Stochastic Calculus for Finance
Course No. MA 372

Course Instructor: Subhamay Saha
Email: saha.subhamay@iitg.ac.in
Text book: Stochastic Calculus for Finance II by Steven Shreve

Contents

1 Basics of Probability Theory

1.1 Infinite Probability Spaces

• Used to model a random experiment with infinitely many possible outcomes.
• Two examples to keep in mind:
i) Choosing a number from the interval $[0, 1]$
ii) Tossing a coin infinitely many times.
For i) the sample space is $[0, 1]$. For ii) the sample space is the set of all infinite sequences of heads and tails. One important issue with infinite sample spaces is that the classical definition fails. For example, what is the probability that you choose a number less than or equal to $1/2$. In this case both the total number of outcomes as well as the number of favorable outcomes are infinite. This leads us to what is called the axiomatic definition of probability. Before that we need one more definition.

Definition 1.1. *Let Ω be a non-empty set and let $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is a $\sigma-$algebra provided:*
i) the empty set ϕ belongs to $\mathcal{F}$
ii) if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$ (closed under complementation)
iii) whenever a sequence of sets $A_1, A_2, \ldots$ belongs to $\mathcal{F}$ then $\cup_{n=1}^{\infty} A_n$ also belongs to $\mathcal{F}$ (closed under countable union).

Two trivial examples $\mathcal{F}_1 = \mathcal{P}(\Omega)$, $\mathcal{F}_2 = \{\phi, \Omega\}$.
Result: Arbitrary intersection of $\sigma-$algebras is again a $\sigma-$algebra.
Proof: Exercise.
Exercise: Is arbitrary union of $\sigma-$algebras again a $\sigma-$algebra ?

Definition 1.2. *Let $\mathcal{A}$ be a collection of subsets of Ω. The $\sigma-$algebra generated by $\mathcal{A}$ is given by*

$$\sigma(\mathcal{A}) \doteq \bigcap \mathcal{F},$$

where the intersection is over all $\sigma-$algebras containing $\mathcal{A}$. Thus this is the smallest $\sigma-$algebra containing $\mathcal{A}$.

Exercise: Let A be a non-empty subset of Ω. Find $\sigma(\{A\})$. Let B be another non-empty subset of Ω such that $A \cap B \neq \phi$ and $A \cup B \neq \Omega$. Find $\sigma(\{A, B\})$.
Exercise: Suppose $\Omega = [0, 1]$. Find the $\sigma-$algebra generated by the collection of singletons.

Definition 1.3 (Axiomatic Definition)**.** *Let Ω be a non-empty set and $\mathcal{F}$ be a $\sigma-$algebra of subsets of Ω. A probability measure $\mathbb{P}$ is a function that to every set $A \in \mathcal{F}$ assigns a number in $[0, 1]$, called the probability of the event A and denoted by $\mathbb{P}(A)$. We require:*
i) $\mathbb{P}(\Omega) = 1$
ii) whenever $A_1, A_2, \ldots$ is a sequence of disjoint sets in $\mathcal{F}$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum \mathbb{P}(A_n)$.

• The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called the probability space.
• The first requirement is just a normalizing property.
• The second requirement is called the countable additivity property.
The following are easy consequences of the definition.

- $\mathbb{P}(\phi) = 0$.
- Finite additivity holds.
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
- If $A \subset B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$. This is called the monotonicity property.
- If $\{A_i\}$ is a sequence of events not necessarily disjoint then $\mathbb{P}(\cup_{i=1}^{\infty})A_i \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ (countable subadditivity).

Result: Let $A_1 \subset A_2 \subset A_3 \ldots$ be an increasing sequence of events. Then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \lim\limits_{n\to\infty} \mathbb{P}(A_n)$. This is called the continuity from below property.
Proof: Exercise.
Result: Let $A_1 \supset A_2 \supset A_3 \ldots$ be a decreasing sequence of events. Then $\mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim\limits_{n\to\infty} \mathbb{P}(A_n)$. This is called the continuity from above property.
Proof: Exercise.
A Probabilistic Model for Tossing a Coin Infinitely Many Times:
• $\Omega = \{\omega = \{\omega_n\} : \omega_n = HorT\}$.
• Let $\mathcal{F}_\infty$ be the $\sigma-$algebra generated by sets which can be described in terms of finitely many coin tosses. Let $A = \{\{\omega_n\} : \omega_1 = H\}$, $B = \{\{\omega_n\} : \omega_1 = H, \omega_2 = H\}$. Clearly $A, B \in \mathcal{F}_\infty$. Let $C = \{\{\omega_n\} : \lim_{n\to\infty} \frac{H_n(\omega)}{n} = 1/2\}$, where $H_n(\omega)$ is the number of heads in first n coin tosses. Clearly C is not determined by finitely many coin tosses. For fixed n and m, define

$$C_{n,m} = \{\omega : \Big|\frac{H_n(\omega)}{n} - 1/2\Big| \leq 1/m\} \,.$$

Clearly $C_{n,m} \in \mathcal{F}_\infty$. From the definition of limit

$$C = \cap_{m=1}^{\infty} \cup_{N=1}^{\infty} \cap_{n=N}^{\infty} C_{n,m} \,.$$

Thus $C \in \mathcal{F}_\infty$.

1.2 Random Variable

Let $\mathcal{B}(\mathbb{R})$ denote the Borel $\sigma-$algebra of $\mathbb{R}$, the $\sigma-$algebra generated by closed intervals.

Definition 1.4. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A function $X : \Omega \to \mathbb{R}$ is said to be a real-valued random variable if $X^{-1}(A) = \{\omega : X(\omega) \in A\} \in \mathcal{F}$ for all $A \in \mathcal{B}(\mathbb{R})$.*

Proposition 1.5. *Let $\mathcal{G}$ be a $\sigma-$algebra generated by a collection $\mathcal{S}$. If $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{S}$, then $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{G}$.*

Proof: Let
$$\mathcal{A} = \{A : X^{-1}(A) \in \mathcal{F}\}.$$
Then $\mathcal{S} \subset \mathcal{A}$. We will now show that $\mathcal{A}$ is a $\sigma-$algebra. $X^{-1}(A^c) = (X^{-1}(A))^c$ Thus if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$. Again, $X^{-1}(\cup A_i) = \cup X^{-1}(A_i)$. Thus if $A_i \in \mathcal{A}$ then $\cup A_i \in \mathcal{A}$. Hence we have shown that $\mathcal{A}$ is a $\sigma-$algebra, as a result $\sigma(\mathcal{S}) = \mathcal{G} \subset \mathcal{A}$. $\square$

Fact: $\sigma-$algebra generated by open rays is $\mathcal{B}(\mathbb{R})$. Thus if $\mathcal{S} = \{(r, \infty) : r \in \mathbb{R}\}$, then $\sigma(\mathcal{S}) = \mathcal{B}(\mathbb{R})$.

Thus using the above proposition X is a random variable iff $X^{-1}(r, \infty) \in \mathcal{F}$ for all $r \in \mathbb{R}$.

Lemma 1.6. *If X and Y are random variables then $X + Y$ is also a random variable.*

Proof:

$$\begin{aligned}
&X + Y > r \\
&\Leftrightarrow X > r - Y \\
&\Leftrightarrow \text{there exists a rational no. q such that } X > q > r - Y.
\end{aligned}$$

Thus $(X + Y)^{-1}(r, \infty) = \cup_{q \in \mathbb{Q}} X^{-1}(q, \infty) \cap Y^{-1}(r - q, \infty)$. $\square$

Exercise: Let A be a subset of Ω. Then show that 1_A is a random variable if and only if $A \in \mathcal{F}$.

Exercise: Prove that if X is a random variable then X^2 is also a random variable. Prove that if X and Y are random variables then XY is also a random variable.

Exercise Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f^{-1}(B) \in \mathcal{B}(\mathbb{R})$ for all $B \in \mathcal{B}(\mathbb{R})$ (such a function is called Borel measurable function). Show that if X is a random variable then $f(X)$ is also a random variable.

Definition 1.7. *Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. The distribution measure of X is the probability measure μ_X on $(\mathbb{R}, \mathcal{B}(\mathbb{R})$ that assigns to each Borel set B of $\mathbb{R}$ the measure $\mu_X(B) = \mathbb{P}(X \in B)$.*

Two examples: Let $\Omega = [0, 1]$, $\mathcal{F} = \sigma-$algebra generated by closed intervals. Let $\mathbb{P}$ be the probability measure which assigns to each interval its length. Define $X(\omega) = \omega$ and $Y(\omega) = 1 - \omega$. Then

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

Now
$$\mu_Y[a, b] = \mathbb{P}(Y \in [a, b]) = \mathbb{P}(a \le 1 - \omega \le b) = \mathbb{P}[1 - b, 1 - a] = b - a.$$

Thus $\mu_X = \mu_Y$ (X and Y are said to be uniformly distributed.) Now suppose we define another probability measure $\tilde{\mathbb{P}}$ by $\tilde{\mathbb{P}}[a, b] = \int_a^b 2\omega d\omega = b^2 - a^2$. Then $\tilde{\mu}_X[a, b] = b^2 - a^2$, whereas $\tilde{\mu}_Y[a, b] = (1 - a)^2 - (1 - b)^2$. Thus under $\tilde{\mathbb{P}}$, X and Y does not have the same distribution.

Definition 1.8. *The distribution function of a random variable X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the function $F : \mathbb{R} \to [0, 1]$ given by*

$$F(x) = \mathbb{P}(X \le x).$$

Proposition 1.9. *The distribution function of a random variable has the following properties:*

(1) $F_X(\cdot)$ *is non-decreasing and hence has only jump discontinuities.*

(2) $\lim_{x\uparrow\infty} F_X(x) = 1$, $\lim_{x\downarrow -\infty} F_X(x) = 0$.

(3) $\lim_{h\downarrow 0} F_X(x+h) = F_X(x), \forall x \in \mathbb{R}$, *thus CDF is right continuous.*

(4) $\lim_{h\downarrow 0} F_X(x-h) = F_X(x) - P(X = x), \forall x \in \mathbb{R}$.

Theorem 1.10. *Let F be a function from $\mathbb{R}$ to $[0,1]$ satisfying the properties of the above proposition, then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable X defined on it whose distribution function is F.*

Two Special Cases

• There exists a non-negative function f on $\mathbb{R}$ such that

$$\mu_X[a,b] = \mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx\,.$$

Thus

$$1 = \mathbb{P}(X \in \mathbb{R}) = \lim_{n\to\infty} \mathbb{P}(-n \leq X \leq n) = \lim_{n\to\infty} \int_{-n}^{n} f(x)dx = \int_{-\infty}^{\infty} f(x)dx\,.$$

• X takes only countably many values x_i. Define $p_i = \mathbb{P}(X = x_i)$. Then

$$\mu_X(B) = \sum_{\{i:x_i \in B\}} p_i\,.$$

In the first case X is said to have an absolutely continuous distribution with probability density function f and in the second case X is said to have a discrete distribution with probability mass function $\{p_i\}$.

Example: Consider the functions:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}\,, N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy\,.$$

Let X be uniformly distributed on $[0,1]$. Notice that N is a strictly increasing function. So it has an inverse N^{-1}. Define the random variable $Z = N^{-1}(X)$. Then

$$\begin{aligned}
\mu_Z[a,b] &= \mathbb{P}(\omega \in \Omega : a \leq Z(\omega) \leq b)\\
&= \mathbb{P}(\omega \in \Omega : a \leq N^{-1}(X)(\omega) \leq b)\\
&= \mathbb{P}(\omega \in \Omega : N(a) \leq NN^{-1}(X)(\omega) \leq N(b))\\
&= \mathbb{P}(\omega \in \Omega : N(a) \leq X(\omega) \leq N(b))\\
&= N(b) - N(a) = \int_a^b \varphi(x)dx\,.
\end{aligned}$$

1.3 Expectation

We want to compute an "average value" of X, where we take the probabilities into account while computing the average.

Definition 1.11. *A random variable X is said to be simple if it is non-negative and it takes only finitely many values. Suppose X takes the values $a_1, a_2, \ldots, a_n$ and $A_i = \{\omega : X(\omega) = a_i\}$. Then X can be written as*

$$X = \sum_{i=1}^{n} a_i 1_{A_i} \,. \tag{1.1}$$

Definition 1.12. *Let X be a simple random variable having the representation* (1.1)*, then the expectation of X is defined to be*

$$\mathbb{E}(X) = \int_\Omega X d\mathbb{P} = \sum_{i=1}^{n} a_i \mathbb{P}(A_i) \,.$$

Definition 1.13. *Let X be any non-negative random variable. We define the expectation of X to be*

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

where the supremum is over all simple random variables Y such that $Y \leq X$. (Note that the supremum can be infinite as well.)

Let X a random variable. Then $X^+ = X \vee 0$ and $X^- = (-X) \vee 0$. Notice that $X = X^+ - X^-$. We say that X is integrable if $\mathbb{E}(X^+) < \infty$ and $\mathbb{E}(X^-) < \infty$.

Definition 1.14. *If X is an integrable random variable then the expectation of X is defined to be*

$$\mathbb{E}(X) = \int_\Omega X d\mathbb{P} = \mathbb{E}(X^+) - \mathbb{E}(X^-) \,.$$

Definition 1.15. *Let X be an integrable random variable or a non-negative random variable and $A \in \mathcal{F}$. Then*

$$\int_A X d\mathbb{P} = \int_\Omega X 1_A d\mathbb{P} \,.$$

Proposition 1.16. *Suppose $X \geq 0$ and $\mathbb{E}(X) = 0$. Then $\mathbb{P}(X = 0) = 1$.*

Proof: Let $E = \{\omega \in \Omega : X(\omega) > 0\}$ and let $E_n = \{\omega \in \Omega : X(\omega) \geq 1/n\}$. Then by definition $0 = \mathbb{E}(X) \geq \int \frac{1}{n} 1_{E_n} d\mathbb{P} = \frac{1}{n}\mathbb{P}(E_n)$. Thus $\mathbb{P}(E_n) = 0$. Hence

$$\mathbb{P}(E) = \mathbb{P}(\cup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(E_n) = 0 \,.$$

Thus the conclusion. $\square$

Exercise: Suppose $X > 0$ on A and $\int_A X d\mathbb{P} = 0$, then show that $\mathbb{P}(A) = 0$.

Theorem 1.17. *Suppose X, Y are either non-negative or integrable random variables. The the following properties hold.*

1. *If $X \leq Y$, then $\mathbb{E}(X) \leq \mathbb{E}(Y)$.*
2. *$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.*
3. *If X is non-negative and $a \geq 0$, or if X is integrable and $a \in \mathbb{R}$, then $\mathbb{E}(aX) = a\mathbb{E}(X)$.*
4. *If X is non-negative and $A \subset B$, then $\int_A X d\mathbb{P} \leq \int_B X d\mathbb{P}$.*

Exercise: Prove properties 1, 3 and 4 of the above theorem.
Exercise: Show that X is integrable if and only if $\mathbb{E}(|X|) < \infty$.

Theorem 1.18 (Monotone Convergence Theorem)**.** *Let $\{X_n\}$ be a sequence of non-negative non-decreasing random variables, i.e., $0 \leq X_1 \leq X_2 \leq X_3 \ldots$. Further suppose that X_n converges to X almost surely, then*

$$\lim_{n\to\infty} \mathbb{E}(X_n) = \mathbb{E}(X)\,.$$

Exercise: Suppose that a non-negative random variable X takes only countably many values $x_0, x_1, \ldots$. Then show that

$$\mathbb{E}(X) = \sum_{k=0}^{\infty} x_k \mathbb{P}(X = x_k)\,.$$

Theorem 1.19 (Dominated Convergence Theorem)**.** *Let $X_1, X_2, \ldots$ be a sequence of random variables converging almost surely to another random variable X. Suppose there exists another non-negative random variable Y such that $\mathbb{E}(Y) < \infty$ and $|X_n| \leq Y$, then*

$$\lim_{n\to\infty} \mathbb{E}(X_n) = \mathbb{E}(X)\,.$$

The space Ω is very abstract and not very pleasant to compute integrals. So a relationship between integration on Ω and integration on $\mathbb{R}$ will be helpful.

Theorem 1.20. *Let X be a random variable and let g be a Borel measurable function. Then*

$$\mathbb{E}(|g(X)|) = \int_{\mathbb{R}} |g(x)| d\mu_X(x)\,,$$

and if the above quantity is finite then

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) d\mu_X(x)\,.$$

Theorem 1.21. *Let X be a random variable having probability density function f and let g be a Borel measurable function. Then*

$$\mathbb{E}(|g(X)|) = \int_{\mathbb{R}} |g(x)| f(x) dx\,,$$

and if the above quantity is finite then

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) f(x) dx\,.$$

1.4 Filtration

Imagine that a random experiment has been performed and the outcome is a particular ω in the set of outcomes Ω. We are given some information, not enough to know the precise ω but to narrow down the possibilities. For example, the true ω may be the outcome of three coin tosses and we are told the outcome of only the first toss. Then we can make a list of sets which for sure contain it and those that do not contain it. These are the sets that are *resolved* by the first toss. Let

$$A_H = \{HHH, HHT, HTH, HTT\}, A_T = \{THH, TTH, THT, TTT\}.$$

It is easy to see these sets are resolved. But $A_{HH} = \{HHT, HHH\}$ is not resolved. Define $\mathcal{F}_1 = \{\phi, \Omega, A_H, A_T\}$. Then this the "information gained by observing the first toss". Similarly define $\mathcal{F}_2 = \{\text{sets resolved by knowing the first and second tosses}\}$. (List the sets.) Once we are told all the three coin tosses we know the precise ω and all the sets are resolved. Thus $\mathcal{F}_3 = \mathcal{P}(\Omega)$. Notice that $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ is an increasing sequence of $\sigma-$algebras.

Definition 1.22. *Let Ω be a non-empty set. Let T be a fixed positive number. Assume that for each $t \in [0,T]$ there is a $\sigma-$algebra $\mathcal{F}_t$ such that for $s < t$, $\mathcal{F}_s \subset \mathcal{F}_t$. Then we call this collection of $\sigma-$algebras $\mathcal{F}_t, t \in [0,T]$ a filtration.*

Example: Suppose $\Omega = C_0[0,T]$, continuous functions on $[0,T]$, with value 0 at the point 0. Then let $\mathcal{F}_t$ be the $\sigma-$algebra of all those sets which are resolved by observing the function upto time t. So the random experiment is choosing an element of $C_0[0,T]$. Let $\bar{\omega}$ be the true outcome. Suppose we know the value of $\bar{\omega}$ for $0 \leq s \leq t$, then the set $\{\omega \in \Omega : \sup_{0 \leq s \leq t} \omega(s) \leq 1\}$ is resolved whereas the set $\{\omega \in \Omega : \omega(T) > 0\}$ is not resolved. The first set belongs to $\mathcal{F}_t$ whereas the second does not.

Definition 1.23. *Let X be a random variable. Then the $\sigma-$algebra generated by X, denoted by $\sigma(X)$ is the collection of all subsets of Ω of the form $X^{-1}(B)$ where B ranges over all Borel subsets of $\mathbb{R}$.*

Definition 1.24. *Let X be a random variable on $(\Omega, \mathcal{F})$. Let $\mathcal{G}$ be a $\sigma-$algebra on Ω. Then X is said to be $\mathcal{G}$ measurable if $\sigma(X) \subset \mathcal{G}$. Thus X is also a random variable on $(\Omega, \mathcal{G})$.*

Thus $\sigma(X)$ is the smallest $\sigma-$algebra with respect to which X is measurable.
Example: Suppose $S_2(HHH) = S_2(HHT) = 10$, $S_2(TTH) = S_2(TTT) = 1$ and $S_2(HTH) = S_2(THT) = S_2(THH) = S_2(THT) = 5$. Then S_2 is $\mathcal{F}_2$ measurable.
Exercise: Suppose X is a constant random variable. Then write down $\sigma(X)$.

Definition 1.25. *Let Ω be a non-empty sample space with a filtration $\mathcal{F}_t, 0 \leq t \leq T$. A sequence of random variables $\{X(t)\}$ indexed by $t \in [0,T]$ is said to be an adapted stochastic process, if for each t, $X(t)$ is $\mathcal{F}_t$ measurable.*

1.5 Independence

Definition 1.26. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two sub-$\sigma-$algebras of $\mathcal{F}$. We say that these two $\sigma-$algebras are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all*

$A \in \mathcal{G}_1$ and $B \in \mathcal{G}_2$. Let X and Y be two random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. We say that X and Y are independent if $\sigma(X)$ and $\sigma(Y)$ are independent. We say that the random variable X is independent of the sub-$\sigma-$algebra $\mathcal{G}$ if $\sigma(X)$ and $\mathcal{G}$ are independent.

Definition 1.27. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}_1, \mathcal{G}_2, \ldots$ be a sequence of sub-$\sigma-$algebras of $\mathcal{F}$. For a fixed n, we say that the n $\sigma-$algebras $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ are independent if $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2) \cdots \mathbb{P}(A_n)$ for all $A_1 \in \mathcal{G}_1, A_2 \in \mathcal{G}_2, \ldots, A_n \in \mathcal{G}_n$. We say that the full sequence of $\sigma-$algebras $\mathcal{G}_1, \mathcal{G}_2, \ldots$ is independent if for any positive integer n, $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ are independent. Similarly, a sequence of random variables $X_1, X_2, \ldots$ is independent if $\sigma(X_1), \sigma(X_2), \ldots$ is independent.*

Theorem 1.28. *Let X and Y be two independent random variables, and let f and g be two Borel measurable functions. Then $f(X)$ and $g(Y)$ are also independent.*

Proof: We need to show that $\sigma(f(X))$ and $\sigma(g(Y))$ are independent. Let $A \in \sigma(f(X))$. Then there exists a Borel set C such that $A = \{\omega \in \Omega : f(X(\omega)) \in C\}$. Let $D = \{x \in \mathbb{R} : f(x) \in C\}$. Then

$$A = \{\omega \in \Omega : f(X(\omega)) \in C\} = \{\omega \in \Omega : X(\omega) \in D\}\,.$$

Thus $A \in \sigma(X)$. Similarly, if we take any $B \in \sigma(g(Y))$, then we can show that $B \in \sigma(Y)$. Hence the result follows. $\square$

Definition 1.29. *A set $A \in \mathbb{R}^n$ is said to be a measurable rectangle if there exist Borel sets $A_1, A_2, \ldots, A_n$ such that $A = A_1 \times A_2 \times \cdots \times A_n$. The sigma$-$algebra on $\mathbb{R}^n$ generated by measurable rectangles is called the Borel $\sigma-$algebra on $\mathbb{R}^n$ and denoted by $\mathcal{B}(\mathbb{R}^n)$.*

Definition 1.30. *Let X and Y be two random variables. The pair (X, Y) takes values in $\mathbb{R}^2$. The joint distribution measure of (X, Y) is given by*

$$\mu_{X,Y}(C) = \mathbb{P}((X, Y) \in C)\ \forall\ C \in \mathcal{B}(\mathbb{R}^2)\,.$$

The joint distribution function of (X, Y) is given by

$$F_{X,Y}(a, b) = \mu_{X,Y}((-\infty, a] \times (-\infty, b]) = \mathbb{P}(X \le a, Y \le b),\ \forall\ a, b \in \mathbb{R}\,.$$

We say that a non-negative, Borel measurable function $f_{X,Y}(\cdot)$ is a joint density for the pair of random variables (X, Y) if

$$\mu_{X,Y}(C) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_C(x, y) f_{X,Y}(x, y)dxdy\ \forall\ C \in \mathcal{B}(\mathbb{R}^2)\,.$$

The above condition holds iff

$$F_{X,Y}(a, b) = \int_{-\infty}^{a}\int_{-\infty}^{b} f_{X,Y}(x, y)dxdy\,.$$

The distribution measures of X and Y can be recovered from the joint distribution in the following way.

$$\mu_X(A) = \mu_{X,Y}(A \times \mathbb{R}),\ \mu_Y(B) = \mu_{X,Y}(\mathbb{R} \times B)\,.$$

μ_X and μ_Y are called the marginal distributions of $\mu_{X,Y}$. If joint densities exist then marginal densities exist as well.

$$\mu_X(A) = \mu_{X,Y}(A \times \mathbb{R}) = \int_A \left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \right) dx,$$
$$\mu_Y(B) = \mu_{X,Y}(\mathbb{R} \times B) = \int_B \left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx \right) dy\,.$$

But the converse is not true.
Counter Example: Let $X \sim N(0,1)$ and $Z \sim Bernoulli(1/2)$ independent of X. Define $Y = XZ$. Now

$$\begin{aligned} F_Y(b) &= \mathbb{P}(Y \le b) \\ &= (1/2)\mathbb{P}(Y \le b | Z = 1) + (1/2)\mathbb{P}(Y \le b | Z = -1) \\ &= (1/2)\mathbb{P}(X \le b) + (1/2)\mathbb{P}(-X \le b) \\ &= (1/2)\left[\int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx + \int_{-b}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \right] \\ &= \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx\,. \end{aligned}$$

Thus Y is again $N(0,1)$. Thus both X and Y have densities. But note that $|X| = |Y|$. So if we define $C = \{(x,y) \in \mathbb{R}^2 : y = \pm x\}$. Then $\mu_{X,Y}(C) = \mathbb{P}((X,Y) \in C) = 1$. But since C has area zero in $\mathbb{R}^2$, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1_C(x,y) f_{X,Y}(x,y) dx dy = 0$ for any f. Hence (X,Y) can not have a joint density.

Theorem 1.31. *Let X and Y be two random variables. The following conditions are equivalent.*

1. *X and Y are independent.*
2. *The joint distribution measure is the product of marginal distributional measures, i.e.,*
$$\mu_{X,Y}(A \times B) = \mu_X(A)\mu_Y(B)\ \forall\ A, B \in \mathcal{B}(\mathbb{R})\,.$$
3. *The joint distribution function factors, i.e.,*
$$F_{X,Y}(a,b) = F_X(a)F_Y(b)\ \forall\ a, b \in \mathbb{R}\,.$$
4. *The joint moment generating function factors, i.e.,*
$$\mathbb{E}\left(e^{uX+vY}\right) = \mathbb{E}\left(e^{uX}\right)\mathbb{E}\left(e^{vY}\right)\ \ \forall\ u, v \in \mathbb{R}\,,$$
for which the expectations are finite.

If there is a joint density then each of the above conditions are equivalent to the following:
5. The joint density factors, i.e.,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)\,.$$

Proof: ($1 \Rightarrow 2$) Assume that X and Y are independent. Then

$$\begin{aligned}\mu_{X,Y}(A \times B) &= \mathbb{P}(X \in A, Y \in B) \\ &= \mathbb{P}(X \in A)\mathbb{P}(Y \in B) = \mu_X(A)\mu_Y(B)\,.\end{aligned}$$

($2 \Rightarrow 3$)

$$F_{X,Y}(a,b) = \mu_{X,Y}((-\infty,a] \times (-\infty,b]) = \mu_X((-\infty,a])\mu_Y((-\infty,b]) = F_X(a)F_Y(b)\,.$$

($3 \Rightarrow 5$) Rewriting the splitting of distribution function in terms of density we get,

$$\int_{-\infty}^{b}\int_{-\infty}^{a} f_{X,Y}(x,y)dxdy = \int_{-\infty}^{b} f_Y(y)dy \int_{-\infty}^{a} f_X(x)dx$$

Differentiating with respect to y we get,

$$\int_{-\infty}^{a} f_{X,Y}(x,b)dx = f_Y(b)\int_{-\infty}^{a} f_X(x)dx\,.$$

Further differentiating with respect to x we get,

$$f_{X,Y}(a,b) = f_X(a)f_Y(b)\,.$$

($5 \Rightarrow 1$)

$$\begin{aligned}\mathbb{P}(X \in A, Y \in B) &= \int_A\int_B f_{X,Y}(x,y)dxdy \\ &= \int_A\int_B f_X(x)f_Y(y)dxdy \\ &= \int_A f_X(x)dx \int_B f_Y(y)dy \\ &= \mathbb{P}(X \in A)\mathbb{P}(Y \in B)\,.\end{aligned}$$

□

Definition 1.32. *Let X be a random variable such that $\mathbb{E}(X^2) < \infty$. Then the variance of X, denoted by $Var(X)$ is given by*

$$Var(X) = \mathbb{E}(X - \mathbb{E}(X))^2\,.$$

The standard deviation of X is given by $\sqrt{Var(X)}$.

Exercise: Show that $Var(X) = 0$ if and only if there exists a constant c such that $\mathbb{P}(X = c) = 1$.

Definition 1.33. *Let X and Y be two random variables such that $Var(X)$ and $Var(Y)$ are finite. Then the co-variance of X and Y is given by*

$$Cov(X,Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right)\,.$$

Suppose X and Y are not constant random variables, then the correlation co-efficient of X and Y is given by

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}\,.$$

X and Y are said to be uncorrelated if $Cov(X, Y) = 0$.
Note that, if X and Y are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ and thus X and Y are uncorrelated. But the converse is not true. For a counter example consider the counter example given to show that joint density may not exist even if marginal densities exist. In that example, we have, $\mathbb{E}(X) = \mathbb{E}(Y) = 0$. Also $\mathbb{E}(XY) = \mathbb{E}(X^2Z) = \mathbb{E}(X^2)\mathbb{E}(Z) = 0$. Thus X and Y are uncorrelated but clearly not independent.

1.6 Few Important Inequalities

Holder's Inequality Let $1 \leq p < \infty$ and $1 \leq q < \infty$ be such that $1/p + 1/q = 1$. Further assume that $\mathbb{E}(|X|^p)$ and $\mathbb{E}(|Y|^q)$ are finite. Then

$$\mathbb{E}(|XY|) \leq (\mathbb{E}(|X|^p))^{\frac{1}{p}} (\mathbb{E}(|Y|^q))^{\frac{1}{q}} .$$

with equality if and only if $X = cY$ for some constant c. The special case where $p = q = 2$ is known as Cauchy Schwartz inequality.
Exercise: Use Cauchy-Schwartz inequality to show that $-1 \leq \rho \leq 1$. Also show that $\rho = \pm 1$ if and only if there exist constants a and b such that $Y = aX + b$.
Jensen's Inequality Let X be a random variable such that $\mathbb{E}|X| < \infty$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function, i.e.,

$$\varphi(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda\varphi(x_1) + (1 - \lambda)\varphi(x_2) .$$

for all $x_1, x_2 \in \mathbb{R}$ and for all $0 \leq \lambda \leq 1$. Also assume that $\mathbb{E}|\varphi(X)| < \infty$. Then

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X)) .$$

Equality occurs if and only if φ is linear.
Proof: We will give the proof under the additional assumption that φ is differentiable at $x = \mathbb{E}(X)$. In this case the tangent at $x = \mathbb{E}(X)$ lies completely below the graph of the function. Let $l(x) = ax + b$ be the equation of the tangent to φ at $x = \mathbb{E}(X)$. Then $\varphi(x) \geq l(x)$ for all x. So

$$\mathbb{E}(\varphi(X)) \geq \mathbb{E}(l(X)) = a\mathbb{E}(X) + b = l(\mathbb{E}(X)) = \varphi(\mathbb{E}(X)) .$$

$\square$

1.7 Conditional Expectation

Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Further let $\mathcal{G}$ be a sub-$\sigma-$algebra. If X is $\mathcal{G}$ measurable then the information in $\mathcal{G}$ is enough to determine the value of X. If X is independent of $\mathcal{G}$ then the information in $\mathcal{G}$ provides no help in determining X. In the intermediate case we can use the information in $\mathcal{G}$ to estimate X but not precisely evaluate X. The conditional expectation of X given $\mathcal{G}$ is such an estimate.

Definition 1.34. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma-$algebra. Let X be a random variable which is either non-negative or integrable. The conditional expectation of X given $\mathcal{G}$ denoted by $\mathbb{E}[X|\mathcal{G}]$ is any random variable that satisfies*
i) (Measurability Property) $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$ measurable.
ii) (Partial Averaging Property) $\int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_A X d\mathbb{P}$, for all $A \in \mathcal{G}$.

Uniqueness of conditional expectation: Suppose there are two random variables Y and Z which satisfy the above two properties. Since both Y and Z are $\mathcal{G}$ measurable the set $A = \{Y - Z > 0\} \in \mathcal{G}$. But by the second property,

$$\int_A Y d\mathbb{P} = \int_A Z d\mathbb{P} = \int_A X d\mathbb{P} .$$

Thus $\int_A (Y - Z)d\mathbb{P} = 0$, implies $\mathbb{P}(A) = 0$. Similarly, we can show that if $B = \{Z - Y > 0\}$, then $\mathbb{P}(B) = 0$. Thus we get $\mathbb{P}(Y = Z) = 1$.
An Important Consequence: Suppose X is integrable then, putting $A = \Omega$ in the partial averaging property we get

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}[X|\mathcal{G}]) .$$

Theorem 1.35. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma-$algebra.*

- *[1] (Linearity) Let X and Y be two random variables and let c_1 and c_2 be two constants, then $\mathbb{E}[c_1 X + c_2 Y|\mathcal{G}] = c_1\mathbb{E}[X|\mathcal{G}] + c_2\mathbb{E}[Y|\mathcal{G}]$.*
- *[2] (Taking out what is known) If X is $\mathcal{G}$ measurable then $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$.*
- *[3] (Tower Property) If $\mathcal{H}$ is a sub-$\sigma-$algebra such that $\mathcal{H} \subset \mathcal{G}$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.*
- *[4] (Independence) If X is independent of $\mathcal{G}$ then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}(X)$*
- *[5] (Conditional Jensen's Inequality) If φ is a convex function then $\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}]$.*

Proof: Proof of [1] is exercise.
Proof of [2]: Measurability is trivial. Suppose that $X = 1_B$ where $B \in \mathcal{G}$. Then for any $A \in \mathcal{G}$ we have

$$\begin{aligned}\int_A X\mathbb{E}[Y|\mathcal{G}]d\mathbb{P} &= \int_A 1_B\mathbb{E}[Y|\mathcal{G}]d\mathbb{P} \\ &= \int_{A\cap B} \mathbb{E}[Y|\mathcal{G}]d\mathbb{P} = \int_{A\cap B} Y d\mathbb{P} = \int_A XY d\mathbb{P} .\end{aligned}$$

Proof of [3]: Again measurability is trivial. Now suppose $A \in \mathcal{H} \subset \mathcal{G}$. Then

$$\begin{aligned}\int_A \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]d\mathbb{P} &= \int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} \\ &= \int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{H}]d\mathbb{P} .\end{aligned}$$

Hence the result.
Proof of [4]: Again measurability is trivial since $\mathbb{E}(X)$ is constant. Now for any $A \in \mathcal{G}$ we have

$$\int_A \mathbb{E}(X) d\mathbb{P} = \mathbb{E}(X)\mathbb{E}(1_A) = \mathbb{E}(X 1_A) = \int_A X d\mathbb{P}\,.$$

Proof of [5]: Same as Jensen's inequality with conditional expectation replacing expectation.
□

Exercise: If $X \geq 0$, then show that $\mathbb{P}(\mathbb{E}[X|\mathcal{G}] \geq 0) = 1$.

Theorem 1.36. *$\mathbb{E}[X|\mathcal{G}]$ is the best estimate (in terms of mean square error) of X given $\mathcal{G}$, i.e.,*

$$\mathbb{E}(X - \mathbb{E}[X|\mathcal{G}])^2 \leq \mathbb{E}(X - Y)^2\,,$$

for any Y, $\mathcal{G}$ measurable.

Proof: Exercise.
Fact: Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let Y be another random variable which is $\sigma(X)$ measurable. Then there exists a Borel measurable function g such that $Y = g(X)$.
Consequence: $\mathbb{E}[Y|X] \doteq \mathbb{E}[Y|\sigma(X)] = g(X)$ for some g Borel measurable.
Example: Suppose X and Y have joint density $f_{X,Y}(\cdot)$. Further assume that the marginal $f_X(\cdot)$ is strictly positive for all x. Now define the function

$$g(x) = \frac{\int_{-\infty}^{\infty} y f_{X,Y}(x,y) dy}{f_X(x)}\,.$$

Then $\mathbb{E}[Y|X] = g(X)$.
Again measurability is trivial. Take $A \in \sigma(X)$. Then there exists B such that $A = \{\omega : X(\omega) \in B\}$. Then

$$\begin{aligned}
\int_A g(X) d\mathbb{P} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_B(x) g(x) f_{X,Y}(x,y) dy dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_B(x) \frac{\int_{-\infty}^{\infty} t f_{X,Y}(x,t) dt}{f_X(x)} f_{X,Y}(x,y) dy dx \\
&= \int_{-\infty}^{\infty} 1_B(x) \frac{\int_{-\infty}^{\infty} t f_{X,Y}(x,t) dt}{f_X(x)} \Big(\int_{-\infty}^{\infty} f_{X,Y}(x,y) dy\Big) dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_B(x) t f_{X,Y}(x,t) dt dx \\
&= \int_{\{X \in B\}} Y d\mathbb{P} = \int_A Y d\mathbb{P}\,.
\end{aligned}$$

Thus we are done.

Definition 1.37. *The conditional variance of X given $\mathcal{G}$ is defined by*

$$Var[X|\mathcal{G}] = \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2|\mathcal{G}]\,.$$

Exercise: i) Prove that $Var(Y) = Var(\mathbb{E}[Y|X]) + \mathbb{E}(Var[Y|X])$.
ii) Prove that X and $Y - \mathbb{E}[Y|X]$ are uncorrelated.

Lemma 1.38. *(Independence Lemma) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma-$algebra. Suppose that the random variables $X_1, X_2, \ldots, X_k$ are $\mathcal{G}$ measurable and $Y_1, \ldots, Y_l$ are independent of $\mathcal{G}$. Then for a Borel measurable function $f : \mathbb{R}^{k+l} \to \mathbb{R}$ define*

$$g(x) = \mathbb{E}(f(x_1, \ldots, x_k, Y_1, \ldots, Y_l))\,.$$

Then

$$\mathbb{E}[f(X_1, \ldots, X_k, Y_1, \ldots, Y_l)|\mathcal{G}] = g(X_1, \ldots, X_k)\,.$$

Definition 1.39. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number and let $\mathcal{F}_t, 0 \le t \le T$ be a filtration of sub-$\sigma-$algebras of $\mathcal{F}$. Consider an adapted stochastic process $M(t), 0 \le t \le T$.*

1. *If $\mathbb{E}[M(t)|\mathcal{F}_s] = M(s)$ for all $0 \le s \le t \le T$ then we say that the process is a martingale.*
2. *If $\mathbb{E}[M(t)|\mathcal{F}_s] \ge M(s)$ for all $0 \le s \le t \le T$ then we say that the process is a sub-martingale.*
3. *If $\mathbb{E}[M(t)|\mathcal{F}_s] \le M(s)$ for all $0 \le s \le t \le T$ then we say that the process is a super-martingale.*

Thus if $M(t)$ is a martingale then $\mathbb{E}(M(t)) = \mathbb{E}(\mathbb{E}[M(t)|\mathcal{F}_0]) = \mathbb{E}(M(0))$ for all $t \ge 0$.

Definition 1.40. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number and let $\mathcal{F}_t, 0 \le t \le T$ be a filtration of sub-$\sigma-$algebras of $\mathcal{F}$. Consider an adapted stochastic process $X(t), 0 \le t \le T$. If for all $0 \le s \le t \le T$ and for every non-negative Borel measurable function f, there exists another Borel measurable function g such that*

$$\mathbb{E}[f(X(t))|\mathcal{F}_s] = g(X(s))\,,$$

then we say that X is a Markov process.

2 Brownian Motion

2.1 Definition and Properties

Let $\{X_n\}$ be a sequence of i.i.d. random variables having distribution

$$X_j = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases} .$$

Define $S_0 = 0$ and $S_n = \sum_{j=1}^{n} X_j, n \geq 1$. Then $\{S_n\}$ is called a simple symmetric random walk. We know that $\{S_n\}$ is a discrete time Markov chain. A simple symmetric random walk also has the following properties:

Independent and Stationary Increments: For non-negative integers $0 = k_0 < k_1 \ldots < k_m$, the random variables

$$S_{k_1} = (S_{k_1} - S_{k_0}), (S_{k_2} - S_{k_1}), \ldots, (S_{k_m} - S_{k_{m-1}})$$

are independent. Moreover, each increment $S_{k_i} - S_{k_{i-1}}$ have mean 0 and variance $k_i - k_{i-1}$.

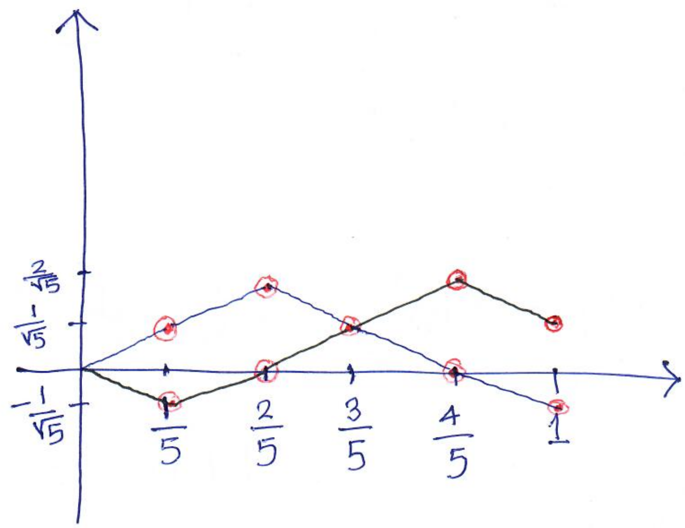
Martingale Property: Let $\mathcal{F}_k = \sigma(X_1, X_2, \ldots, X_k)$. Then clearly S_k is $\mathcal{F}_k$ measurable. For $k < l$ we have,

$$\mathbb{E}[S_l | \mathcal{F}_k] = \mathbb{E}[S_l - S_k + S_k | \mathcal{F}_k] = \mathbb{E}(S_l - S_k) + S_k = S_k \,.$$

Scaled Random Walk: A simple symmetric random walk is a discrete time stochastic process. But using this we can construct a continuous-time stochastic process as follows. Fix a positive integer n. Define

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} S_{nt} \,,$$

if nt is an integer. If nt is not an integer, we define $W^{(n)}(t)$ by linear interpolation of its values at the nearest points s and u to the left and right of t for which ns and nu are integers. Thus you walk n times faster but your step sizes are $\frac{1}{\sqrt{n}}$ times smaller. Brownian Motion is the limit “in a certain sense” of scaled random walks.



Definition 2.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A continuous-time stochastic process $W(t), t \geq 0$ is called a standard Bownian motion if*

- $\mathbb{P}(W(0) = 0) = 1$
- $\mathbb{P}\{\omega : W(\cdot, \omega)$ *is continuous*$\} = 1$.
- *For all $0 = t_0 < t_1 < \ldots < t_m$ the increments $W(t_1) = W(t_1) - W(t_0), W(t_2) - W(t_1), \ldots, W(t_m) - W(t_{m-1})$ are independent and normally distributed with mean 0 and variance $t_i - t_{i-1}, i = 1, \ldots, m$.*

For a fixed ω, the function $W(\cdot, \omega)$ is called a sample path of the Brownian motion.

Thus for any fixed t, $W(t) \sim N(0, t)$. For any two time points $0 \leq s < t$, the covariance of $W(t)$ and $W(s)$ is given by

$$\begin{aligned} Cov(W(t), W(s)) = \mathbb{E}(W(t)W(s)) &= \mathbb{E}\left(W(s)(W(t) - W(s)) + W^2(s)\right) \\ &= \mathbb{E}(W(s))\mathbb{E}(W(t) - W(s)) + \mathbb{E}(W^2(s)) \\ &= 0 + s = t \wedge s\,. \end{aligned}$$

Filtration for Brownian Motion:

Definition 2.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $W(t), t \geq 0$ be a Brownian motion defined on it. A filtration for Brownian motion is a collection of $\sigma-$algebras $\mathcal{F}_t, t \geq 0$ satisfying*

1. *$\mathcal{F}_s \subset \mathcal{F}_t$ for $0 \geq s < t$.*
2. *For each $t \leq 0$, $W(t)$ is $\mathcal{F}_t$ measurable.*
3. *For $0 \leq t < u$, the increment $W(u) - W(t)$ is independent of $\mathcal{F}_t$.*

Example of a Brownian filtration is given by the collection of $\sigma-$algebras $\mathcal{F}_t = \sigma\{W(s) : 0 \leq s \leq t\}$.

Theorem 2.3 (Martingale Property)**.** *Brownian motion is a martingale with respect to a Brownian filtration.*

Proof:

$$\begin{aligned}
\mathbb{E}[W(t)|\mathcal{F}_s] &= \mathbb{E}[W(t) - W(s) + W(s)|\mathcal{F}_s] \\
&= \mathbb{E}[W(t) - W(s)|\mathcal{F}_s] + W(s) \\
&= \mathbb{E}(W(t) - W(s)) + W(s) \\
&= 0 + W(s) = W(s)\,.
\end{aligned}$$

□

Exercise: Show that $W^2(t) - t$ is a martingale with respect to a Brownian filtration.

Theorem 2.4 (Markov Property)**.** *Let $W(t), t \geq 0$ be a Brownian motion and let $\mathcal{F}_t, t \geq 0$ be a filtration for this Brownian motion. Then $W(t), t \geq 0$ is a Markov process.*

Proof: Take $0 \leq s < t$ and let f be a non-negative Borel measurable function. We need to show that there exists another Borel measurable function g such that

$$\mathbb{E}[f(W(t))|\mathcal{F}_s] = g(W(s))\,.$$

$$\mathbb{E}[f(W(t))|\mathcal{F}_s] = \mathbb{E}[f(W(t) - W(s) + W(s))|\mathcal{F}_s]\,.$$

Since $W(t) - W(s)$ is independent of $\mathcal{F}_s$ and $W(s)$ is measurable with respect to $\mathcal{F}_s$, we have by Independence Lemma,

$$\mathbb{E}[f(W(t))|\mathcal{F}_s] = g(W(s))\,,$$

where

$$\begin{aligned}
g(x) &= \mathbb{E}(f(W(t) - W(s) + x)) \\
&= \frac{1}{\sqrt{2\pi(t-s)}} \int f(w + x) e^{-\frac{w^2}{2(t-s)}} dw \\
&= \frac{1}{\sqrt{2\pi\tau}} \int f(y) e^{-\frac{(y-x)^2}{2\tau}} dy\,,
\end{aligned}$$

where $\tau = (t - s)$. □

Define the transition density of Brownian motion by

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(y-x)^2}{2\tau}}\,.$$

Then in terms of the transition density

$$g(x) = \int f(y) p(\tau, x, y) dy\,.$$

Exercise: For $\mu \in \mathbb{R}$, consider the Brownian motion with drift μ:

$$X(t) = W(t) + \mu t\,.$$

Prove that $X(t), t \geq 0$ is a Markov process with respect to the Brownian filtration. Further show that the transition density is given by

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(y-x-\mu\tau)^2}{2\tau}}\,.$$

Exercise: For $\nu \in \mathbb{R}$ and $\sigma > 0$, consider the geometric Brownian:

$$X(t) = e^{\sigma W(t) + \nu t}\,.$$

Prove that $X(t), t \geq 0$ is a Markov process with respect to the Brownian filtration. Further show that the transition density is given by

$$p(\tau, x, y) = \frac{1}{\sigma y\sqrt{2\pi\tau}} e^{-\frac{(\log(y/x)-\nu\tau)^2}{2\sigma^2\tau}}\,.$$

Fact: A very important and interesting property of Brownian motion is that Brownian motion sample paths are almost surely no-where differentiable, i.e.,

$$\mathbb{P}\{\omega : W(\cdot, \omega) \text{ is no-where differentiable}\} = 1.$$

Exponential Martingale:

Theorem 2.5. *Let $W(t), t \geq 0$ be a Brownian motion with filtration $\mathcal{F}_t, t \geq 0$ and let σ be a constant. Then the process*

$$Z(t) = e^{\sigma W(t) - \frac{1}{2}\sigma^2 t}\,,$$

is a martingale with respect to $\mathcal{F}_t$.

Proof: We have for $s < t$,

$$\begin{aligned}
\mathbb{E}[Z(t)|\mathcal{F}_s] &= \mathbb{E}[e^{\sigma W(t) - \frac{1}{2}\sigma^2 t}|\mathcal{F}_s] \\
&= \mathbb{E}[e^{\sigma(W(t)-W(s)) + \sigma W(s) - \frac{1}{2}\sigma^2 t}|\mathcal{F}_s] \\
&= \mathbb{E}[e^{\sigma(W(t)-W(s))}] e^{\sigma W(s) - \frac{1}{2}\sigma^2 t} \\
&= e^{\frac{1}{2}\sigma^2(t-s)} e^{\sigma W(s) - \frac{1}{2}\sigma^2 t} = e^{\sigma W(s) - \frac{1}{2}\sigma^2 s}\,.
\end{aligned}$$

Hence the proof. □

Definition 2.6. *(Stopping Time) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a filtration $\mathcal{F}_t, t \geq 0$. A non-negative extended real-valued random variable $T : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ is said to be a stopping time if $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.*

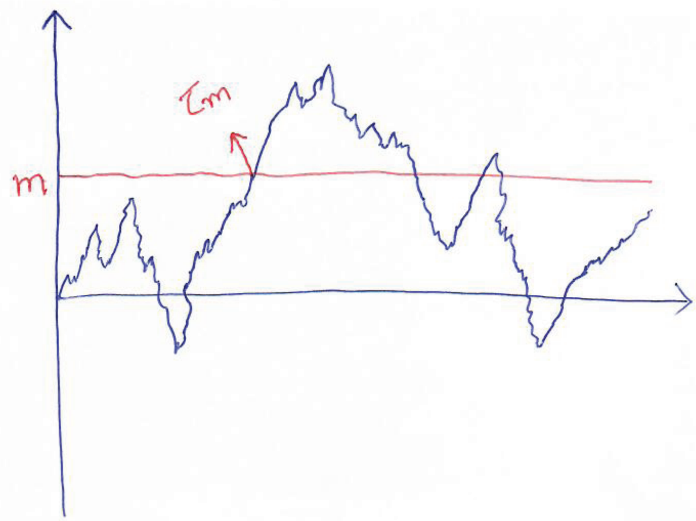
Example of a Stopping time: Let $W(t), t \geq 0$ be a Brownian motion with filtration $\mathcal{F}_t, t \geq 0$. Fix $m \in \mathbb{R}$. Define

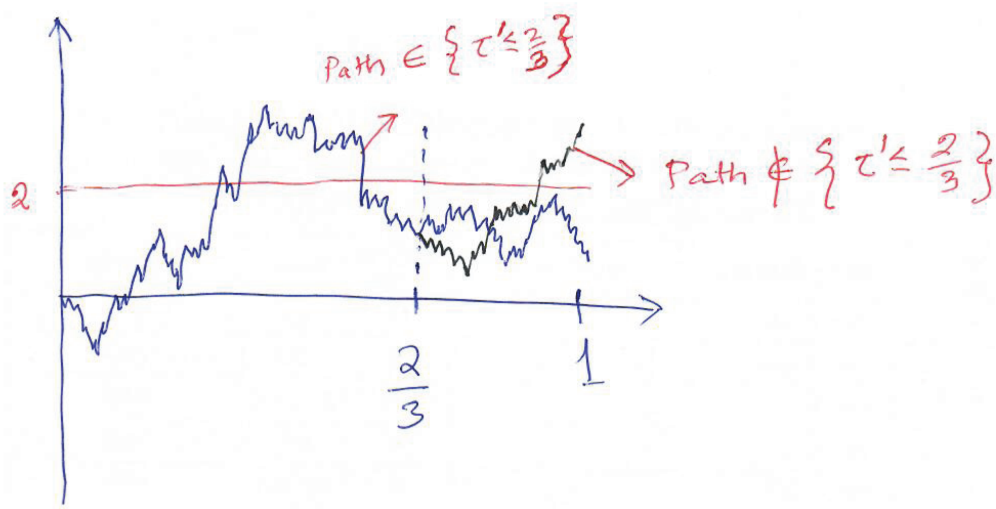
$$\tau_m = \min\{t \geq 0 : W(t) = m\}\,.$$

Then τ_m is a stopping time.
Example of not a Stopping time: Let $W(t), t \geq 0$ be a Brownian motion with filtration $\mathcal{F}_t, t \geq 0$. Fix $m \in \mathbb{R}$. Define

$$\tau' = \sup\{0 \leq t \leq 1 : W(t) = 2\}\,.$$

Then τ' is not a stopping time.





Theorem 2.7. *Let $M(t), t \geq 0$ be a martingale with respect to a filtration $\mathcal{F}_t, t \geq 0$ and let*

T be a stopping time. Then define the stopped process

$$M(t \wedge T) = \begin{cases} M(t) & \forall t < T \\ M(T) & \forall t \geq T. \end{cases}$$

Then $M(t \wedge T), t \geq 0$ is again a martingale.

Fix $m > 0$. Then since Brownian motion is a martingale with respect to the Brownian filtration, the stopped process $W(t \wedge \tau_m)$ is also a martingale, where τ_m is as in example. Thus for $\sigma > 0$ we have

$$1 = \mathbb{E}(Z(0)) = \mathbb{E}(Z(0 \wedge \tau_m)) = \mathbb{E}(Z(t \wedge \tau_m)) = \mathbb{E}\left(e^{\sigma W(t\wedge\tau_m) - \frac{1}{2}\sigma^2(t\wedge\tau_m)}\right). \tag{2.1}$$

Now $0 \leq e^{\sigma W(t\wedge\tau_m) - \frac{1}{2}\sigma^2(t\wedge\tau_m)} \leq e^{\sigma m}$. So if $\tau_m < \infty$, the term $e^{-\frac{1}{2}\sigma^2(t\wedge\tau_m)}$ converges to $e^{-\frac{1}{2}\sigma^2\tau_m}$ as $t \to \infty$. On the other hand if $\tau_m = \infty$, then $e^{-\frac{1}{2}\sigma^2(t\wedge\tau_m)} = e^{-\frac{1}{2}\sigma^2 t} \to 0$ as $t \to \infty$. Thus

$$\lim_{t\to\infty} e^{-\frac{1}{2}\sigma^2(t\wedge\tau_m)} = 1_{\{\tau_m<\infty\}} e^{-\frac{1}{2}\sigma^2\tau_m}.$$

Again if $\tau_m < \infty$, then $e^{\sigma W(t\wedge\tau_m)} \to e^{\sigma W(\tau_m)} = e^{\sigma m}$ as $t \to \infty$. On the otherhand, if $\tau_m = \infty$, then we do not know what happens to $e^{\sigma W(t\wedge\tau_m)}$ but we know that it is bounded. Thus

$$\lim_{t\to\infty} e^{\sigma W(t\wedge\tau_m) - \frac{1}{2}\sigma^2(t\wedge\tau_m)} = 1_{\{\tau_m<\infty\}} e^{\sigma m - \frac{1}{2}\sigma^2\tau_m}.$$

Thus by taking limit $t \to \infty$ on both sides of (2.1), and using dominated convergence theorem we get

$$1 = \mathbb{E}(1_{\{\tau_m<\infty\}} e^{\sigma m - \frac{1}{2}\sigma^2\tau_m}),$$

which implies

$$\mathbb{E}(1_{\{\tau_m<\infty\}} e^{-\frac{1}{2}\sigma^2\tau_m}) = e^{-\sigma m}.$$

Now by MCT we have

$$\begin{aligned} &\lim_{\sigma\downarrow 0} \mathbb{E}(1_{\{\tau_m<\infty\}} e^{-\frac{1}{2}\sigma^2\tau_m}) = 1 \\ &\Rightarrow \mathbb{P}(\tau_m < \infty) = 1. \end{aligned}$$

Thus

$$\mathbb{E} e^{-\frac{1}{2}\sigma^2\tau_m} = e^{-\sigma m}.$$

Theorem 2.8. *Let $m \in \mathbb{R}$, then the first passage time to level m is finite almost surely, and the laplace transform of its distribution is given by*

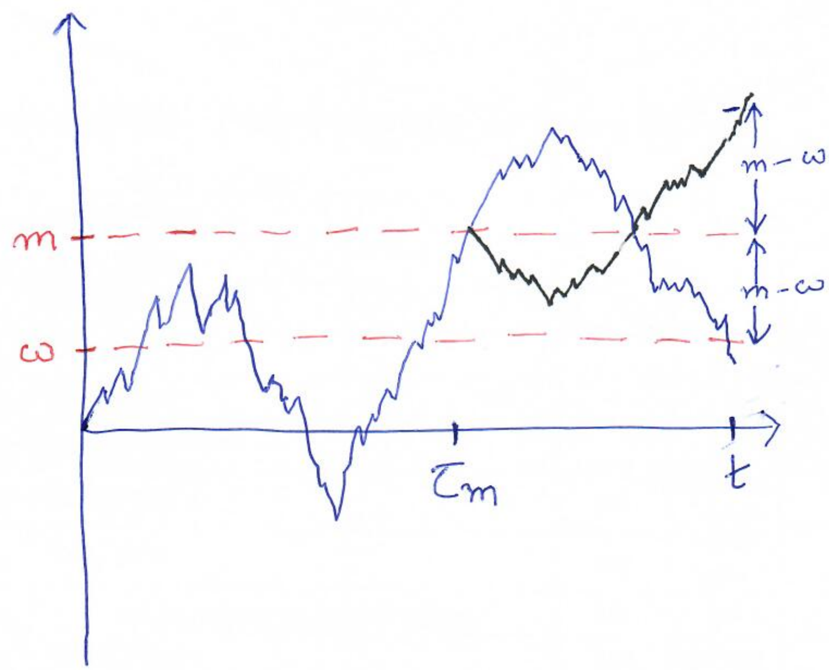
$$\mathbb{E}(e^{-\alpha\tau_m}) = e^{-|m|\sqrt{2\alpha}} \quad \textit{for } \alpha > 0.$$

Proof: First assume that $m > 0$. Then by above we have $\mathbb{P}(\tau_m < \infty) = 1$. Define $\sigma = \sqrt{2\alpha}$. Then we get the result. For $m < 0$, we have by symmetry of Brownian motion ($W(t)$ has same distribution as $-W(t)$), that τ_m has the same distribution as τ_{-m}. Hence we have the complete result. $\square$

Reflection Identity:

$$\mathbb{P}(\tau_m \leq t, W(t) \leq w) = \mathbb{P}(\tau_m \leq t, W(t) \geq 2m - w) = \mathbb{P}(W(t) \geq 2m - w)$$

for $m > 0$ and $w \leq m$.



Theorem 2.9. *For all $m \neq 0$, the random variable τ_m has distribution function*

$$\mathbb{P}(\tau_m \leq t) = \frac{2}{\sqrt{2\pi}} \int_{\frac{|m|}{\sqrt{t}}}^{\infty} e^{(-y^2/2)} dy\,,$$

and pdf

$$f_{\tau_m}(t) = \frac{|m|}{t\sqrt{2\pi t}} e^{\frac{-m^2}{2t}}, t \geq 0\,.$$

Proof: We prove it for $m > 0$, and for $m < 0$ the proof follows by symmetry. Putting $w = m$ in the reflection identity we get,

$$\mathbb{P}(\tau_m \leq t, W(t) \leq m) = \mathbb{P}(W(t) \geq m)\,.$$

Also we have trivially,

$$\mathbb{P}(\tau_m \leq t, W(t) \geq m) = \mathbb{P}(W(t) \geq m)\,.$$

Adding the above two we get,

$$\begin{aligned}\mathbb{P}(\tau_m \leq t) &= 2\mathbb{P}(W(t) \geq m) \\ &= \frac{2}{\sqrt{2\pi t}} \int_m^{\infty} e^{-\frac{x^2}{2t}} dx\end{aligned}$$

Taking $y = \frac{x}{\sqrt{t}}$, $dx = \frac{dx}{\sqrt{t}}$, we get,

$$\mathbb{P}(\tau_m \leq t) = \frac{2}{\sqrt{2\pi}} \int_{\frac{|m|}{\sqrt{t}}}^{\infty} e^{(-y^2/2)} dy \,.$$

Differentiation with respect to t we get,

$$f_{\tau_m}(t) = \frac{|m|}{t\sqrt{2\pi t}} e^{\frac{-m^2}{2t}}, t \geq 0 \,.$$

$\square$

Define $M(t) = \max_{0 \leq s \leq t} W(t)$, to be the running maximum of Brownian motion. Now for $m > 0$, $\tau_m \leq t$ if and only if $M(t) \geq m$. Thus the reflection identity can be rewritten as

$$\mathbb{P}(M(t) \geq t, W(t) \leq w) = \mathbb{P}(W(t) \geq 2m - w) \,.$$

Theorem 2.10. *The joint density of $(M(t), W(t))$ is given by*

$$f_{M(t),W(t)}(m, w) = \frac{2(2m - w)}{t\sqrt{2\pi t}} e^{\frac{-(2m-w)^2}{2t}}, w \leq m, m > 0 \,.$$

Proof: Writing the reflection principle in terms of density,

$$\int_m^{\infty} \int_{-\infty}^{w} f_{M(t),W(t)}(x, y) dy dx = \frac{1}{\sqrt{2\pi t}} \int_{2m-w}^{\infty} e^{-y^2} 2t dy \,.$$

Differentiation with respect to m we get,

$$\int_{-\infty}^{w} f_{M(t),W(t)}(x, y) dy = \frac{2}{\sqrt{2\pi t}} e^{\frac{-(2m-w)^2}{2t}} \,.$$

Further differentiation with respect to w, we get,

$$f_{M(t),W(t)}(m, w) = \frac{2(2m - w)}{t\sqrt{2\pi t}} e^{\frac{-(2m-w)^2}{2t}} \,.$$

$\square$

2.2 Quadratic Variation of Brownian Motion

Definition 2.11. *Let $\Pi = \{t_0, t_1, \ldots, t_n\}$, where $0 = t_0 < t_1 \ldots < t_n$ be a partition of $[0, T]$. Let $||\Pi|| = \max_{j=1,2,\ldots,n}(t_j - t_{j-1})^2$. Let $f(\cdot)$ be a function defined on $[0, T]$. For $p \geq 1$, the pth variation of f upto time T is defined as*

$$[f, f]_p(T) = \lim_{||\Pi|| \to 0} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)|^p \,.$$

The limit is taken as the number of partition points n goes to infinity and the length of the longest sub-interval goes to 0.

Convention: For $p = 2$, we will write $[f, f]_2$ simply as $[f, f]$.

Theorem 2.12. *Let $W(\cdot)$ be a Brownian motion. Then $[W, W](T) = T$, for all $T \geq 0$, where the limit is in L^2 sense.*

Proof: Let $\Pi = \{t_0, t_1, \ldots, t_n\}$, where $0 = t_0 < t_1 \ldots < t_n$ be a partition of $[0, T]$. Define

$$Q_\Pi = \sum_{j=0}^{n-1} |W(t_{j+1}) - W(t_j)|^2 .$$

We will show that $\mathbb{E}(Q_\Pi - T)^2$ goes to 0 as $||\Pi||$ goes to 0. Now

$$\mathbb{E}(Q_\Pi) = \sum_{j=0}^{n-1} \mathbb{E}(W(t_{j+1}) - W(t_j))^2 = \sum_{j=0}^{n-1} t_{j+1} - t_j = T .$$

Thus

$$\mathbb{E}(Q_\Pi - T)^2 = \mathbb{E}(Q_\Pi - \mathbb{E}(Q_\Pi))^2 = Var(Q_\Pi) .$$

Now

$$\begin{aligned} Var(Q_\Pi) &= \sum_{j=0}^{n-1} Var(W(t_{j+1}) - W(t_j))^2 \\ &= \sum_{j=0}^{n-1} \left[\mathbb{E}(W(t_{j+1}) - W(t_j))^4 - \{\mathbb{E}(W(t_{j+1}) - W(t_j))^2\}^2\right] \\ &= 3 \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 - \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \\ &= 2 \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \leq 2||\Pi||T \to 0 \end{aligned}$$

as $||\Pi||$ goes to 0. Hence the proof. □

Let us also compute the cross variation of $W(t)$ and t and the quadratic variation of t.

$$\lim_{||\Pi|| \to 0} \sum_{j=0}^{n-1} (W(t_{j+1}) - W(t_j))(t_{j+1} - t_j) \leq \lim_{||\Pi|| \to 0} \max_{0 \leq j \leq n-1} |W(t_{j+1}) - W(t_j)| T .$$

Since Brownian motion sample paths are continuous almost surely $\max_{0 \leq j \leq n-1} |W(t_{j+1}) - W(t_j)| \to 0$ as $||\Pi|| \to 0$. For quadratic variation of t,

$$\lim_{||\Pi|| \to 0} \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \leq \lim_{||\Pi|| \to 0} ||\Pi|| T = 0 .$$

Symobolically the above three results are written as

$$dW_t dW_t = dt, \quad dW_t dt = 0, \quad dt dt = 0 .$$

What about the first variation of $W(t)$? Before answering that question we will state a fact.
Fact If $\{X_n\}$ converges to X in L^2, then there exists a subsequence $\{X_{n_k}\}$ which converges to X almost surely.

Theorem 2.13. *Brownian motion sample paths have infinite first variation almost surely, i.e.*

$$\mathbb{P}\{\omega : W(\cdot,\omega) \text{ has infinite first variation on } [0,T]\} = 1\,.$$

Proof:

$$\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2 \leq \max_{0\leq j\leq n-1} |W(t_{j+1}) - W(t_j)| \sum_{j=0}^{n-1} |W(t_{j+1}) - W(t_j)|^2\,.$$

Since $\lim_{||\Pi||\to 0} \sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2 = T$ and $\lim_{||\Pi||\to 0} \max_{0\leq j\leq n-1} |W(t_{j+1}) - W(t_j)| = 0$, we have

$$\lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} |W(t_{j+1}) - W(t_j)| = \infty\,.$$

$\square$

Infinite first variation is a testimony of the fact that Brownian motion sample paths "fluctuate very wildy".

3 Stochastic Integration

3.1 Motivation

We are interested in defining the integral of the form

$$\int_0^T f(t)dW_t\,,$$

where f is a stochastic process satisfying suitable assumptions. Defining it as

$$\int_0^T f(t)W'(t)dt$$

is not possible since Brownian motion paths are nowhere differentiable almost surely. So let us try in the Riemann-Stieltjes sense for the process $f(t) = W(t)$, i.e., we are trying to define the integral

$$\int_0^T W(t)dW_t\,.$$

Take a partition $\Pi = \{t_0, t_1, \dots, t_n\}$ of $[0, T]$. Then

$$L_n = \sum_{i=1}^{n} W(t_{i-1})(W(t_i) - W(t_{i-1}))\,,$$

and

$$R_n = \sum_{i=1}^{n} W(t_i)(W(t_i) - W(t_{i-1}))\,.$$

If the integral has to exist in the Riemann-Stieltjes sense then the difference $R_n - L_n$ should go to 0 as $||\Pi||$ goes to 0. But

$$\lim_{||\Pi||\to 0}(R_n - L_n) = \lim_{||\Pi||\to 0}\sum_{i=1}^{n}(W(t_i) - W(t_{i-1}))^2 = T\,.$$

Ideally we would like the stochastic integral to be a martingale. Now

$$\begin{aligned} R_n + L_n &= \sum_{i=1}^{n}(W(t_i) + W(t_{i-1}))(W(t_i) - W(t_{i-1})) \\ &= \sum_{i=1}^{n}(W^2(t_i) - W^2(t_{i-1})) \\ &= W^2(T) - W^2(0)\,. \end{aligned}$$

Thus

$$R_n = \frac{1}{2}(R_n + L_n + R_n - L_n)R_n = \frac{1}{2}(R_n + L_n + R_n - L_n)\,.$$

So,

$$\lim_{||\Pi||\to 0} R_n = \frac{1}{2}(W^2(T) + T)\,,\ \lim_{||\Pi||\to 0} L_n = \frac{1}{2}(W^2(T) - T)\,.$$

In the second case it is indeed a martingale.

3.2 Definition and Properties

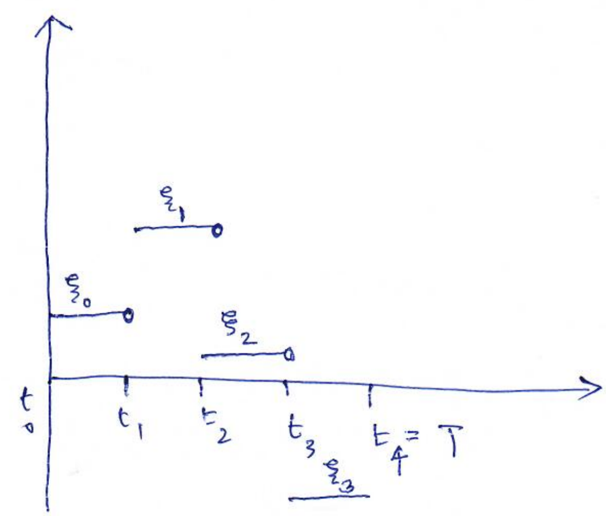
Definition 3.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $W(\cdot)$ be a Brownian motion defined on it. Let $\mathcal{F}_t$ be a Brownian filtration. For $T > 0$, define the space $L^2_{ad}([0,T] \times \Omega)$ to denote the space of all stochastic processes $f(t,\omega), 0 \le t \le T, \omega \in \Omega$ satisfying the following two properties:*

- *$f(t,\omega)$ is adapted to the filtration $\{\mathcal{F}_t\}$.*
- *$\int_0^T \mathbb{E}|(f(t)|^2)dt < \infty$.*

We will construct the integral in several steps like we did for expectation.
step 1: f is a step stochastic process in $L^2_{\text{ad}}([0,T] \times \Omega)$, i.e.,

$$f(t,\omega) = \sum_{i=1}^{n} \xi_{i-1} 1_{[t_{i-1},t_i)}(t)\,,$$

where $\{t_0, t_1, \ldots, t_n\}$ is a partition of $[0,T]$ and ξ_{i-1} is $\mathcal{F}_{t_{i-1}}$ adapted and $\mathbb{E}(\xi^2_{i-1}) < \infty$.



In this case we define

$$I(f) = \sum_{i=1}^{n} \xi_{i-1}(W(t_i) - W(t_{i-1}))\,. \tag{3.1}$$

Exercise: If f and g are two step processes, then show that $I(af + bg) = aI(f) + bI(g)$ for any $a, b \in \mathbb{R}$.

Lemma 3.2. *(Ito's Isometry) Let $I(f)$ be as in (3.1), then $\mathbb{E}(I(f)) = 0$ and $\mathbb{E}(I(f))^2 = \int_0^T \mathbb{E}|(f(t)|^2)dt$.*

Proof: For each $1 \leq i \leq n$ in (3.1),

$$\begin{aligned}
&\mathbb{E}(\xi_{i-1}(W(t_i) - W(t_{i-1})) \\
&= \mathbb{E}(\mathbb{E}[\xi_{i-1}(W(t_i) - W(t_{i-1})|\mathcal{F}_{t_{i-1}}]) \\
&= \mathbb{E}(\xi_{i-1}\mathbb{E}((W(t_i) - W(t_{i-1}))) = 0\,.
\end{aligned}$$

Thus $\mathbb{E}(I(f)) = 0$. Now

$$(I(f))^2 = \sum_{i,j=1}^{n} \xi_{i-1}\xi_{j-1}(W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1}))\,.$$

For $i \neq j$, suppose $i < j$, then

$$\begin{aligned}
&\mathbb{E}\left(\xi_{i-1}\xi_{j-1}(W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1}))\right) \\
&= \mathbb{E}\left(\mathbb{E}[\xi_{i-1}\xi_{j-1}(W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1}))|\mathcal{F}_{t_{j-1}}]\right) \\
&= \mathbb{E}(\xi_{i-1}\xi_{j-1}(W(t_i) - W(t_{i-1}))\mathbb{E}(W(t_j) - W(t_{j-1}))) = 0\,.
\end{aligned}$$

For $i = j$, we have

$$\begin{aligned}
&\mathbb{E}\left(\xi_{i-1}^2(W(t_i) - W(t_{i-1}))^2\right) \\
&= \mathbb{E}\left(\mathbb{E}[\xi_{i-1}^2\xi_{j-1}(W(t_i) - W(t_{i-1}))^2|\mathcal{F}_{t_{i-1}}]\right) \\
&= \mathbb{E}(\xi_{i-1}^2\mathbb{E}(W(t_i) - W(t_{i-1}))^2) \\
&= \mathbb{E}(\xi_{i-1}^2(t_i - t_{i-1}))\,.
\end{aligned}$$

Thus $\mathbb{E}(I(f))^2 = \sum_{i=1}^{n} \mathbb{E}(\xi_{i-1}^2)(t_i - t_{i-1}) = \int_0^T \mathbb{E}|(f(t)|^2)dt$. □

Theorem 3.3. *Define the process $I_t(f) = \int_0^t f(s)dW(s)$ for $0 \leq t \leq T$. Then $I_t(f)$ is a martingale.*

Proof: Let $0 \leq s \leq t \leq T$. We assume that s and t are in different sub-intervals, i.e., there are t_l and t_k with $t_l < t_k$ and $s \in [t_l, t_{l+1})$, $t \in [t_k, t_{k+1})$. If they are in the same interval the following proof simplifies and is left as an exercise. Now

$$I_t(f) = \sum_{i=0}^{l-1} \xi_i(W(t_{i+1} - W(t_i)) + \xi_l(W(t_{l+1} - W(t_l)) + \sum_{i=l+1}^{k-1} \xi_i(W(t_{i+1} - W(t_i)) + \xi_k(W(t) - W(t_k))\,.$$

$$\mathbb{E}[\sum_{i=0}^{l-1} \xi_i(W(t_{i+1} - W(t_i))|\mathcal{F}_s] = \sum_{i=0}^{l-1} \xi_i(W(t_{i+1} - W(t_i))\,.$$

$$\begin{aligned}
\mathbb{E}[\xi_l W(t_{l+1} - W(t_l))|\mathcal{F}_s] &= \xi_l\mathbb{E}[W(t_{l+1} - W(t_l))|\mathcal{F}_s] \\
&= \xi_l\mathbb{E}[W(t_{l+1} - W(s) + W(s) - W(t_l))|\mathcal{F}_s] \\
&= \xi_l(W(s) - W(t_l))\,.
\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\sum_{i=l+1}^{k-1} \xi_i(W(t_{i+1}-W(t_i))|\mathcal{F}_s] &= \sum_{i=l+1}^{k-1} \mathbb{E}[\xi_i(W(t_{i+1}-W(t_i))|\mathcal{F}_s] \\ &= \sum_{i=l+1}^{k-1} \mathbb{E}[\mathbb{E}[\xi_i(W(t_{i+1}-W(t_i))|\mathcal{F}_{t_i}]|\mathcal{F}_s] \\ &= \sum_{i=l+1}^{k-1} \mathbb{E}[\xi_i\mathbb{E}(W(t_{i+1}-W(t_i))|\mathcal{F}_s] = 0\,.\end{aligned}$$

Similarly,

$$\mathbb{E}[\xi_k(W(t)-W(t_k))|\mathcal{F}_s] = 0\,.$$

Thus adding all together we get,

$$\mathbb{E}[I_t(f)|\mathcal{F}_s] = \sum_{i=0}^{l-1} \xi_i(W(t_{i+1}-W(t_i)) + \xi_l(W(s)-W(t_l)) = I_s(f)\,.$$

□

Theorem 3.4. *The quadratic variation of the Ito integral on* $[0,T]$ *is* $\int_0^T f^2(u)du$.

Proof: We compute the quadratic variation of the Ito integral on one of the sub-intervals $[t_j, t_{j+1}]$ on which f is constant. For that choose partition points $t_j = s_0 < s_1 < \ldots < s_m = t_{j+1}$ and consider

$$\begin{aligned}&\sum_{i=0}^{m-1}[I(s_{i+1})-I(s_i)]^2 \\ &= \sum_{i=0}^{m-1}[\xi_{t_j}(W(s_{i+1})-W(s_i))]^2 \\ &= \xi_{t_j}^2 \sum_{i=0}^{m-1}(W(s_{i+1})-W(s_i))^2\,.\end{aligned}$$

Now as $m \to \infty$ and the norm of the partition goes to 0,

$$\sum_{i=0}^{m-1}(W(s_{i+1})-W(s_i))^2 \to t_{j+1}-t_j$$

in L^2. Hence

$$\lim_{||\Pi||\to 0} \sum_{i=0}^{m-1}[I(s_{i+1})-I(s_i)]^2 = \xi_{t_j}^2(t_{j+1}-t_j) = \int_{t_j}^{t_{j+1}} \xi_{t_j}^2\,dt\,.$$

The result now follows by adding up the sub-intervals. □

Symbolically we have, $dI(t) = f(t)dW(t)$ and $dI(t)dI(t) = f^2(t)dt$.

Now what about general integrands.

Lemma 3.5. *Suppose $f \in L^2_{ad}([0,T] \times \Omega)$. Then there exists a sequence of step stochastic processes $\{f_n\}$ such that*

$$\lim_{n\to\infty} \int_0^T \mathbb{E}|f_n(t) - f(t)|^2 dt = 0 \,.$$

Fact: Let L^2 be the space of all random variables X such that $\mathbb{E}(X^2) < \infty$. Then L^2 is complete, i.e., every Cauchy sequence converges. So if $\{X_n\}$ is a sequence of random variables such that $\mathbb{E}|X_n - X_m|^2$ goes to 0 as $m, n \to \infty$, then there exists a random variable X in L^2 such that $\mathbb{E}|X_n - X|^2$ goes to 0 as $n \to \infty$.
Let $\{f_n\}$ and f be as in Lemma 3.5. Then we have $I(f_n) - I(f_m) = I(f_n - f_m)$. So by Ito's isometry,

$$\mathbb{E}(I(f_n) - I(f_m))^2 = \int_0^T \mathbb{E}(f_n(t) - f_m(t))^2 dt \to 0 \,,$$

as $m, n \to \infty$. Thus $\{I(f_n)\}$ is Cauchy in L^2 and hence has a limit in L^2. So define $I(f)$ to be

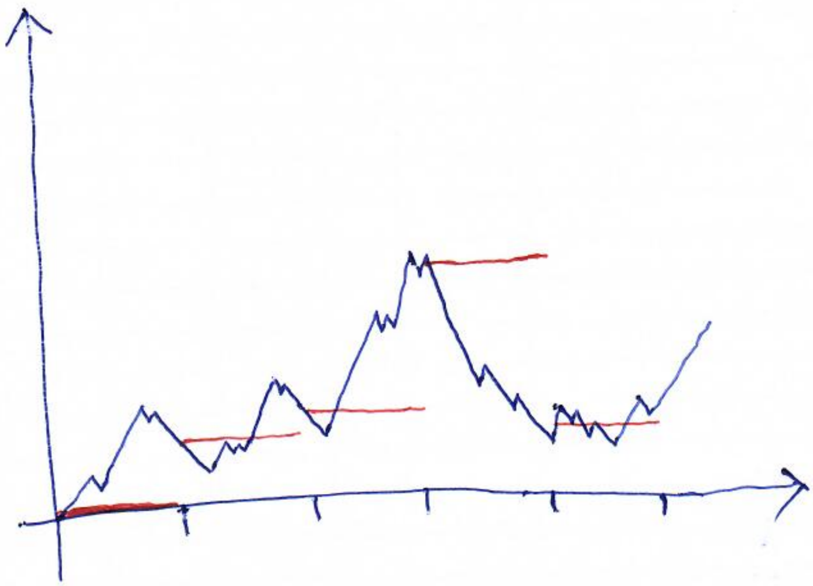
$$I(f) = \lim_{n\to\infty} I(f_n) \,.$$

Theorem 3.6. *(Properties of Ito Integral) Let T be a positive constant. Let $f, g \in L^2_{ad}([0,T] \times \Omega)$. The the following are true:*

1. *$I_t(af + bg) = aI_t(f) + bI_t(g)$ for all real constants a, b and for all $t \in [0,T]$.*
2. *$I_t(f), 0 \le t \le T$ has almost continuous sample paths almost surely.*
3. *For each t, $I_t(f)$ is $\mathcal{F}_t$ measurable, where $\mathcal{F}_t$ is a filtration for Brownian motion.*
4. *$I_t(f), 0 \le t \le T$ is a martingale.*
5. *(Ito's Isometry) $\mathbb{E}(I_t(f)^2) = \int_0^t \mathbb{E}f^2(u)du$ for all $t \in [0,T]$.*
6. *The quadratic variation is given by $[I(f), I(f)](t) = \int_0^t f^2(u)du$.*

Example: Find $\int_0^T W^2(t)dW(t)$. Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0,T]$. Define the step stochastic process $f_n(t,\omega) = W^2(t_{i-1}, \omega)$ for $t \in [t_{i-1}, t_i)$, $i = 1, 2, \ldots, n$. Then

$$\lim_{n\to\infty} \int_0^T \mathbb{E}|f_n(t) - f(t)|^2 dt = 0 \,.$$



So by definition

$$\int_0^T W^2(t)dW(t) = \lim_{n\to\infty}\int_0^T f_n(t)dW(t) = \lim_{n\to\infty}\sum_{i=1}^n W^2(t_{i-1})(W(t_i) - W(t_{i-1}),$$

where the limit is in L^2 sense. It is easy to check that

$$3\sum_{i=1}^n W^2(t_{i-1})(W(t_i) - W(t_{i-1}) =$$

$$W^3(T) - W^3(0) - \sum_{i=1}^n (W(t_i) - W(t_{i-1}))^3 - 3\sum_{i=1}^n W(t_{i-1})(W(t_i) - W(t_{i-1}))^2.$$

Now

$$\mathbb{E}(\sum_{i=1}^n (W(t_i) - W(t_{i-1}))^3)^2 = \mathbb{E}\sum_{i=1}^n (W(t_i) - W(t_{i-1})^6 = 15\sum_{i=1}^n (t_i - t_{i-1})^3 \le 15||\Pi||^2 T \to 0.$$

Thus $\sum_{i=1}^n (W(t_i) - W(t_{i-1}))^3$ converges to 0 in L^2. For the second summation term,

$$\mathbb{E}\left|\sum_{i=1}^n W(t_{i-1})(W(t_i) - W(t_{i-1})^2 - \sum_{i=1}^n W(t_{i-1})(t_i - t_{i-1})\right|^2$$

$$= \sum_{i=1}^n 2t_{i-1}(t_i) - t_{i-1})^2 \le 2T(T)||\Pi|| \to 0.$$

From the above we can say that $\sum_{i=1}^n W(t_{i-1})(W(t_i) - W(t_{i-1}))^2$ converges to $\int_0^T W(t)dt$ in L^2. Thus we have

$$\int_0^T W^2(t)dW(t) = \frac{1}{3}W^3(T) - \int_0^T W(t)dt.$$

3.3 Ito's Formula

Chain rule in ordinary calculus tells us that if f and g are differentiable functions then, $\frac{d}{dt}f(g(t)) = f'(g(t))g'(t)$. Thus by Fundamental Theorem of Calculus,

$$f(g(T)) - f(g(0)) = \int_0^T f'(g(t))g'(t)dt\,.$$

If $g(t) = W(t)$, then $W'(t)$ does not make sense but if we write $dW(t)$ in place of $W'(t)dt$ then we get,

$$f(W(T)) - f(W(0)) = \int_0^T f'(W(t))dW(t)\,.$$

So in Ito Calculus is this formula correct? Let us take $f(t) = t^2$. By Taylor's Theorem,

$$f(y) - f(x) = f'(x)(y-x) + \frac{1}{2}f''(x)(y-x)^2\,.$$

Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0,T]$. Then,

$$\begin{aligned}
f(W(T)) - f(W(0)) &= \sum_{j=0}^{n-1} f(W(t_{j+1})) - f(W(t_j)) \\
&= \sum_{j=0}^{n-1} f'(W(t_j))(W(t_{j+1}) - W(t_j)) + \frac{1}{2}\sum_{j=0}^{n-1} f''(W(t_j))(W(t_{j+1}) - W(t_j))^2 \\
&= \sum_{j=0}^{n-1} 2(W(t_j))(W(t_{j+1}) - W(t_j)) + \sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2
\end{aligned}$$

Thus

$$\begin{aligned}
f(W(T)) - f(W(0)) &= \sum_{j=0}^{n-1} f''(W(t_j))(W(t_{j+1}) - W(t_j))^2 \\
&= \lim_{||\Pi||\to 0}\sum_{j=0}^{n-1} 2(W(t_j))(W(t_{j+1}) - W(t_j)) + \lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2 \\
&= \int_0^T 2W(t)dW(t) + T = \int_0^T f'(W(t))dW(t) + \frac{1}{2}\int_0^T f''(W(t))dt\,.
\end{aligned}$$

So the blue term is the extra term, it is sometimes referred to as Ito's correction term.

Theorem 3.7. *(Ito-Doeblin Formula for Brownian Motion) Let $f(t,x)$ be a function such that the partial derivatives $f_t(t,x), f_x(t,x)$ and $f_{xx}(t,x)$ all exist and are continuous. Let $W(\cdot)$ be a Brownian motion. Then for every $T \geq 0$,*

$$f(T, W(T)) - f(0, W(0)) = \int_0^T f_t(t, W(t))dt + \int_0^T f_x(t, W(t))dW(t) + \frac{1}{2}\int_0^T f_{xx}(t, W(t))dt\,.$$

In differential form,

$$df(t, W(t)) = f_t(t, W(t))dt + f_x(t, W(t))dW(t) + \frac{1}{2}f_{xx}(t, W(t))dt\,.$$

Example: Use Ito's formula to show that $W^3(t) - 3\int_0^t W(t)dt$ is a martingale.
Using Ito's formula to the function $f(x) = x^3$, we get

$$W^3(t) - W^3(0) = \int_0^t 3W^2(t)dW(t) + 3\int_0^t W(t)dt\,.$$

Now using the fact that stochastic integral is a martingale we are done.

Definition 3.8. *Let $W(t), t \geq 0$ be a Brownian motion and let $\mathcal{F}_t, t \geq 0$ be a filtration for the Brownian motion. An Ito process is a stochastic process of the form*

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du\,,$$

where $\Delta(u)$ and $\Theta(u)$ are adapted stochastic processes such that $\int_0^t \mathbb{E}\Delta^2(u)du < \infty$ and $\int_0^t |\Theta(u)|du < \infty$ for all $t \geq 0$.

Lemma 3.9. *The quadratic variation of the above Ito process is*

$$[X, X](t) = \int_0^t \Delta^2(u)du\,.$$

Remark: Symbolically, $dX(t) = \Delta(t)dW(t) + \Theta(t)dt$, so

$$dX_t dX_t = \Delta^2(t)dW_t dW_t + \Theta^2(t)dtdt + 2\Delta(t)\Theta(t)dW_t dt = \Delta^2(t)dt\,.$$

Proof: Suppose $I(t) = \int_0^t \Delta(u)dW(u)$ and $R(t) = \int_0^t \Theta(u)du$. Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0, t]$. Then,

$$\begin{aligned}\sum_{j=0}^{n-1}[X(t_{j+1}) - X(t_j)]^2 &= \sum_{j=0}^{n-1}[I(t_{j+1}) - I(t_j)]^2 + \sum_{j=0}^{n-1}[R(t_{j+1}) - R(t_j)]^2 \\ &\quad + \sum_{j=0}^{n-1}[I(t_{j+1}) - I(t_j)][R(t_{j+1}) - R(t_j)]\,.\end{aligned}$$

The first term converges to $\int_0^t \Delta^2(u)du$ as $||\Pi||$ goes to 0. For the second term,

$$\begin{aligned}\sum_{j=0}^{n-1}[R(t_{j+1}) - R(t_j)]^2 &\leq \max_{0\leq k\leq n-1} |R(t_{j+1}) - R(t_j)| \sum_{j=0}^{n-1} |R(t_{j+1}) - R(t_j)| \\ &= \max_{0\leq k\leq n-1} |R(t_{j+1}) - R(t_j)| \sum_{j=0}^{n-1} |\int_{t_j}^{t_{j+1}} \Theta(u)du| \\ &\leq \max_{0\leq k\leq n-1} |R(t_{j+1}) - R(t_j)| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} |\Theta(u)|du \\ &= \max_{0\leq k\leq n-1} |R(t_{j+1}) - R(t_j)| \int_0^t |\Theta(u)|du \to 0\end{aligned}$$

as $||\Pi|| \to 0$, since $R(\cdot)$ is continuous. The third term is bounded by

$$
\begin{aligned}
2 \max_{0\le k\le n-1} |I(t_{j+1}) - I(t_j)| \sum_{j=0}^{n-1} |R(t_{j+1}) - R(t_j)| &= \max_{0\le k\le n-1} |I(t_{j+1}) - I(t_j)| \sum_{j=0}^{n-1} |\int_{t_j}^{t_{j+1}} \Theta(u)du| \\
&\le \max_{0\le k\le n-1} |I(t_{j+1}) - I(t_j)| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} |\Theta(u)|du \\
&= \max_{0\le k\le n-1} |I(t_{j+1}) - I(t_j)| \int_0^t |\Theta(u)|du \to 0
\end{aligned}
$$

as $||\Pi|| \to 0$, since $I(\cdot)$ is continuous. Hence we have the result. □

Definition 3.10. *Let $X(t), t \ge 0$ be an Ito process. Let $\Gamma(t), t \ge 0$ be an adapted process. Further suppose that $\int_0^t \mathbb{E}\Gamma^2(u)\Delta^2(u)du$ and $\int_0^t |\Gamma(u)\Theta(u)|du$ are finite for all $t \ge 0$. We define the integral with respect to an Ito process by,*

$$\int_0^t \Gamma(u)dX(u) = \int_0^t \Gamma(u)\Delta(u)dW(u) + \int_0^t \Gamma(u)\Theta(u)du\,.$$

Theorem 3.11. *(Ito-Doeblin Formula for Ito process) Let $X(t), t \ge 0$ be an Ito process. Let $f(t, x)$ be a function such that the partial derivatives $f_t(t, x), f_x(t, x)$ and $f_{xx}(t, x)$ all exist and are continuous. Then for every $T \ge 0$,*

$$
\begin{aligned}
f(T, X(T)) - f(0, W(0)) &= \int_0^T f_t(t, X(t))dt + \int_0^T f_x(t, X(t))dX(t) \\
&+ \frac{1}{2}\int_0^T f_{xx}(t, X(t))dX(t)dX(t) \\
&= \int_0^T f_t(t, X(t))dt + \int_0^T f_x(t, X(t))\Delta(t)dW(t) + \int_0^T f_x(t, X(t))\Theta(t)dt \\
&+ \frac{1}{2}\int_0^T f_{xx}(t, X(t))\Delta^2(t)dt\,.
\end{aligned}
$$

Exercise: Let $W(\cdot)$ be a Brownian motion and let $\sigma(t)$ be a non random function. Define $I(t) = \int_0^t \sigma(u)dW(u)$. Then show that $I(t)$ is a normal random variable with mean 0 and variance $\int_0^t \sigma^2(u)du$.

3.4 Black-Scholes Market

In a Black-Scholes market an agent can invest in a money market(risk free) account which pays a constant rate of interest r and in a stock(risky) modeled by a geometric Brownian motion

$$S(t) = S(0)e^{\sigma W(t) + (\alpha - \frac{1}{2}\sigma^2)t}\,.$$

Thus $S(t)$ has log normal distribution. An amount x invested in the money market account gives a guaranteed return xe^{rt} after time t. The return on the stock is random. But $\mathbb{E}(S(t)) = S(0)e^{\alpha t}$. Thus α is the mean rate of return. Now,

$$S(t+1) = S(t)e^{\sigma(W(t+1) - W(t)) + (\alpha - \frac{1}{2}\sigma^2)}\,.$$

Thus
$$\log(S(t+1)/S(t)) = \sigma(W(t+1) - W(t)) + (\alpha - \frac{1}{2}\sigma^2)\,.$$
Thus $\sqrt{Var(\log(S(t+1)/S(t)))} = \sigma$. The parameter σ is called the volatility of the stock.

Theorem 3.12. *Fix $t \geq 0$. As $n \to \infty$, the distribution of the scaled random walk $W^n(t)$ at the point t converges to the normal distribution with mean 0 and variance t.*

Consider a multi-period Binomial model on $[0, t]$, such that the stock price takes n steps per unit time. Assume that n and t are so chosen so that nt is an integer. So this is basically a nt period Binomial model. Suppose the up factor is $u_n = 1 + \sigma/\sqrt{n}$ and the down factor is $d_n = 1 - \sigma/\sqrt{n}$, where $\sigma > 0$. And suppose that the probabilities of going up and down are 1/2 each. Then $S_n(t) = S(0)u_n^{H_{nt}} d_n^{T_{nt}}$, where H_{nt} is the number of upward movements and T_{nt} is the number of downward movements. Then $H_{nt} + T_{nt} = nt$ and $M_{nt} = H_{nt} - T_{nt}$ is the position of a simple random walk after nt steps. Thus $H_{nt} = \frac{1}{2}(nt + M_{nt})$ and $T_{nt} = \frac{1}{2}(nt - M_{nt})$. Thus
$$S_n(t) = (1 + \sigma/\sqrt{n})^{\frac{1}{2}(nt+M_{nt})}(1 - \sigma/\sqrt{n})^{\frac{1}{2}(nt-M_{nt})}\,.$$
Claim: As $n \to \infty$, the distribution of $S_n(t)$ converges to the distribution of $S(t) = S(0)e^{\sigma W(t) - \frac{1}{2}\sigma^2 t}$, where $W(t)$ is normal with mean 0 and variance t.
Proof of claim: Enough to show that $\log S_n(t)$ converges in distribution to $\log S(t)$. Now $\log(1 + x) = x - x^2/2 + O(x^3)$. Then taking $x = \sigma/\sqrt{n}$ and $-\sigma/\sqrt{n}$ we get,
$$\begin{aligned}
\log S_n(t) &= \log S(0) + \frac{1}{2}(nt + M_{nt})(\sigma/\sqrt{n} - \sigma^2/(2n) + O(n^{-3/2})) \\
&+ \frac{1}{2}(nt - M_{nt})(-\sigma/\sqrt{n} - \sigma^2/(2n) + O(n^{-3/2})) \\
&= \log S(0) + nt(-\sigma^2/(2n) + O(n^{-3/2})) + M_{nt}(\sigma/\sqrt{n} + O(n^{-3/2})) \\
&= \log S(0) - \frac{\sigma^2 t}{2} + O(n^{-1/2}) + \sigma W^n(t) + O(n^{-1})W^n(t) \\
&\Rightarrow \log S(0) - \frac{\sigma^2 t}{2} + \sigma W(t) = \log S(t)\,.
\end{aligned}$$
By Ito's formula we have,
$$dS(t) = \alpha S(t)dt + \sigma S(t)dW(t)\,.$$
Suppose an investor has a portfolio with value $X(t)$ at time t, of which $\delta(t)S(t)$ is invested in the stock and the remaining $X(t) - \Delta(t)S(t)$ in the money market account. Then the evolution of the portfolio is given by,
$$\begin{aligned}
dX(t) &= \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt \\
&= rX(t)dt + (\alpha - r)\Delta(t)S(t)dt + \sigma\Delta(t)dW(t)
\end{aligned}$$
The three terms appearing above can be interpreted as follows:

- 1st term is an average rate of return r on the portfolio.
- 2nd term is the risk premium for investing in the stock.
- 3rd term is the volatility term proportional to the stock investment.

We shall often be interested in the discounted stock price $e^{-rt}S(t)$ and discounted portfolio value of an agent $e^{-rt}X(t)$. By Ito's formula,

$$\begin{aligned} d(e^{-rt}S(t)) &= -re^{-rt}S(t)dt + e^{-rt}dS(t) \\ &= (\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t)\,. \end{aligned}$$

$$\begin{aligned} d(e^{-rt}X(t)) &= -re^{-rt}X(t)dt + e^{-rt}dX(t) \\ &= \Delta(t)[(\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t)] \\ &= \Delta(t)d(e^{-rt}S(t))\,. \end{aligned}$$

Discounting reduces the mean rate of return on the stock from α to $\alpha - r$. And discounting the portfolio value removes the average underlying rate of return. Basically, this makes the interest rate 0. So the change in discounted portfolio value is solely due to change in discounted stock price.

Definition 3.13. *An European call option on the stock $S(\cdot)$ is an agreement which gives its holder the right(but no obligation) to buy one unit of stock at time T(time of maturity) at a price K(strike price) from the seller(or writer) of the option.*

The payoff of an European call option is $(S(T) - K)^+$. So what should be the price of such an option? Black-Scholes and Merton argued that the option price should depend on the stock price, time to maturity, the model parameters and K. Out of these only two quantities are variable, stock price and time. For this reason we let $c(t, x)$ denote the value of the call option at time t, if the stock price at time t is $S(t) = x$. Our goal is to determine the function $c(t, x)$. By Ito's formula,

$$\begin{aligned} dc(t, S(t)) &= c_t(t, S(t))dt + c_x(t, S(t))dS(t) + \frac{1}{2}c_{xx}(t, S(t))dS(t)dS(t) \\ &= c_t(t, S(t))dt + c_x(t, S(t))\alpha S(t)dt + c_x(t, S(t))\sigma S(t)dW(t) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t)dt \\ &= [c_t(t, S(t)) + c_x(t, S(t))\alpha S(t) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t)]dt + c_x(t, S(t))\sigma S(t)dW(t) \end{aligned}$$

We next compute $d(e^{-rt})c(t, S(t))$. Again by Ito's formula,

$$\begin{aligned} &d(e^{-rt})c(t, S(t)) = -re^{-rt}c(t, S(t))dt + e^{-rt}dc(t, S(t)) \\ &= e^{-rt}[-rc(t, S(t)) + c_t(t, S(t)) + c_x(t, S(t))\alpha S(t) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t)]dt \\ &+ e^{-rt}c_x(t, S(t))\sigma S(t)dW(t)\,. \end{aligned}$$

Now the option price should be such that starting with that amount, the seller of the option should be able to maintain a portfolio which will be equal to the option price at all times $t \in [0,T]$. Such a portfolio is called a hedging portfolio or a replicating portfolio. So if $X(t)$ denotes the value of the hedging portfolio at time t, then we want $X(t) = c(t, S(t))$ for all $t \in [0,T]$. This happens if and only if $e^{-rt}X(t) = e^{-rt}c(t, S(t))$. One way to ensure this equality is to make make that $d(e^{-rt}X(t)) = d(e^{-rt}c(t, S(t)))$ for all $t \in [0,t)$ and $X(0) = c(0, S(0))$. This is because, integrating we get,

$$e^{-rt}X(t) - X(0) = e^{-rt}c(t, S(t)) - c(0, S(0)) \quad \forall\, t \in [0,t)\,.$$

Since $X(0) = c(0, S(0))$, we get the desired equality. The equality at $t = T$ is taken care of by continuity. So comparing the evolution of discounted portfolio and discounted option price we get

$$\begin{aligned}&\Delta(t)[(\alpha - r)S(t)dt + \sigma S(t)dW(t)]\\ &= [-rc(t, S(t)) + c_t(t, S(t)) + c_x(t, S(t))\alpha S(t) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t)]dt + c_x(t, S(t))\sigma S(t)dW(t)\,.\end{aligned}$$

Equating $dW(t)$ terms we get,

$$\Delta(t) = c_x(t, S(t)) \quad \forall\, t \in [0,t)\,.$$

Equating dt terms we get,

$$\alpha\Delta(t)S(t) - r\Delta(t)S(t) = -rc(t, S(t)) + c_t(t, S(t)) + c_x(t, S(t))\alpha S(t) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t)\,.$$

Implies,

$$rc(t, S(t)) = c_t(t, S(t)) + rS(t)c_x(t, S(t)) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t) \quad \forall\, t \in [0,t)\,.$$

Thus we need a continuous function $c(t,x)$ that is a solution of the Black-Scholes-Merton PDE

$$c_t(t, S(t)) + rS(t)c_x(t, S(t)) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t) = rc(t,x) \quad \forall\, t \in [0,t)\,,$$

with terminal condition $C(T,x) = (x - K)^+$.

Suppose we find a solution to the above pde, then if an investor starts with the initial capital $X(0) = c(0, S(0))$ and invests $\Delta(t)S(t)$ in stocks, with $\Delta(t)$ given by $c_x(t, S(t))$ and the remaining in the money market account, then his portfolio satisfies $X(t) = c(t, S(t))$ for all $t \in [0,t)$. Taking limit $t \uparrow T$ and using the fact that $X(t)$ and $c(t, S(t))$ are continuous we get,

$$X(T) = c(T, S(T)) = (S(T) - K)^+\,.$$

Thus the seller is perfectly hedged.

The BSM PDE is a partial differential equation of the type called bachward parabolic. For such an equation, in addition to the terminal condition, one needs boundary conditions at

$x = 0$ and $x = \infty$. For the boundary condition at $x = 0$, we take limit $x \downarrow 0$ in the BSM PDE to get,

$$rc(t, 0) = c_t(t, 0)\,.$$

This is an ODE in t, whose solution is

$$c(t, 0) = e^{rt} c(0, 0)\,.$$

Now $c(T, 0) = (0 - K)^+ = 0$. Thus $c(0, 0) = 0$, implies $c(t, 0) = 0$, for all $t \in [0, T]$. As $x \to \infty$, the function $c(t, x)$ grows without bound. In such a case we specify the boundary condition at $x = \infty$ by specifying the rate of growth. The boundary condition at $x \to \infty$ is given by

$$\lim_{x\to\infty}[c(t, x) - x - e^{-r(T-t)}K] = 0 \quad \forall\, t \in [0, T]\,.$$

For large x, the call is deep in money and very likely to end in money. In this case, the call option is like a forward contract. Now what is the price of a forward contract? Start with an initial capital of $x - Ke^{-rT}$, where $S(0) = x$. Borrow Ke^{-rT} and use the entire money to buy one unit stock. At time T, the value of this portfolio is $S(T) - K$ which is also the payoff of a forward contract. Thus by no-arbitrage principle the value of a forward contract at time t should be $S(t) - Ke^{-r(T-t)}$.

Theorem 3.14. *The solution to the BSM PDE with the specified terminal and boundary conditions is given by*

$$c(t, x) = xN(d_+(T - t, x)) - Ke^{-r(T-t)}N(d_-(T - t, x)) \quad\quad 0 \le t < T,\ x > 0\,,$$

where $d_\pm(T-t, x) = \frac{1}{\sigma\sqrt{\tau}}[\log(x/K)+(r\pm\frac{\sigma^2}{2})\tau]$ *and N is the CDF of $N(0, 1)$. Note that $c(t, x)$ is not defined for $t = T$ and $x = 0$. But $c(t, x)$ is defined in such a way that* $\lim_{t\to T} c(t, x) = (x - K)^+$ *and* $\lim_{x\downarrow 0} c(t, x) = 0$.

Exercise:

1. Verify that $Ke^{-r(T-t)}N'(d_-) = xN'(d_+)$.
2. Prove that $c_x(t, x) = N(d_+(T - t, x))$.
3. Prove that $c_t(t, x) = -rKe^{-r(T-t)}N'(d_-(T - t, x)) - \frac{\sigma x}{2\sqrt{T-t}}N(d_+(T - t, x))$.
4. Prove that $c_{xx}(t, x) = \frac{1}{\sigma x\sqrt{T-t}}N(d_+(T - t, x))$.
5. Use the above formulas to show that $c(\cdot, \cdot)$ satisfies the BSM PDE.

c_x is called the delta of the option, c_t is called the theta of the option and c_{xx} is called the gamma of the option. Since both N and N' are always positive, the option price is increasing in stock price, decreasing in time and convex in x.

Definition 3.15. *An European put option on the stock $S(\cdot)$ is an agreement which gives its holder the right(but no obligation) to sell one unit of stock at time T(time of maturity) at a price K(strike price) to the seller(or writer) of the option.*

Thus the payoff of an European put option is $(K - S(T))^+$. Now again the question is what should be the price of a put option. The call price and the put price are related by the following relation.

Put-Call Parity: Let $p(t,x)$ denote the option price at time t when the stock price is $S(t) = x$. Then

$$c(t,x) - p(t,x) = x - Ke^{-r(T-t)}\,.$$

Exercise: Using Put-Call parity show that the put price is given by

$$p(t,x) = Ke^{-r(T-t)}N(-d_-(T-t,x)) - xN(-d_+(T-t,x))\,.$$

3.5 Multivariable Stochastic Calculus

Definition 3.16. *A $d-$dimensional Brownian motion is a process $W(t) = (W_1(t), W_2(t), \ldots, W_d(t))$ with the following properties:*
a) Each $W_i(t)$ is a Brownian motion.
b) They are independent as stochastic processes.

Associated with a $d-$dimensional Brownian motion, the filtration $\{\mathcal{F}_t\}$ has the following properties:
i) For each t, the random vector $W(t)$ is $\mathcal{F}_t$ measurable.
ii) For $0 \le t < u$, the vector of increments $W(t) - W(u)$ is independent of $\mathcal{F}_t$.
Since $W_i(t)$ is a Brownian motion for each i, $dW_i(t)dW_i(t) = dt$. What is $dW_i(t)dW_j(t)$ for $i \neq j$. Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0,T]$. For $i \neq j$, define

$$C_\Pi = \sum_{k=0}^{n-1}(W_i(t_{k+1}) - W_i(t_k))(W_j(t_{k+1}) - W_j(t_k))\,.$$

Now

$$\begin{aligned}\mathbb{E}(C_\Pi) &= \mathbb{E}\left(\sum_{k=0}^{n-1}(W_i(t_{k+1}) - W_i(t_k))(W_j(t_{k+1}) - W_j(t_k))\right)\\ &= \sum_{k=0}^{n-1}\mathbb{E}(W_i(t_{k+1}) - W_i(t_k))\mathbb{E}(W_j(t_{k+1}) - W_j(t_k)) = 0\,.\end{aligned}$$

Now

$$\begin{aligned}C_\Pi^2 &= \sum_{k=0}^{n-1}(W_i(t_{k+1}) - W_i(t_k))^2(W_j(t_{k+1}) - W_j(t_k))^2\\ &\quad + 2\sum_{l<k=0}^{n-1}(W_i(t_{l+1}) - W_i(t_l))(W_j(t_{l+1}) - W_j(t_l))(W_i(t_{k+1}) - W_i(t_k))(W_j(t_{k+1}) - W_j(t_k))\end{aligned}$$

All the increments appearing in the cross term are independent of one another and have mean 0. Thus

$$
\begin{aligned}
Var(C_\Pi) = \mathbb{E}(C_\Pi^2) &= \mathbb{E}\left[\sum_{k=0}^{n-1}(W_i(t_{k+1}) - W_i(t_k))^2(W_j(t_{k+1}) - W_j(t_k))^2\right] \\
&= \sum_{k=0}^{n-1}\mathbb{E}(W_i(t_{k+1}) - W_i(t_k))^2\mathbb{E}(W_j(t_{k+1}) - W_j(t_k))^2 = \sum_{k=0}^{n-1}(t_{k+1} - t_k)^2 \leq ||\Pi||T \to 0\,,
\end{aligned}
$$

as $||\Pi|| \to 0$. Thus $[W_i, W_j](t) = 0$, in differential form, $dW_i(t)dW_j(t) = 0$.
Let $X(t)$ and $Y(t)$ be Ito processes, which means they are processes of the form,

$$
\begin{aligned}
X(t) &= X(0) + \int_0^t \Theta_1(u)du + \int_0^t \sigma_{11}(u)dW_1(u) + \int_0^t \sigma_{12}(u)dW_2(u) \\
Y(t) &= Y(0) + \int_0^t \Theta_1(u)du + \int_0^t \sigma_{21}(u)dW_1(u) + \int_0^t \sigma_{22}(u)dW_2(u)\,,
\end{aligned}
$$

where Θ_i and σ_{ij} are adapted stochastic processes satisfying the required properties. The following theorem generalizes one dimensional Ito-Doeblin formula.

Theorem 3.17. *Let $f(t, x, y)$ be a function with continuous partial derivatives $f_t, f_x, f_y, f_{xx}, f_{yy}, f_{xy}$. Let $X(t)$ and $Y(t)$ be Ito processes as above. Then the differential form of the two dimensional Ito-Doeblin formula is given by,*

$$
\begin{aligned}
&df(t, X(t), Y(t)) = f_t(t, X(t), Y(t))dt + f_x(t, X(t), Y(t))dX(t) + f_y(t, X(t), Y(t))dY(t) \\
&+ \frac{1}{2}f_{xx}(t, X(t), Y(t))dX(t)dX(t) + \frac{1}{2}f_{yy}(t, X(t), Y(t))dY(t)dY(t) + f_{xy}(t, X(t), Y(t))dX(t)dY(t)\,.
\end{aligned}
$$

Exercise: Write down the formula in integral form by expanding $dX(t)$, $dY(t)$, $dX(t)dX(t)$, $dY(t)dY(t)$ and $dX(t)dY(t)$.

Corollary 3.18. *(Ito's Product Rule) Let $X(t)$ and $Y(t)$ be Ito processes. Then*

$$
d(X(t)Y(t)) = Y(t)dX(t) + X(t)dY(t) + dX(t)dY(t)\,.
$$

A Brownian motion is a martingale with continuous sample paths and whose quadratic variation is $[W, W](t) = t$. It turns out that these properties characterize Brownian motion.

Theorem 3.19. *(Levy) Let $M(t), t \geq 0$ be a martingale relative to a filtration $\mathcal{F}_t$. Assume that $M(0) = 0$, $M(t)$ has continuous sample paths and $[M, M](t) = t$ for all $t \geq 0$. The $M(t)$ is a Brownian motion.*

Theorem 3.20. *(Levy) Let $M_1(t)$, $M_2(t)$ be martingales relative to a filtration $\mathcal{F}_t$. Assume that, for $i = 1, 2$, we have $M_i(0) = 0$, $M_i(t)$ has continuous sample paths, $[M_i, M_i](t) = t$ for all $t \geq 0$. If in addition, $[M_1, M_2](t) = 0$ for all $t \geq 0$ then (M_1, M_2) is a two dimensional Brownian motion.*

Exercise: Let $W_1(t)$ and $W_2(t)$ be two independent Brownian motions and $-1 < \rho < 1$. Show that $W_3(t)$ defined by

$$W_3(t) = \rho W_1(t) + \sqrt{1-\rho^2}W_2(t)$$

is a Brownian motion. Use Ito's product rule to find $\mathbb{E}(W_1(t)W_3(t))$.
Exercise: Let $W(t)$ be a Brownian motion and define

$$B(t) = \int_0^t sgn(W(s))dW(s)\,,$$

where

$$sgn(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0. \end{cases}$$

Show that $B(t)$ is a Brownian motion.

4 Risk Neutral Pricing

4.1 Change of Measure

Fact: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let Z be a random variable such that $\mathbb{P}(Z > 0) = 1$ and $\mathbb{E}(Z) = 1$. Define a set function $\tilde{\mathbb{P}} : \mathcal{F} \to [0, 1]$ by

$$\tilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P} .$$

Then $\tilde{\mathbb{P}}$ is a probability measure. Further $\mathbb{P}(A) = 0$ if and only if $\tilde{\mathbb{P}}(A) = 0$. Two such measures are called equivalent.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be probability space. Let X be a random variable defined on it such that $X \sim N(0, 1)$. Define another random variable Y by $Y = X + \theta$, $\theta \in \mathbb{R}$. Then under $\mathbb{P}$, $Y \sim N(\theta, 1)$. Define a new probability measure $\tilde{\mathbb{P}}$ by $\tilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P}$, where $Z = e^{-\theta X - \frac{1}{2}\theta^2}$. What is the distribution of Y under $\tilde{P}$.

$$\tilde{\mathbb{E}}(e^{tY}) = \mathbb{E}(e^{tY} Z) = \mathbb{E}(e^{t(X+\theta)} e^{-\theta X - \frac{1}{2}\theta^2}) = e^{t\theta - \frac{1}{2}\theta^2} \mathbb{E}(e^{(t-\theta)X}) = e^{t\theta - \frac{1}{2}\theta^2} e^{\frac{1}{2}(t-\theta)^2} = e^{\frac{t^2}{2}} .$$

Thus $Y \sim N(0, 1)$ under $\tilde{\mathbb{P}}$. Thus we can change the mean of a random variable by changing the measure appropriately.Now let us try to do such an exercise for a stochastic process. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a filtration $\mathcal{F}_t, 0 \le t \le T$. Further suppose Z be a random variable such that $\mathbb{P}(Z > 0) = 1$ and $\mathbb{E}(Z) = 1$. Define a new probability measure $\tilde{\mathbb{P}} : \mathcal{F} \to [0, 1]$ by

$$\tilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P} .$$

Now define the Radon-Nikodym derivative process $Z(t) = \mathbb{E}[Z|\mathcal{F}_t]$. Now for $s < t$,

$$\mathbb{E}[Z(t)|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[Z|\mathcal{F}_s] = Z(s) .$$

Thus $Z(\cdot)$ is a martingale.

Lemma 4.1. *Let t satisfying $0 \le t \le T$ be given and let Y be an $\mathcal{F}_t$ measurable random variable. Then $\tilde{\mathbb{E}}(Y) = \mathbb{E}(Y Z(t))$.*

Proof:

$$\tilde{\mathbb{E}}(Y) = \mathbb{E}(YZ) = \mathbb{E}(\mathbb{E}[YZ|\mathcal{F}_t]) = \mathbb{E}(Y\mathbb{E}[Z|\mathcal{F}_t]) = \mathbb{E}(YZ(t)) .$$

$\square$

Lemma 4.2. *Let s and t satisfying $0 \le s \le t \le T$ be given and let Y be an $\mathcal{F}_t$ measurable random variable. Then*

$$\tilde{\mathbb{E}}[Y|\mathcal{F}_s] = \frac{1}{Z(s)} \mathbb{E}[YZ(t)|\mathcal{F}_s] .$$

Proof: It is clear that $\frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}_s]$ is $\mathcal{F}_s$. Thus in order to show the above we need to show that,

$$\int_A \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}_s]d\tilde{\mathbb{P}} = \int_A Y d\tilde{\mathbb{P}}\,,$$

for all $A \in \mathcal{F}_s$. The left hand side is equal to

$$\begin{aligned}\tilde{\mathbb{E}}\left(1_A\frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}_s]\right) &= \mathbb{E}\left(1_A\mathbb{E}[YZ(t)|\mathcal{F}_s]\right)\\ &= \mathbb{E}\left(1_A YZ(t)\right) = \tilde{\mathbb{E}}(1_A Y) = \int_A Y d\tilde{\mathbb{P}}\,,\end{aligned}$$

where the first and third equalities are by Lemma above.

Theorem 4.3. *(Girsanov) Let $W(t), t \in [0,T]$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_t, t \in [0,T]$ be a filtration for this Brownian motion. Let $\theta(t), t \in [0,T]$ be an adapted process. Define*

$$Z(t) = \exp\left\{-\int_0^t \theta(u)dW(u) - \frac{1}{2}\int_0^t \theta^2(u)du\right\} \quad \textit{and}$$

$$\tilde{W}(t) = W(t) + \int_0^t \theta(u)du\,.$$

Assume that $\int_0^T \mathbb{E}(Z^2(t)\theta^2(t))dt < \infty$. Set $Z = Z(T)$. Then $\mathbb{E}(Z) = 1$, and under the probability measure $\tilde{\mathbb{P}}$ given by $\tilde{\mathbb{P}}(A) = \int_A Zd\mathbb{P}$, $\tilde{W}(t)$ is a Brownian motion.

Proof: By Levy's theorem we need to show that $\tilde{W}$ is a $\tilde{\mathbb{P}}$ martingale and $[\tilde{W},\tilde{W}](t) = t$.

$$\begin{aligned} d\tilde{W}(t) &= dW(t) + \theta(t)dt\\ \Rightarrow d\tilde{W}(t)d\tilde{W}(t) &= dW(t)dW(t) = dt\,.\end{aligned}$$

Thus it remains to show that $\tilde{W}$ is a martingale under $\tilde{\mathbb{P}}$. Now by Ito's formula,

$$dZ(t) = -\theta(t)Z(t)dW(t)\,.$$

Thus $Z(t)$ is a martingale. Hence $\mathbb{E}(Z) = \mathbb{E}(Z(T)) = \mathbb{E}(Z(0)) = 1$. Since $Z(t)$ is a martingale,

$$z(t) = \mathbb{E}[Z(T)|\mathcal{F}_t] = \mathbb{E}[Z|\mathcal{F}_t]\,.$$

Thus $Z(t)$ is a Radon Nikodym derivative process and the above two lemmas are applicable. Now by Ito's product rule,

$$\begin{aligned} d(\tilde{W}(t)Z(t)) &= \tilde{W}(t)dZ(t) + Z(t)d\tilde{W}(t) + dZ(t)d\tilde{W}\\ &= -theta(t)\tilde{W}(t)Z(t)dW(t) + Z(t)dW(t) + Z(t)\theta(t)dt - Z(t)\theta(t)dt\\ &= (-\tilde{W}(t)\theta(t) + 1)Z(t)dW(t)\,.\end{aligned}$$

Thus $\tilde{W}(t)Z(t)$ is a martingale under $\mathbb{P}$. Hence using the lemma above,

$$\tilde{\mathbb{E}}[\tilde{W}(t)|\mathcal{F}_s] = \frac{1}{Z(s)}\tilde{\mathbb{E}}[\tilde{W}(t)Z(t)|\mathcal{F}_s] = \frac{1}{Z(s)}\tilde{W}(s)Z(s) = \tilde{W}(s)\,.$$

Hence the proof. $\square$

4.2 Black Scholes Market with Single Stock

Let $W(t), t \in [0,T]$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_t, t \in [0,T]$ be a filtration for this Brownian motion. Consider a stock price process satisfying,

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t)\,,$$

where $\alpha(t)$ and $\sigma(t)$ are appropriate adapted processes. Such a process is called a generalized geometric Brownian motion.

Exercise Using Ito's formula show that $S(t)$ is given by

$$S(t) = S(0)\exp\left\{\int_0^t \sigma(s)dW(s) + \int_0^t (\alpha(s) - \frac{\sigma^2(s)}{2})ds\right\}\,.$$

In addition assume that we have an adapted interest rate process $R(t)$. We define the discount process by

$$D(t) = e^{-\int_0^t R(s)ds}\,.$$

Thus $dD(t) = -R(t)D(t)dt$. Thus by Ito's product rule the discounted stock price process is

$$D(t)S(t) = S(0)\exp\left\{\int_0^t \sigma(s)dW(s) + \int_0^t (\alpha(s) - R(s) - \frac{\sigma^2(s)}{2})ds\right\}\,,$$

and its differential is

$$\begin{aligned} d(D(t)S(t)) &= (\alpha(t) - R(t))D(t)S(t)dt + \sigma(t)D(t)S(t)dW(t) \\ &= \sigma(t)D(t)S(t)(\theta(t)dt + dW(t)\,, \end{aligned}$$

where $\theta(t) = \frac{\alpha(t)-R(t)}{\sigma(t)}$. Now using this $\theta(t)$ we define the measure $\tilde{\mathbb{P}}$ via Girsanov's theorem. Under the new measure

$$d\tilde{W}(t) = dW(t) + \theta(t)dt$$

is a Brownian motion. Thus

$$d(D(t)S(t)) = \sigma(t)D(t)S(t)d\tilde{W}(t)\,.$$

Hence under $\tilde{\mathbb{P}}$, $D(t)S(t)$ is a martingale. This measure $\tilde{\mathbb{P}}$ is called the risk neutral measure. Replacing $W(t)$ by $\tilde{W}(t)$ we see that $S(t)$ satisfies

$$dS(t) = R(t)S(t)dt + \sigma(t)S(t)dtildeW(t)\,,$$

or equivalently

$$S(t) = S(0)\exp\left\{\int_0^t \sigma(s)d\tilde{W}(s) + \int_0^t (R(s) - \frac{\sigma^2(s)}{2})ds\right\}\,.$$

Consider an agent who begins with an initial capital $X(0)$ and at each time t, $0 \leq t \leq T$, holds $\Delta(t)$ shares of stock, investing or borrowing at the interest rate $R(t)$ as necessary to finance this. The differential of the porfolio is give by,

$$\begin{aligned} dX(t) &= \Delta(t)dS(t) + R(t)(X(t) - \Delta(t)S(t))dt \\ &= R(t)X(t)dt + \Delta(t)(\alpha(t) - R(t))S(t)dt + \Delta(t)\sigma(t)S(t)dW(t) \\ &= R(t)X(t)dt + \Delta(t)\sigma(t)S(t)(\theta(t(dt + dW(t)) \,. \end{aligned}$$

So by Ito's product rule,

$$d(D(t)X(t)) = \Delta(t)\sigma(t)D(t)S(t)(\theta(t)dt + dW(t)) = \Delta(t)\sigma(t)D(t)S(t)d\tilde{W}(t) \,.$$

Thus the discounted portfolio value process is a martingale under $\tilde{\mathbb{P}}$.
While deriving the Black-Scholes-Merton equation for the price of a European call option, we asked what initial capital $X(0)$ and portfolio process $\Delta(t)$ an agent would need in order to hedge a short position in the call, i.e., in order to have $X(T) = (S(T) - K)^+$. We generalise the question in this chapter. Let $V(T)$ be an $\mathcal{F}_T$ measurable random variable representation the payoff at time T of an European derivative security. We are interested to know what initial capital $X(0)$ and portfolio process $\Delta(t)$ an agent would need in order to hedge a short position in this derivative security, i.e., in order to have $X(T) = V(T)$. So the question is whether this is at all possible. Now if this can be done then the fact that the discounted portfolio process is a martingale under the risk neutral measure $\tilde{\mathbb{P}}$ implies,

$$D(t)X(t) = \tilde{\mathbb{E}}[D(T)X(T)|\mathcal{F}_t] = \tilde{\mathbb{E}}[D(T)V(T)|\mathcal{F}_t] \,.$$

Now since $X(t)$ is the value of the hedging portfolio at time t, by no-arbitrage argument this should be the price of the derivative security at time t. Thus if we denote the price of the derivative security at time t, by $V(t)$, the $V(t)$ must be given by

$$V(t) = \frac{1}{D(t)}\tilde{\mathbb{E}}[D(T)V(T)|\mathcal{F}_t] \,.$$

This is known as the risk neutral valuation formula. Now let us use this formula to re-obtain the BSM price of an European call. Thus for this part we assume constant volatility σ and constant interest rate r. Thus by the risk neutral valuation formula the call price should be

$$c(t, S(t)) = \tilde{\mathbb{E}}[e^{-r(T-t)}(S(T) - K)^+|\mathcal{F}_t] \,,$$

where $S(t)$ satisfies

$$S(t) = S(0)\exp\{\sigma\tilde{W}(t) + (r - \sigma^2/2)t\} \,.$$

Thus

$$\begin{aligned} S(T) &= S(t)\exp\{\sigma(\tilde{W}(T) - \tilde{W}(t)) + (r - \sigma^2/2)\tau\} \\ &= S(t)\exp\{\sigma(-\sigma\sqrt{\tau}Y + (r - \sigma^2/2)\tau\} \end{aligned}$$

Where Y is the standard normal random variable $Y = -\frac{\tilde{W}(T)-\tilde{W}(t)}{\sqrt{T-t}}$ and $\tau = T-t$. Thus $S(T)$ is the product of $\mathcal{F}_t$ measurable random variable $S(t)$ and the random variable $\exp\{\sigma(-\sigma\sqrt{\tau}Y + (r-\sigma^2/2)\tau\}$ which is independent of $\mathcal{F}_t$. Thus by Independence lemma,

$$\begin{aligned}
c(t,x) &= \tilde{\mathbb{E}}[e^{-r\tau}(x\exp\{\sigma(-\sigma\sqrt{\tau}Y + (r-\sigma^2/2)\tau\} - K)^+] \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-r\tau}(x\exp\{\sigma(-\sigma\sqrt{\tau}y + (r-\sigma^2/2)\tau\} - K)^+ e^{-y^2/2}dy\,.
\end{aligned}$$

Now

$$\begin{aligned}
& x\exp\{\sigma(-\sigma\sqrt{\tau}y + (r-\sigma^2/2)\tau\} - K > 0 \\
\Longleftrightarrow & y < \frac{1}{\sigma\sqrt{\tau}}[\log(x/K) + (r-\sigma^2/2)\tau] = d_(\tau,x)\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
c(t,x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_(\tau,x)} e^{-r\tau}(x\exp\{\sigma(-\sigma\sqrt{\tau}y + (r-\sigma^2/2)\tau\} - K)^+ e^{-y^2/2}dy \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_(\tau,x)} xe^{-\frac{y^2}{2}-\sigma\sqrt{\tau}y-\frac{\sigma^2\tau}{2}}dy - = \frac{K}{\sqrt{2\pi}}\int_{-\infty}^{d_(\tau,x)} e^{-r\tau-\frac{y^2}{2}}dy \\
&= \frac{x}{\sqrt{2\pi}}\int_{-\infty}^{d_(\tau,x)} e^{-\frac{(y+\sigma\sqrt{tau})^2}{2}}dy - e^{-r\tau}KN(d_(\tau,x)) \\
&= \frac{x}{\sqrt{2\pi}}\int_{-\infty}^{d_(\tau,x)+\sigma\sqrt{\tau}} e^{-\frac{z^2}{2}}dz - e^{-r\tau}KN(d_(\tau,x)) \\
&= xN(d+(\tau,x)) - e^{-r\tau}KN(d_(\tau,x))\,,
\end{aligned}$$

where $d_+(\tau,x) = d_(\tau,x) + \sigma\sqrt{\tau} = \frac{1}{\sigma\sqrt{\tau}}[\log(x/K) + (r+\sigma^2/2)\tau]$. Thus we have obtained the same formula.

Exercise: Using risk neutral valuation formula find the price of a forward contract on the stock price $S(\cdot)$ with strike price K and maturity T. (Recall the payoff at maturity is $S(T)-K$.)

Risk neutral evaluation formula was derived under the assumption that if an agent begins with the correct initial capital, then there exists a portfolio process $\delta(t)$ such that the agent's portfolio value at time T will be $V(T)$. We will now verify the assumption. The existence of a hedging portfolio depends on the following theorem.

Theorem 4.4. *(Martingale Representation Theorem) Let $W(t), 0 \le t \le T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_t, 0 \le t \le T$ be the filtration generated by this Brownian motion. Let $M(t), 0 \le t \le T$ be a martingale with respect to $\mathcal{F}_t$. Then there is an adapted process $\Gamma(t), 0 \le t \le T$, such that*

$$M(t) = M(0) + \int_0^t \Gamma(u)dW(u), \quad 0 \le t \le T\,.$$

Exercise: Let $\mathcal{F}_t$ be the filtration generated by the Brownian motion $W(\cdot)$. Find the martingale representation for the following martingales:

- $M_t = \mathbb{E}[W^2(T)|\mathcal{F}_t]$, $t \in [0,T]$.
- $M_t = \mathbb{E}[W^3(T)|\mathcal{F}_t]$, $t \in [0,T]$.
- $M_t = \mathbb{E}[e^{B(T)}|\mathcal{F}_t]$, $t \in [0,T]$.

Now we return to the hedging problem. Define $V(t)$ by the risk neutral evaluation formula,

$$D(t)V(t) = \tilde{\mathbb{E}}[D(T)V(T)|\mathcal{F}_t]\,.$$

Then $D(t)V(t)$ is a martingale with respect to $\mathcal{F}_t$. Now it is also known that for any portfolio value process $X(t)$ we have,

$$D(t)X(t) = X(0) + \int_0^t \Delta(u)\sigma(u)D(u)S(u)d\tilde{W}(u)\,.$$

By MRT, there exists $\Gamma(t), 0 \le t, \le T$ such that

$$D(t)V(t) = V(0) + \int_0^t \Gamma(u)d\tilde{W}(u)\,.$$

So if we want $X(t) = V(t)$ for all $t \in [0,T]$, we choose $X(0) = V(0)$ and $\Delta(t)$ satisfying,

$$\Delta(t) = \frac{\Gamma(t)}{\sigma(t)D(t)S(t)}\,.$$

4.3 Black Scholes Market with Multiple Stocks

Now we will extend our market to the case of multiple stocks driven by multiple Brownian motions.

Theorem 4.5. *(Girsanov, multiple dimensions) Let $W(t) = (W_1(t), \ldots, W_d(t)), t \in [0,T]$ be a d-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_t, t \in [0,T]$ be a filtration for this Brownian motion. Let $\Theta(t) = (\Theta_1(t), \ldots, \Theta_d(t)), t \in [0,T]$ be an adapted d-dimensional process. process. Define*

$$Z(t) = \exp\left\{-\int_0^t \Theta(u)\cdot dW(u) - \frac{1}{2}\int_0^t ||\Theta||^2(u)du\right\} \quad \textit{and}$$

$$\tilde{W}(t) = W(t) + \int_0^t \Theta(u)du\,.$$

Assume that $\int_0^T \mathbb{E}(Z^2(t)||\Theta||^2(t))dt < \infty$. Set $Z = Z(T)$. Then $\mathbb{E}(Z) = 1$, and under the probability measure $\tilde{\mathbb{P}}$ given by $\tilde{\mathbb{P}}(A) = \int_A Zd\mathbb{P}$, $\tilde{W}(t)$ is a d-dimensional Brownian motion.

Theorem 4.6. *(MRT, multiple dimensions) Let $\mathcal{F}_t, 0 \le t \le T$ be the filtration generated by the $d-$dimensional Brownian motion. Let $M(t), 0 \le t \le T$ be a martingale with respect to $\mathcal{F}_t$. Then there is a d-dimensional adapted process $\Gamma(t) = (\Gamma_1(t), \ldots, \Gamma_d(t)), 0 \le t \le T$, such that*

$$M(t) = M(0) + \int_0^t \Gamma(u)\cdot dW(u), \quad 0 \le t \le T\,.$$

Consider a market with m stocks, each satisfying the stochastic differential equation

$$dS_i(t) = \alpha_i(t)S_i(t)dt + S_i(t)\sum_{j=1}^{d} \sigma_{ij}(t)dW_j(t)\,,$$

for $i = 1, 2, \ldots, m$ and where $W = (W_1, W_2, \ldots, W_d)$ is a $d-$dimensional Brownian motion. Set $\sigma_i(t) = \sqrt{\sum_{j=1}^{d} \sigma_{ij}^2(t)}$, which we assume is never zero. Define $B_i(t) = \sum_{j=1}^{d} \int_0^t \frac{\sigma_{ij}(u)}{\sigma_i(u)} dW_j(u)$, $i = 1, 2, \ldots, m$. Each $B_i(t)$ is a continuous martingale and

$$dB_i(t)dB_i(t) = \sum_{j=1}^{d} \frac{\sigma_{ij}^2}{\sigma_i^2(t)}dt = dt.$$

Thus each $B_i(t)$ is a Brownian motion. In terms of $B_i(t)$ we have,

$$dS_i(t) = \alpha_i(t)S_i(t)dt + S_i(t)\sigma_i(t)dB_i(t)\,.$$

Exercise: Use Ito's product rule to show that $Cov(B_i(t), B_k(t)) = \mathbb{E}\left(\int_0^t \frac{\sum_{j=1}^{d} \sigma_{ij}(u)\sigma_{kj}(u)}{\sigma_i(u)\sigma_k(u)} du\right)$.

Thus $S_i(t)$s are also correlated. We assume an adapted interest rate process and define the discount process by $D(t) = e^{-\int_0^t R(u)du}$ or in differential form, $dD(t) = -D(t)R(t)dt$. So by Ito's product rule,

$$\begin{aligned} d(D(t)S_i(t)) &= D(t)S_i(t)[(\alpha_i(t) - R(t))dt + \sigma_i(t)dB_i(t)] \\ &= D(t)S_i(t)[(\alpha_i(t) - R(t))dt + \sum_{j=1}^{d} \sigma_{ij}(t)dW_j(t)]\,. \end{aligned}$$

Definition 4.7. *A probability measure $\tilde{\mathbb{P}}$ is said to be a risk neutral measure if*

1. *$\mathbb{P}$ and $\tilde{\mathbb{P}}$ are equivalent.*
2. *Under $\tilde{\mathbb{P}}$, the discounted stock price process $D(t)S_i(t)$ is a martingale for all $i = 1, 2, \ldots, m$.*

If we can rewrite

$$d(D(t)S_i(t)) = D(t)S_i(t)[(\alpha_i(t) - R(t))dt + \sum_{j=1}^{d} \sigma_{ij}(t)dW_j(t)]$$

as

$$d(D(t)S_i(t)) = D(t)S_i(t)\sum_{j=1}^{d} \sigma_{ij}(t)[\Theta_j(t)dt + dW_j(t)]$$

for some Θ_j, then we can use multi-dimensional Girsanov theorem to construct an equivalent probability measure $\tilde{\mathbb{P}}$ under which $\tilde{W} = (\tilde{W}_1, \ldots, \tilde{W}_d)$ where $d\tilde{W}_j(t) = \Theta_j(t)dt + dW_j(t)$, is a Brownian motion. Thus under $\tilde{\mathbb{P}}$

$$d(D(t)S_i(t)) = D(t)S_i(t)\sum_{j=1}^{d} \sigma_{ij}(t)d\tilde{W}_j(t)\,,$$

and hence $D(t)S_i(t)$ is a martingale.Thus equating dt terms we see that finding a risk neutral measure boils down to finding processes $\Theta_j(t)$ that satisfy

$$\alpha_i(t) - R(t) = \sum_{j=1}^{d} \sigma_{ij}(t)\Theta_j(t), i = 1, 2, \dots, m.$$

These are called the market price of risk equations. (m equations in d unknown processes.) If it is not possible to solve the market price of risk equations, then there is an arbitrage opportunity lurking in the model. We will not see a proof of this result but we will see an example illustrating this. Before coming to the example let us define arbitrage mathematicaly.

Definition 4.8. *An arbitrage is a portfolio value process $X(t)$satisfying $X(0) = 0$ and there exists some $T > 0$ such that,*

$$\mathbb{P}(X(T) \geq 0) = 1, \quad \mathbb{P}(X(T) > 0) > 0\,.$$

Exercise: (i) Suppose the market has an arbitrage. So there is a portfolio value process satisfying $X_1(0) = 0$ and $\mathbb{P}(X(T) \geq 0) = 1, \quad \mathbb{P}(X(T) > 0) > 0$ for some $T > 0$. Show that if $X_2(0)$ is positive, then there exists a portfolio value process $X_2(t)$ starting at $X_2(0)$ and satisfying

$$\mathbb{P}(X_2(T) \geq \frac{X_2(0)}{D(T)}) = 1, \quad \mathbb{P}(X_2(T) > \frac{X_2(0)}{D(T)}) > 0\,.$$

(ii) Suppose that the market has a portfolio process $X_2(t)$ such that $X_2(0)$ is positive and the above holds. Then show that the market has an arbitrage.
Example: Suppose there are two stocks (m=2) and one Brownian motion(d=1) and suppose further that all co-efficients are constants. Thus

$$dS_1(t) = \alpha_1 S_1(t)dt + \sigma_1 S_1(t)dW(t) \quad \text{and}$$

$$dS_2(t) = \alpha_2 S_2(t)dt + \sigma_2 S_2(t)dW(t)\,.$$

Then the market price of risk equations are

$$\alpha_1 - R = \sigma_1\Theta, \alpha_2 - R = \sigma_2\Theta\,.$$

These have a solution if and only if,

$$\frac{\alpha_1 - R}{\sigma_1} = \frac{\alpha_2 - R}{\sigma_2}\,.$$

Suppose this is not the case. Suppose

$$\frac{\alpha_1 - R}{\sigma_1} > \frac{\alpha_2 - R}{\sigma_2}\,.$$

Define

$$\mu = \frac{\alpha_1 - R}{\sigma_1} - \frac{\alpha_2 - R}{\sigma_2} > 0\,.$$

Suppose that at each time t an agent holds $\Delta_1(t) = \frac{1}{S_1(t)\sigma_1}$ shares of stock 1 and $\Delta_2(t) = -\frac{1}{S_2(t)\sigma_2}$ shares of stock 2, borrowing or investing as necessary at the interest rate R to setup and maintain this portfolio. The initial capital required to take the stock positions is $\frac{1}{\sigma_1} - \frac{1}{\sigma_2}$. If this is positive then we borrow and if this is negative then we invest. So the initial capital required to setup this portfolio is 0, i.e., $X(0) = 0$. Now

$$\begin{aligned} d(X(t)) &= \Delta_1(t)dS_1(t) + \Delta_2(t)dS_2(t) + R(X(t) - \Delta_1(t)S_1(t) - \Delta_2(t)S_2(t))dt \\ &= \frac{\alpha_1 - R}{\sigma_1}dt + dW(t) - \frac{\alpha_2 - R}{\sigma_2}dt - dW(t) + RX(t)dt \\ &= \mu dt + RX(t)dt\,. \end{aligned}$$

The differential of the discounted portfolio is

$$d(e^{-Rt}X(t)) = \mu e^{-Rt}dt\,.$$

Hence

$$\begin{aligned} e^{-Rt}X(t) &= \frac{\mu}{R}(1 - e^{-Rt}) \\ \text{implies}\quad X(t) &= \frac{\mu}{R}(e^{Rt} - 1)\,. \end{aligned}$$

Thus this is an arbitrage opportunity.
Now consider an agent who begins with an initial capital of $X(0)$ and at each time t, holds $\Delta_i(t)$ shares of stock S_i, investing and borrowing from the market as necessary. Thus the differential of the portfolio is given by

$$\begin{aligned} dX(t) &= \sum_{i=1}^{m} \Delta_i(t)dS_i(t) + R(t)(X(t) - \sum_{i=1}^{m} \Delta_i(t)S_i(t))dt \\ &= R(t)X(t)dt + \sum_{i=1}^{m} \Delta_i(t)(dS_i(t) - R(t)S_i(t)dt) \\ &= R(t)X(t)dt + \sum_{i=1}^{m} \frac{\Delta_i(t)}{D(t)}(D(t)dS_i(t) - D(t)R(t)S_i(t)dt) \\ &= R(t)X(t)dt + \sum_{i=1}^{m} \frac{\Delta_i(t)}{D(t)}d(D(t)S_i(t))\,. \end{aligned}$$

Thus

$$\begin{aligned} d(D(t)X(t)) &= \sum_{i=1}^{m} \Delta_i(t)d(D(t)S_i(t)) \\ &= \sum_{i=1}^{m} \Delta_i(t)D(t)S_i(t)\sum_{j=1}^{d} \sigma_{ij}(t)d\tilde{W}_j(t) \\ &= \sum_{j=1}^{d} \left(\sum_{i=1}^{m} \Delta_i(t)D(t)S_i(t)\sigma_{ij}(t) \right) d\tilde{W}_j(t)\,. \end{aligned}$$

Thus under the risk neutral measure $\tilde{\mathbb{P}}$, the discounted portfolio process is also a martingale.

4.4 Fundamental Theorems of Asset Pricing

Theorem 4.9. *(First Fundamental Theorem of Asset Pricing) If a market model has a risk neutral measure then it does not admit any arbitrage.*

Proof: If a market model has a risk neutral measure $\tilde{\mathbb{P}}$, then every discount portfolio value process is a martingale under $\tilde{\mathbb{P}}$. In particular, every portfolio value process satisfies $\tilde{\mathbb{E}}(D(T)X(T)) = X(0)$ for all $T > 0$. Let $X(t)$ be a portfolio value process with $X(0) = 0$. Suppose there exists $T > 0$ such that $\mathbb{P}(X(T) \geq 0) = 1$, i.e., $\mathbb{P}(X(T) < 0) = 0$. Since $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are equivalent, $\tilde{\mathbb{P}}(X(T) < 0) = 0$. Thus $\tilde{\mathbb{P}}(X(T) \geq 0) = 1$.
Claim: $\tilde{\mathbb{P}}(X(T) > 0) = 0$. If not, then $\tilde{\mathbb{P}}(X(T) > 0) > 0$, implies $\tilde{\mathbb{E}}(D(T)X(T)) > 0$, which is a contradiction. Hence the claim. By equivalence of $\mathbb{P}$ and $\tilde{\mathbb{P}}$, $\mathbb{P}(X(T) > 0) = 0$. Since $X(t)$ was any portfolio, there cannot exist an arbitrage opportunity. □

Definition 4.10. *A market model is said to be complete if every derivative security can be hedged.*

Suppose that the market model has a risk neutral measure. That means, we have been able to solve the market price of risk equations, used the resulting Θ_is to define the risk neutral measure $\tilde{P}$ via Girsanov's theorem. Further suppose that the filtration is generated by the d-dimensional Brownian motion $W(t)$. Let $V(T)$ be an $\mathcal{F}_T$ measurable random variable representing the payoff of some time T maturity derivative security. We want to know whether it is possible to hedge a short position in this derivative security. Define the process $V(t), 0 \leq t \leq T$ by

$$D(t)V(t) = \tilde{\mathbb{E}}[D(T)V(T)|\mathcal{F}_t]\,.$$

Then $D(t)V(t)$ is a martingale under $\tilde{\mathbb{P}}$ and so by martingale representation theorem there exists a d-dimensional process $\Gamma(t) = (\Gamma_1(t), \ldots, \Gamma_d(t))$ such that for all $t \in [0, T]$,

$$D(t)V(t) = V(0) + \sum_{j=1}^{d} \int_0^t \Gamma_j(u) d\tilde{W}_j(u)\,.$$

Now for any portfolio value process $X(t)$ we have

$$d(D(t)X(t)) = \sum_{j=1}^{d} \left(\sum_{i=1}^{m} \Delta_i(t) D(t) S_i(t) \sigma_{ij}(t) \right) d\tilde{W}_j(t)\,.$$

Thus

$$D(t)X(t) = X(0) + \sum_{j=1}^{d} \int_0^t \left(\sum_{i=1}^{m} \Delta_i(u) D(u) S_i(u) \sigma_{ij}(u) \right) d\tilde{W}_j(u)\,.$$

So if we start with an initial capital of $X(0) = V(0)$ and is able to choose portfolio processes $\Delta_1(t), \ldots, \Delta_m(t)$ such that the hedging equations

$$\frac{\Gamma_j(t)}{D(t)} = \sum_{i=1}^{m} \Delta_i(t) S_i(t) \sigma_{ij}(t)$$

are satisfied for $j = 1, \ldots, d$ then we get,

$$D(t)X(t) = V(0) + \sum_{j=1}^{d} \int_0^t \Gamma_j(u) d\tilde{W}_j(u) = D(t)V(t)$$

for all $t \in [0, T]$. Thus $X(t) = V(t)$ for all $t \in [0, T]$, or in other words, $X(t)$ is a hedging portfolio.

Theorem 4.11. *(Second Fundamental Theorem of Asset Pricing) Consider a market model that has a risk neutral measure. The model is complete if and only if the risk neutral measure is unique.*

Proof: We first assume that the model is complete. Suppose the model has two risk neutral probability measures $\tilde{\mathbb{P}}_1$ and $\tilde{\mathbb{P}}_2$. Let A be a set in $\mathcal{F}_T = \mathcal{F}$. Consider the derivative security with payoff $V(T) = 1_A \frac{1}{D(T)}$. Because the model is complete, a short position in this derivative security can be hedged, i.e., there exists a portfolio value process $X(t)$ with some initial condition $X(0)$ and satisfies $X(T) = V(T)$. Since both $\tilde{\mathbb{P}}_1$ and $\tilde{\mathbb{P}}_2$ are risk neutral measures, the discounted portfolio value process $D(t)X(t)$ is a martingale under both $\tilde{\mathbb{P}}_1$ and $\tilde{\mathbb{P}}_2$. Thus

$$\begin{aligned}\tilde{\mathbb{P}}_1(A) &= \tilde{\mathbb{E}}_1(D(T)V(T)) = \tilde{\mathbb{E}}_1(D(T)X(T)) \\ &= X(0) = \tilde{\mathbb{E}}_2(D(T)X(T)) = \tilde{\mathbb{E}}_2(D(T)V(T)) = \tilde{\mathbb{P}}_2(A)\,.\end{aligned}$$

Since A was an arbitrary set in $\mathcal{F}$, we have that the measures are equal.
For the converse, suppose there is only one risk neutral measure. Thus the market price of risk equations has a unique solution. These equations are of the form $Ax = b$ where A is the $m \times d$ dimensional matrix

$$A = \begin{bmatrix} \sigma_{11}(t) \ldots \sigma_{1d}(t) \\ \vdots \\ \sigma_{m1}(t) \ldots \sigma_{md}(t) \end{bmatrix},$$

x is the d-dimensional column vector

$$x = \begin{bmatrix} \Theta_1(t) \\ \vdots \\ \Theta_d(t) \end{bmatrix},$$

and b is the m-dimensional column vector

$$b = \begin{bmatrix} \alpha_1(t) - R(t) \\ \vdots \\ \alpha_m(t) - R(t) \end{bmatrix}.$$

Since by assumption this system of equation has a unique solution, so $KerA = \{x \in \mathbb{R}^d : Ax = \mathbf{0}\}$ must be trivial, i.e., $KerA = \{\mathbf{0}\}$. Thus the columns of A are linearly independent.

Thus $rankA = d = rankA^t$. Now in order to show that the market is complete we need to show that the hedging equations always has a solution. The hedging equations can be written in the form $A^t y = c$ where y is the m-dimensional column vector

$$y = \begin{bmatrix} \Delta_1(t)S_1(t) \\ \vdots \\ \Delta_m(t)S_m(t) \end{bmatrix},$$

and c is the d-dimensional column vector

$$b = \begin{bmatrix} \frac{\Gamma_1(t)}{D(t)} \\ \vdots \\ \frac{\Gamma_d(t)}{D(t)} \end{bmatrix}.$$

Now since $rankA^t = d$, we have $rangeA^t = \mathbb{R}^d$ and hence the completeness follows. $\square$

5 Stochastic Differential Equations

Consider a SDE of the form

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \quad X(0) = Z\,.$$

Question 1: Does there exist a solution? And if there is a solution then is it unique?
Question 2: How to solve such a SDE?
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $W(\cdot)$ be a Brownian motion defined on it. Let $\mathcal{F}_t$ be the filtration generated by $W(t)$ and Z, i.e., $\mathcal{F}_t = \sigma\{Z, W(s), s \le t\}$.

Definition 5.1. *A solution of the SDE above is a continuous stochastic process $X(t), 0 \le t \le T$ with the following properties:*

1. *$X(t)$ is adapted to the filtration $\mathcal{F}_t$.*
2. *$\mathbb{P}(X(0) = Z) = 1$.*
3. *$\int_0^T \mathbb{E}(|b(t, X(t))|)dt < \infty$, $\int_0^T \mathbb{E}(|\sigma(t, X(t))|^2)dt < \infty$.*
4. *$X(t) = X(0) + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s)$, $0 \le t \le T$ almost surely.*

Definition 5.2. *The SDE above is said to have a unique solution, if X and $\tilde{X}$ are two solutions, then $\mathbb{P}(X(t) = \tilde{X}(t), 0 \le t \le T) = 1$.*

Theorem 5.3. *(Existence and Uniqueness) Suppose the co-efficients $b(t, x)$ and $\sigma(t, x)$ satisfy the global lipschitz and linear growth conditions*

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \le K|x - y| \quad \textit{and}$$

$$|b(t, x)| + |\sigma(t, x)| \le K(1 + |x|)$$

for some positive constant K. Further suppose that $\mathbb{E}(Z^2) < \infty$. Then the SDE has a unique solution. Further $X(t)$ satisfies $\mathbb{E}\int_0^T |X(t)|^2 dt < \infty$.

Consider the deterministic differential equations

$$dX(t) = X^2(t)dt \quad \text{and} \quad dX(t) = 3X^{2/3}(t)dt\,.$$

For the first one $B(t, x) = x^2$ does not satisfy the linear growth condition. For the second one, $b(t, x) = 3x^{2/3}$ does not satisfy the lipschitz condition. In the first case, the solution (unique) is $X(t) = \frac{1}{1-t}$. But this "explodes" as $t \uparrow 1$. Thus linear growth condition ensures that the solution does not "explode" in finite time.
For the second case, there are infinitely many solutions, in fact for any $a > 0$,

$$X(t) = \begin{cases} 0 & \text{for } t \le a \\ (t-a)^3 & \text{for } t \ge a\,, \end{cases}$$

is a solution. Thus lipschitz property ensures uniqueness.

Lemma 5.4. *(Gronwall's Inequality) Let $f(\cdot)$ be a continuous function such that*

$$f(t) \leq C + K\int_0^t f(s)ds \quad \text{for } t \in [0,T]\,,$$

where C is a constant and K is a positive constant. Then $f(t) \leq Ce^{Kt}$ for $t \in [0,T]$.

Proof: Define $W(t) = C + K\int_0^t f(s)ds$. Then $W(t) \geq f(t)$ for all $t \in [0,T]$. Now bt Fundamental Theorem of Calculus,

$$\begin{aligned}
&W'(t) = Kf(t) \leq KW(t)\\
\Rightarrow & e^{-Kt}W'(t) - Ke^{-Kt}W(t) \leq 0\\
\Rightarrow & \frac{d}{dt}(e^{-Kt}W(t)) \leq 0\\
\Rightarrow & e^{-Kt}W(t) - W(0) \leq 0\\
\Rightarrow & W(t) \leq Ce^{Kt}\\
\Rightarrow & f(t) \leq Ce^{Kt}\,.
\end{aligned}$$

□**Proof of Uniqueness:** Suppose there exists two solutions $X_1(t)$ and $X_2(t)$. Thus

$$X_1(t) = Z + \int_0^t b(s, X_1(s))ds + \int_0^t \sigma(s, X_1(s))dW(s)\,,$$

$$X_2(t) = Z + \int_0^t b(s, X_2(s))ds + \int_0^t \sigma(s, X_2(s))dW(s)\,.$$

Thus,

$$\begin{aligned}
&\mathbb{E}|X_1(t) - X_2(t)|^2 = \mathbb{E}\left(\int_0^t [b(s,X_1(s)) - b(s,X_2(s))]ds + \int_0^t [\sigma(s,X_1(s)) - \sigma(s,X_2(s))]dW(s)\right)^2\\
&\leq 2\left[\mathbb{E}\left(\int_0^t [b(s,X_1(s)) - b(s,X_2(s))]ds\right)^2 + \mathbb{E}\left(\int_0^t [\sigma(s,X_1(s)) - \sigma(s,X_2(s))]dW(s)\right)^2\right]\\
&\leq 2\left[t\int_0^t \mathbb{E}\left(b(s,X_1(s)) - b(s,X_2(s))\right)^2 ds + \int_0^t \mathbb{E}\left(\sigma(s,X_1(s)) - \sigma(s,X_2(s))\right)^2 ds\right]\\
&\leq 2K^2(1+t)\int_0^t \mathbb{E}|X_1(s) - X_2(s)|^2 ds\,.
\end{aligned}$$

Thus

$$\mathbb{E}|X_1(t) - X_2(t)|^2 \leq 2K^2(1+T)\int_0^t \mathbb{E}|X_1(s) - X_2(s)|^2 ds\,,$$

for all $t \in [0,T]$. So by Gronwall's inequality,

$$\mathbb{E}|X_1(t) - X_2(t)|^2 = 0\,.$$

Thus $\mathbb{P}(X_1(t) = X_2(t)) = 1$ for each $t \in [0, T]$. Thus $\mathbb{P}(X_1(t) = X_2(t)\ \forall\ t \in \mathbb{Q} \cap [0, T]) = 1$. By continuity, $\mathbb{P}(X_1(t) = X_2(t)\ \forall\ t \in [0, T]) = 1$. Hence we have the uniqueness. $\square$

Consider the SDE

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t), \quad S(0) = S_0\,,$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are deterministic functions satisfying $|\mu(t)| + |\sigma(t)| \leq K$. Thus $b(t, x) = \mu(t)x$ and $\sigma(t, x) = \sigma(t)x$. Both the conditions for uniqueness and existence are satisfied. Now by Ito's formula,

$$\begin{aligned} d(\log(S(t)) &= \frac{1}{S(t)}dS(t) - \frac{1}{2S^2(t)}dS(t)dS(t) \\ &= \frac{1}{S(t)}dS(t)[\mu(t)S(t)dt + \sigma(t)S(t)dW(t)] - \frac{1}{2S^2(t)}[\sigma^2(t)S^2(t)] \\ &= \sigma(t)dW(t) + (\mu(t) - \frac{1}{2}\sigma^2(t))dt\,. \end{aligned}$$

Thus

$$\log\frac{S(t)}{S_0} = \int_0^t \sigma(s)dW(s) + \int_0^t (\mu(s) - \frac{1}{2}\sigma^2(s))ds$$

$$\Rightarrow S(t) = S_0 \exp\{\int_0^t \sigma(s)dW(s) + \int_0^t (\mu(s) - \frac{1}{2}\sigma^2(s))ds\}\,.$$

Consider a first order ordinary differential equation

$$\frac{dx(t)}{dt} = f(t)x(t) + g(t), \quad x(0) = x\,,$$

where f is a continuous function. Then we know the integrating factor method for solving the above equation. The integrating factor is given by $h(t) = e^{-\int_0^t f(s)ds}$. The solution is given by,

$$x(t) = (h(t))^{-1}x + (h(t))^{-1}\int_0^t h(s)g(s)ds\,.$$

By linear SDE, we mean a SDE of the form,

$$dX(t) = \{\phi(t)X(t) + \theta(t)\}dW(t) + \{f(t)X(t) + g(t)\}dt \quad X(0) = Z\,.$$

Define $H(t)$ as follows:

$$H(t) = e^{-Y(t)}, \quad Y(t) = \int_0^t \phi(s)dW(s) + \int_0^t f(s)ds - \frac{1}{2}\int_0^t \phi^2(s)ds\,.$$

Now $d(H(t)X(t)) = X(t)dH(t) + H(t)dX(t) + dH(t)dX(t)$

$$\begin{aligned} dH(t) &= -e^{-Y(t)}dY(t) + \frac{1}{2}e^{-Y(t)}dY(t)dY(t) \\ &= -H(t)[f(t)dt + \phi(t)dW(t) - \frac{1}{2}\phi^2(t)] + \frac{1}{2}H(t)\phi^2(t)dt \\ &= -H(t)[f(t)dt + \phi(t)dW(t) - \phi^2(t)dt]\,. \end{aligned}$$

Thus

$$d(X(t)dH(t) = -H(t)\phi(t)[\phi(t)X(t) + \theta(t)]dt\,.$$

So we get,

$$\begin{aligned} d(X(t)H(t)) = H(t)[&dX(t) - X(t)f(t)dt - X(t)\phi(t)dW(t) + \\ &\phi^2(t)X(t)dt - \phi^2(t)X(t)dt - \theta(t)\phi(t)dt] \\ = H(t)[&\theta(t)dW(t) + g(t)dt - \theta(t)\phi(t)dt]\,. \end{aligned}$$

Thus

$$H(t)X(t) = Z + \int_0^t H(s)\theta(s)dW(s) + \int_0^t H(s)\{g(s) - \theta(s)\phi(s)\}ds\,.$$

$$\Rightarrow X(t) = Ze^{Y(t)} + \int_0^t e^{Y(t)-Y(s)}\theta(s)dW(s) + \int_0^t e^{Y(t)-Y(s)}\{g(s) - \theta(s)\phi(s)\}ds\,.$$

Example:

$$dX(t) = \mu X(t)dt + \sigma dW(t), \quad X(0) = Z\,.$$

Thus $f(t) = \mu$, $g(t) = 0$, $\phi(t) = 0$, $\theta(t) = \sigma$. Thus $Y(t) = \mu t$. Thus the solution is given by

$$X(t) = Ze^{\mu t} + \int_0^t e^{\mu(t-s)}\sigma dW(s)\,.$$

Exercise:
i) $d(X)(t) = -X(t)dt + e^{-t}dW(t)$, $X(0) = Z$.
ii) $d(X)(t) = rdt + \alpha X(t)dW(t)$, $X(0) = Z$.
iii) $d(X)(t) = (m - X(t))dt + \sigma dW(t)$, $X(0) = Z$.
iv) $d(X)(t) = \frac{1}{2}X(t)dt + X(t)dW(t)$, $X(0) = 1$.
v) $d(X)(t) = \frac{b-X(t)}{1-t}dt + dW(t)$, $X(0) = a$.

Consider the SDE,

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t)\,.$$

Let $h(\cdot)$ be a Borel measurable function. Define the function

$$g(t, x) = \mathbb{E}[h(X(T))|X(t) = x] = \mathbb{E}_{t,x}[h(X(T))]\,. \tag{5.1}$$

Theorem 5.5. *Let $X(u), u \geq 0$, be a solution to the SDE above with some initial condition at 0. Then for any $0 \leq t \leq T$,*

$$\mathbb{E}[h(X(T))|\mathcal{F}_t] = g(t, X(t))\,.$$

Corollary 5.6. *Solutions to stochastic differential equations are Markov processes.*

The following theorem relates SDEs and PDEs.

Theorem 5.7. *(Feynman-Kac) Consider the SDE,*

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t) .$$

Let $h(\cdot)$ be a Borel measurable function. Let $g(t,x)$ be as in (5.1)*. Then $g(t,x)$ satisfies the PDE*

$$g_t(t,x) + b(t,x)g_x(t,x) + \frac{1}{2}\sigma^2(t,x)g_{xx}(t,x) = 0 ,$$

with terminal condition $g(T,x) = h(x)$ for all x.

Proof: Claim: $g(t, X(t))$, $0 \le t \le T$ is a martingale. Now by previous theorem $g(t, X(t)) = \mathbb{E}[h(X(T))|\mathcal{F}_t]$. Thus for $s < t$,

$$\mathbb{E}[g(t, X(t))|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[h(X(T))|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[h(X(T))|\mathcal{F}_s] = g(s, X(s)) .$$

Hence the claim. Now

$$\begin{aligned} dg(t, X(t)) &= g_t(t, X(t))dt + g_x(t, X(t))dX(t) + \frac{1}{2}g_{xx}(t, X(t)) \\ &= (g_t(t, X(t)) + b(t, X(t))g_x(t, X(t)) + \frac{1}{2}\sigma^2(t, X(t))g_{xx}(t, X(t)))dt + g_x(t, X(t))\sigma(t, X(t))dW(t) \end{aligned}$$

Since $g(t, X(t))$ is a martingale, so the the dt term must be equal to 0. Thus we must have

$$g_t(t,x) + b(t,x)g_x(t,x) + \frac{1}{2}\sigma^2(t,x)g_{xx}(t,x) ,$$

for all x. □

Theorem 5.8. *(Discounted Feynman-Kac) Consider the SDE,*

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t) .$$

Let $h(\cdot)$ be a Borel measurable function. Define $f(t,x) = \mathbb{E}_{t,x}[e^{-r(T-t)}h(X(T))]$ Then $f(t,x)$ satisfies the PDE

$$f_t(t,x) + b(t,x)f_x(t,x) + \frac{1}{2}\sigma^2(t,x)f_{xx}(t,x) = rf(t,x) ,$$

with terminal condition $f(T,x) = h(x)$ for all x.

Proof: By a similar argument as in the previous theorem it can be shown that $e^{-rt}f(t, X(t))$ is a martingale. Now,

$$\begin{aligned} &d(e^{-rt}f(t), X(t)) = e^{-rt}df(t, X(t)) - re^{-rt}f(t, X(t)) \\ &= e^{-rt}(-rf(t, X(t)) + f_t(t, X(t)) + b(t, X(t))f_x(t, X(t)) + \frac{1}{2}\sigma^2(t, X(t))f_{xx}(t, X(t)))dt \\ &+ e^{-rt}f_x(t, X(t))\sigma(t, X(t))dW(t) . \end{aligned}$$

So in order to have the dt term equal to 0 we must have,

$$f_t(t,x) + b(t,x)f_x(t,x) + \frac{1}{2}\sigma^2(t,x)f_{xx}(t,x) = rf(t,x)\,.$$

$\square$

Application to BSM model: The risk neutral valuation of an European call is given by

$$V(t) = \tilde{E}[e^{-r(T-t)}(S(T)-K)^+|\mathcal{F}_t]\,,$$

where $S(t)$ satisfies,

$$dS(t) = rS(t)dt + \sigma S(t)d\tilde{W}(t)\,.$$

Thus $S(t)$ is a Markov process. Hence $V(t) = c(t, S(t))$ where $c(t,x) = \tilde{\mathbb{E}}_{t,x}[(S(T)-K)^+]$. So by discounted Feynman-Kac formula, $c(t,x)$ satisfies

$$c_t(t,x) + rxc_x(t,x) + \frac{1}{2}\sigma^2 x^2 c_{xx}(t,x) = rc(t,x)\,,$$

with terminal condition $c(T,x) = (x-K)^+$.