# Stochastic Processes

Labix

May 8, 2025

**Abstract**

## Contents

# 1 Introduction to Stochastic Processes

## 1.1 Basic Terminology

**Definition 1.1.1: Stochastic Processes**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. A stochastic process is a collection of random variables

$$\{X(t) : \Omega \to S \mid t \in T\}$$

indexed by some set $T$. In this case we write the stochastic process as $X : \Omega \times T \to S$.

**Definition 1.1.2: Terminology**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $X : \Omega \times T \to S$ be a stochastic process.

- Define the state space of the stochastic process to be $S$.
- Write $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$.

**Definition 1.1.3: Discrete Time Stochastic Processes**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $X : \Omega \times T \to S$ be a stochastic process. We say that the stochastic process has discrete time if $T \subseteq \mathbb{N}$.

**Definition 1.1.4: Filtrations**

Let $\mathcal{F} = \{\mathcal{F}_t \mid t \in T\}$ be a family of $\sigma$-algebras. We say that $\mathcal{F}$ is a filtration if $\mathcal{F}_t \leq \mathcal{F}_s$ for $t \leq s$.

**Definition 1.1.5: Adapted Stochastic Process**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $X : \Omega \times T \to S$ be a stochastic process. Let $\mathcal{F} = \{\mathcal{F}_t \mid t \in T\}$ be a filtration. We say that $X$ is adapted to $\mathcal{F}$ if $X_t$ is $\mathcal{F}_t$-measurable for all $t \in T$.

## 1.2 Hitting Time and Stopping Time

**Definition 1.2.1: Hitting Time**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $X : \Omega \times T \to S$ be a stochastic process. Let $A \in \Sigma$ be a measurable set. Define the hitting time of $X$ to be the random variable $H_A : \Omega \to T$ given by

$$H_A(\omega) = \inf\{t \in T \mid X(\omega, t) \in A\}$$

Recall that for a $\sigma$-algebra $\mathcal{F}$ and a subset $Y \subseteq \mathcal{F}$, $\sigma(Y)$ is the smallest $\sigma$-algebra that contains $Y$.

**Definition 1.2.2: Stopping Time**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $X : \Omega \times T \to S$ be a stochastic process. Let $S : \Omega \to T$ be a random variable. We say that $S$ is a stopping time if $S$ has the property that

$$\{\omega \in \Omega \mid S(\omega) \leq t\} \subseteq \sigma(X_t)$$

for all $t \in T$.

In other words, events determined by $S$ up to some time $t \in T$ is dictated entirely using the variables $X_0, \dots, X_t$.

**Lemma 1.2.3**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $X : \Omega \times T \to S$ be a stochastic process. Let $S : \Omega \to T$ be a random variable. Then $S$ is a stopping time if and only if $S$ has the property that

$$\{\omega \in \Omega \mid S(\omega) = t\} \subseteq \sigma(X_t)$$

for all $t \in T$.

**Theorem 1.2.4: Début Theorem**

# 2 Markov Chains

## 2.1 Markov Chains

**Definition 2.1.1: Stochastic Matrix**

Let $P = (p_{i,j}) \in M_n(\mathbb{R})$ be a matrix. We say that $P$ is a stochastic matrix if the following are true.

- $0 \leq p_{i,j} \leq 1$ for $1 \leq i, j \leq n$.
- For any fixed $1 \leq k \leq n$, we have

$$\sum_{j=1}^{n} p_{k,j} = 1$$

**Definition 2.1.2: Markov Chain**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $I = \{1, \dots, k\} \subseteq \mathbb{N}$. Let $\{X_n : \Omega \to I \mid n \in \mathbb{N}\}$ be a discrete stochastic process. Let $\lambda : I \to [0,1]$ be a probability distribution. Let $P = (p_{i,j})$ be a stochastic matrix. We say that $(X_n)_{n\geq 0}$ is a Markov chain with initial distribution $\lambda$ and transition matrix $P$ if the following are true.

- $P_{X_0} = \lambda$
- For any $i_0, \dots, i_{n+1} \in I$, we have

$$\mathbb{P}(X_{n+1} = i_{n+1} \mid X_k = i_k \text{ for } 0 \leq k \leq n) = \lambda(i_0) \cdot \prod_{j=0}^{n} p_{i_j, i_{j+1}}$$

In this case we say that $(X_n)_{n\geq 0}$ is Markov$(\lambda, P)$.

In other words, Markov chains are a sequence of random variables where the next step does not depends on what states you visited previously, but only on the state now. There is an important property that makes studying Markov chains worth while.

**Proposition 2.1.3**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $I = \{1, \dots, k\}$. Let $\{X_n : \Omega \to I \mid n \in \mathbb{N}\}$ be a sequence of random variables. Let $\lambda : I \to [0,1]$ be a probability distribution. Then $(X_n)_{n\geq 0}$ is a Markov chain if and only if

$$\mathbb{P}(X_{n+1} = i_{n+1} \mid X_k = i_k \text{ for } 0 \leq k \leq n) = \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n)$$

for any $i_0, \dots, i_{n+1} \in I$. In this case, the transition matrix of the Markov chain is given by

$$P = \begin{pmatrix} \mathbb{P}(X_1 = 1 \mid X_0 = 1) & \cdots & \mathbb{P}(X_1 = k \mid X_0 = 1) \\ \vdots & \ddots & \vdots \\ \mathbb{P}(X_1 = 1 \mid X_0 = k) & \cdots & \mathbb{P}(X_1 = k \mid X_0 = k) \end{pmatrix}$$

**Theorem 2.1.4: Markov Property**

Let $(X_n)_{n\geq 0}$ be a markov chain. Suppose that $X_m = E_m$ is given. Then $(X_{m+n})_{n\geq 0}$ is a Markov chain.

Notice that the transition probabbility $\mathbb{P}(X_{n+1} = j | X_n = i)$ still depedns on $i, j, n$. We further restrict out study of Markov chains to the following type.

**Definition 2.1.5: Homogenous Markov Chains**

We say that a Markov chain $(X_n)_{n\geq 0}$ is homogenous if

$$\mathbb{P}(X_{n+1} = j|X_n = i) = \mathbb{P}(X_1 = j|X_0 = i)$$

This means that homogenous Markov chains are time independent.

**Lemma 2.1.6**

The transition matrix $P$ of a homogenous Markov chain is a Stochastic matrix.

**Theorem 2.1.7**

Let $(X_n)_{n\geq 0}$ be Markov. Then for all $m, n \geq 0$, we have

- $\mathbb{P}(X_n = j|X_0 = i) = \mathbb{P}(X_1 = j|X_0 = i)^n$
- $\mathbb{P}(X_n = j) = \sum_{i\in I} \mathbb{P}(X_0 = i)\mathbb{P}(X_n = j|X_0 = i)$

This theorem is saying that $\mathbb{P}(X_n = j|X_0 = i)$ is equal to the $i, j$th entry of the matrix $P^n$ if $P$ is the transition matrix. To find out the total probability of reaching the $j$th state at the $n$the step regardless of the starting point, you sum up the $\mathbb{P}(X_n = j|X_0 = i)$ multiplying the probability that you start at state $i$.

We often like to create new Markov chains from old (at least from exams). If we want to show that a squence $(X_n)_{n\geq 0}$ of random vriables is Markov, try and write $X_{n+1} = X_n +$ some term only related to $n+1$.

We use

$$h_i^A = \mathbb{P}(H^A < \infty|X_0 = i)$$

to denote the corresponding probability. We use

$$k_i^A = E_i[H_A] = \sum_{k=0}^{\infty} k\mathbb{P}(H_A = k|X_0 = i) + \infty\mathbb{P}(H_A = \infty|X_0 = i)$$

to denote the expectation.

**Proposition 2.1.8**

We can find the probabilities of all hitting times by solving the system of linear equations

$$\begin{cases} \mathbb{P}(H_A < \infty|X_0 = i) = 1 & i \in A \\ \mathbb{P}(H_A < \infty|X_0 = i) = \sum_{j\in I} \mathbb{P}(X_1 = j|X_0 = i)\mathbb{P}(H_A < \infty|X_0 = j) & x \notin A \end{cases}$$

**Proposition 2.1.9**

We can find the expected number of hitting times by solving the system of linear equations

$$\begin{cases} E_i[H_A] = 0 & i \in A \\ E_i[H_A] = 1 + \sum_{j\notin A} \mathbb{P}(X_1 = j|X_0 = i)E_j[H_A] & i \notin A \end{cases}$$

**Theorem 2.1.10: Strong Markov Property**

Let $(X_n)_{n\geq 0}$ be Markov$(\lambda, P)$ and $T$ a stopping time of $(X_n)_{n\geq 0}$. Then conditional on both $\{X_T = i\}$ and $\{T < \infty\}$, we have $(X_{T+n})_{n\geq 0}$ is Markov$(\delta_i, P)$ and independent of $X_0, \dots, X_T$.

## 2.2 Communicating Classes

**Definition 2.2.1: Talks and Communicates**

Let $i, j$ be states. We say that $i$ talks to $j$ which is $i \to j$ if there exists $n \in \mathbb{N}$ such that

$$\mathbb{P}(X_n = j | X_0 = i) > 0$$

We say that two states $i$ and $j$ communicate if $i \leftrightarrow j$.

**Definition 2.2.2: Communicating Class**

Let $C \subseteq I$. $C$ is a communicating class if $\forall i, j \in C$, $i \leftrightarrow j$ and $\forall i \in C$ and $\forall k \in I \setminus C$, $i \not\leftrightarrow k$

**Definition 2.2.3: Closed and Absorbing**

We say that a class is closed if no states in the class talks to any states outside of that class. We say that a class is absorbing if it forms a closed class by itself.

**Definition 2.2.4: Irreducible Markov Chains**

A Markov chain is said to be irreducible if for all $i, j \in I$, $i \leftrightarrow j$.

## 2.3 Recurrence and Transcience

**Definition 2.3.1: Recurrence and Transcience**

A state $i \in I$ is recurrent if $\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 1$. A state $i \in I$ is transient if $\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 0$.

**Definition 2.3.2: $k$-th Passage Time**

The first passage time of state $i$ is a stopping time such that

$$T_i(\omega) = \inf\{n \geq 1 | X_n(\omega) = i\}$$

The $k$-th passage time is a stopping time such that

$$T_i^{(k)}(\omega) = \inf\{n \geq T_i^{(k-1)} | X_n(\omega) = i\}$$

The $k$-th excursion time is defined by

$$S_i^{(k)} = \begin{cases} T_i^{(k)} - T_i^{(k-1)} & \text{if } T_i^{(k-1)} < \infty \\ \infty & \text{otherwise} \end{cases}$$

Intuitively, $T_i^{(k)}$ outputs the time it takes for the Markov chain to reach state $i$ for the $k$th time and $S_i^k$ outputs the number of steps taken between consecutive visits.

**Lemma 2.3.3**

For $k = 2, 3, \dots$ and if $T_i^{(k-1)} < \infty$, then $S_i^{(k)}$ is independent of $\{X_m : m \leq T_i^{(k-1)}\}$ and

$$\mathbb{P}(S_i^{(k)} = n | T_i^{(k-1)} < \infty) = \mathbb{P}_i(T_i = n)$$

**Definition 2.3.4: Visit Counting Function**

Let $i \in I$ be a state. Define
$$V_i = \sum_{k=0}^{\infty} 1_{\{X_n=i\}}$$
the number of visits ever to state $i$.

**Lemma 2.3.5**

If $V_i$ counts the number of vists to state $i$, then
$$E_i(V_i) = \sum_{k=0}^{\infty} P_{ii}^{(k)}$$

**Lemma 2.3.6**

For $k = 0, 1, 2, \dots,$ we have
$$\mathbb{P}_i(V_i > k) = (\mathbb{P}_i(T_i < \infty))^k$$

This lemma means that $V_i$ asserts a geometric distribution.

**Theorem 2.3.7: Characteristic of Recurrent and Transient States**

Let $(X_n)_{n\geq 0}$ be a Markov chain. Let $T_i$ be the first passage time of state $i$. If $\mathbb{P}_i(T_i < \infty) = 1$, then $i$ is recurrent and $\sum_{k=1}^{\infty} \mathbb{P}(X_1 = k|X_0 = k)^n = \infty$. If $\mathbb{P}_i(T_i < \infty) < 1$, then $i$ is transient and $\sum_{k=1}^{\infty} \mathbb{P}(X_1 = k|X_0 = k)^n < \infty$.

**Theorem 2.3.8**

Let $C$ be a communicating class. Then either all states in $C$ are transient or all are recurrent.

**Theorem 2.3.9**

Let $C \subset I$ be a class of a Markov chain. Then

- Every recurrent class is closed
- Every finite closed class is recurrent

**Theorem 2.3.10**

If $P$ is irreducible and recurrent then for all $j \in I$,
$$\mathbb{P}(T_j < \infty) = 1$$

**Definition 2.3.11: Component Independent Simple Random Walk on $\mathbb{Z}^d$**

A component independent simple random walk on $\mathbb{Z}^d$ for $d \in \mathbb{N}$ is defined as
$$X_{n+1} = X_n + Z_{n+1}$$
where $Z_{n+1} = (Z_{n+1}^1, \dots, Z_{n+1}^d)$ with
$$Z_m^j = \begin{cases} 1 & \text{with probability } p_j \\ -1 & \text{with probabilty } q_j \end{cases}$$
where $Z_m^j$ are independent for all $j, m$.

**Proposition 2.3.12**

Let $(X_n)_{n\geq 0}$ be a CISRW on $\mathbb{Z}^d$. If there exists $j \in \{1, \dots, d\}$ such that $p_j \neq \frac{1}{2}$ then $(X_n)$ is transient.

**Proposition 2.3.13**

Let $(X_n)_{n\geq 0}$ be a CISRW on $\mathbb{Z}^d$ and $p_j = q_j = \frac{1}{2}$ for all $j$. Then

- If $d \leq 2$ then all states are recurrent
- If $d \geq 3$ then all states are transient

## 2.4 Branching Process

**Definition 2.4.1: Branching Process**

Let $(X_n)_{n\geq 0}$ be the number of individuals in a population at time $n$ where $X_n \in \{0\} \cup \mathbb{N}$. At every time step each individual in the population gives birth to a random number of off-spring. Thus $X_{n+1}$ is defined to be

$$X_{n+1} = \sum_{k=1}^{X_n} Z_k^n$$

where $Z_k^n \sim Z$ and $\mathbb{P}(Z \geq 0) = 1$. Then $(X_n)_{n\geq 0}$ is said to be a branching process.

**Proposition 2.4.2**

Define $G(s) = E[s^Z]$ and $F_{n+1}(s) = E[S^{X_{n+1}} | X_0 = 1]$ for $s \in (0, 1)$. Then

$$F_n(s) = G(F_{n-1}(s))$$

**Proposition 2.4.3**

For a braching process $(X_n)_{n\geq 0}$ with off-spring distribution $Z$ where $E[Z] = \mu$, we have

$$E[X_n] = \mu^n$$

**Theorem 2.4.4: Extinction Probability**

The extinction probability is the smallest non-negative root of $G(\alpha) = \alpha$.

**Theorem 2.4.5**

Suppose that $G(0) > 0$, $\mu = E[Z] = G'(1)$. Then $\mu \leq 1$ implies certain extiction. $\mu > 1$ implies uncertain extinction.

## 2.5 Invariant Distributions on a Markov Chain

**Definition 2.5.1: Invariant Distribution**

A measure $\lambda = (\lambda_i : i \in I)$ with non-negative entries is said to be an invariant distribution if

- $\lambda P = P$
- $\pi_i \in [0, 1]$ for all $i$
- $\sum_{i\in I} \pi_i = 1$

The following theorem shows that after applying a Markov process to an invariant distribution, the new distribution will be the same as the old one. Notice here that right multiplication of a distribution gives the next step on the Markov chain instead of the usual left multiplication.

**Theorem 2.5.2**

Let $(X_n)_{n\geq 0}$ be Markov$(\pi, P)$, where $\pi$ is an invariant distribution for $P$. Then $\forall m \geq 0$, $(X_{n+m})_{n\geq 0}$ is Markov$(\pi, P)$.

*Proof.* By the Markov property, the new sequence will be Markov. But we do not yet know its distribution. We have that

$$\begin{aligned} \mathbb{P}(X_m = i) &= (\pi P^m)_i && \text{(this denotes the } i\text{th entry of } P) \\ &= (\pi P P^{m-1})_i \\ &= (\pi P^{m-1})_i \end{aligned}$$

Thus by induction, we see that $\mathbb{P}(X_m = i) = \pi_i$ thus we are done. □

**Theorem 2.5.3**

Let $I$ be finite. Suppose that for some $i \in I$,

$$P_{ij}^n \to \pi_j$$

for all $j \in I$. Then $\pi = (\pi_j : j \in I)$ is an invariant distribution.

Let us look at an example.

**Example 2.5.4: Gene Mutation**

Recall the gene mutation example that has transition matrix $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$. Recall that

$$P^n = \begin{pmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n \\ ? & ? \end{pmatrix}$$

As $n \to \infty$, observe that the first entry tends to $\frac{\beta}{\alpha+\beta}$ the second entry tends to $\frac{\alpha}{\alpha+\beta}$. This means that $\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$ is an invariant distribution for the gene mutation model.

We can also find the invariant distribution directly: Let $\pi = (\pi_1, \pi_2)$ be an invariant distribution for $P$. We have that

$$\begin{aligned} \pi P &= \pi \\ \begin{pmatrix} \pi_1(1-\alpha) + \pi_2\beta & \pi_1\alpha + \pi_2(1-\beta) \end{pmatrix} &= \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \\ \begin{pmatrix} \pi_2\beta - \pi_1\alpha & \pi_1\alpha - \pi_2\beta \end{pmatrix} &= 0 \end{aligned}$$

Notice that these two equations mean the same thing. (In general if you have an $n \times n$ transition matrix, only the first $n-1$ equations will be useful and the last one will be redundant.) This is why we need to use the fact that $\pi_1 + \pi_2 = 1$ as our final equation.

Now solving it, we have that $\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$ which is the exact same answer as the one given in the first method.

A third way to think about it is that notice that $P^T \pi^T = \pi^T$ is the equation that we are attempting to solve. Recall that 1 is necessarily an eigenvector of a transtition matrix which means that this equation really is just a question of finding eigenvectors from the eigenvalue 1.

The above examples gives a lot of practical information when calculating the invariant distribution of a transition matrix.

**Definition 2.5.5: Expected Time Spent**

Define the expected time spent in state $i$ in between visits to state $k$ by

$$\gamma_i^k = E_k\left(\sum_{n=0}^{T_k-1}\mathcal{X}_{X_n=i}\right)$$

This uses $k$ as a reference, and as we keep going back to $k$, this records how much time we have spent in $i$ in the process of leaving and returning to $k$.

**Theorem 2.5.6**

Let $P$ be irreducible and recurrent. Then

- $\gamma_k^k = 1$
- $\gamma^k = (\gamma_i^k : i \in I)$ is an invariant measure satisfying $\gamma^k P = \gamma^k$
- $0 < \gamma_i^k < \infty\ \forall i \in I$

**Theorem 2.5.7**

Let $P$ be irreducible and $\lambda$ and invariant measure and $\lambda_k = 1$ for some $k$. Then $\lambda \geq \gamma^k$. If we also have that $P$ is recurrent, then $\lambda = \gamma^k$

This theorem assesrts that as long as $P$ is irreducible and recurrent, then any invariant measure is exactly equal to the $\gamma^k$ we defined, which means that there is really only one invariant measure up to rescaling.

**Definition 2.5.8: Positive Recurrent**

A state $i \in I$ is positive recurrent if it is recurrent and

$$E_i[T_i] < \infty$$

A state which is not positive recurrent but is recurrent is called null recurrent.

**Theorem 2.5.9**

Let $P$ be irreducible. Then the following are equivalent.

- Every state is positive recurrent
- Some state $i$ is positive recurrent
- $P$ has an invariant distribution $\pi$. More over, $\pi_i = \frac{1}{E_i[T_i]}$

# 3 Brownian Motion

**Definition 3.0.1: Brownian Motion**

Let $(\Omega, E, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $B : \Omega \times [0, T] \to S$ be continuous stochastic processes. We say that $B$ is a Brownian motion if the following are true.

- $B_0 = 0$.
- For any $0 \leq t_1 < t_2 < t_3 \leq T$,
$$B_{t_2} - B_{t_1} \quad \text{and} \quad B_{t_3} - B_{t_2}$$
are independent variables.
- For any $0 \leq t_1 < t_2 \leq T$,
$$B_{t_2} - B_{t_1} \sim N(0, t_2 - t_1)$$
- For any $\omega \in \widetilde{\Omega}$ such that $P(\widetilde{\Omega}) = 1$, we have that $B(\omega, -) : T \to S$ is a continuous function.

# 4 Martingale Processes

## 4.1 Basic Definitions

**Definition 4.1.1: Martingale Process**

Let $(\Omega, E, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $M : \Omega \times [0, T] \to S$ be a continuous stochastic process adapted to a filtration $\mathcal{F} = \{\mathcal{F}_t \mid t \in [0, T]\}$. We say that $M$ is Martingale with respect to $\mathcal{F}$ if the following are true.

- $E[|M_k|] < \infty$ for all $k$.
- $E[M_{t_2} \mid \mathcal{F}_{t_1}] = M_{t_1}$ for all $0 \leq t_1 < t_2 \leq T$.

**Definition 4.1.2: Continuous Martingale Process**

Let $(\Omega, E, P)$ be a probability space. Let $(S, \Sigma)$ be a measurable space. Let $M : \Omega \times [0, T] \to S$ be a continuous stochastic process. We say that $M$ is continuous if for all $\omega \in \widetilde{\Omega}$ such that $P(\widetilde{\Omega}) = 1$, the function

$$M(\omega, -) : [0, T] \to S$$

is continuous.

**Definition 4.1.3: Standard Brownian Filtration**

## 4.2 Martingale Transformations

# 5 Stochastic Calculus

## 5.1 Ito's Integral

Given a Martingale process $M : \Omega \times [0,T] \to \mathbb{R}$ on some filtered probability space, we may ask to integrate $M$ over the product measure of the measurable space $\Omega \times [0,T]$ because $M$ is in particular a measurable function. Ito's integral gives a construction of such an integral under certain assumption on $M$.

**Definition 5.1.1: Simple Functions**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Define $\mathcal{H}_0^2 \subseteq L^2(\Omega \times [0,T])$ to be the vector subspace whose measurable functions are of the form

$$f(\omega,t) = \sum_{i=0}^{n-1} a_i(\omega) 1_{(t_i, t_{i+1}]}$$

**Definition 5.1.2: Ito Integral of Simple Functions**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $f \in \mathcal{H}_0^2$. Define the Ito integral of $f(\omega,t) = \sum_{i=0}^{n-1} a_i(\omega) 1_{(t_i, t_{i+1}]}$ by the formula

$$I(f)(\omega) = \sum_{t=0}^{n-1} a_i(\omega)(B_{t_{i+1}} - B_{t_i})$$

In particular, $I$ is a function $I : \mathcal{H}_0^2 \to L^2(\Omega)$.

**Lemma 5.1.3: Ito's Isometry I**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. The map $I : \mathcal{H}_0^2 \to L^2(\Omega)$ is a linear isometry.

**Definition 5.1.4: Ito Integrable Functions**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Define the space of Ito integrable functions $\mathcal{H}^2 \subseteq L^2(\Omega \times [0,T])$ to be the vector subspace whose measurable functions $f \in L^2(\Omega \times [0,T])$ satisfies the fact that

$$E\left[\int_{[0,T]} f^2(\omega,t)\,dt\right] = \int_\Omega \int_{[0,T]} f^2(\omega,t)\,dt\,dP < \infty$$

**Lemma 5.1.5**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Then $\mathcal{H}_0^2$ is dense in $\mathcal{H}^2$.

**Definition 5.1.6: Ito Integral of Integrable Functions**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $f \in \mathcal{H}^2$. Let $(f_n)_{n \in \mathbb{N} \setminus \{0\}}$ be a sequence such that $f_n \to f$ in $L^2(\Omega \times [0,T])$. Then define the Ito integral of $f$ by

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

In particular, $I$ is a function $I : \mathcal{H}^2 \to L^2(\Omega)$.

**Lemma 5.1.7: Ito's Isometry II**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. The map $I : \mathcal{H}^2 \to L^2(\Omega)$ is a linear isometry.

## 5.2 The Integral as a Martingale

## 5.3 Ito's Lemma

## 5.4 Localization

# 6 Stochastic Differential Equations

## 6.1 Basic Definitions

**Definition 6.1.1: Stochastic Differential Equations**

Methods of solving differential equations can also be applied to stochastic differential equations.

## 6.2 Geometric Brownian Motions

**Definition 6.2.1: Geometric Brownian Motion**

Let $(\Omega, E, P)$ be a probability space. Let $S : \Omega \times [0, T] \to \mathbb{R}$ be a continuous Martingale process with respect to the standard Brownian filtration $\mathcal{B} = \{B_t \mid t \in [0, T]\}$. We say that $S$ follows a geometric Brownian motion if it satisfies the following stochastic differential equation

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

where $\mu$ is called the percentage drift and $\sigma$ is called the percentage volitality.

Recall that $dS_t$ really means the Radon–Nikodym derivative $\frac{dS_t}{d\mu}$ where $\mu$ here (not the same one as above) refers to the standard measure on $\mathbb{R}$, while $dt$ means the genuine infinitesimally small increment in $\mathbb{R}$.

**Lemma 6.2.2**

Let $(\Omega, E, P)$ be a probability space. Let $S : \Omega \times [0, T] \to \mathbb{R}$ be a continuous Martingale process with respect to the standard Brownian filtration $\mathcal{B} = \{B_t \mid t \in [0, T]\}$. Suppose that $S$ follows a geometric Brownian motion. Then the solution of the stochastic differential equation is given by

$$S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma B_t}$$

**Proposition 6.2.3**

Let $(\Omega, E, P)$ be a probability space. Let $S : \Omega \times [0, T] \to \mathbb{R}$ be a continuous Martingale process with respect to the standard Brownian filtration $\mathcal{B} = \{B_t \mid t \in [0, T]\}$. Suppose that $S$ follows a geometric Brownian motion. Then the following are true.

- For each $t$, $S_t$ has a log normal distrbution.
- For each $t$, we have
$$E[S_t] = S_0 e^{\mu t}$$
- For each $t$, we have
$$\text{Var}(S_t) = S_0^2 e^{2\mu t}(e^{\sigma^2 t} - 1)$$

We also provide a discrete approach to solving the differential equation for easy simulation using the Monte Carlo method.

**Proposition 6.2.4**

The discretization of the defining stochastic differential equation of a geometric Brownian motion is given by

$$S_{t+dt} - S_t = \mu S_t dt + \sigma S_t N \sqrt{dt}$$

where $dt$ denotes a small change in time, and $N \sim N(0, 1)$ refers a random variable with a normal distribution.

## 6.3 Systems of SDEs