t\to 0^+}\mathsf{P}\left(|X_t^{(n)} - X_t| > \varepsilon\right) = 0$

, then X is a Lévy process.

Proof. We just check whether it satisfies the three properties:

(a) Since $\forall\, n \in N, X_0^{(n)} = 0$ a.s., we have $X_0 = 0$ a.s.

(b) From first assumption, we have, $\forall\, 0 \leq t_1 < \cdots < t_k$, then $\forall\, 1 \leq j \leq k-1$:

$$X^{(n)}_{t_{j+1}} - X^{(n)}_{t_j} \xrightarrow{p.} X_{t_{j+1}} - X_{t_j}, \text{ as } n \to \infty$$

But this is not enough, then notice $\exists$ subsequence $\{n_r : r \in N\}$ s.t:

$$X^{(n_r)}_{t_{j+1}} - X^{(n_r)}_{t_j} \xrightarrow{a.s.} X_{t_{j+1}} - X_{t_j}, \text{ as } n_r \to \infty$$

Now, since limiting process has independent increment, so as the X has. Next:

$$\mathsf{E}e^{iu^\top (X_t - X_s)} = \lim_{n_r \to \infty} \mathsf{E}e^{iu^\top (X^{(n_r)}_t - X^{(n_r)}_s)} = \lim_{n_r \to \infty} \mathsf{E}e^{iu^\top X^{n_r}_{t-s}} = \mathsf{E}e^{iu^\top X_{t-s}}$$

This exactly finishes the stationary independence.

(c) Similarly, we just splite the event, $\forall\, \varepsilon > 0, \forall\, t \geq 0$:

$$\begin{aligned}\mathsf{P}(|X_t| > \varepsilon) &= \mathsf{P}(|X_t - X^{(n)}_t + X^{(n)}_t| > \varepsilon) \leq \mathsf{P}(|X_t - X^{(n)}_t| + |X^{(n)}_t| > \varepsilon) \\ &\leq \mathsf{P}(|X_t - X^{(n)}_t| > \varepsilon/2) + \mathsf{P}(|X^{(n)}_t| > \varepsilon/2)\end{aligned}$$

Using condition with stochastic continuity for limiting process to complete.

As showed, it satisfies the properties then we completed the proof. □

From above, we could see the importance of stochastic continuity, then following is another to depict this property, in the sense of probability measure.

Proposition 6.12. *If $X = \{X_t : t \geq 0\}$ is a stochastic process, such that $X_0 = 0$ a.s. Let P_t be the distribution of X_t. Then X is stochastically continous at $t = 0$ iff P_t is weakly convergent to δ_0 as $t \to 0^+$.*

Proof. To recall, if we want to show it is weakly convergent, it suffice to show:

$$\forall\, f \in C_B(\mathbb{R}^d), \lim_{t \to 0^+} \int_{\mathbb{R}^d} f(y)\mathsf{P}(\mathrm{d}y) = \int_{\mathbb{R}^d} f(y)\delta_0(\mathrm{d}y) = f(0)$$

(1) Let's first see forward direction, we control by (i) continuity of f:

$$\forall\, \varepsilon > 0, \exists\, \delta > 0, \text{ s.t. } \forall\, |x| < \delta, |f(x) - f(0)| < \varepsilon$$

(ii) the stochastic continuity, $\exists\, \delta' > 0$, s.t. $\forall t \in (0, \delta')$:

$$\int_{B_\delta(0)^c} \mathsf{P}_t(\mathrm{d}x) = \mathsf{P}(|X_t| > \delta) < \varepsilon$$

Now, same trick as before, we splite the event using this δ distance, $\forall\, t \in (0, \delta')$:

$$\begin{aligned}\left|\int_{\mathbb{R}^d} f(x)\mathsf{P}_t(\mathrm{d}x) - f(0)\right| &= \left|\int_{\mathbb{R}^d} (f(x) - f(0))\mathsf{P}_t(\mathrm{d}x)\right| \leq \int_{\mathbb{R}^d} |f(x) - f(0)|\, \mathsf{P}_t(\mathrm{d}x) \\ &= \int_{B_\delta(0)} |f(x) - f(0)|\, \mathsf{P}_t(\mathrm{d}x) + \int_{B_\delta(0)^c} |f(x) - f(0)|\, \mathsf{P}_t(\mathrm{d}x) \\ &\leq \varepsilon \int_{B_\delta(0)} \mathsf{P}_t(\mathrm{d}x) + 2\|f\|_{L^\infty} \int_{B_\delta(0)^c} \mathsf{P}_t(\mathrm{d}x) \leq (1 + 2\|f\|_{L^\infty})\varepsilon\end{aligned}$$

Since we strict to bounded space, then this is control by ε.

(2) $\forall\, \varepsilon > 0, r > 0$, we choose g supported on $B_r(0)$ s.t.$0 \le g \le 1, g(0) > 1-\varepsilon$, then:

$$\exists\, \delta > 0, \text{ s.t. } \left| \int_{\mathbb{R}^d} g(x)\mathsf{P}_t(\mathrm{d}x) - g(0) \right| < \varepsilon, \forall\, t \in (0,\delta)$$

Next, we could investigate, $\forall\, t \in (0,\delta)$:

$$\mathsf{P}(|X_t| > r) = 1 - \int_{B_r(0)} \mathsf{P}_t(\mathrm{d}x) \le 1 - \int_{B_r(0)} g(x)\mathsf{P}_t(\mathrm{d}x) \le 1 - (g(0) - \varepsilon) < 2\varepsilon$$

Then we completed the proof from both directions. □

Remark 6.13. Stochastically continous can be described by measure.

Then we naturally talk about some interesting stuff about measure.

Definition 6.9. A family of probability measure $\mathsf{P}_t, t \ge 0$ with $\mathsf{P}_0 = \delta_0$ is called **convolution semigroup** if $\mathsf{P}_{s+t} = \mathsf{P}_s * \mathsf{P}_t\ \forall\, s,t \ge 0$.

The semigroup is weakly continous if P_t is weakly convergent to δ_0 as $t \to 0^+$.

Remark 6.14. A convolution semigroup $\{\mathsf{P}_t, t \ge 0\}$ is weakly continous iff

$$\lim_{s \to t^+} \int_{\mathbb{R}^d} f(y)\mathsf{P}_s(\mathrm{d}y) = \int_{\mathbb{R}^d} f(y)\mathsf{P}_t(\mathrm{d}y), \forall\, f \in C_B(\mathbb{R}^d), \forall\, t \ge 0.$$

And the follwong corollary states out the connection with Lévy process:

Corollary 6.13. *If $X = \{X_t, t \ge 0\}$ is a Lévy process, P_t is the distribution of $X_t, \forall\, t \ge 0$, then $\{\mathsf{P}_t, t \ge 0\}$ is weakly continous convolution semigroup.*

Proof. Here we only present rough idea of proof:

(1) Stationary of Increment $\implies$ P_t is the distribution of increment.

(2) Independence of Increment $\implies$ Convolution.

(3) Stochastic Continuity $\implies$ Weak continuity of P_t.

We only present the most important three properties' justification. □

Then naturally, we want to ask how to reverse this process?

Theorem 6.14. *If $\{\mathsf{P}_t, t \ge 0\}$ is a weakly continous convolution semigroup of probability measures, then there exists a Lévy process $X = \{X_t, t \ge 0\}$ such that P_t is the distribution of $X_t, \forall\, t \ge 0$.*

Proof. We only give the rough idea of proof:

(1) Set $\Omega = \{\omega : \mathbb{R}^+ \mapsto \mathbb{R}^d \text{ with } \omega(0) = 0\}$. We consider sylinder sets of the form:

$$I_{t_1,\cdots,t_n}^{A_1,\cdots,A_n} = \{\omega \in \Omega : w(t_1) \in A_1, \cdots, \omega(t_n) \in A_n\}$$

, $\forall\ A_1, \cdots, A_n \in \mathcal{B}(\mathbb{R}^d), 0 \le t_1 < \cdots < t_n, \forall\ n \in N$, and set $\mathcal{F}$ the σ-algebra generated by such cylinder sets. We define P on $\mathcal{F}$ by:

$$\mathsf{P}\left(I_{t_1,\cdots,t_n}^{A_1,\cdots,A_n}\right) = \int_{(\mathbb{R}^d)^n} \mathbf{1}_{A_1}(y_1)\cdots \mathbf{1}_{A_n}(y_1 + \ldots y_n)\mathsf{P}_{t_1}(\mathrm{d}y_1)\cdots \mathsf{P}_{t_n - t_{n-1}}(\mathrm{d}y_n)$$

(2) We define $X_t(\Omega) = \omega(t)$ for $t \geq 0, \omega \in \Omega$. Then the distribution of X_t is P_t. Moreover, $X_0 = 0$ and X_t is stochastically continuous by 6.12. The stationary and independence of the increments are deduced via charateristic function.

And the details of the proof is left for interested readers. □

Remark 6.15. The process X above constructed is called the cannonical Lévy process. Reversely, one can recover Lévy by an infinitely divisible probability measure.

Corollary 6.15. *If μ is an infinitely divisible probability measure on $\mathbb{R}^d$ with Lévy symbol η, then there exists a Lévy process X such that the distribution of X_1 is μ.*

Proof. Firstly, from 6.2, there exists a triplet (b, A, ν) such that: $\phi_\mu(u) = e^{\eta(u)}$, besides let $\phi_{\mu,t}(u) = e^{t\eta(u)}$ be the characteristic function of an infinitely divisible probability measure of P_t. Then we only need to check it is weakly convergent to δ_0, to see this: $\phi_{\mu,0}(u) = 1 = \phi_{\delta_0}(u)$, then using convolution property, we completed the continuity. □

6.3 Subordinator

Next, we introduce a group of special process under Lévy.

Definition 6.10. A subordinator is an 1-dimensional Lévy process and is increasing a.s.

Following theorem depicts subordinator in "Lévy" symbol:

Theorem 6.16. *If T is a subordinator, then its Lévy symbol has the form:*

$$\eta_T(u) = ibu + \int_0^\infty \left(e^{iuy} - 1\right) \nu(\mathrm{d}y)$$

, where $b \geq 0$, and ν is Lévy measure satisfying:

(a) ν *supported on* $[0, \infty)$.

(b) $\int_0^\infty y \wedge 1 \nu(\mathrm{d}y) < \infty$.

Proof. The proof is omitted, here we just present this theorem. □

Remark 6.16. 1. The pair (b, ν) is called the characteristic of T.

2. The Laplace transform of T is $\mathsf{E}e^{-uT_t} = e^{-t\psi(u)}$, where $\psi(u) = -\eta_T(iu)$.

Even though it seems like a new concepts, the following examples are all subordinator:

Example 6.2. The following two examples' verification is left as homework.

1. A poisson process is always a subordinator.
2. A compound poisson process is a subordinator iff $Z_n \geq 0 \ \forall\, n \in \mathbb{N}\ a.s.$

And we have similar result in recovering subordinator from probability measure:

Theorem 6.17. *If μ is an infinitely divisible probability measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, then there exists a subordinator $T = \{T_t : t \geq 0\}$, s.t. μ is the distribution of T_1.*

Proof. From 6.15, there exists a Lévy process, X, s.t. μ is the distribution of X_1. It remains to show it is increasing a.s., key is the measure is on $\mathbb{R}_+$, then:

$$X_1 \geq 0 \text{ a.s.} \overset{\text{inf div}}{\Longrightarrow} X_{1/n} \geq 0 \text{ a.s. } \forall\, n \in \mathbb{N} \implies X_t \geq 0 \text{ a.s. } \forall\, t \in \mathbb{Q}$$

By stochastic continuity, we have convergence in probability, then using subsequence that is convergence almost surely, therefore, to complete the irrational points on real line. $\square$

The aim to introduce subordinator is its increasing, then its kind of "time":

Theorem 6.18. ***(Time Changing)***
Let X be a Lévy process, T be a subordinator on $(\Omega, \mathcal{F}, \mathsf{P})$ such that X and T are independent. Set $Z_t = X_{T_t}$. Then process $Z = \{Z_t : t \geq 0\}$ is a Lévy process.

Proof. Similarly, just need to check conditions one by one:

(1) $Z_0 = X_{T_0} = X_0 = 0$ a.s., as X and T are Lévy processes.

(2) Indepent increment uses characteristic function, we omit here, then for proving stationary increment of Z, considering $\forall\, 0 \leq t_1 < t_2, \forall A \in \mathcal{B}(\mathbb{R}^d)$, we have:

$$\begin{aligned}
&\mathsf{P}(Z_{t_2} - Z_{t_1} \in A) = \mathsf{P}(X_{T_{t_2}} - X_{T_{t_1}} \in A) = \mathsf{E}\mathbf{1}_{\{X_{T_{t_2}} - X_{T_{t_1}} \in A\}} \\
&= \mathsf{E}\left\{\mathsf{E}\left[\mathbf{1}_{\{X_{T_{t_2}} - X_{T_{t_1}} \in A\}} \mid T_{t_1}, T_{t_2}\right]\right\} = \mathsf{E}\left\{\mathsf{E}\left[\mathbf{1}_{\{X_{T_{t_2} - T_{t_1}} \in A\}} \mid T_{t_1}, T_{t_2}\right]\right\} \\
&= \mathsf{E}\mathbf{1}_{\{X_{T_{t_2} - T_{t_1}} \in A\}} = \mathsf{E}\left\{\mathsf{E}\left[\mathbf{1}_{\{X_{T_{t_2} - T_{t_1}} \in A\}} \mid X\right]\right\} = \mathsf{E}\left\{\mathsf{E}\left[\mathbf{1}_{\{X_{T_{t_2 - t_1}} \in A\}} \mid X\right]\right\} \\
&= \mathsf{E}\mathbf{1}_{\{X_{T_{t_2 - t_1}} \in A\}} = \mathsf{P}(Z_{t_2 - t_1} \in A)
\end{aligned}$$

, where we use independence of X and T, and stationary of X.

(3) By the stochastical continuity of X and T, $\forall\, \epsilon > 0, \eta > 0$, $\exists\, \delta > 0$, s.t. $\forall\, h \in (0, \delta)$, it holds that $\mathsf{P}(|X_h| > \eta) < \epsilon$, and $\exists\, \delta' > 0$, s.t. $\forall\, h' \in (0, \delta')$, it holds that $\mathsf{P}(T_{h'} \geq \delta) < \epsilon$. Then $\forall h \in (0, \min\{\delta, \delta'\})$:

$$\begin{aligned}
\mathsf{P}(|Z_h| > \eta) = \mathsf{P}(|X_{T_h}| > \eta) &= \mathsf{P}(|X_{T_h}| > \eta, T_h \geq \delta) + \mathsf{P}(|X_{T_h}| > \eta, T_h < \delta) \\
&\leq \mathsf{P}(T_h \geq \delta) + \sup_{u \in [0,\delta)} \mathsf{P}(|X_u| > \eta) < 2\epsilon.
\end{aligned}$$

Therefore, we completed the proof. $\square$

Corollary 6.19. $\forall\, u \in \mathbb{R}^d, \eta_Z(u) = -\psi_T(-\eta_X(u))$.

Proof. By independence of X and T, we know:

$$e^{t\eta_Z(u)} = \mathsf{E}e^{iu^\top X_{T_t}} = \mathsf{E}\left[\mathsf{E}\left(e^{iu^\top X_{T_t}} \mid T_t\right)\right] = \mathsf{E}e^{T_t \eta_X(u)} = e^{-t\psi(-\eta_X(u))}$$

Then we finished the proof. $\square$

6.4 Recurrence and Transience

Finishing the ordinator, let's introduce another concepts, recurrent and transient:

Definition 6.11. A Lévy process $X = \{X_t : t \geq 0\}$ is:

(a) recurrent (at the origin) if $\liminf_{t\to\infty} |X_t| = 0$ *a.s.*

(b) transient if $\lim_{t\to\infty} |X_t| = \infty$ *a.s.*

From definition, one can see it is not easy to show whether a Lévy process is recurrent or transient, then the following criterion helps us to understand the concept:

Theorem 6.20. *A Lévy process $X = \{X_t : t \geq 0\}$ is recurrent iff:*

(a) For $a > 0$, $\lim_{q\to 0^+} \int_{B_a(0)} \Re\left(\frac{1}{q-\eta(u)}\right) \mathrm{d}u = \infty$. (Chung-Fuchs criterion)

(b) $\forall\, a > 0$, $\int_{B_a(0)} \Re\left(\frac{1}{-\eta(u)}\right) \mathrm{d}u = \infty$.(Spitzer criterion)

Proof. The proof is omitted since its complexity. $\square$

Theorem 6.21. *On $(\Omega, \mathcal{F}, \mathsf{P})$, $X = \{X_t : t \geq 0\}$ is a Lévy process with Lévy symbol η. $\forall\, u \in \mathbb{R}^d$, we set:*

$$M_t^u = \exp\left(iu^\top X_t - t\eta(u)\right), \forall\, t \geq 0$$

Fix $u \in \mathbb{R}^d$, $M = \{M_t : t \geq 0\}$ is a ***complex martingale*** *with respect to $\mathcal{F}^X$.*

Proof. (1) Integrability: $\mathsf{E}|M_t^u| = \mathsf{E}e^{-t\Re(\eta(u))} < \infty, \forall\, u \in \mathbb{R}^d, \forall\, t \geq 0$.

(2) $\forall\, u \in \mathbb{R}^d, \forall\, 0 \leq s \leq t$, notice we could rewrite:

$$M_t^u = M_s^u \cdot \exp\left(iu^\top (X_t - X_s) - (t-s)\eta(u)\right)$$

Then, by independence and stationary increment of X, we have:

$$\begin{aligned}\mathsf{E}\left(M_t^u \mid \mathcal{F}_s^X\right) &= M_s^u \cdot \mathsf{E}\left[\exp\left(iu^\top (X_t - X_s) - (t-s)\eta(u)\right) \mid \mathcal{F}_s^X\right] \\ &= M_s^u \cdot \mathsf{E}\exp(iu^\top X_{t-s} - (t-s)\eta(u)) = M_s^u\end{aligned}$$

Then we completed the proof. $\square$

Theorem 6.22. *On $(\Omega, \mathcal{F}, \mathsf{P})$, every Lévy process, $X = \{X_t : t \geq 0\}$ has a modification, Y, that is RCLL, and also a Lévy process.*

Proof. From 6.21, $\forall\, u \in \mathbb{R}^d, M^u$ is a martingale so $\forall\, t \geq 0$:

$$M_{t-}^u = \lim_{\substack{s\to t- \\ s\in\mathbb{Q}}} M_s^u \text{ and } M_{t+}^u = \lim_{\substack{s\to t+ \\ s\in\mathbb{Q}}} M_s^u \quad \exists \text{ a.s.}$$

Now denote $\Omega_u' = \{\omega : \text{at least one of the limits fail to exist}\}$, then $\mathsf{P}(\Omega_u') = 0$, also:

$$\Omega' := \bigcup_{u\in\mathbb{Q}^d} \Omega_u' \implies \mathsf{P}(\Omega') = 0$$

Then we are going to work on Ω', $\forall\, \omega \in \Omega'^c, \forall\, t \geq 0$, choose increasing sequence $\{s_n : n \in \mathbb{N}\} \subseteq \mathbb{Q}^+$, such that, $\lim_{n\to\infty} s_n = t$.We want to show that the limit is unique on this set (since by construction is already exist), then assume $\{X_{s_n}(\omega) : n \in \mathbb{N}\}$ has accumulation points $X^{(1)} := \lim_{i\to\infty} X_{s_{n_i}}$ and $X^{(2)} := \lim_{j\to\infty} X_{s_{n_j}}$. Then the existence of limits give:

$$\lim_{s\to t} M_s^u \,\exists \implies \lim_{n\to\infty} \exp(iu^\top X_{s_n})\,\exists \implies \frac{1}{2\pi} u^\top (X^{(1)} - X^{(2)}) \in \mathbb{N}, \forall\, u \in \mathbb{Q}^d \text{ a.s.}$$

This forces $X^{(1)} = X^{(2)}$ a.s. Then X has left (right) limit on $\mathbb{Q}^+$. Now define:

$$Y_t(\omega) = \begin{cases} \lim_{\substack{s\to t+ \\ s\in\mathbb{Q}^+}} X_t(\omega) & ,\omega \notin \Omega' \\ 0 & ,\omega \in \Omega' \end{cases}$$

One could verify Y is a mofication of X and also a RCLL Lévy process. $\square$

Following theorem shows a new way to construct Lévy process:

Theorem 6.23. *On $(\Omega, \mathcal{F}, \mathsf{P})$, $X = \{X_t : t \geq 0\}$ be a RCLL Lévy process adapted to a filtration $\{\mathcal{F}_t : t \geq 0\}$. Let T be a bounded stopping time relative to $\{\mathcal{F}_t : t \geq 0\}$. Set: $X^{(T)} = X_{T+t} - X_T, \forall\, T \geq 0$, then:*

(a) $X^{(T)}$ is a Lévy process and is independent of $\mathcal{F}_T$.

(b) $\forall\, t \geq 0, X_t^{(T)} \overset{d.}{=} X_t$ pointwisely.

(c) $X^{(T)}$ is RCLL and is adapted to $\{\mathcal{F}_{T+t} : t \geq 0\}$.

Omitted since just similar to before.

Proof. (a) For $\forall A \in \mathcal{F}_T$, $\forall\, u_j \in \mathbb{R}^d$ with $1 \leq j \leq n$, $\forall 0 = t_0 \leq t_1 < \cdots < t_n$, we have:

$$\mathsf{E}\left[\mathbf{1}_A \exp\left(i \sum_{j=1}^n u_j^\top (X_{T+t_j} - X_{T+t_{j-1}})\right)\right] = \mathbb{E}\left[\mathbf{1}_A \prod_{j=1}^n \frac{M_{T+t_j}^{u_j}}{M_{T+t_{j-1}}^{u_j}} \cdot \prod_{j=1}^n \phi_{t_j - t_{j-1}}(u_j)\right]$$

, where $\phi_t(u) = \mathsf{E}e^{iu^\top X_t}$ and $M_t^u = \exp(iu^\top X_t - t\eta(u))$. Using:

$$\mathsf{E}\left[\mathbf{1}_A \frac{M_{T+b}^u}{M_{T+a}^u}\right] = \mathsf{E}\left[\mathbf{1}_A \frac{1}{M_{T+a}^u} \mathsf{E}\left(M_{T+b}^u \mid \mathcal{F}_{T+a}\right)\right] = \mathsf{P}(A)$$

, to take conditional expectation w.r.t. $\mathcal{F}_{T+t_{n-1}}, \mathcal{F}_{T+t_{n-2}}, \ldots, \mathcal{F}_{T+t_1}$ to obtain:

$$\mathsf{E}\left[\mathbf{1}_A \exp\left(i \sum_{j=1}^n u_j^\top (X_{T+t_j} - X_{T+t_{j-1}})\right)\right] = \mathsf{P}(A) \cdot \prod_{j=1}^n \phi_{t_j - t_{j-1}}(u_j)$$

Then, take $A = \Omega, n = 1$ to get $\mathsf{E}e^{iu^\top (X_{T+t} - X_T)} = \mathsf{E}e^{iu^\top X_t}$.

(b) We have $X_0^T = 0$. Take $A = \Omega$ to obtain the stationary and independence of increments. The stochastically continuity follows directly from the stochastically continuity of X. By change of variable $v_n = u_n, v_k = u_k - u_{k+1}, 1 \leq k \leq n-1$:

$$\mathsf{E}\left[\mathbf{1}_A \exp\left(i \sum_{j=1}^n v_j^\top X_{t_j}^T\right)\right] = \mathsf{P}(A)\mathsf{E}\left[\exp\left(i \sum_{j=1}^n v_j^\top X_{t_j}^T\right)\right] = \mathsf{E}\left\{\mathbf{1}_A \mathsf{E}\left[\exp\left(i \sum_{j=1}^n v_j^\top X_{t_j}^T\right)\right]\right\}$$

As $A \in \mathcal{F}_T$ is arbitrary, we have

$$\mathsf{E}\left[\exp\left(i\sum_{j=1}^{n} v_j^\top X_{t_j}^T\right) \mid \mathcal{F}_T\right] = \mathsf{E}\left[\exp\left(i\sum_{j=1}^{n} v_j^\top X_{t_j}^T\right)\right]$$

We can deduce that $\exp(i\sum_{j=1}^{n} v_j^\top X_{t_j}^T)$ is independent of $\mathcal{F}_T$.

(c) Immediately comes from RCLL of X.

Finally, we completed the proof. □

6.5 Poisson Random Measure

Before we introduce poisson process, we first define jump process:

Definition 6.12. Let $X = \{X_t : t \geq 0\}$ be a RCLL Lévy process. Set:

$$\Delta X_t = X_t - X_{t-}, \ \forall \, t \geq 0$$

Then the process:

$$\Delta X = \{\Delta X_t : t \geq 0\}$$

is called the **jump process** of X.

Now, the poisson process is certain type of Lévy in the sense that:

Theorem 6.24. *Let N be a $\mathbb{N}$-valued Lévy process. If N is increasing a.s. and $\Delta N \in \{0, 1\}$. Then N is a **poisson process**.*

Proof. The difficulty is to recover the distribution, now we define a set of stopping times:

$$T_0 = 0, T_n = \inf\{t > T_{n-1} : N_t - N_{T_{n-1}} \neq 0\} = T_{n-1} + \inf\{t > 0 : N_{t+T_{n-1}} - N_{T_{n-1}} \neq 0\}$$

This is the time that the process jumps, from construction, $\{T_n - T_{n-1}\}$ are i.i.d. Then:

$$\begin{aligned}\mathsf{P}(T_1 > s + t) &= \mathsf{P}(N_s = 0, N_{t+s} - N_s = 0) = \mathsf{P}(N_s = 0)\mathsf{P}(N_{t+s} - N_s = 0) \\ &= \mathsf{P}(N_s = 0)\mathsf{P}(N_t = 0) = \mathsf{P}(T_1 > s)\mathsf{P}(T_1 > t)\end{aligned}$$

, $\forall \, t, t \geq 0$, then similar to 6.8, $\mathsf{P}(T_1 > t) = \mathsf{P}(N_t = 0) = \mathsf{P}(N_1 = 0)^t$. Since we want the probability finite, $\mathsf{P}(N_1 = 0) \in (0, 1)$, then we denote as $\mathsf{P}(N_1 = 0) = e^{-\lambda}, \lambda > 0$. Then:

$$\mathsf{P}(T_1 > t) = e^{-\lambda t} \implies \mathsf{P}(N_t = n) = \mathsf{P}(T_{n+1} > t, T_n \leq t) = \mathsf{P}(T_{n+1} > t) - \mathsf{P}(T_n > t)$$

Lastly, using the gamma distribution to recover the distribution (do computation). □

More interesting, if we fix at any t, the process will not jump:

Lemma 6.25. *If X is a Lévy process, then fixed any $t > 0$, then $\Delta X_t = 0$ a.s.*

Proof. WLOG, assume X is RCLL, (or use its modification), then there exists:

$$\{t_n \in \mathbb{R}^+ : n \in \mathbb{N}\} \text{ with } t_n \uparrow t \text{ s.t. } X_{t_n} \xrightarrow{p.} X_t$$

This gives us a subsequence $Y_{t_{n_k}}$ that converges to Y_t a.s. then $Y_{t-} = Y_t$ a.s. □

Remark 6.17. 1. This doesn't mean Lévy process don't have jump. Since we fixed t. If we want to study the jump: $\mathsf{P}(\exists\, t, \Delta X_t \neq 0) \approx \cup_t \mathsf{P}(\Delta X_t \neq 0)$, the summation of inifnite zeros may result in nontrivial.

2. It is possible that, $\sum_{0 \le s \le t} |\Delta X_s| = \infty, a.s.$

3. Moreover, for stopping time, $\Delta X_T = 0$ a.s is not guaranteed.

Now, based on jump process, we introduce random measure:

Definition 6.13. Let X be a RCLL Lévy process in $\mathbb{R}^d$. $\forall\, t \ge 0, \forall\, A \in \mathcal{B}(\mathbb{R}^d/\{0\})$, we define:

$$N(t, A) = |\{0 \le s \le t : \Delta X_s(\omega) \in A\}|$$

Then $\mathsf{E}N(t, \cdot)$ is a Borel measure on $\mathcal{B}(\mathbb{R}^d/\{0\})$.

In particular, we call $\mu(\cdot) = \mathsf{E}N(1, \cdot)$ the **intensity measure** of X.

Remark 6.18. 1. The function of N is to study the jumps of X.

2. If $A \in \mathcal{B}(\mathbb{R}^d/\{0\})$ is bounded away from 0, i.e. $0 \notin \overline{A}$, then $\forall\, t \ge 0$, $N(t, A) < +\infty, a.s.$ The proof can be seen from textbook lemma 2.3.4.

Then, using this definition, we could extend our result:

Theorem 6.26. *For any $A \in \mathcal{B}(\mathbb{R}^d/\{0\})$ that is bounded away from 0, then:*

$$\{N(t, A) : t \ge 0\}$$

*is a **poisson process** with intensity $\mu(A)$.*

Proof. Firstly, notice $N(\cdot, A)$ is increasing for every A, also $\Delta N(\cdot, A) \in \{0, 1\}$, follow from 6.24, we only need to show $\{N(t, A) : t \ge 0\}$ is a Lévy process:

(1) Since $N(0, A) = \mathbf{1}_{\{X_0 \in A\}}$, and A is bounded away from 0, then $N(0, A) = 0$.

(2) Note that $N(t, A) - N(s, A) \ge n$ iff:

$$\exists\, s < t_1 < \cdots < t_n < t, \text{ s.t. } \Delta X_{t_j} \in A, \forall\, 1 \le j \le n$$

This implies $N(t, A) - N(s, A) \in \sigma(X_u - X_v : s < v < u < t)$. Thus its independence of increments follow from X. Similarly, we have stationary of increments.

(3) Lastly, by independence and stationary of increments:

$$\begin{aligned}
\mathsf{P}(N(t, A) = 0) &= \mathsf{P}\left(N\left(\frac{kt}{n}, A\right) - N\left(\frac{(k-1)t}{n}\right) = 0, \forall\, 1 \le k \le n\right) \\
&= \prod_{k=1}^{n} \mathsf{P}\left(N\left(\frac{kt}{n}, A\right) - N\left(\frac{(k-1)t}{n}\right) = 0\right) \\
&= \mathsf{P}\left(N\left(\frac{t}{n}, A\right) = 0\right)^n, \forall\, n \in \mathbb{N},\ \forall\, t > 0
\end{aligned}$$

For any fixed t, set $n \to \infty$ to obtain $\lim_{n\to\infty} \mathsf{P}(N(t/n, A) = 0) = 0$ or 1. However, if $\forall\, t > 0, \mathsf{P}(N(t, A) = 0) = 0$, then $N(t, A) > 0$ a.s., which contradicts right-continuity. Then we recover the stochastic continuity at 0.

Therefore, we completed the proof, following 6.24. $\square$

Remark 6.19. If $A_1, \cdots, A_m$ are disjoint in $\mathcal{B}(\mathbb{R}^d/\{0\})$ and $s_1, \cdots, s_m$ are disjoint, then $N(s_1, A_1), \cdots, N(s_m, A_m)$ are independent.

After all the preparation, we come to random measure:

Definition 6.14. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space, S be a set. And $\mathcal{A}$ is an algebra of subsets of S. A **random measure**, $\mathfrak{M}$, on $(S, \mathcal{A})$ is a collection of r.v.s $\{\mathfrak{M}(B) : B \in \mathcal{A}\}$ if:

(a) $\mathfrak{M}(\emptyset) = 0$.

(b) $\mathfrak{M}(A \cup B) = \mathfrak{M}(A) + \mathfrak{M}(B)$ if $A \cap B = \emptyset$.

Remark 6.20. A random measure is σ-addictivity if (b) is replaced by closed by countable infinitely addictivity.

Since its related to r.v., then we also need independence:

Definition 6.15. A random measure $\mathfrak{M}$ is independently scattered if for disjoint family $B_1, \cdots, B_n$ in $\mathcal{A}$, the r.v.s $\mathfrak{M}(B_1), \cdots, \mathfrak{M}(B_n)$ are independent.

More special, we are concerning with poisson random measure:

Definition 6.16. Let S be a set, $\mathcal{S}$ be a σ-algebra of subsets of S, $\mathcal{A} \subseteq \mathcal{S}$ be an algebra. An independent scattered, σ-finite random measure $\mathfrak{M}$ on $(S, \mathcal{S})$ is called **poisson random measure** if:

$$\forall\ B \in \mathcal{A}, \mathfrak{M}(B) < +\infty, \text{ and } \mathfrak{M}(B) \sim \text{Poi}(\cdot)$$

Remark 6.21. $\lambda(B) = \mathsf{E}\mathfrak{M}(B)$ on $B \in \mathcal{A}$ can be extended to $(S, \mathcal{S})$ to a σ-finite measure.

And then following theorem gives us why poisson random measure is important:

Theorem 6.27. *Let S be a set, $\mathcal{S}$ be a σ-algebra, λ be a σ-finite measure on $(S, \mathcal{S})$. Then there exists a poisson random measure $\mathfrak{M}$ on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ s.t:*

$$\forall\ A \in \mathcal{S},\ \lambda(A) = \mathsf{E}\mathfrak{M}(A)$$

In this case $\mathcal{A} = \{A \in \mathcal{S} : \lambda(A) < +\infty\}$.

Proof. The proof is intentionally omitted. $\square$

Remark 6.22.

1. X be a RCLL Lévy process, $S = \mathbb{R}^d/\{0\}$, $\mathcal{S} = \mathcal{B}(S)$, $\mathcal{A} =$ algebra generated by subsets of S that is bounded away from 0. Then $\forall\ t \geq 0, \mathfrak{M}_t(A) = N(t, A)$ is a poisson measure with $\lambda(\cdot) = t\mu(\cdot)$, $(\mu(A) = \mathsf{E}N(1, A))$.

2. Set $\mathfrak{M}([s, t] \times A) = \mathfrak{M}_t(A) - \mathfrak{M}_s(A), \forall\ s \leq t$, then $\mathfrak{M}$ can be extended to a measure on $(\mathbb{R}_+ \times (\mathbb{R}^d/\{0\}), \mathcal{B}(\mathbb{R}_+ \times (\mathbb{R}^d/\{0\})))$. (Also a σ-additivity poisson random measure)

3. Note the following relation: $\lambda(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}t\mu(\mathrm{d}x)$.

4. $\tilde{N}(t, A) = N(t, A) - t\mu(A)$ is called compensated poisson random measure.

6.6 Poisson Integral

After defining measure, the natural extension is to check integral:

Definition 6.17. Let X be a RCLL Lévy process, N be the associated poisson random measure. For Borel measurable mapping $f : \mathbb{R}^d \mapsto \mathbb{R}^d$, and A that is bounded away from 0. Then we define, the **poisson integral**:

$$\int_A f(x)N(t, \mathrm{d}x) = \sum_{x\in A} f(x)N(t, \{x\})$$

, provided that the summation is finite a.s.

Then the difference between the normal integral is that N has poisson distribution.

Remark 6.23. Quick recall:

$$N(t, A) = |\{0 \le s \le t : \Delta X_s(\omega) \in A\}| = \sum_{0\le s\le t} \mathbf{1}_{\{X_s(\omega)\in A\}}$$

Therefore, we are surely curious about how to compute it?

Proposition 6.28. *Let X be RCLL Lévy process, N be the associated poisson random measure, A is bounded away from 0. We define a measure, $\forall\, B \in \mathcal{B}(\mathbb{R}^d)$:*

$$\mu_{t,A}(B) = \mu\left(A \cap f^{-1}(B)\right)$$

Therefore, we have following conclusions:

(a) $\forall\, t \ge 0$, *if* $\int_A |f|\mu(\mathrm{d}x) < \infty$, *then:*

$$\int_A f(x)N(t, \mathrm{d}x)$$

has compounded poisson distribution with characteristic function:

$$\mathsf{E}\exp\left(iu^\top \int_A f(x)N(t, \mathrm{d}x)\right) = \exp\left(t\int_{\mathbb{R}^d} (e^{iu^\top x} - 1)\mu_{t,A}(\mathrm{d}x)\right)$$

(b) $\forall\, t \ge 0$, *if* $\int_A |f|\mu(\mathrm{d}x) < \infty$, *then:*

$$\mathsf{E}\int_A f(x)N(t, \mathrm{d}x) = t\cdot\int_A f(x)\mu(\mathrm{d}x)$$

(c) $\forall\, t \ge 0$, *if* $\int_A |f|\mu(\mathrm{d}x) < \infty, \int_A |f|^2\mu(\mathrm{d}x) < \infty$, *then:*

$$\mathrm{Var}\left(\int_A f(x)N(t, \mathrm{d}x)\right) = t\cdot\int_A |f(x)|^2\mu(\mathrm{d}x)$$

Proof. Notice once (a) proved, (b)-(c) can be derived by taking derivative. Then we focus on (a), consider case $f = \sum_{j=1}^n c_j \mathbf{1}_{A_j}$ with disjoint A_j. Using independently scattered:

$$\begin{aligned}
&\mathsf{E}\exp\left(iu^\top \int_A f(x)N(t,\mathrm{d}x)\right) = \mathsf{E}\exp\left(iu^\top \sum_{j=1}^n c_j N(t,A_j)\right) \\
&= \prod_{j=1}^n \mathsf{E}\exp\left(iu^\top c_j N(t,A_j)\right) = \prod_{j=1}^n \exp\left[t\left(e^{iu^\top c_j}-1\right)\mu(A_j)\right] \\
&= \exp\left[t\int_A \left(e^{iu^\top f(x)}-1\right)\mu(\mathrm{d}x)\right]
\end{aligned}$$

Then for $f \in L^1(A, \mu|_A)$ and $f \geq 0$, one can approximate using simple functions. And for general function, we split into $f = f^+ + f^-$, therefore we completed the proof. $\square$

The reason why we introduce these tools, is to better understand Lévy process:

Theorem 6.29. ***(Lévy-Itô Decomposition)***
If X is a d-dimensional Lévy process, then there exists, $b \in \mathbb{R}^d$, a d-dimensional Brownian motion, $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ of some filtration $\{\mathcal{F}_t : t \geq 0\}$, a $d\times d$ matrix Q, and an independent poisson random measure N on $\mathbb{R}^+ \times \mathbb{R}^+/\{0\}$, s.t. $\forall\, t \geq 0$:

$$X_t = bt + Q\mathsf{B}_t + \int_{|x|<1} x\tilde{N}(t,\mathrm{d}x) + \int_{|x|>1} xN(t,\mathrm{d}x)$$

Proof. Omitted since its complexity. $\square$

Remark 6.24. $Q\mathsf{B}_t$ is a Brownian motion with covariance matrix $QQ^\top$.

6.7 Stable Random Variable

Lastly, we introduce one more concept related to poisson distribution:

Definition 6.18. A r.v. X is **stable** if there exists a positive sequence $\{c_n : n \in \mathbb{N}\}$ and a sequence $\{d_n : n \in \mathbb{N}\}$ such that, $\forall\, n \in \mathbb{N}$:

$$\sum_{k=1}^n X_k \overset{d.}{=} c_n X + d_n$$

, where $X_1, \cdots, X_n$ are i.i.d. copies of X.

Moreover, X is said to be strictly stable if $d_n = 0, \forall\, n \in \mathbb{N}$.

Remark 6.25.
1. If exists, then $c_n = \sigma n^{1/\alpha}$ for some $\alpha \in (0,2]$, $\sigma > 0$, here α is called the **index of stability**.
2. For normal r.v., we know: $c_n \approx \sigma\sqrt{n}$, i.e. $(\alpha = 2)$.
3. Any stable r.v. are always infinitely divisible.

Surprisingly, an infinitely divisible r.v. can always be characterised:

Theorem 6.30. *Let X be a real-valued stable r.v. with characteristic triplet (a, A, ν) then one of followings holds:*

(a) If $\alpha = 2, \nu = 0$, then $X \sim \mathcal{N}(a, A)$.

(b) If $\alpha \in (0, 2), A = 0$, then:

$$\nu(\mathrm{d}x) = c_1 x^{-1-\alpha} \mathbf{1}_{(0,+\infty)}(x)\mathrm{d}x + c_2 |x|^{-1-\alpha} \mathbf{1}_{(-\infty,0)}(x)\mathrm{d}x$$

, with $c_1, c_2 \geq 0, c_1 + c_2 > 0$.

Proof. Omitted. □

Remark 6.26. In second case, we have $\mathsf{P}(|X| > y) \sim y^{-\alpha}$ as $y \to +\infty$. Moreover:

$$\mathsf{E}|X|^p = \begin{cases} < +\infty & , p < \alpha \\ = -\infty & , p \geq \alpha \end{cases}$$

One can also add stable r.v. into Lévy process:

Definition 6.19. A stable Lévy process (sometimes called stable process) is a Lévy process, $X = \{X_t : t \geq 0\}$ such that $\forall\, t \geq 0$, X_t is a stable r.v.

Remark 6.27. In some literature stable Lévy process is assumed to be **rotationally invariant**. In this case, $\eta(u) = -\sigma^\alpha |u|^\alpha$ with $\alpha \in (0, 2]$ and $\sigma > 0$. Moreover, if X is strictly stable, then X has self-similarity in the sense that $X_{at} \overset{d.}{=} a^{1/\alpha} X_t, \forall\, t \geq 0$.

6.8 Homeworks

The followings are exercises for this section: Note that all vectors are column vectors.

Problem 6.1. Let $X = X_1 + X_2$, where X_1 and X_2 are independent r.v.s taking values in $\mathbb{R}^d$ with $X_1 \sim \mathcal{N}(a, \Sigma)$ and $X_2 \sim \text{Poisson}(\lambda, \mu)$. Prove that the characteristic function of X is:

$$\phi_X(u) = \exp\left(i a^T u - \frac{1}{2} u^T \Sigma u + \int_{\mathbb{R}^d} \lambda \left(e^{i u^T y} - 1 \right) \mu(dy) \right),$$

where the integral is over $\mathbb{R}^d$ with respect to the measure μ.

Problem 6.2. Prove that every Lévy measure is σ-finite.

Problem 6.3. Let X be a Lévy process with the characteristic triple (a, A, ν). Prove that:

1. the measure $\tilde{\nu}$ given by $\tilde{\nu}(A) = \nu(-A)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ is also a Lévy measure;
2. $-X$ is also a Lévy process with the characteristic triple $(-a, A, \tilde{\nu})$.

Problem 6.4. For the standard d_2-dimensional Brownian motion $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$, determine the Lévy symbol and the characteristic triple of $X = \{at + P\mathsf{B}_t : t \geq 0\}$, where $a \in \mathbb{R}^{d_1}$ and P is a $d_1 \times d_2$ matrix.

Problem 6.5. Prove the following statements:

1. A Poisson process is always a subordinator.
2. A compound Poisson process is a subordinator if $Z_n \geq 0$ almost surely for all $n \in \mathbb{N}$.

Problem 6.6. If X is a compound Poisson process, prove that for any $t > 0$,

$$\sum_{0 \leq s \leq t} |\Delta X_s| < \infty, \quad \text{a.s.}$$

7 Brownian Sheets

Having journeyed from discrete to continuous time, and explored Markov and Lévy processes, we now extend our gaze beyond one-dimensional time. The Brownian sheet is the natural multi-dimensional analogue of Brownian motion, a Gaussian random field indexed by $\mathbb{R}_+^d$ whose covariance factorizes as $\prod \min(s_j, t_j)$. This chapter navigates the richer landscape of random fields, where linear order gives way to partial order, and familiar tools like filtrations and martingales adapt into new, multi-parameter forms. We will see how the Brownian sheet emerges as both a strong martingale and an integrated white noise, offering a bridge between the one-parameter theory and the spatial randomness that underlies fields like stochastic PDEs and spatial statistics.

7.1 Motivation

Recall that all we studied before is stochastic process, $X = \{X_t : t \geq 0\}$, so naturally the index, or time, should be nonnegative. However, what if we want to extend this into more general spaces? Say $\mathbb{R}^d$? Is there any differences? And also what kind of problems we have when defining this "process". This is actually called **random field**. Let's start with the easiest stochastic process, we have studied previously, the Brownian Motion:

Definition 7.1. Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration on $(\Omega, \mathcal{F}, \mathsf{P})$. Then the continous stochastic process $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ adapted to $\{\mathcal{F}_t : t \geq 0\}$ is called $d-$dimensional **Brownian Motion** if:

(1) $\mathsf{B}_0 = \mathbf{0}_d, a.s.$

(2) $\forall\, 0 \leq s \leq t, \mathsf{B}_t - \mathsf{B}_s \sim \mathcal{N}(\mathbf{0}_d, (t-s)\mathbf{I}_d)$ and independent of $\mathcal{F}_s$

But in this definition, there is a hugh problem if we extend into higher dimension: the comparision, how do we compare two time vectors in $\mathbb{R}^d$? Then check another candidates:

Definition 7.2. A continous process $\mathsf{B} = \{\mathsf{B}_t : t \geq 0\}$ adapted to filtration $\{\mathcal{F}_t : t \geq 0\}$ is a standard brownian motion if:

(a) B is a zero-mean/centered Gaussian process.

(b) $\mathsf{E}\mathsf{B}_t\mathsf{B}_s = \min\{s, t\}. \forall\, t, s \geq 0$

This only asks us to check the distribution information, and also the covariance, which are eaiser to be done in higher dimensions.

7.2 Formalisation

Then alongside above reasoning, we first give the definition of **random field**.

Definition 7.3. A collection, $X = \{X_t : t \in \mathbb{R}^d\}$, of r.v.s is called a **random field**.

Remark 7.1. In case of $d = 1$, random field becomes stochastic process.

Recall how we studied brownian motion, and also remember our obstacle, **order**:

$$\text{Filtration} \longmapsto \text{Martingale} \longmapsto \text{Brownian Motion}$$

And besides, we are going to concentrate into $\mathbb{R}_+^2$ since extension from this to higher/full dimension is simple, but generalise from $\mathbb{R}_+$ is the really difficulty.

Remark 7.2. We are going to use following notation, called **partial order**, $\preceq$:

$$(s,t) \preceq (s',t') \overset{\text{def.}}{\Longleftrightarrow} s \le s', t \le t'$$
$$(s,t) \prec (s',t') \overset{\text{def.}}{\Longleftrightarrow} s < s', t < t'$$

Therefore, we first establish the filtration:

Definition 7.4. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space, and $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ be a family of sub-σ-algebra of $\mathcal{F}$ and satisfying:

(a) If $z \preceq z'$, then $\mathcal{F}_z \subseteq \mathcal{F}_{z'}$.

(b) $\mathcal{F}_0$ contains all null sets of $\mathcal{F}$.

(c) $\forall\, z \in \mathbb{R}_+^2$, $\mathcal{F}_z = \bigcap_{z \preceq z'} \mathcal{F}_{z'}$.

(d) $\forall\, z = (s,t) \in \mathbb{R}_+^2$, for σ-algebras:

$$\mathcal{F}_z^{(1)} = \mathcal{F}_{s\infty} = \sigma\left(\mathcal{F}_{sv} : v \in \mathbb{R}_+\right), \mathcal{F}_z^{(2)} = \mathcal{F}_{\infty t} = \sigma\left(\mathcal{F}_{ut} : u \in \mathbb{R}_+\right)$$

, then $\mathcal{F}_z^{(1)} \perp \mathcal{F}_z^{(2)} \mid \mathcal{F}_z$.

(d') An alternative of (d) is: $\forall$ bounded r.v. X, $\forall\, z \in \mathbb{R}_+^2$:

$$\mathsf{E}\left(X \mid \mathcal{F}_z\right) = \mathsf{E}\left[\mathsf{E}\left(X \mid \mathcal{F}_z^{(1)}\right) \mid \mathcal{F}_z^{(2)}\right]$$

Be aware that for stochastic prcess, the 7.4's(c) be can derived from 7.4'(a), which is not true for random field. And following remark reveal the relation of this three σ-algebras:

Remark 7.3. For $A \in \mathcal{F}_z^{(1)} \cap \mathcal{F}_z^{(2)}$, choose $X = \mathbf{1}_A$, then from 7.4'(d'):

$$\mathsf{E}\left(\mathbf{1}_A \mid \mathcal{F}_z\right) = \mathsf{E}\left[\mathsf{E}\left(\mathbf{1}_A \mid \mathcal{F}_z^{(1)}\right) \mid \mathcal{F}_z^{(2)}\right] = \mathsf{E}\left(\mathbf{1}_A \mid \mathcal{F}_z^{(2)}\right) = \mathbf{1}_A$$

Then it means: $A \in \mathcal{F}_z^{(1)} \cap \mathcal{F}_z^{(2)} \implies \mathbf{1}_A \in \mathcal{F}_z \implies A \in \mathcal{F}_z \implies \mathcal{F}_z^{(1)} \cap \mathcal{F}_z^{(2)} \subseteq \mathcal{F}_z$. For another inclusion, it can be seen from 7.4's (d), then $\mathcal{F}_z = \mathcal{F}_z^{(1)} \cap \mathcal{F}_z^{(2)}$.

Until now, a natural question is that is there "filtration" satisfied those conditions? Since proving them rigourously is too demanding, we show its existence by providing example:

Example 7.1. On $(\Omega, \mathcal{F}, \mathsf{P})$, considering two independent filtration: $\{\mathcal{F}_s^{(1)} : s \in \mathbb{R}_+\}$, $\{\mathcal{F}_t^{(2)} : t \in \mathbb{R}_+\}$, then $\forall\, z = (s,t) \in \mathbb{R}_+^2$:

$$\mathcal{F}_z = \sigma\left(\mathcal{F}_s^{(1)} \cup \mathcal{F}_t^{(2)} \cup \{A : \mathsf{P}(A) = 0\}\right)$$

, then this clearly satisfies 7.4's (a)-(d).

Here we give a more concrete example,

Example 7.2. Let $\{X(A) : A \text{ is rectangle in } \mathbb{R}_+^2\}$ be a collection of r.v.s satifying: $X(A_1), \cdots, X(A_n)$ are independent if $A_1, \cdots, A_n$ are mutually disjoint, then set:

$$\mathcal{F}_z = \sigma\left(\sigma\left(X(A) : A \preceq z\right) \cup \{A : \mathsf{P}(A) = 0\}\right)$$

, which also satisfies 7.4's (a)-(d).

Confirmed with the filtration, we could further define the adaption:

Definition 7.5. A random field of $X = \{X_t : t \in \mathbb{R}_+^2\}$ is adapted with respect to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ if $\forall\, z \in \mathbb{R}_+^2, X_z \in \mathcal{F}_z$.

Moreover, measurable and progressively measurable process can also be formulated:

Definition 7.6. A random field of $X = \{X_t : t \in \mathbb{R}_+^2\}$, adapted with respect to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$, is called measurable if the function:

$$(z, w) \mapsto X_z(\omega)$$

, is $\mathcal{B}(\mathbb{R}_+^2) \otimes \mathcal{F}$ measurable.

Similarly for progressively measurable, since we won't use it further then leave for reader to verify. Next, we are ready to "try" to define the martingale:

Definition 7.7. A random field, $M = \{M_z : z \in \mathbb{R}_+^2\}$ is a martingale relative to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ if followings hold:

(a) M is adapted with respect to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$.

(b) $\forall\, z \in \mathbb{R}_+^2$, M is integrable, i.e. $\mathsf{E}|M_z| < +\infty$.

(c) $\forall\, z \preceq z', \mathsf{E}\left(M_{z'} \mid \mathcal{F}_z\right) = M_z$.

To proceed, we use following convention, $\forall\, z = (s, t) \preceq (s', t') = z' \in [0, +\infty)^2$, denote:

$$(z, z'] := (s, s'] \times (t, t']$$

Then getting this tool, we provide two remarks without narratively stating and proving:

Remark 7.4. 1. The increment of a random field X over rectangle $(z, z']$ is:

$$X_{(z,z']} = X_{s',t'} - X_{s',t} - X_{s,t'} + X_{s,t}$$

2. A random field, X, defined as above, induces a **random signed measure**.

All of above seems smooth, but indeed, they are not "real" martingale that we want, to modify some properties, we are more interested into weak/strong martingale:

Definition 7.8. Let $X = \{X_z : z \in \mathbb{R}_+^2\}$ be a random field that is integrable:

(a) X is a **weak martingale** relative to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ if:

  (1) X is adapted with respect to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$.

  (2) $\forall\, z \preceq z'$, $\mathsf{E}\left(X_{(z,z']} \mid \mathcal{F}_z\right) = 0$.

(b) X is a **martingale** relative to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ if:

  (1) X is adapted with respect to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$.

  (2) $\forall\, z \preceq z'$, $\mathsf{E}\left(X_{z'} - X_z \mid \mathcal{F}_z\right) = 0$.

(c) X is a **strong martingale** relative to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ if:

  (1) X is adapted with respect to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$.

  (2) $\forall\, z \preceq z'$, $\mathsf{E}\left[X_{(z,z']} \mid \sigma\left(\mathcal{F}_z^{(1)}, \mathcal{F}_z^{(2)}\right)\right] = 0$.

  (3) $M_z = 0$ if $z = (s, 0)$ or $(0, t)$.

Beside, we can also introduce "martingale" with only one filtration:

Definition 7.9. Let $X = \{X_z : z \in \mathbb{R}_+^2\}$ be a random field s.t. M_z is integrable $\forall\, z \in \mathbb{R}_+^2$. For $i = 1, 2$, M is an i-**martingale** relative to $\left\{\mathcal{F}_z^i : z \in \mathbb{R}_+^2\right\}$ if:

(a) M is adapted to $\left\{\mathcal{F}_z^i : z \in \mathbb{R}_+^2\right\}$.

(b) $\forall\, z \preceq z', \mathsf{E}\left[X_{(z,z']} \mid \mathcal{F}_z^i\right] = 0$.

Then a natural question is to ask the relations between these "martingales":

Proposition 7.1. *(a) A strong martingale is a martingale.*

(b) A martingale is both a 1- and 2-martingale.

(c) A random field, that is both 1- and 2- martingale, is a weak martingale.

Proof. Let $X = \{X_z : z \in \mathbb{R}_+^2\}$ be a random field, then we proceed the proof:

(a) If X is a strong martingale w.r.t $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$, then $X \in \mathcal{F}_z, \forall\, z \in \mathbb{R}_+^2$, notice:

$$\begin{aligned} X_{s',t'} - X_{s,t} &= (X_{s',t'} - X_{s,t'} - X_{s',t} + X_{s,t}) + (X_{s',t} - X_{s,t}) + (X_{s,t'} - X_{s,t}) \\ &= X_{(z,z']} + X_{((s,0),(s',t)]} + X_{((0,t),(s,t')]} \end{aligned}$$

Writing all of them in difference, we could use the proerties of strong martingale:

  (1) For $X_{(z,z']}$, we simply use the property, note that $\mathcal{F}_z \subseteq \sigma\left(\mathcal{F}_z^{(1)}, \mathcal{F}_z^{(2)}\right)$:

$$\mathsf{E}\left[X_{(z,z']} \mid \mathcal{F}_z\right] = \mathsf{E}\left[\mathsf{E}\left[X_{(z,z']} \mid \sigma\left(\mathcal{F}_z^{(1)}, \mathcal{F}_z^{(2)}\right)\right] \mid \mathcal{F}_z\right] = \mathsf{E}\left[0 \mid \mathcal{F}_z\right] = 0$$

  (2) For $X_{((s,0),(s',t)]}$, note that $\mathcal{F}_{s,t} \subseteq \mathcal{F}_{s,0}^{(1)} \subseteq \sigma\left(\mathcal{F}_z^{(1)}, \mathcal{F}_z^{(2)}\right)$:

$$\mathsf{E}\left[X_{((s,0),(s',t)]} \mid \mathcal{F}_z\right] = \mathsf{E}\left[\mathsf{E}\left[X_{((s,0),(s',t)]} \mid \sigma\left(\mathcal{F}_{s,0}^{(1)}, \mathcal{F}_{s,0}^{(2)}\right)\right] \mid \mathcal{F}_z\right] = 0$$

(3) Similar result for $X_{((0,t),(s,t')]}$.

Then we finished proof of (a), since:

$$\mathsf{E}\left[X_{z'} - X_z \mid \mathcal{F}_z\right] = \mathsf{E}\left[X_{(z,z']} + X_{((s,0),(s',t)]} + X_{((0,t),(s,t')]} \mid \mathcal{F}_z\right] = 0 + 0 + 0 = 0$$

(b) Let X be a martingale w.r.t $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$, then $\forall\, z \in \mathbb{R}_+^2, X_z \in \mathcal{F}_z \subseteq \mathcal{F}_z^{(1)}, \mathcal{F}_z^{(2)}$:

$$\begin{aligned}
\mathsf{E}\left[X_{(z,z']} \mid \mathcal{F}_{s,t}^{(1)}\right] &= \mathsf{E}\left[X_{s',t'} - X_{s',t} \mid \mathcal{F}_{s,t}^{(1)}\right] - \mathsf{E}\left[X_{s,t'} - X_{s,t} \mid \mathcal{F}_{s,t}^{(1)}\right] \\
&= \mathsf{E}\left[\mathsf{E}\left[X_{s',t'} - X_{s,t'} \mid \mathcal{F}_{s,t'}^{(2)}\right] \mid \mathcal{F}_{s,t'}^{(1)}\right] - \mathsf{E}\left[\mathsf{E}\left[X_{s',t} - X_{s,t} \mid \mathcal{F}_{s,t}^{(2)}\right] \mid \mathcal{F}_{s,t}^{(1)}\right] \\
&= \mathsf{E}\left[X_{s',t'} - X_{s,t'} \mid \mathcal{F}_{s,t'}\right] - \mathsf{E}\left[X_{s',t} - X_{s,t} \mid \mathcal{F}_{s,t}\right] = 0 - 0 = 0
\end{aligned}$$

This follows X is 1-martingale, then similar to 2-martingale we finish proof of (b).

(c) Let X is both 1- and 2-martingale, then for adaptivity:

$$\forall\, z \in \mathbb{R}_+^2, \begin{matrix} X_z \in \mathcal{F}_z^{(1)} \implies \sigma(X_z) \subseteq \mathcal{F}_z^{(1)} \\ X_z \in \mathcal{F}_z^{(2)} \implies \sigma(X_z) \subseteq \mathcal{F}_z^{(2)} \end{matrix} \implies \sigma(X_z) \subseteq \mathcal{F}_z^{(1)} \bigcap \mathcal{F}_z^{(2)} = \mathcal{F}_z$$

, so $X_z \in \mathcal{F}_z$, and integrability follows from construction, lastly:

$$\mathsf{E}\left[X_{(z,z']} \mid \mathcal{F}_z\right] = \mathsf{E}\left[\mathsf{E}\left[X_{(z,z']} \mid \mathcal{F}_z^{(1)}\right] \mid \mathcal{F}_z\right] = \mathsf{E}\left[0 \mid \mathcal{F}_z\right] = 0$$

Therefore, we completed the proof. □

Remark 7.5. Generally: 1- and 2-martingale $\not\Rightarrow$ Martingale $\not\Rightarrow$ Strong Martingale.

After all preparation, we could add distribution into random field, now we introduce:

Definition 7.10. Let $X = \{X_t : t \in \mathbb{R}_+^d\}$ be a random field. If $\forall\, n \in \mathbb{N}$, $\forall$ disjoint, $t^{(1)}, \cdots, t^{(n)} \in \mathbb{R}_+^d$, the joint distribution of $X_{t^{(1)}}, \cdots, X_{t^{(n)}}$ is **Gaussian**, then X is called a **Gaussian Random Field**.

Recall 5.3, for BM = adapted + centered gaussian + $\mathsf{E}B_t\mathsf{B}_s = \min\{s,t\}$, then:

Definition 7.11. A centered Gaussian random field, $\mathscr{B} = \{\mathscr{B}_t : t \in \mathbb{R}_+^d\}$ with convariance function, $\forall\, s = (s_1, \cdots, s_d), t = (t_1, \cdots, t_d) \in \mathbb{R}_+^d$:

$$\mathsf{E}\mathscr{B}_t\mathscr{B}_s = \prod_{j=1}^{d} \min\{s_j, t_j\}$$

, is called a **Brownian sheet (d, 1)**.

Remark 7.6. Notice that if B and B' are two independent Brownian motion, and since $\mathsf{E}\mathsf{B}_t\mathsf{B}_s = \min\{s,t\}$ and $\mathsf{E}\mathsf{B}'_{t'}\mathsf{B}'_{s'} = \min\{s',t'\}$, then define $X_t = \mathsf{B}_t\mathsf{B}'_t$ is not a Brownian sheet (since it is not Gaussian).

Also recall 5.1, BM = adapted + independence increment + $\mathsf{B}_t - \mathsf{B}_s \sim \mathcal{N}(0, t-s)$:

Definition 7.12. Let $(S, \mathcal{S}, \nu)$ is a σ-finite measure space. A random set function W on $\{A \in \mathcal{S} : \nu(A) < +\infty\}$, satisfying:

(a) $W(A) \sim \mathcal{N}(0, \nu(A))$.

(b) If $A \cap B = \emptyset$, then $W(A) \perp W(B)$ and $W(A \cup B) = W(A) + W(B)$.

, is called a **white noise** based on ν.

Remark 7.7. If W is a white noise with $\nu =$ Lebs, set $X_t = W\left((0, t_1] \times \cdots \times (0, t_d]\right)$, here $t = (t_1, \cdots, t_d) \in \mathbb{R}_+^d$, then $\{X_t : t \in \mathbb{R}_+^d\}$ is a Brownian sheet w.r.t $\left\{\mathcal{F}_z^X : z \in \mathbb{R}_+^d\right\}$.
A natural question is to ask can we recover brownian motion through brownian sheet:

Proposition 7.2. *Let $\mathscr{B} = \{\mathscr{B}_t : t \in \mathbb{R}_+^d\}$ be a Brownian sheet w.r.t $\{\mathcal{F}_z : z \in \mathbb{R}_+^d\}$:*

(a) $\mathscr{B}_t = 0$ a.s. if $t_1 \cdots t_d = 0$.

(b) Set $X_{t_1} = (t_2 \cdots t_d)^{-1/2} \mathscr{B}_t$ for $t = (t_1, \cdots, t_d) \in \mathbb{R}_+^d$, if we fix $t_2, \cdots, t_d$, then $X = \{X_t : t_1 \in \mathbb{R}_+\}$ is a standard Brownian motion w.r.t $\{\mathcal{F}_t : t_1 \in \mathbb{R}_+\}$.

Proof. The proof is simple, since we just need to check conditions

(a) $0 \le \mathsf{E}\mathscr{B}_t^2 = \prod_{j=1}^d t_j = 0 \implies \mathscr{B}_t^2 = 0$ a.s.

(b) $\mathsf{E}X_{t_1}X_{t_1'} = (t_2 \cdots t_d)^{-1} \mathsf{E}\mathscr{B}_{t_1, \cdots, t_d}\mathscr{B}_{t_1', \cdots, t_d} = \min\{t_1, t_1'\}$

Then we finished the proof easily. □

Besides, we want to introduce a sample path property of Brownian sheet:

Proposition 7.3. *A Brownian sheet has a modification, which has continuous sample paths.*

Proof. The key tool is Kolmogorov Continuity Theorem for Random field, and the idea of the proof is to find:

$$\mathsf{E}|\mathscr{B}_t - \mathscr{B}_s|^\alpha \le C \cdot |t - s|^?$$

Then recover the modification so that continuous sample paths. □

Lastly, we present the theorem connect Brownian sheet to martingale:

Theorem 7.4. *A Brownian sheet $\mathscr{B}$ is a strong martingale w.r.t. $\left\{\mathcal{F}_z^{\mathscr{B}} : z \in \mathbb{R}_+^d\right\}$, where $\mathcal{F}_z^{\mathscr{B}} = \sigma\left(\sigma\left(\mathscr{B}_{z'} : z' \prec z\right) \bigcup \mathcal{F}_{null}\right)$, here $\mathcal{F}_{null} = \{A \in \mathcal{F} : \mathsf{P}(A) = 0\}$.*

Proof. We only present the idea of the proof:

(1) For disjoint rectangle $A = (z, z'], \tilde{A} = (\tilde{z}, \tilde{z'}]$, check $\mathsf{E}\mathscr{B}_A\mathscr{B}_{\tilde{A}} = 0 \implies \mathscr{B}_A \perp \mathscr{B}_{\tilde{A}}$

(2) $\mathscr{B}_t = 0, t = (t_1, t_2)$ with $t_1 = t_2 = 0$.

(3) Check $\mathsf{E}\left[\mathscr{B}_{(z,z']} \mid \sigma\left(\mathcal{F}_z^{(1)}, \mathcal{F}_z^{(2)}\right)\right] = 0, \forall\ z \preceq z'$.

And readers are encouraged to complete the proof. □

Remark 7.8. Lévy process can also be extended to random field but too complicated.

References

[1] K. L. Chung, *A Course in Probability Theory*. 3rd ed., Academic Press, San Diego, CA, 2001. xviii+419 pp.

[2] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*. 2nd ed., Springer-Verlag, New York, 1991. xxiv+470 pp.

[3] D. Applebaum, *Lévy Processes and Stochastic Calculus*. 2nd ed., Cambridge University Press, Cambridge, 2009. xxx+460 pp.

[4] R. Durrett, *Probability: Theory and Examples*. 5th ed., Cambridge University Press, Cambridge, 2019.

[5] D. Williams, *Probability with Martingales*. Cambridge University Press, Cambridge, 1991.

[6] J.-F. Le Gall, *Brownian Motion, Martingales, and Stochastic Calculus*. Springer, Cham, 2016.