A Mathematical Approach to Black-Scholes Formula

A thesis Submitted in
Partial Fulfillment of the
Requirements for the
Degree of

MASTER OF SCIENCE

by

Rishabh Sareen

to the

School of Mathematical Sciences

National Institute of Science Education and Research

Bhubaneswar
India

April 2013

Dedicated to my dear friends

ACKNOWLEDGEMENTS

I am indebted to my supervisor Dr. Vardharajan Muruganandam for his useful advice and suggestions in the course of this project. I am grateful to him for his patience in attending my lectures pertaining to this project work, clearing out my doubts and giving insights wherever necessary. Finally I would like to thank my dear friends for their valuable discussions.

ABSTRACT

In this report we present an application of probability theory in finance. We start with defining the concept of fair game principle and then move on to define some financial terms like options, hedging, arbitrage situation etc. We then use the fair game principle to develop hedging portfolios and study arbitrage situations. After this we move into deeper mathematical concepts such as conditional expectations and martingales. We use these mathematical concepts to create strategy for hedging any claim on the options. We shall see that conditional expectations is used for creating the strategy, which is latter generalized by using martingales. In other words martingales can be seen as the mathematical formulation of a sequence of fair games. Then we look at the behavior of the share price, which leads to the study of Brownian motion.

We end up in deriving the Black-Scholes Formula. For this we develop the theory of Stochastic or Itô's Calculus. Here we look into concepts like Itô's Integration, Stochastic Integral and Differential Equations, Itô's Processes and Itô's Lemma, Stochastic version of Fundamental Theorem of Calculus, etc. We then look into two different approaches, one is by the first principle approach i.e using the fair game principle and second is by using the martingale properties of the share price as a Brownian motion. Aggregating the above theories we derive the Black-Scholes Equation which leads to Black-Scholes formula, that can hedge any claim on the option.

Contents

Chapter 1

Introduction

"There are very few things which we know, which are not capable of being reduced to a Mathematical Reasoning: and when they cannot, it's a sign our knowledge of them is very small and confused; and where a mathematical reasoning can be had, it's as great a folly to make use of any other, as to grope for a thing in the dark, when you have a candle standing by you."

John Arbuthnot,1692,

Preface, Of the Laws of chance

A *financial market* is a market in which people and entities can trade financial securities or commodities at low transaction costs and at prices that reflect *supply and demand*. Securities include stocks and bonds, and commodities include precious metals or agricultural goods. In economics, typically, the term market means the aggregate of possible buyers and sellers of a certain good or service and the transactions between them.

During the 1980s and 1990s, a major growth sector in financial markets was the trade in so called *derivative products, or derivatives* for short. For example a share holder can trade with a buyer for a share or an *option*, such as the buyer has an option either to buy or not the share after a certain time interval for a fixed decided price. Naturally if the share price rises above the decided price after the given time interval then buyer will buy it and if it falls then he wont. Now the seller will be in loss if the option is to be exercised. So he would charge for the option such as to compensate the loss. The theory of charging the option is known as *Option Pricing*.

Such an option to buy is called *call option*. Similarly an option to sell is called a *put option*. We will study in details the theory of option pricing and the factors affecting it in the report.

In the financial markets, stock prices, bond prices, currency rates, interest rates and dividends go up and down, creating risk. Derivative products are financial products which are used to control risk or paradoxically *exploit risk*. Such studies are called *financial economics*.

In finance, as discussed earlier, an option is a derivative financial instrument that specifies a contract between two parties for a future transaction on an asset at a reference price called *the strike*. The buyer of the option gains the right, but not the obligation, to engage in that transaction, while the seller incurs the corresponding obligation to fulfill the transaction. The price of an option derives from the difference between the reference price and the value of the underlying asset, commonly a stock, a bond, a currency or a futures contract, plus a premium based on the time remaining until the expiration of the option.

An option which conveys the right to buy something at a specific price is called a *call option*; an option which conveys the right to sell something at a specific price is called a *put option*.

In 1973 **Fischer Black** and **Myron Scholes** articulated a model known as the Black-Scholes model in the paper **"The Pricing of Options and Corporate Liabilities"**, published in the **Journal of Political Economy**. They derived a partial differential equation, now called the *Black-Scholes Equation*, which governs the price of the option over time. They transformed the option pricing problem into the task of solving a (parabolic) partial differential equation (PDE) with a final condition. The main conceptual idea of Black and Scholes lies in the construction of a riskless portfolio taking positions in bonds (cash), option, and the underlying stock. Such an

approach strengthens the use of the no-arbitrage principle as well. In other words, the key idea behind the derivation was to hedge, i.e. reduce any substantial losses/gains, perfectly the option by buying and selling the underlying asset in just the right way and consequently "eliminate risk".
This hedge, in turn, implies that there is only one right price for the option, as returned by the Black-Scholes formula derived from a partial differential equation. The partial differential equation, we will study latter, is of the form

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \quad where$$

S stands for price of the stock; $V(S,t)$ is the price of a derivative as a function of time and stock price; σ is the volatility of the stock's returns; (in finance) volatility is a measure for variation of price of a financial instrument over time; and r is the annualized risk-free interest rate, continuously compounded.

Solving the above Black-Scholes equation for the corresponding terminal and boundary conditions we get the Black-Scholes formula. The Black-Scholes formula calculates the price of European put and call options. This price is consistent with the Black-Scholes equation.

Chapter 2

Fair Games and Call Option 1

[↑]

Gambling is one of mankind's oldest and favourite pastimes, whether for profit or pleasure, and provide the early problems that prompted the development of probability theory.

Buying shares is a form of gambling and it is not surprising that the mathematical tools used to analyze the games of chance can be adapted to study the movement of share price. Both the areas have basic concepts, such as risk, reward and hedging in common. By examining simple games we can uncover basic principles which transfer to finance.

Before moving to these ideas lets have a brief look on the relationship between money and time.

2.1 Money and Time

There exists a relationship between *money and time*. When prices are stable, those with money feel financially secured. However, prices do change depending on *supply and demand*. The rate of change over the time in the price of a commodity is called *inflation*. For example, if a share cost Rs 10 last year and now Rs 12 then it has $\frac{12-10}{10} \times 100\% = 20\%$ annual inflation.

Now if the real value of money is what it is capable of buying, then the presence of inflation means that the real value of money is a function of time.

Now inflation creates problem for those with money. In its absence they can estimate their financial obligations and requirements. The presence of inflation reduces their financial securities and forces them to confront an intrinsic problem: *how to maintain the future real value of money?*. Hence the system of renting comes into existence. And the price of renting money is called *interest*.

Interest rates and inflation rates are distinct processes, one increasing the *nominal value* of the money, the other reducing the *real value*. Transactions such as renting money for a price do not involve any *risks*. But other ways, such as investing in business etc, involves risks. Thus risk involvement leads to the birth of a new concept called *speculation*. Since risk is involved, the profit gained, or loss incured, increases manifold times than the usual. So if one can speculate the possibilities of favorable situations, and willing to take chance then he can gain great profits, but with a chance of also losing huge amount. Or he can safely avoid any risk and involves in riskless deals, with minimal profits. Hence we can identify two groups according to their different approaches to the management of money.

Hedgers are those who wish to eliminate risk as much as possible, while *Speculators* are those who are willing to take risks in the expectations of higher profits. [ch 2][↑]

2.2 Fair Game

We consider a simple betting game between 2 players, J and M. Let W_J and W_M be the winnings of J and M respectively. Now if the game is favorable to J, then his *expected winning* is greater than that of M, i.e. $\mathbb{E}[W_J] \geq \mathbb{E}[W_M]$, and vice-versa. It is reasonable to say that the game is fair if it is favorable to both the players, i.e.

$$\mathbb{E}[W_J] = \mathbb{E}[W_M]. \tag{2.2.0.1}$$

Now if S_J and S_M are the stakes or bets placed by J and M respectively, then the game is *Zero-Sum Game* if total input(bets) is equal to total output(winning=S_J +

S_M). In such games,

$$\mathbb{E}[W_J] + \mathbb{E}[W_M] = output - Input = 0. \tag{2.2.0.2}$$

Thus by (2.2.0.1) and (2.2.0.2)

$$\mathbb{E}[W_J] = \mathbb{E}[W_M] = 0.$$

The converse is also true.

Proposition 2.2.1 *A zero-sum game is a fair game iff the expected winnings of each player is zero.*

Now we see what happens if the parameter is changed.

Example 2.2.1

Consider the game of tossing a coin. J wins if head(H) turns and M wins is tail(T) turns. Each of them bets Rs 5 and winner gets Rs 10.

Now Is this a fair ? Yes, provided each outcome is equally likely since then

$$\mathbb{E}[W_J] = n \times \frac{1}{2} \times (10 - 5) + n \times \frac{1}{2} \times (0 - 5) = 0$$

Similarly $\mathbb{E}[W_M] = 0$. And hence the game is fair.

Thus

$$\begin{aligned}\mathbb{E}[W] = \sum_{outcomes} (number\ of\ games) \quad &\times \quad (probability\ of\ this\ outcome) \\ &\times \quad (winnings\ on\ this\ outcome)\end{aligned}$$

Now we change the parameters

(a) J bets Rs 3 and M bets Rs 7.

(b) P(H)= 80%

For case (a) with unbiased coin

$$\mathbb{E}[W_J] = 1/2(10-3) + 1/2(0-3) = 2$$

and for (b)

$$\mathbb{E}[W_J] = 4/5(10-5) + 1/5(0-5) = 3$$

and for (a)+(b)

$$\mathbb{E}[W_J] = 4/5(10-3) + 1/5(0-3) = 5.$$

Thus changing benefits J.

Working backwards i.e. how to change parameters to make a game fair. Suppose J bets Rs x and has winning probability p. Then

$$\mathbb{E}[W_J] = p(10-x) + (1-p)(0-x)$$

and in a fair game $\mathbb{E}[W_J] = 0$. Now for case (a) p=3/10 and for (b) x=8 makes the game fair.

Thus we see that for case (a) or (b) or (a)+(b) J's expected winning is more and the game is in his favor, hence he will be tempted play. But still there is *risk* if M is lucky and wins most of the games. Thus **Risk** enters in the game, and this example clearly distinguishes expected profits from guaranteed profits. [ch 2][↑]

2.3 Hedging and Arbitrage

As seen earlier there are two kinds of traders, the *hedgers*, who try by eliminating risks to maintain the real value of their assets, and speculators, who take risks in hope of large profits. The strategy obtained by the hedgers to eliminate risks is known as *hedging*. There is a third category of traders known as *arbitrageurs*, who move in when they see an opportunity to make riskless profits. We will look at their approach from the following examples.

Example 2.3.1

Consider another betting game on a Horse Race.

Let J bets Rs 400 on horse L with a bookmaker and M bets Rs 100 on Mc with the same bookmaker. Total amount wagered is Rs 500, bookmaker profit is 10% i.e Rs 50 and winner takes Rs 450. Thus J's profit is Rs 50 if L wins and M's profit is Rs 350 if Mc wins. Assuming it to be fair game, if p is the probability of L winning, then

$$\mathbb{E}[W_J] = 50p + (-400)(1-p) = 0$$

implies p=8/9. And if q is probability of Mc winning then

$$\mathbb{E}[W_M] = 350q + (-100)(1-q) = 0$$

implies q=2/9.

Here we observe that J accepts 8/9 as the probability that L will win, while M assumes 7/9 as the probability for the same event. There are two different games being played, one between J and bookmaker and other between M and bookmaker.

The bookmaker makes different terms, a win for Mc will result in 7 win for every 2 bet on it and hence the odds on Mc will be given as 7 to 2. Similarly for L, odds is 1 to 8.

So far the bookmaker has no risk and is guaranteed Rs 50 profit. But now if another bet of Rs 300 is placed on Mc at odds 7 to 2. Then W_B, winning of bookmaker, depends on outcome of the race, i.e.

$$W_B = 50 + 300 = Rs\ 350 \text{ if L wins}$$

$$W_B = 50 - 1050 = -Rs\ 1000 \text{ if Mc wins .}$$

So we encounter a conditional situation, i.e. winning is conditional.

Roughly, Conditional Expectation can be seen as

$$\mathbb{E}[W_B|L \text{ wins}] = 350$$

$$\mathbb{E}[W_B|Mc \text{ wins}] = -1000$$

Now he has a risk of losing money if Mc wins. But he can reduce the risk by following steps :

the total amount wagered is Rs 800 on each horse. If all these bets had been placed initially, then for a profit of 10% i.e Rs 80, winner takes Rs 720, gives

$$\mathbb{E}[W_J] = 0 = 320p + (-400)(1-p)$$

p=5/9. Thus odds should be placed as 4 to 5 on each horse.

The bookmaker changes the odds in line with the new bet, and hopes that these will attract further bets on L. The change in odds will apply only to new bets; and if the next bet is very large and on Mc, then he would run the risk of even more substantial loss. Then he may refuse it or limit the bets, etc.

Another approach is to lay off part of the bet. Here bookmaker becomes a punter, one who places bets, and places a bet on Mc, since Mc wins was leading him to a loss. Suppose another bookmaker gives odds 3 to 1, which is marginally worse than 7 to 2. If he bets x on Mc, then

$$\mathbb{E}[W_B|L \text{ wins}] = 350 - x$$

$$\mathbb{E}[W_B|Mc \text{ wins}] = -1000 + 3x$$

If he is a reluctant hedger and wishes to remove in advance of the outcome any uncertainty regarding his final situation. To achieve this, W_B should not depends on the outcome of the race, and hence x should be such that

$$\mathbb{E}[W_B|L \text{ wins}] = \mathbb{E}[W_B|Mc \text{ wins}]$$

$$350 - x = -1000 + 3x$$

implies x = 337.50

Hence if he bets Rs 337.50 at odds 3 to 1, then $\mathbb{E}[W_B] = 350 - 337.50 = Rs$ 12.50 regardless who wins the race.

By sacrificing an uncertain, but possibly large, profit, the bookmaker removed the risk of uncertainty and has a guaranteed profit.

Now we see from M point of view. His original bet Rs 100 on Mc at odds 7 to 2. We wins Rs 350 if Mc wins and loose Rs 100 otherwise. Similarly if M bets on L at Rs 90, as calculated as in case of the bookmaker, with odds 4 to 1, then

$$\mathbb{E}[W_M|L \text{ wins}] = 360 - 100 = 260$$

$$\mathbb{E}[W_M|Mc \text{ wins}] = 350 - 90 = 260$$

Thus M is guaranteed to win Rs 260 regardless of who won the race. This an *arbitrage* situation.

Thus we observe two significant hedging principle arises

(1) To remove the uncertainty associated with unpredictable future events, equate the associated results and develop a hedging strategy by working backwards.

(2) To reduce the potential loss due to an unfavorable occurring, place a bet in favor of the event happening

Latter we will utilize similar techniques to hedge the claim made in pricing the call options.

Now using the basic ideas on interest rates, fair games and intuitive notion of expected value, as a weighted average, we present the *binomial model* for pricing *options.*

[ch 2][↑]

2.4 Call option 1

A *call option* is an option to buy a certain asset, the underlying security, on or before a certain date, the *maturity date* or the *exercised date*, for a certain price.

If the call option is for a fixed quantity of shares, the price per at maturity, if the option is taken up or exercised, is called the *strike price* or *exercised price*.

An option to cell is called a *put option*, and when the option is replaced by an obligation to buy or sell, it is called a *forward contract* or *future contract*.

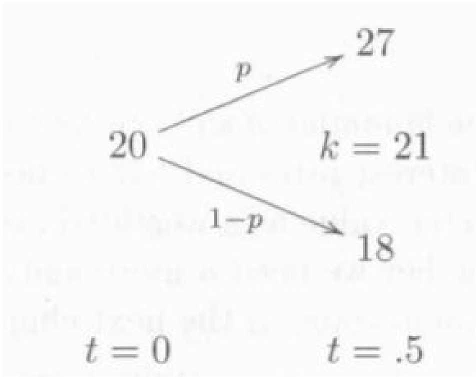
If the option can only be exercised at the maturity date, it is called a *European Option*; while if it can be exercised at any time prior to the maturity date, it is called an *American Option*. We only consider European options and termed them as only options.

Options can be viewed as a measure of transferring risk, either to hedge and reduce exposure to risk or to speculate and, in the process, accept exposure to risk.

The following example gives an idea to formulate our first *binomial model* for call options.

Example 2.4.1

Suppose a certain stock's share price is Rs 20 today. If the share price can take one of the two values Rs 18 or Rs 27 in 6 months time and the seller takes no commission, we want to calculate what is the fair price per share for a call option with strike price Rs 21 and maturity date 6 months. Consider the rate of interest as 12% continuously compounded.



The buyer has a choice. He can buy shares or buy a call option. He assumes the market price for shares is fair and that a fair price for a call option will be based on the fair share price.

If he buys a share today at Rs 20, then he can sell it in 6 months time at either Rs 18 or Rs 27. The discounted value, today, of a single share will be either Rs $27e^{-.12\times.5}$ or Rs $18e^{-.12\times.5}$. And for a fair game, the expected discounted value from buying shares should be equal to the initial investment

$$27e^{-.12\times.5}p + 18e^{-.12\times.5}(1-p) = 20$$

which gives p = 0.3596 .

This established for the buyer a standard to judge the actual price of a call option. If the share price increases the buyer, of the call option, will receive $(27-21)e^{-.12\times.5}$, on discounting back, and 0 if it decreases, as the option won't be exercised.

Thus the total expected return should be equal to the initial investment, i.e. the price of call option which equals

$$(27-21)\times e^{-.12\times.5}p + 0\times e^{-.12\times.5}(1-p) = 2.032$$

Hence from the buyers point of view, Rs $2.032(\equiv f)$ is a fair price for the call option.

Now we consider from the seller's point of view. We assume that he takes no commission, expects no profit and is unwilling to take any risk or any loss. He lays of the bet, as the bookmaker, by buying a quantity $\triangle$ of stocks. Risk is avoided if his financial status is independent of the share price at all times.

Suppose he starts with 0 capital and that the call option is for just for one share. Let g be the price he charges for the option. Hence he borrows Rs$(20 \triangle -g)$ to buy those shares. At the end of 6 months if the share price rises, then his status will be

$$27 \triangle -6 - (20 \triangle -g)e^{0.06}$$

and if the price falls then

$$18 \triangle -(20 \triangle -g)e^{0.06}.$$

For a no-risk no-gain situation, both must be equal and equate the starting sum i.e.

$$27 \triangle -6 - (20 \triangle -g)e^{0.06} = 0$$
$$18 \triangle -(20 \triangle -g)e^{0.06} = 0.$$

This implies $\triangle = 2/3$ and g = Rs 2.032 .

Hence we notice that the buyer's and seller's option price coincides, regardless their approaches are different.

Let us now use the hedging principle, as used by the bookmaker, when parameters are changed.

Suppose the option price rises to Rs 2.20. Then seller borrows Rs $(\frac{2}{3} \times 20 - 2.20)$= Rs 11.13 to buy 2/3 of a share. If the share price goes up, then he end up with

$$\frac{2}{3} \times 27 - 6 - 11.13e^{0.06} = 0.18$$

and if the price falls down then he ends up in

$$\frac{2}{3} \times 18 - 11.13e^{0.06} = 0.18.$$

Thus in either case he is guaranteed a profit of Rs 0.18 . Hence any other price above the fair price leads to arbitrage.

Now if the option price falls to Rs 1.80, the buyer borrows some x shares, sells it for Rs y, buy a call option and invest the remaining Rs$(y-1.80)$ in bonds. Now if he wishes to be in an arbitrage situation, then his at all time his financial status should be invariant with respect to the price change. Thus at maturity date, he receives $(y-1.80)e^{0.06}+6-27x$ if price rises and $(y-1.80)e^{0.06}-18x$ if price falls. Hence equating

$$(y-1.80)e^{0.06}+6-27x=(y-1.80)e^{0.06}-18x$$

we get x=2/3. That is if he borrows 2/3 of a share and sells at any price(y), then he would land in arbitrage position.

Hence we conclude that Rs 2.032 is an arbitrage-free price for a call option on one share.

A put option will be exercised only if the share price is below the strike price at maturity. The expected return is

$$p.(0).e^{-0.06}+(1-p).(3).e^{-0.06}=1.809$$

Then the fair price of put option is Rs 1.809 and it is arbitrage free and seller can hedge it, as in case of call option.

These are some commonly used terminology used in finance.

The *payoff* of an option is the amount received when the option is exercised. For example, on a call option the payoff will be the difference between the share price and strike price if this is positive otherwise it is 0.

A *claim* is a payment which may be demanded at some future time according to a contract. Thus the buyer of a call option is really buying the right to make a claim, while the seller is making a contract.

The seller of the option, as in the previous example, sets up a portfolio at time 0 consisting of risky and non-risky assets to hedge the claim. The risky assets were the shares and non-risky assets were the borrowings, which are obtained in the forms of bonds rather than bank loans.

An individual is said to have a *long position* on shares owned and a *short position* on shares not owned but contracted to provide if required.

The portfolio is risk-less if, at all times, its value is independent of the changes in the price of the risky assets.

An arbitrage opportunity exists if it is possible to construct a portfolio with value V_t at time t such that $V_0 = 0,\ V_t \geq 0$ for all t and $V_T > 0$ for some $T > 0$.

Proposition 2.4.1 *Suppose the interest rate is r, the share price of a certain stock is S at time 0 and that at a future time T it will either be Su or Sd where* $0 < d < 1 < e^{rT} < u$.[1] *The risk neutral probability p that the share price will go up is*

$$p = \frac{e^{rT} - d}{u - d} \tag{2.4.1.1}$$

The arbitrage-free price for a call option, C_T, with strike price k, $Sd < k < Su$, and maturity date T is

$$C_T = \frac{Su - k}{u - d}.(1 - e^{-rT}d). \tag{2.4.1.2}$$

The seller's portfolio for hedging the call option consists of Δ shares and B bonds where

$$\Delta = \frac{Su - k}{Su - Sd} \text{ and } B = -de^{-rT}\left(\frac{Su - k}{u - d}\right). \tag{2.4.1.3}$$

[1] u and d represents the fraction by which the share price will either move up or fall down at time T. In normal times it is reasonable to suppose that $d < 1 < u$. We require only $d < u$, keeping $d < 1 < u$, we can write about the share price going up or down. If Rs 1 is deposited in bank then it will amount to e^{rt} by time T. And if Rs 1 worth share is brought then it will either increase to u or fall to d. And Since deposits are risk-less while shares are not, we may suppose $u > e^{rt}$

The arbitrage-free price for a put option, P_T, with strike price k, $Sd < k < Su$, and maturity date T is

$$P_T = \frac{Sd - k}{u - d}(1 - e^{-rT}u) = \frac{k - Sd}{u - d}.(e^{-rT}u - 1). \tag{2.4.1.4}$$

The call-put parity formula

$$C_T - P_T = S - ke^{-rT} \tag{2.4.1.5}$$

gives the relationship between the price of the call and put options and the price of a contract to buy a share at time T at price k.

Proof The proof follows the similar pattern as described in the preceding example and the call-put parity is calculated by simply substituting the values. [ch 2][↑]

Chapter 3

Conditional Expectation

[↑]

In call option 1, the price of call option, from buyer's point of view, was calculated by first calculating the risk-neutral probabilities, that were found by *Fair Game Principle*, i.e by assuming that the expected return on investing directly on shares would, when discounted back to the present, be equal to initial investment.

3.1 Call Option 2

Similar principle applied now on a slightly more complicated model, i.e. with information available and trading allowed at one intermediate step.

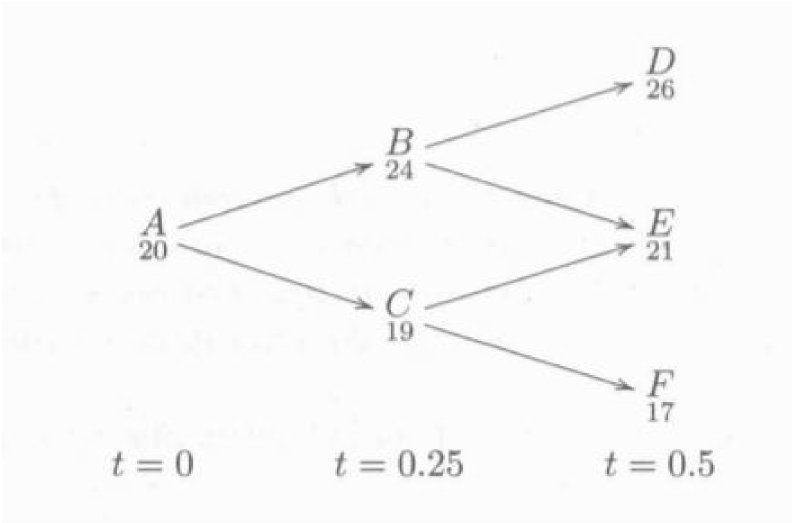
Example 3.1.1

Consider the following call option: strike price is Rs 22, maturity date is six months, interest rate is 12% per annum continuously compounded, share price is Rs 20 today, in three months time it will be either Rs 24 or Rs 19 and in six months time it will be either go from Rs 24 to Rs 21 or Rs 26 or from Rs 19 to Rs 21 or Rs 17. The new feature in this example is the introduction of supplementary information at an intermediate time. We make two assumptions :

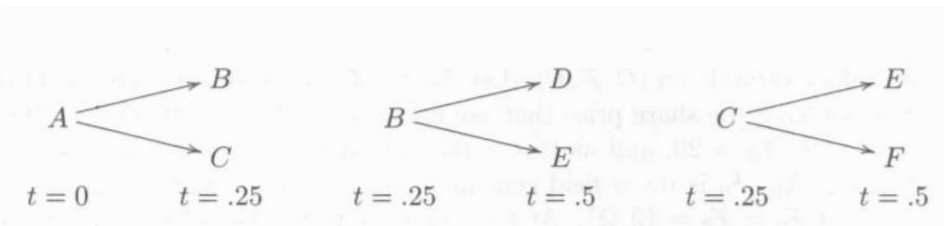
1. the future depends only on the present and is independent of the past.

2. the option can be brought or sold at $t = .25$

The following is the *binomial tree diagram* of the given data.



The points A,B,C,D,E and F are called *Nodes.* Information is received at nodes and shares and options are brought or sold at nodes. Over here the sample space, Ω, is the set of all paths that the share price can follow from $t = 0$ to $t = .5$ i.e. $\Omega = \{ABD, ABE, ACE, ACF\}$ and let $\mathcal{F} = 2^{\Omega}$. For a $\omega \in \Omega$ i.e. a path is an example of what may happen to the share price over the duration of the option i.e from $t = 0$ to $t = .5$. Each pair $(t, \omega), \omega \in \Omega$, determines a unique node and hence any function on the nodes can be pulled back to define a function of (t, ω). Let X denote the share price and let $X_t, t = 0, .25, .5$ be the share price at time t. Decomposing above figure we obtain the following figure

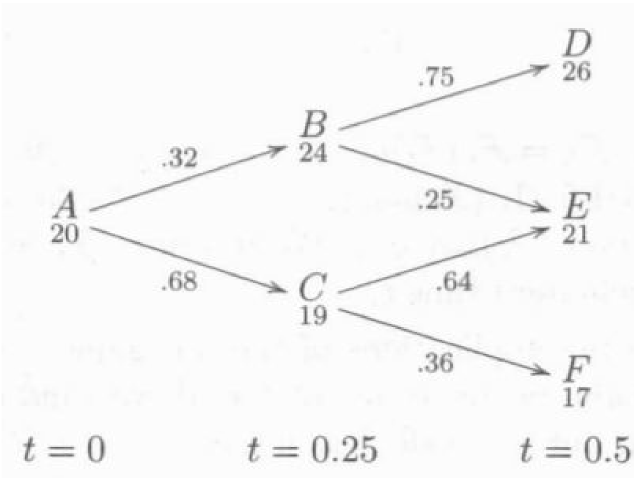


By Proposition 2.4.1.2

$$p = \frac{\text{initial price}\times e^{.25\times .12} - \text{lower price}}{\text{higher price } - \text{lower price}}$$

is the risk-neutral probability that the share price will rise in each of the three phase, i.e. from t=0 to t=0.25 (A to B or C); t=0.25 to t=0.5 (B to D or E); and t=0.25 to t=0.5 (C to E or F).

Hence $p_A = \frac{20e^{.03}-19}{24-19} = .32$ and similarly $p_B = .75$, $p_C = .64$ where the initial node is used as a subscript for identification. Thus we get the following figure



And using the independence assumption (1), we assign probability to all paths in Ω

$$\begin{aligned} P[ABD \text{ occurs }] &= P[AB \textit{ and } BD \text{ occurs }] \\ &= P[AB \text{ occurs }]P[BD \text{ occurs }] \\ &= (.32)(.75) = .24 \end{aligned}$$

Similarly $P[ABD \text{ occurs }] = .08$, $P[ACE \text{ occurs }] = .44$, $P[ACF \text{ occurs }] = .24$. With these probabilities, we have constructed risk-neutral probability space $(\Omega, \mathcal{F}, P)$.

Next we interpret $X_t,\ t = 0, .25, .5$, as random variables on $(\Omega, \mathcal{F}, P)$. Since on each path the share price achieves a unique price at time t, it defines X_t as a function on Ω and since $\mathcal{F} = 2^\Omega$, each X_t is a random variable on $(\Omega, \mathcal{F}, P)$.

Let $\mathcal{F}_t,\ t = 0, .25, .5$ denote the $\sigma-$field of events involving the share price that are known or have been decided by time t.

At $t = 0$, we have only information X_0, hence $\mathcal{F}_0 = \sigma-$field generated by X_0 and as $X_0(\omega) = 20\ \ \forall\omega \in \Omega$, this implies $\mathcal{F}_0 = \mathcal{F}_\phi = \{\phi, \Omega\}$. At $t = .25$, we have precise information about X_0 *and* $X_{.25}$ and nothing else. Hence $\mathcal{F}_{.25}$ is the smallest $\sigma-$field for which X_0 *and* $X_{.25}$ are measurable. And since

$$X_{.25}^{-1}(24) = \{ABD,\ ABE\}$$

$$X_{.25}^{-1}(19) = \{ACE,\ ACF\}$$

$\mathcal{F}_{.25}$ is generated by partition $\{ABD,\ ABE\}$ and $\{ACE,\ ACF\}$ of Ω.

At $t = .5$, all the events are known and this implies $\mathcal{F}_{.5} = \mathcal{F} = 2^\Omega$. On the other hand $\sigma-$field generated by $X_0,\ X_{.25,\ X.5}$ contains sets $\{ABD,\ ABE\}$, $\{ACE,\ ACF\}$ and $X_{.25}^{-1}(21) = \{ABE,\ ACE\}$ and hence contains all subsets of Ω.

This shows that $\mathcal{F}_{.5}$ is generated by $X_0,\ X_{.25},\ X_{.5}$. Thus $\mathcal{F}_t$ is the $\sigma-$field generated by $(X_s)_{s\leq t}$ for $t = 0,\ .25,\ .5$.

Since $\mathcal{F}_0 \subset \mathcal{F}_{.25} \subset \mathcal{F}_{.5} = \mathcal{F}$, $(\mathcal{F}_t)_{t=0,.25,.5}$ is a discrete filtration of $(\Omega, \mathcal{F}, P)$ and $(X_t)_{t=0,.25,.5}$, is a stochastic process adapted to filtration $(\mathcal{F}_t)_{t=0,.25,.5}$. We interpret $\mathcal{F}_t$ as the history of the process up to (and including) time t.

Now the fair game principle, used to find risk-neutral probability, in terms of above random variables is

$$X_0 = e^{-.25\times.12}\mathbb{E}[X_{.25}] \tag{3.1.0.1}$$

The other two calculations were *conditional* on what happened at $t = .25$.

The first says that the discounted expectation of $X_{.25}$ is 24 if the share price increased in the first period, i.e given that either path ABD or ABE is followed. That is

$$e^{-.25\times.12}\mathbb{E}[X_{.5}|\{ABD,\ ABE\}] = 24 \tag{3.1.0.2}$$

$$e^{-.25\times.12}\mathbb{E}[X_{.5}|\{ACE,\ ACF\}] = 19 \tag{3.1.0.3}$$

this implies that the above expectation values are not constant rather a function of path being followed. Now since events on which it is conditional, generates $\mathcal{F}_{.25}$, we write

$$e^{-.25\times.12}\mathbb{E}[X_{.5}|\mathcal{F}_{.25}](\omega) = 24 = X_{.25}(\omega)\ \ if\ \omega \in \{ABD,\ ABE\}$$
$$e^{-.25\times.12}\mathbb{E}[X_{.5}|\mathcal{F}_{.25}](\omega) = 19 = X_{.25}(\omega)\ \ if\ \omega \in \{ACE,\ ACF\}$$

We call $\mathbb{E}[X_{.5}|\mathcal{F}_{.25}]$ the conditional expectation of $X_{.5}$ given $\mathcal{F}_{.25}$. By definition it is a random variable on $(\Omega,\ \mathcal{F},\ P)$. Combining (3.1.0.2) and (3.1.0.3),we can write as a single equation involving two random variables

$$e^{-.25\times.12}\mathbb{E}[X_{.5}|\mathcal{F}_{.25}] = X_{.25} \tag{3.1.0.4}$$

Note that the risk neutral probabilities are derived from (3.1.0.1) and (3.1.0.4).

Let V denote the buyer's price for the option and let V_t denote the price for the option at time t. We are interested in calculating V_0. Since V_0 will not change with time, identify V_0 with a constant random variable defined by $V_0(\omega) = V_0\ \forall\ \omega \in \Omega$. This means V_0 is an $\mathcal{F}_0$ measurable random variable on $(\Omega,\ \mathcal{F},\ P)$. Since strike price is Rs 22, the option will be exercised only if the share price (X_t) follows path ABD, i.e moves up twice. Hence at $t = .5$, $V_{.5} = 4$ if $X_t = 26$, otherwise it is worthless.

Thus for each path ω, assign unique value to $V_{.5}$, $V_{.5}(ABD) = 4,\ V_{.5}(ABE) = V_{.5}(ACE) = V_{.5}(ACF) = 0$. Hence $V_{.5}$ is an $\mathcal{F}_{.5}$ measurable random variable on

$(\Omega,\ \mathcal{F},\ P)$. Thus relation between X and V random variables at $t = .5$ is

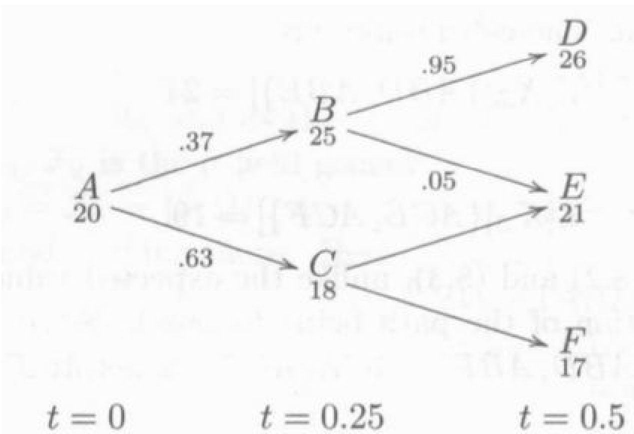
$$V_{.5} = (X_{.5} - 22)^{+} \tag{3.1.0.5}$$

By fair game principle, V_0 is discounted expectation value of $V_{.5}$ i.e

$$V_0 = e^{-.5\times.12}\mathbb{E}[V_{.5}] = e^{-.5\times.12}\mathbb{E}[(X_{.5} - 22)^{+}]. \tag{3.1.0.6}$$

As $p(\{ABD\}) = .32 \times .75 = .24$, we get $\mathbb{E}[V_{.5}] = 4 \times .24$ and $V_0 = 4 \times .24 \times e^{-.5\times.12} = .90$.

We now consider the impact of changing the information provided at the intermediate stage while keeping the same initial and final data. For example, let us consider the following binomial tree diagram for a call option with already calculated risk neutral probability



The option price here is *Rs* 1.34 > *Rs* 0.90 (original option price). Thus intermediate information, over which the buyer has no control, may affect option price. This should, assuming that the price is *arbitrage free*, be balanced by something over which buyer has control.

This appears as the buyer's right to sell the option at $t = .25$. We now wish to find this price for original data i.e $V_{.25}$. At $t = 0$, one cannot make definite statements about the situation at $t = .25$, but he can make a number of *if...then* statements. If the share price is Rs 24 at $t = .25$, then the expected discounted return on buying the option at $t = .25$ is $.75 \times 4 \times e^{-.25\times.12} + .25 \times 0 \times e^{-.25\times.12} = 2.90$. If the share price is Rs 19 at $t = .25$, then the option is worthless. Now we can consider $V_{.25}$ as a random variable on $(\Omega,\ \mathcal{F},\ P)$ by associating a unique value to it for each $\omega \in \Omega$. Therefore

$$V_{.25}(ABD) = V_{.25}(ABE) = 2.90 \qquad and$$

$$V_{.25}(ACE) = V_{.25}(ACF) = 0.$$

Thus $V_{.25}$ is $\mathcal{F}_{.25}$ measurable.

As with X, $(V_t)_{t=0,.25,.5}$ is a finite discrete stochastic process adapted to the filtration $(\mathcal{F}_t)_{t=0,.25,.5}$. Rephrasing it

$$e^{-.25\times.12}\mathbb{E}[V_{.5}|\{ABD,\ ABE\}] = 2.90 \tag{3.1.0.7}$$

$$e^{-.25\times.12}\mathbb{E}[V_{.5}|\{ACE,\ ACF\}] = 0 \tag{3.1.0.8}$$

and using $\sigma-$field $\mathcal{F}_{.25}$, we get

$$e^{-.25\times.12}\mathbb{E}[V_{.5}|\mathcal{F}_{.25}](\omega) = 2.90,\ \ if\ \omega \in \{ABD,\ ABE\}$$

$$e^{-.25\times.12}\mathbb{E}[V_{.5}|\mathcal{F}_{.25}](\omega) = 0,\ \ if\ \omega \in \{ACE,\ ACF\}$$

We call $\mathbb{E}[V_{.5}|\mathcal{F}_{.25}]$ the conditional expectation of $V_{.5}$ given $\mathcal{F}_{.25}$. By definition it is a random variable on $(\Omega,\ \mathcal{F},\ P)$. Combining (3.1.0.7), (3.1.0.8) into single equation of two random variables

$$e^{-.25\times.12}\mathbb{E}[V_{.5}|\mathcal{F}_{.25}] = V_{.25} \tag{3.1.0.9}$$

Thus equation (3.1.0.6) and (3.1.0.9) summarize all relevant information regarding option price.

Example 3.1.2

This example is an abstract version of the previous example. Suppose we have a call option with maturity date T, strike price k, intermediate information provided at time t, $0 < t < T$, and interest rate r. By (3.1.0.1) and (3.1.0.4) we can see that the risk neutral probabilities are calculated from the following two equations :

$$\mathbb{E}[e^{-rt}X_t] = X_0 \tag{3.1.0.10}$$

$$\mathbb{E}[e^{-rt}X_T|\mathcal{F}_t] = e^{-rt}X_T \tag{3.1.0.11}$$

By (3.1.0.6) and (3.1.0.9) the prices of the call option at times 0 and t are calculated from the following equations:

$$\mathbb{E}[e^{-rt}V_T] = V_0 \tag{3.1.0.12}$$

$$\mathbb{E}[e^{-rt}V_T|\mathcal{F}_t] = e^{-rt}V_t \tag{3.1.0.13}$$

and from $V_T = (X_T - k)^+$ and (3.1.0.6) we obtain

$$V_0 = \mathbb{E}[e^{-rt}(X_T - k)^+]. \tag{3.1.0.14}$$

The similarity between the two pairs of equations is not mere coincidence and latter we will see using martingales that they form part of a larger picture.

In our analysis we introduce special cases of a new concept: *conditional expectation.* This allows us to express in a concise fashion all relevant information connected with pricing a call option. Conditional expectations are a modified version of expected values. The crucial difference is that conditional expectations are random variables. [ch 3][↑]

3.2 Conditional Expectation

If X denotes the number of heads that appear when a fair coin is tossed three times in succession, then $\mathbb{E}[X] = 1.5$. If, however, it is known that the first toss resulted in a head, then $\mathbb{E}[X|$ first toss is a head $] = 2$ and similarly $\mathbb{E}[X|$ first toss is a tail $] = 1$. Our expectation changes as new information becomes available. More generally, if X is a random variable on $(\Omega, \mathcal{F}, P)$, where Ω is finite, $\mathcal{F} = 2^{\Omega}$, $P(\{\omega\}) > 0$ for $\omega \in \Omega$ and $A \in \mathcal{F}$, $0 < P(A) < 1$, we let

$$\mathbb{E}[X|A] = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\}|A).$$

Since

$$P(\{\omega\}|A) = \frac{P(\{\omega\} \cap A)}{P(A)} = \begin{cases} \frac{P(\{\omega\})}{P(A)} & if\ \omega \in A \\ 0 & if\ \omega \notin A \end{cases}$$

Hence

$$\mathbb{E}[X|A] = \sum_{\omega \in A} X(\omega) \frac{P(\{\omega\})}{P(A)} := \frac{1}{P(A)} \int_A X dP$$

and similarly

$$\mathbb{E}[X|A^c] = \sum_{\omega \in A^c} X(\omega) \frac{P(\{\omega\})}{P(A^c)} := \frac{1}{P(A^c)} \int_{A^c} X dP$$

The $\sigma-$field generated by A, $\mathcal{A}$, consists of the sets $\{\phi, A, A^c, \Omega\}$. Hence, if $B \in \mathcal{A}$ and $P(B) > 0$, then

$$\mathbb{E}[X|B] = \frac{1}{P(B)} \int_B X dP.$$

We write this as a function on Ω. If $B \in \mathcal{A},\ P(B) > 0\ and\ \omega \in B$, let

$$\mathbb{E}[X|\mathcal{A}](\omega) = \frac{1}{P(B)} \int_B X dP.$$

This minor change of notation defines an $\mathcal{A}$ measurable random variable on $(\Omega, \mathcal{F}, P)$ and allows us to apply the theory already developed for random variables. Motivated by the above we define conditional expectations with respect to a $\sigma-$field generated by a countable partition. Afterwards we define the general concept and prove its main properties, but confine our proofs to the countably generated case.

Definition 3.2.1 *Let $(\Omega, \mathcal{F}, P)$ denote a probability space and let $\mathcal{G}$ denote a $\sigma-$field on Ω generated by a countable partition $(G_n)_{n=1}^{\infty}$ of Ω. Suppose $\mathcal{G} \subset \mathcal{F}$ and $P(G_n) > 0$ for all n. If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$, let*

$$\mathbb{E}[X|\mathcal{G}](\omega) = \frac{1}{P(G_n)} \int_{G_n} X dP \tag{3.2.1.1}$$

for all n and $\omega \in G_n$. We call $\mathbb{E}[X|\mathcal{G}]$ the conditional expectation of X given $\mathcal{G}$. If $\mathcal{G}$ is generated by a random variable Y on $(\Omega, \mathcal{F}, P)$, we also write $\mathbb{E}[X|Y]$ in place of $\mathbb{E}[X|\mathcal{F}_Y]$

Since $(G_n)_{n=1}^{\infty}$ partition Ω, each ω lies in precisely one G_n and (3.2.1.1) defines $\mathbb{E}[X|\mathcal{G}](\omega)$ for all $\omega \in \Omega$. The mapping

$$\mathbb{E}[X|\mathcal{G}] : \Omega \to \mathbb{R} \tag{3.2.1.2}$$

is a constant on each G_n and hence $\mathcal{G}$ measurable, and, as $\mathcal{G} \subset \mathcal{F}$, it is also $\mathcal{F}$ measurable and $\mathbb{E}[X|\mathcal{G}]$ is a random variable on $(\Omega, \mathcal{F}, P)$. By (3.2.1.1), if $\omega_n \in G_n$, then $|\mathbb{E}[X|\mathcal{G}](\omega_n)| \leq \int_{G_n} |X| dP / P(G_n)$ and

$$\begin{aligned}
\int_{\Omega} |\mathbb{E}[X|\mathcal{G}]| dP &= \sum_{n=1}^{\infty} \int_{G_n} |\mathbb{E}[X|\mathcal{G}]| dP = \sum_{n=1}^{\infty} |\mathbb{E}[X|\mathcal{G}](\omega_n)|.P(G_n) \\
&\leq \sum_{n=1}^{\infty} \int_{G_n} |X| dP = \int_{\Omega} |X| dP = \mathbb{E}[|X|]
\end{aligned}$$

Hence $\mathbb{E}[X|\mathcal{G}]$ is integrable.

The trivial $\sigma-$field $\mathcal{F}_{\phi}$ is countably generated and, since $P(\Omega) = 1$, it follows $\mathbb{E}[X|\mathcal{F}_{\phi}](\omega) = \mathbb{E}[X]$ for all $\omega \in \Omega$. Hence we may identify $\mathbb{E}[X]$ and the constant random variable $\mathbb{E}[X|\mathcal{F}_{\phi}]$ and regard the expectation as a special case of conditional expectation.

Our next result characterizes conditional expectations in the countably generated case and extends it with slightly weaker form of uniqueness to a arbitrary conditional expectations.

Proposition 3.2.2 *Let $(\Omega, \mathcal{F}, P$ denote a probability space and let $\mathcal{G}$ denote a $\sigma-$field on Ω generated by a countably partition $(G_n)_{n=1}^{\infty}$ of Ω. We suppose $\mathcal{G} \subset \mathcal{F}$ and $P(G_n) > 0$ for all n. If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$, then $\mathbb{E}[X|\mathcal{G}]$ is the unique $\mathcal{G}$ measurable random variable on $(\Omega, \mathcal{F}, P)$ satisfying*

$$\int_A \mathbb{E}[X|\mathcal{G}]dP = \int_A XdP \tag{3.2.2.1}$$

for all $A \in \mathcal{G}$.

Proof Let n be arbitrary and let $\omega \in G_n$. Since $\mathbb{E}[X|\mathcal{G}]$ is constant on each G_n, it is $\mathcal{G}$ measurable and

$$\begin{aligned}
\int_{G_n} \mathbb{E}[X|\mathcal{G}]dP &= \mathbb{E}[X|\mathcal{G}](\omega). \int_{G_n} dP \\
&= \left(\frac{1}{P(G_n)} \int_{G_n} XdP \right) .P(G_n) \\
&= \int_{G_n} XdP
\end{aligned}$$

If $A \in \mathcal{G}$, then $A = \bigcup_{n \in M} G_n$ for some $M \subset \mathbb{N}$. Hence

$$\begin{aligned}
\int_A \mathbb{E}[X|\mathcal{G}]dP &= \int_{\cup_{n\in M} G_n} \mathbb{E}[X|\mathcal{G}]dP \\
&= \sum_{n \in M} \int_{G_n} \mathbb{E}[X|\mathcal{G}]dP \\
&= \sum_{n \in M} \int_{G_n} XdP \\
&= \int_{\cup_{n\in M} G_n} XdP = \int_A Xdp.
\end{aligned}$$

The uniqueness is established from the fact that if X and Y are integrable random variable on $(\Omega, \mathcal{F}, P)$ and $\int_A XdP = \int_A YdP$ for all $A \in \mathcal{F}$, then $X = Y$ almost surely.

If $A = \Omega$ in (3.2.2.1), then

$$\begin{aligned} \mathbb{E}[\mathbb{E}[X|\mathcal{G}]] &= \int_\Omega \mathbb{E}[X|\mathcal{G}]dP \\ &= \int_\Omega XdP = \mathbb{E}[X], \end{aligned} \tag{3.2.2.2}$$

which says that the average of the average is the average.

Conditional expectation is a rather subtle concept and a much more powerful tool. We consider $\mathbb{E}[X|\mathcal{G}]$ as our expectation of X with all possible information about X, that can be derived from $\mathcal{G}$ events, incorporated. When $\mathcal{G}$ is countably generated we can interpret $\mathbb{E}[X|\mathcal{G}]$ *pointwise*. But this is not true in general case. We will shortly look into it.

Now if $\mathcal{G}$ is countably generated, as in definition 3.2.1 and X is a random variable on $(\Omega, \mathcal{F}, P)$, then, for all n and all $\omega \in G_n$, we have

$$\begin{aligned} \mathbb{E}[X|\mathcal{G}](\omega) &= \text{Expected value of X given that } G_n \text{ has occurred.} \\ &= \text{Expected value of X with respect to the probability measure } P(\cdot|G_n) \\ &= \text{Average of X over } G_n, \\ &= \text{Average of } \mathbb{E}[X|\mathcal{G}] \text{ over } G_n. \end{aligned}$$

Now we prepare for the general definition of conditional expectation. Let X denote a positive integrable random variable on $(\Omega, \mathcal{F}, P)$. If $\mathcal{G}$ is a σ-field on Ω and $\mathcal{G} \subset \mathcal{F}$, let $Q(A) = \int_A XdP$ for all $A \in \mathcal{G}$. Then it can be shown that Q is a finite measure on the measurable space $(\Omega, \mathcal{G})$ and $P(A) = 0$ implies $Q(A) = 0$. If $A \in \mathcal{G},\ P(A) > 0$ and for every $\mathcal{G}$ measurable set B with $A \bigcap B \neq \phi$ we have $A \subset B$, then A cannot be partitioned into smaller $\mathcal{G}$ measurable sets, and it is natural to let $\mathbb{E}[X|\mathcal{G}](\omega) = \frac{1}{P(A)} \int_A XdP$ for all $\omega \in A$. In particular, if $A = \{\omega\} \in \mathcal{G}$ and $P(\{\omega\}) > 0$, then $\int_{\{\omega\}} XdP = X(\omega) \cdot P(\{\omega\})$ and

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

However, this may not always be the case and we may have $\{\omega\} \in \mathcal{G}$ with $P(\{\omega\}) = 0$. If there exists a decreasing sequence of $\mathcal{G}$ measurable sets $(A_n)_{n=1}^{\infty}$, with $P(A_n) > 0$ for all n such that $\lim_{n\to\infty} A_n = \{\omega\}$, then $\lim_{n\to\infty} P(A_n) = 0 = \lim_{n\to\infty} Q(A_n)$. Now in view of definition 3.2.1 it is reasonable to attempt to define

$$\mathbb{E}[X|\mathcal{G}](\omega) = \lim_{n\to\infty}\left(\frac{1}{P(A_n)}\int_{A_n} XdP\right) = \lim_{n\to\infty}\frac{Q(A_n)}{P(A_n)}.$$

This limit may not exists at all points but, by *Radon-Nikodým Theorem*, it can be shown to exists almost surely and to define a $\mathcal{G}$ measurable random variable on Ω. Writing $\frac{\Delta Q_n}{\Delta P_n}$ in place of $\frac{Q(A_n)}{P(A_n)}$, we can see why the almost sure limit is denoted by $\frac{dQ}{dP}$ in The Radon-Nikodým Theorem. We now present the *Radon-Nikodým Theorem.*

Proposition 3.2.3 (The Radon-Nikodým Theorem) *If P and Q are finite measures on the measurable space $(\Omega, \mathcal{F})$ and Q(A)=0 whenever $A \in \mathcal{F}$ and $P(A) = 0$, then there exists a positive measurable function Y on Ω such that*

$$Q(A) = \int_A YdP \tag{3.2.3.1}$$

for all $A \in \mathcal{F}$. Moreover, any $\mathcal{F}$ measurable function Z on Ω satisfying (3.2.3.1) for all $A \in \mathcal{F}$ is equal to Y almost everywhere. Here the function Y is represented by $\frac{dQ}{dP}$.

The following proposition states the main result on the existence of generalized conditional expectation, i.e. when the σ-field is uncountably generated. The case of countably generated also follows from here as a special case.

Proposition 3.2.4 *If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$ and $\mathcal{G}$ is a σ-field on Ω such that $\mathcal{G} \subset \mathcal{F}$, then there exits a $\mathcal{G}$ measurable integrable random variable on $(\Omega, \mathcal{F}, P)$, $\mathbb{E}[X|\mathcal{G}]$, such that*

$$\int_A \mathbb{E}[X|\mathcal{G}]dP = \int_A XdP \tag{3.2.4.1}$$

for all $A \in \mathcal{G}$. Moreover, if Y is any $\mathcal{G}$ measurable random variable satisfying

$$\int_A Y dP = \int_A X dP \tag{3.2.4.2}$$

for all $A \in \mathcal{G}$, then $Y = \mathbb{E}[X|\mathcal{G}]$ almost surely in $(\Omega, \mathcal{F}, P)$. When Y is a random variable we let $\mathbb{E}[X|Y] = \mathbb{E}[X|\mathcal{F}_Y]$. We call $\mathbb{E}[X|\mathcal{G}]$ and $\mathbb{E}[X|Y]$ the conditional expectations of X given $\mathcal{G}$ and Y respectively.

Conditional expectation satisfy the usual laws such as $\mathbb{E}[X + Y|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$. In addition to these, it also follows the following properties.

Proposition 3.2.5 *Let X and Y denote integrable random variable on the probability space $(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}$ and $\mathcal{H}$ denote σ-fields on Ω where $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$.*

(a) ***Taking out what is known.*** *If $X.Y$ is integrable and X is $\mathcal{G}$ measurable, then*

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

(b) ***Independence drops out.*** *If X and $\mathcal{G}$ are independent, that is $\mathcal{F}_X \perp \mathcal{G}$, then*

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

(c) ***tower Law.***

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

If X is $\mathcal{G}$ measurable, then $\mathcal{F}_X \subset \mathcal{G}$ and $\mathbb{E}[X|\mathcal{G}]$ contains complete information about all events involving X. Thus in (a) we are taking out what is known. A useful particular case arises when we replace X by $X.1_\Omega$ and suppose X is $\mathcal{G}$ measurable. We then have almost surely

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X.1_\Omega|\mathcal{G}] = X.\mathbb{E}[1_\Omega] = X.$$

The opposite occurs when X, or equivalently $\mathcal{F}_X$, and $\mathcal{G}$ are independent. Then no $\mathcal{G}$ event contains information about X and, by (b), independence drops out. Thus the conditional expectation becomes a constant random variable and the constant is given by $\mathbb{E}[X]$. A useful special case, that we have previously considered from a different point of view, occurs when $\mathcal{F} = \mathcal{F}_\phi = \{\phi, \Omega\}$. For any event A we have $P(A \cap \Omega) = P(A) = P(A).P(\Omega)$ and $P(A \cap \phi) = P(\phi) = 0 = P(A).P(\phi)$. Hence $\mathcal{F}_X$ and $\mathcal{F}_\phi$ are independent and (b) implies $\mathbb{E}[X|\mathcal{F}_\phi] = \mathbb{E}[X]$.

The tower law says, when $\mathcal{H} \subset \mathcal{G}$, that averaging first over $\mathcal{G}$ and then over $\mathcal{H}$ is the same as just averaging over $\mathcal{H}$. In particular

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{F}_\phi] = \mathbb{E}[X|\mathcal{F}_\phi] = \mathbb{E}[X].$$

Example 3.2.1

This example is a continuation of example 3.1.2, and we use the same notation : X_t denotes the share price and V_t the call option price at time t. We call $Y_t := e^{-rt}X_t$ the *discounted share price* and $Z_t := e^{-rt}V_t$ the *discounted option price* at time t. Rewriting (3.1.0.10) and (3.1.0.11) with this notation, we obtain $\mathbb{E}[Y_t|\mathcal{F}_0] = Y_0$ and $\mathbb{E}[Y_T|\mathcal{F}_t] = Y_t$. Applying the tower law we obtain, since Y_0 is a constant random variable and $\mathcal{F}_0 = \mathcal{F}_\phi \subset \mathcal{F}_t$,

$$\mathbb{E}[Y_T|\mathcal{F}_0] = \mathbb{E}[\mathbb{E}[Y_T|\mathcal{F}_t]|\mathcal{F}_0] = \mathbb{E}[Y_t|\mathcal{F}_0] = Y_0. \tag{3.2.5.1}$$

Similarly (3.1.0.12) and (3.1.0.13) can be rewritten as $\mathbb{E}[Z_T|\mathcal{F}_0] = Z_0$ and $\mathbb{E}[Z_T|\mathcal{F}_t] = Z_t$ and, using the tower law, we obtain

$$\mathbb{E}[Z_t|\mathcal{F}_0] = \mathbb{E}[\mathbb{E}[Z_T|\mathcal{F}_t]|\mathcal{F}_0] = \mathbb{E}[Z_T|\mathcal{F}_0] = Z_0. \tag{3.2.5.2}$$

On taking out what is known, $\mathbb{E}[Y_u|\mathcal{F}_u] = Y_u$ and $\mathbb{E}[Z_u|\mathcal{F}_u] = Z_u$ for $u \in \{0, t, T\}$. We summarize both these equations and (3.1.0.10),(3.1.0.11),(3.1.0.12),(3.1.0.13),(3.2.5.1) and (3.2.5.2) as follows:

$$\mathbb{E}[Y_v|\mathcal{F}_u] = Y_u, \quad when\; u, v \;\in \{0, t, T\}, \; u \leq v \tag{3.2.5.3}$$

$$\mathbb{E}[Z_v|\mathcal{F}_u] = Z_u, \quad when\; u, v \;\in \{0, t, T\}, \; u \leq v \tag{3.2.5.4}$$

Equation (3.2.5.3) is used to find the call option. Our simple model now contains a probability space, a filtration, two finite stochastic processes, and two sets of equations using conditional expectation. [ch 3][↑]

3.3 Hedging

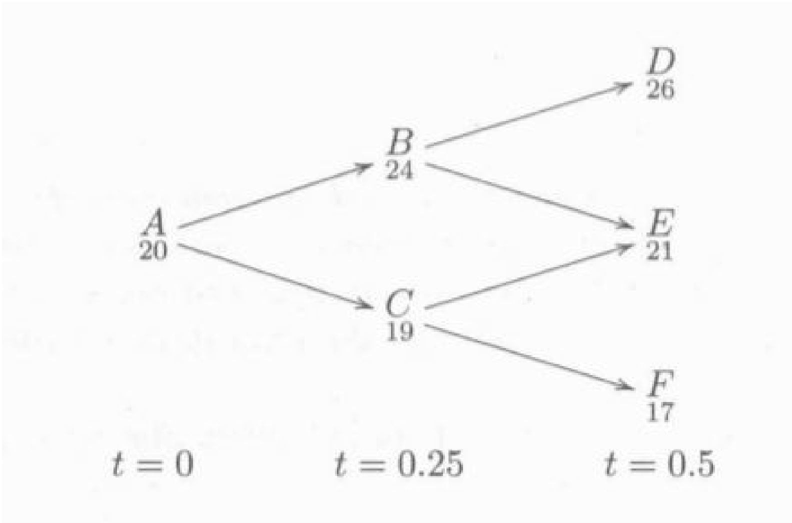
Now we will show that the seller can set up a portfolio to cover or hedge any claim on the option in Example 3.1.1.

Example 3.3.1

We suppose, as previously, that the seller receives no commission and that the portfolio consists of shares and bonds. The portfolio will need to match the claim at each node. The option has value 4 at node D and value 0 at nodes C,E,F. This leaves only A and B. Let v_0 be sellers option at $t = 0$ and $v_{.25}$ at $t = .25$.

Our aim is to show v_0 coincides with V_0, the buyers option price at $t = 0$, and $v_{.25}$ coincides with $V_{.25}$ and that these prices allow the seller to hedge any claim on the option.

We consider the portfolio at each stage



Now if the share price goes to C, the option will not be exercised no matter what happens at $t = .5$ and hence third case need not be considered.

So, let the seller's portfolio contains $\triangle_0$ shares at $t = 0$ and $\triangle_{.25}$ shares at $t = .25$. At $t = 0$ the seller will receive v_0 and needs to borrow $20 \triangle_0 - v_0$ to buy the shares. At $t = .25$ the portfolio will be worth $24 \triangle_0 - e^{.25 \times .12}(20 \triangle_0 - v_0)$ if the share price goes up and be worth $19 \triangle_0 - e^{.25 \times .12}(20 \triangle_0 - v_0)$ if the price goes down. Hence

$$24 \triangle_0 - e^{.25 \times .12}(20 \triangle_0 - v_0) = v_{.25} \tag{3.3.0.5}$$

$$19 \triangle_0 - e^{.25 \times .12}(20 \triangle_0 - v_0) = 0 \tag{3.3.0.6}$$

and hence $\triangle_0 = v_{.25}/5$.

At node B the number of shares will change from $\triangle_0$ to $\triangle_{.25}$ and require additional borrowings of $24(\triangle_{.25} - \triangle_0)$, if $\triangle_{.25} > \triangle_0$. But it may happen that $\triangle_{.25} < \triangle_0$ in which case shares are sold. And equating the portfolio with the claim at D,

$$26 \triangle_{.25} - 24(\triangle_{.25} - \triangle_0)e^{.25 \times .12} - (20 \triangle_0 - v_0)e^{.5 \times .12} = 4 \tag{3.3.0.7}$$

and at node E

$$21 \triangle_{.25} - 24(\triangle_{.25} - \triangle_0)e^{.25\times.12} - (20 \triangle_0 - v_0)e^{.5\times.12} = 0 \tag{3.3.0.8}$$

and hence, $\triangle_{.25} = .8$

By (3.3.0.6),

$$19 \triangle_0 e^{.03} = e^{.06}(20 \triangle_0 - v_0)$$

and on substituting this into (3.3.0.8) with $\triangle_{.25} = .8$ we obtain

$$.8(21 - 24e^{.03}) = \triangle_0(19e^{.03} - 24e^{.03})$$

Hence $\triangle_0 = .58$ and $v_{.25} = 2.90 = V_{.25}$. By (3.3.0.6),

$$v_0 = (20 - 19e^{.03})\triangle_0 = .90 = V_0.$$

So the seller's hedging portfolio at $t = 0$ consists of .58 shares and borrowings of *Rs* 10.70 worth of bonds. If the share price increases at $t = .25$, the seller rebalances the portfolio by buying an extra .22 *i.e* $(\triangle_{.25} - \triangle_0)$ shares and increasing his borrowings to *Rs* 15.98; otherwise he sells his shares and repays his borrowings.

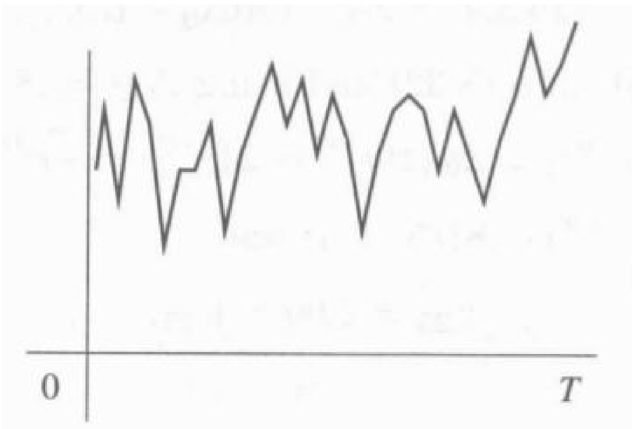
Two different approaches led to the same agreed price. On the buyer's side the fair game principle was used to construct the risk-neutral probability spaces, which was then used to price the option. The two steps were summarized by similar formula (3.2.5.3) and (3.2.5.4).

On the seller's side a hedging portfolio was constructed at each node and gave the seller's fair price for the option.

The next step, in obtaining a more realistic *model*, for pricing and hedging an option can be to partition $[0, T]$ into a large number of small intervals, providing information at the end points of every interval, and extending the previously used methods. By providing information at n different times, the share price can take any of $(i+1)$ different values at the i^{th} intermediate time and any of $(n+1)$ values at the maturity date. This extension contains no surprises and is only complicated by the increased amount of data to be processed.

Formulae (3.2.5.3) and (3.2.5.4) will be developed latter to overcome this problem. This involves the use of conditional expectation to define martingales. The martingales themselves will be seen as a mathematical formulation of a set of fair games. Afterwards, it is natural to take limits by letting the mesh of partition tend to zero, and this leads to *Black Scholes Formula* for pricing the option and covers the buyer's side of the story.

The sample space for the binomial model with information provided at times $\{0, T/n, ..., T\}$ consists of paths similar to the one featured in the following graph



In the limits the sample space is similar to as for a *Wiener Process*, which we discuss in later chapters. Each point in this space is a potential path for the share price over $[0, T]$ and defines a mapping from $[0, T]$ into $\mathbb{R}$. It is important to know how smooth the paths are in the sample space. Almost surely all the paths are continuous, but also almost surely all paths are no-where differentiable. This means we can integrate but not differentiate along paths and we briefly consider the implications for hedging strategy.

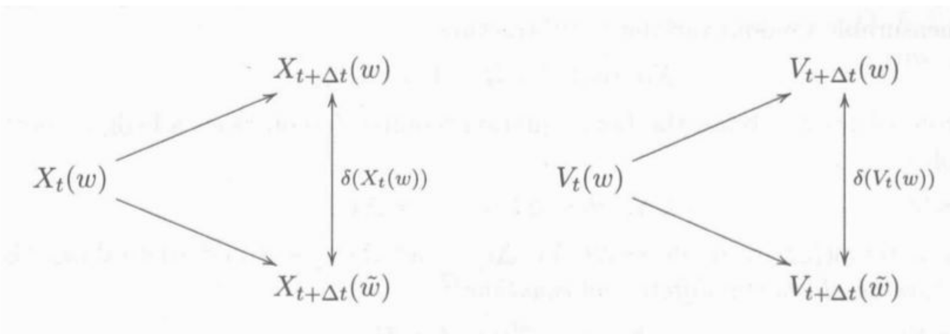
Suppose information is given a finite number of times in $[0,T]$ and that $[t,t+\Delta t]$ is one such interval, where information is given at t and $t+\Delta t$ and nowhere in between. The portfolio that hedges the claim over the interval has to be placed at time t, using the information available at time t, and must match the price of the call option at both ends of the interval. Suppose the portfolio consists of θ_t shares and β_t units of a $risk-less\ bond$ at time t. We let $B(t)$ denote the value of the bond at time t. In the usual case of a continuous compounded constant interest rate r with unit cost for the bond at $t=0$, we have $B(t)=e^{rt}$. If the portfolio hedges the claim at all times, then

$$V_t = \theta_t X_t + B(t)\beta_t \tag{3.3.0.9}$$

where V_t is the value of the claim at time t. Let $\omega \in \Omega$ be arbitrary and let $\tilde{\omega}$ denote any path such that

$$X_t(\omega) = X_t(\tilde{\omega}) \quad and \quad X_{t+\Delta t}(\omega) \neq X_{t+\Delta t}(\tilde{\omega})$$

i.e



If the share price follows the path ω up to time t, then at time $t+\Delta t$ the share price will either be $X_{t+\Delta t}(\omega)$ with corresponding claim $V_{t+\Delta t}(\omega)$ or $X_{t+\Delta t}(\tilde{\omega})$ with corresponding claim $V_{t+\Delta t}(\tilde{\omega})$. Since the portfolio is unchanged over the interval $(t+\Delta t)$ and must match either of these possibilities at time $t+\Delta t$, we have by (3.3.0.9)

$$X_{t+\Delta t}(\omega)\theta_t(\omega) + B(t+\Delta t)\beta_t(\omega) = V_{t+\Delta t}(\omega) \tag{3.3.0.10}$$

$$X_{t+\Delta t}(\tilde{\omega})\theta_t(\omega) + B(t+\Delta t)\beta_t(\omega) = V_{t+\Delta t}(\tilde{\omega}) \tag{3.3.0.11}$$

Let $\delta(X_t(\omega)) := X_{t+\Delta t}(\omega) - X_{t+\Delta t}(\tilde{\omega})$ denote the gap between the prices the stock may achieve at $t+\Delta t$ if it followed path ω up to time t, and let $\delta(V_t(\omega)) := V_{t+\Delta t}(\omega) - V_{t+\Delta t}(\tilde{\omega})$ denote the corresponding gap in the value of the claim.

From (3.3.0.10) and (3.3.0.11), we get

$$\theta_t(\omega) = \theta_t(\tilde{\omega}) = \frac{\delta(V_t(\omega))}{\delta(X_t(\omega))} = \frac{\delta(V_(\tilde{\omega}))}{\delta(X_t(\tilde{\omega}))}. \tag{3.3.0.12}$$

Clearly $V_{t+\Delta t}(\omega) > V_{t+\Delta t}(\tilde{\omega})$ if and only if $X_{t+\Delta t}(\omega) > X_{t+\Delta t}(\tilde{\omega})$ and hence $\theta_t(\omega) > 0$. If $X_{t+\Delta t}(\omega) = X_{t+\Delta t}(\tilde{\omega})$, then no hedging is necessary. Hence we get

$$\begin{aligned} \beta_t(\omega) &= B(t+\Delta t)^{-1}(V_{t+\Delta t}(\omega) - X_{t+\Delta t}(\omega)\theta_t(\omega)) \\ &= B(t+\Delta t)^{-1}(V_{t+\Delta t}(\tilde{\omega}) - X_{t+\Delta t}(\tilde{\omega})\theta_t(\omega)) \end{aligned} \tag{3.3.0.13}$$

As both θ_t and β_t are independent of what happens after time t, both are $\mathcal{F}_t$ measurable and we have an abstract version of the material in Example 3.3.1. Working backwards from T, we can set up a portfolio which hedges the claim.

Let $\Delta X_t = X_{t+\Delta t} - X_t$, $\Delta B(t) = B(t+\Delta t) - B(t)$, $\Delta V(t) = V_{t+\Delta t} - V_t$. If $\mathcal{F}_t$ denotes the $\sigma - field$ of events known by time t, then ΔX_t and ΔV_t are $\mathcal{F}_{t+\Delta t}$ measurable random variable.

Subtracting $X_t(\omega)\theta_t(\omega) + B(t)\beta_t(\omega) = V_t(\omega)$ from (3.3.0.10) we obtain the basic equation required to construct a hedging portfolio:

$$\Delta X_t\theta_t + B(t)\beta_t = V_t \tag{3.3.0.14}$$

It may be tempting to divide (3.3.0.14) by Δt, and letting $\Delta t \to 0$ and write down the following *Stochastic Differential equation*

$$X_t'\theta_t + B(t)'\beta_t = V_t' \tag{3.3.0.15}$$

hoping it will provide, in the limit, a continuous hedging strategy.

However, we should be able to interpret $\lim\limits_{\Delta t \to 0} \frac{\Delta X_t(\omega)}{\Delta t}$ for any path ω, as $\frac{\Delta X_t}{\Delta t}$ is a random variable. Since, almost surely, paths are nowhere differentiable, this limit will very rarely exists and (3.3.0.15) has a symbolic meaning.

There is an alternative approach, i.e , in place of taking limits in (3.3.0.14), add them together to form Riemann Sums

$$\sum_{i=1}^{n} \theta_{t_i} \Delta X_{t_i} + \sum_{i=1}^{n} \beta_{t_i} \Delta B_{t_i} = \sum_{i=1}^{n} \Delta V_{t_i} \tag{3.3.0.16}$$

and take a limit to obtain, the *Stochastic Integral equation*

$$\int_0^T \theta_t dX_t + \int_0^T \beta_t dB(t) = \int_0^T dV_t = V_T - V_0 \tag{3.3.0.17}$$

[ch 3][↑]

Chapter 4

Martingales

[↑]

A martingale is a stochastic process satisfying a condition that removes bias. There are two basic types of martingales: discrete and continuous. In the following section we define and discuss discrete martingales. Equations (3.2.5.3) and (3.2.5.4) motivate the definition and provide simple examples.

4.1 Discrete Martingales

We first define filtration for a given measurable space.

Definition 4.1.1 *Let $(\Omega, \mathcal{F})$ be a measurable space, then*

(a) A discrete filtration on $(\Omega, \mathcal{F})$ is an increasing sequence of σ-fields $(\mathcal{F}_n)_{n=1}^{\infty}$ such that

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_i \subset \cdots \mathcal{F}.$$

(b) A continuous filtration on $(\Omega, \mathcal{F})$ is a set of σ-fields $(\mathcal{F}_t)_{t \in I}$, where I is an interval in $\mathbb{R}$, such that for all t,s $\in I$, $t < s$, we have

$$\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}.$$

We call $\mathcal{F}_n$(respectively $\mathcal{F}_t$) the history up to time n (respectively time t).

Definition 4.1.2 *Let $(\mathcal{F}_n)_{n=1}^{\infty}$ be a filtration on the probability space $(\Omega, \mathcal{F}, P)$. A discrete martingale on $(\Omega, \mathcal{F}, P, (\mathcal{F}_n)_{n=1}^{\infty})$ is a sequence $(X_n)_{n=1}^{\infty}$ of integrable random* *variable on* $(\Omega, \mathcal{F}, P)$*, adapted to the filtration* $(\mathcal{F}_n)_{n=1}^{\infty}$*; that is* X_n *is* $\mathcal{F}_n$ *measurable for all* n*, such that*

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n \tag{4.1.2.1}$$

for all $n \geq 1$*.*

We can think n as measuring time, for example days or weeks, associated with one experiment occurring in each time interval. The random variable X_n is associated with the n^{th} experiment, and $\mathcal{F}_n$ is the history of the first n periods or the information available at the end of the n^{th} time period.

Let $\mathcal{G}_n := \sigma(X_1, ..., X_n)$ denote the σ-field generated by $(X_n)_{n=1}^{n}$; that is $\mathcal{G}_n$ is the smallest σ-field on Ω for which X_i, $1 \leq i \leq n$, are measurable. Since $\mathcal{F}_{X_n} \subset \mathcal{F}_n \subset \mathcal{F}$, $\mathcal{G}_n \subset \mathcal{F}_n$ and $(\mathcal{G}_n)_{n=1}^{\infty}$ is a filtration on $(\Omega, \mathcal{F}, P)$. Taking the conditional expectation in (4.1.2.1) with respect to $\mathcal{G}_n$, we obtain

$$\mathbb{E}[\mathbb{E}[X_{n+1}|\mathcal{F}_n]|\mathcal{G}_n] = \mathbb{E}[X_n|\mathcal{G}_n]. \tag{4.1.2.2}$$

On taking out what is known from R.H.S of (4.1.2.2) and applying the tower law to L.H.S, we obtain

$$\mathbb{E}[X_{n+1}|\mathcal{G}_n] = X_n.$$

This shows, since $(X_n)_{n=1}^{\infty}$ is adapted to the filtration $(\mathcal{G}_n)_{n=1}^{\infty}$, that $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P, (\mathcal{G}_n)_{n=1}^{\infty})$. In many examples we have $\mathcal{F}_n = \mathcal{G}_n$ for all n, and if specifically not mentioned it is understood this is the case.

Example 4.1.1

This example shows that discrete martingales generalize the concept of a sequence of independent integrable variables. Let $(X_n)_{n=1}^{\infty}$ denote a sequence of independent integrable random variables on the probability space $(\Omega, \mathcal{F}, P)$ and suppose $\mathbb{E}[X_n] = 0$ for all n. If $\mathcal{F}_n$ is the σ-field generated by $X_1, ..., X_n$, then $\mathcal{F}_n \subset \mathcal{F}$ for all n and $(\mathcal{F}_n)_{n=1}^{\infty}$ is a filtration on $(\Omega, \mathcal{F}, P)$. Let $Y_n = \sum_{i=1}^{n} X_i$ for all n. Since $|Y_n| \leq \sum_{i=1}^{n} |X_i|$ and each X_i is integrable, this implies that Y_n is integrable. The random variable Y_n, as a finite sum of $\mathcal{F}_n$ measurable functions, is $\mathcal{F}_n$ measurable. If $i \leq n$, the random variable X_{n+1} is independent of X_i, and hence also of $\mathcal{F}_n$. Since independence drops out, $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_{n+1}] = 0$. On taking out what is known, we obtain

$$
\begin{aligned}
\mathbb{E}[Y_{n+1}|\mathcal{F}_n] &= \mathbb{E}[X_{n+1} + Y_n|\mathcal{F}_n] \\
&= \mathbb{E}[X_{n+1}|\mathcal{F}_n] + \mathbb{E}[Y_n|\mathcal{F}_n] \\
&= Y_n
\end{aligned}
$$

and $(Y_n)_{n=1}^{\infty}$ is a martingale adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$.

Example 4.1.2

This example is the key to presenting the price of a call option as a martingale. Let $(\mathcal{F}_n)_{n=1}^{\infty}$ denote a filtration on $(\Omega, \mathcal{F}, P)$ and let X denote an integrable random variable. For each n let $X_n = \mathbb{E}[X|\mathcal{F}_n]$. By proposition 3.2.4, X_n is integrable and, by the definition of conditional expectation, $\mathcal{F}_n$ measurable. Hence $(X_n)_{n=1}^{\infty}$ is adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$. Since $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, the tower law implies

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_{n+1}]|\mathcal{F}_n] = \mathbb{E}[X|F_n] = X_n$$

and $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P)$ adapted to filtration $(\mathcal{F}_n)_{n=1}^{\infty}$.

The following proposition provides a test to verify whether a sequence is martingale or not.

Proposition 4.1.3 *If $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P, (\mathcal{F}_n)_{n=1}^{\infty})$, then*

$$\mathbb{E}[X_n] = \mathbb{E}[X_m]$$

for all n and m

Proof $\mathbb{E}[X_n] = \int_\Omega X_n dP = \int_\Omega \mathbb{E}[X_{n+1|\mathcal{F}_n}]dP = \int_\Omega X_{n+1}dP = \mathbb{E}[X_{n+1}]$.

Hence $\mathbb{E}[X_n] = \mathbb{E}[X_{n+1}] = ... = \mathbb{E}[X_m]$ for all n and m.

Example 4.1.3

Martingales are the mathematical formulation of a sequence of fair games. Let X_n denote the winnings per unit stake on the n^{th} game in a sequence of fair games. Then $\mathbb{E}[X_n] = 0$ and let $Y_n = \sum_{i=1}^{n} X_n$ are the winnings accumulated by the end of n^{th} game. In mathematical setting let Ω_n denote the set of outcomes of the n^{th} game, and suppose there exists a σ-field $\mathcal{G}_n$ on Ω_n and a probability measure P_n on $(\Omega_n, \mathcal{G}_n)$ such that X_n is a random variable on $(\Omega_n, \mathcal{G}_n, P_n)$. If $\omega_n \in \Omega_n$, then $X_n(\omega_n)$ are the winnings per unit stake on the n^{th} game when ω_n is the outcome. Since the random variables $(X_n)_{n=1}^{\infty}$ are not all defined on the same probability space, we cannot say whether Y_n is a martingale. So we have to construct a product probability space. Here we have an infinite set of random variables and an infinite product of probability measures on an infinite product of measurable spaces.

We confine our discussion to an outline of the main highlights of one approach and omit the non-trivial technical details.

Since the sample space must include all information about all possible games, let $\Omega = \prod_{n=1}^{\infty} \Omega_n := \{(\omega_n)_{n=1}^{\infty} : \omega_n \in \Omega_n \ for\ all\ n\}$. A typical point in Ω represents one possible set of results from a full set of games. Hence each point in Ω is called a path. Identify $X_n : \Omega_n \to \mathbb{R}$ with $X_n : \Omega \to \mathbb{R}$ where $X_n((\omega_m)_{m=1}^{\infty}) = X_n(\omega_n)$. Let $\mathcal{F}_n$ denote the σ-field generated by the random variables $(X_i)_{i=1}^{n}$. Then $\mathcal{F}_n := \mathcal{G}_1 \times \mathcal{G}_2 \times ... \times \mathcal{G}_n$ is generated by $A := A_1 \times A_2 \times ... \times A_n \times \Omega_{n+1} \times \Omega_{n+1} \times ...$ and A contains all possible futures of all points in $A_1 \times A_2 \times ...A_n$. Let $\mathcal{F}_\infty = \bigcup_{n=1}^{\infty} \mathcal{F}_n$ and let $\mathcal{F}$ denote the

σ-field generated by $(X_i)_{i=1}^{\infty}$. For all n, $\mathcal{F}_n \subset \mathcal{F}_\infty \subset \mathcal{F}$. In general, $\mathcal{F}_\infty$ is not a σ-field, but it does have some useful properties such as the finite union of sets in $\mathcal{F}_\infty$ still lies in $\mathcal{F}_\infty$; and if $A \in \mathcal{F}_\infty$, then $A^c \in \mathcal{F}_\infty$. Since $\mathcal{F}_1 \subset \mathcal{F}_2 \subset ...\mathcal{F}_n... \subset \mathcal{F}$, $(\mathcal{F}_n)_{n=1}^{\infty}$ is a filtration on $(\Omega, \mathcal{F})$. As $\mathcal{F}_n$ contains all events from the first n games, it is the history of the process up to the end of the n^{th} game. The σ-field $\mathcal{F}$ contains all events from all games.

Then there exists a measure Q_n on $(\Omega, \mathcal{F}_n)$ such that

$$Q_n(A_1 \times A_2 \times ... \times A_n \times \Omega_{n+1} \times \Omega_{n+2}...) = P(A_1)P(A_2)...P(A_n)$$

whenever $A_i \in \mathcal{G}_i$ for $i = 1, 2, ..., n$. Moreover, we have the consistency relationship

$$Q_{n+1}(A) = Q_n(A) \;\; for\ all\ A \in \mathcal{F}_n$$

This implies that there exists a mapping $P : \mathcal{F}_\infty \to [0, 1]$ such that $P(\Omega) = 1$, $P(\bigcup\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} P(A_i)$ for any finite sequence of pairwise disjoint sets in $\mathcal{F}_\infty$ and $P(A^c) = 1 - P(A)$ for all $A \in \mathcal{F}_\infty$. P also satisfies the countably additive property and P can be extended to $\mathcal{F}$,retaining countable additivity. Thus the required probability space is $(\Omega, \mathcal{F}, P)$.

On this probability space the sequence $(X_n)_{n=1}^{\infty}$ becomes a sequence of independent random variables, and also $(Y_n)_{n=1}^{\infty}$ is a martingale adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$.

Now the natural question arises can a player adopt some strategy to make the game favorable ? For instance can he decide to quit when certain amount has been won or loss or to no bet after certain successive wins etc. A system will consists of betting different amounts on different games with the amount B_n to be placed on the n^{th} game, chosen at any time prior to the start the of the n^{th} game and thus with the full knowledge of what happened in the $(n-1)$ previous games. This means B_n is $\mathcal{F}_{n-1}$ measurable for $n > 1$. Suppose there is a limit on the amount that can be placed on any one game; i.e. there exists $M \in \mathbb{R}$ such that $|B_n| \leq M$ for all n. The accumulated winnings W_n at the end of n^{th} game are given by

$$W_n = B_1.X_1 + ... + B_n.X_n.$$

Since X_i and B_i are $\mathcal{F}_n$ measurable for all $i \leq n$, W_n is $\mathcal{F}_n$ measurable; and as X_i is integrable and B_i is bounded, W_n is integrable. Hence, using the facts that B_{n+1}, W_n are both $\mathcal{F}_n$ measurable and X_{n+1} is independent of $\mathcal{F}_n$.

$$\begin{aligned}
\mathbb{E}[W_{n+1}|\mathcal{F}_n] &= \mathbb{E}[W_n + B_{n+1}.X_{n+1}|\mathcal{F}_n] \\
&= \mathbb{E}[W_n|\mathcal{F}_n] + \mathbb{E}[B_{n+1}.X_{n+1}|\mathcal{F}_n] \\
&= W_n + B_{n+1}.\mathbb{E}[X_{n+1}|\mathcal{F}_n] \\
&= W_n + B_{n+1}.\mathbb{E}[X_{n+1}] \\
&= W_n
\end{aligned}$$

and $(W_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$. Thus proposition 4.1.3, $\mathbb{E}[W_n] = \mathbb{E}[W_1] = \mathbb{E}[B_1.X_1] = B_1.\mathbb{E}[X_1] = 0$. Hence no strategy can turn a fair game into a favorable or unfavorable game.

Example 4.1.4

Let $(X_n)_{n=1}^{\infty}$ denote a sequence of independent random variables on $(\Omega, \mathcal{F}, P)$ where $\Omega = \{-1, +1\}$, $\mathcal{F} = 2^{\Omega}$, $P(\{1\}) = P(\{-1\}) = 1/2$ and X_n takes the values ± 1 with probability 1/2. Clearly, $\mathbb{E}[X_n] = 0$ and $Var(X_n) = 1$. The sequence is called a symmetric random walk. Now we need to construct a probability space $(\Omega, \mathcal{F}, P)$ on which each X_n is defined and such that $(X_n)_{n=1}^{\infty}$ is a sequence of independent random variables. We have $\Omega = 2^{\mathbb{N}}$, $\mathcal{F}$ is the σ-field generated by $(X_n)_{n=1}^{\infty}$ and P is the product measure. Let $Y_n = \sum\limits_{i=1}^{n} X_i$ for n a positive integer. By Example 4.1.1, that the sequence (Y_n) is a martingale.

We now consider the sequence $(Z_n)_{n=1}^{\infty}$ where $Z_n = Y_n^2 - n$ for all n. Since $0 \leq Y_n^2 \leq n^2$ the sequence $(Z_n)_{n=1}^{\infty}$ is a stochastic process of bounded, and hence integrable, random variables adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$. Since $\mathbb{E}[X_n] = 0$ and $\mathbb{E}[X_n^2] = 1$ we have $\mathbb{E}[Y_n] = \sum_{i=1}^{n} \mathbb{E}[X_n] = 0$ and, $\mathbb{E}[Y_n^2] = Var(Y_n) = n$. Hence $\mathbb{E}[Z_n] = \mathbb{E}[Y_n^2] - n = 0$. It is possible, by proposition 4.1.3, that the sequence $(Z_n)_{n=1}^{\infty}$ is a martingale. We have

$$\begin{aligned}
\mathbb{E}[Z_{n+1}|\mathcal{F}_n] &= \mathbb{E}[Y_{n+1}^2|\mathcal{F}_n] - (n+1) \\
&= \mathbb{E}[Y_n^2 + 2Y_n X_{n+1} + X_{n+1}^2|\mathcal{F}_n] - n - 1 \\
&= \mathbb{E}[Y_n^2|\mathcal{F}_n] + 2\mathbb{E}[Y_n X_{n+1}|\mathcal{F}_n] + \mathbb{E}[X_{n+1}^2|\mathcal{F}_n] - n - 1 \\
&= Y_n^2 + 2Y_n\mathbb{E}[X_{n+1}|\mathcal{F}_n] + \mathbb{E}[X_{n+1}^2|\mathcal{F}_n] - n - 1 \\
&\quad (\textit{taking out what is known}) \\
&= Y_n^2 + 2Y_n\mathbb{E}[X_{n+1}] + \mathbb{E}[X_{n+1}^2] - n - 1 \\
&\quad (\textit{independence drops out}) \\
&= Y_n^2 + 1 - n - 1 \\
&= Z_n
\end{aligned}$$

and $(Z_n)_{n=1}^{\infty}$ is a martingale. [ch 3][↑]

4.2 Martingale Convergence

If $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$, then $(X_n(\omega))_{n=1}^{\infty}$ is a sample path for each $\omega \in \Omega$. We investigate if this stabilizes with time, that is whether or not $\lim_{n\to\infty} X_n(\omega)$ exists.

Our point of departure is a rather novel approach to the convergence of sequence of real numbers and an even more novel application to martingales. We know that a sequence of real numbers $(a_n)_{n=1}^{\infty}$ converges if and only if $-\infty < \liminf_{n\to\infty} a_n =$

$\limsup_{n\to\infty} a_n < +\infty$. Hence the sequence does not converges if and only if there exists a pair of real numbers a and b with $a < b$ such that

$$\liminf_{n\to\infty} a_n < a < b < \limsup_{n\to\infty} a_n. \tag{4.2.0.1}$$

We call the set of points obtained by joining (n, a_n) to $(n+1, a_{n+1})$, $n = 1, 2, ...$ the graph of $(a_n)_{n=1}^{\infty}$. If (4.2.0.1) holds, then the graph contains an infinite number of points above the horizontal line through b and an infinite number of points below the horizontal line through a, and the graph crosses the horizontal strip $[a, b]$ an infinite number of times.

An *upcrossing* occurs when the graph crosses the horizontal strip $[a, b]$ from below to above.

We summarizes the above

Lemma 4.2.1 *A sequence of real number converges if and only if U(a,b), the number of upcrossing over [a,b], is finite for every pair of real numbers a and b, a < b.*

The space of integrable random variable is denoted by $\mathbf{L}^1(\Omega, \mathcal{F}, P)$. A collection of integrable random variables, $(X_\alpha)_{\alpha\in\Gamma}$ is $\mathbf{L}^1 -$ *bounded* if

$$\sup_{\alpha\in\Gamma} \mathbb{E}[|X_\alpha|] < \infty.$$

If $(X_n)_{n=1}^{\infty}$ is a sequence of integrable random variables and there exists an integrable random variable X such that $\lim_{n\to\infty} \mathbb{E}[|X_n - X|] = 0$, then the sequence $(X_n)_{n=1}^{\infty}$ is said to converge to X in $\mathbf{L}^1(\Omega, \mathcal{F}, P)$ i.e $X_n \xrightarrow{\mathbf{L}^1} X$ as $n \to \infty$.

Proposition 4.2.2 *If $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P)$ and $(X_n)_{n=1}^{\infty}$ is $\boldsymbol{L}^1$-bounded, then there exits an integrable random variable X on $(\Omega, \mathcal{F}, P)$ such that $\lim_{n\to\infty} X_n = X$ almost surely.*

Proof Consider $A := \{\omega \in \Omega : (X_n(\omega))_{n=1}^{\infty} converges\}$. If $\omega \notin A$, then

$$\liminf_{n\to\infty} X_n(\omega) < \limsup_{n\to\infty} X_n(\omega)$$

and since the rationals are dense in the real, there exists rational number p and q, $p < q$, such that

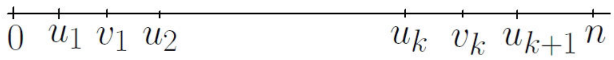
$$\omega \in A_{p,q} := \{\omega \in \Omega : \liminf_{n\to\infty} X_n(\omega) < p < q < \limsup_{n\to\infty} X_n(\omega)\}$$

Hence, if $P(A_{p,q}) = 0$ for all $p, q \in \mathbb{Q}, p < q$, then

$$P(A^c) = P(\bigcup_{p,q\in\mathbb{Q};p<q} A_{p,q}) \leq \sum_{p,q\in\mathbb{Q};p<q} P(A_{p,q}) = 0$$

and the sequence $(X_n(\omega))_{n=1}^{\infty}$ converges almost surely. Thus now it is suffices to prove $P(A_{p,q}) = 0$ for all $p, q \in \mathbb{Q};\ p < q$.

For this purpose we consider two players, Sara and Stephen, taking part in a sequence of fair games. We suppose Sara bets one unit on each game. Stephen feels that when Sara's accumulated winnings fall below p her luck will change for the better and that when they rise above q it will change for worse. His strategy is to observe Sara's accumulated winnings at the end of each game. He places his first bet when her winnings fall below p and keeps betting until her winnings rise above q. He then stops playing and only plays again when her winnings fall below p. He continues with this strategy and always bets one unit. Sara's winnings are given by martingale $(X_n)_{n=1}^{\infty}$ where X_n denotes her accumulated winnings at the end of n^{th} game. If Y_n denotes Stephen's accumulated winnings at the end of n^{th} game, then $(Y_n)_{n=1}^{\infty}$ is also a martingale by Example 4.1.3. Stephen's betting will consist of continuous playing periods followed by waiting periods. We call a *cycle* for Stephen one full playing period followed by a waiting period. Let U_n denote the number of full cycles that Stephen has completed at the end of the n^{th} game. Since U_n is known by the end of n^{th} game, U_n is $\mathcal{F}_n$-measurable and $(U_n)_{n=1}^{\infty}$ is an increasing sequence of random variables. If $U_n = k$, then, over the interval of the first n games we have the following landmarks where u_i is the



first game of the i^{th} cycle and v_i is the first game that Stephen does not play in that cycle. Stephen will play the games $\{u_i, ..., v_i - 1\}$ and place the same bets as Sara on these games and will not play games $\{v_i, ..., u_{i+1} - 1\}$. Since Sara's accumulated winnings are less than p at u_i and greater than q at the beginning of the game v_i, Stephen wins at least $(q - p)$ over every completed cycle. We consider the final, possibly uncompleted, cycle. We have two possibilities, as in the figure.

In the first case, $u_{k+1} \leq n < v_{k+1}$ and Stephen plays all these games. His winnings for these games will match those of Sara and amount to at least $X_n - p$. This may be negative. Hence Stephen's accumulated winnings by the end of n^{th} game, Y_n, will be at least $k(q - p) - |X_n - p|$.

In the second case we have $v_{k+1} \leq n < u_{k+2}$, and Stephen will have completed the playing period of the $(k + 1)^{th}$ cycle. In this case his accumulated winnings will be at least $(k + 1)(q - p)$. Since $(k + 1)(q - p) \geq k(q - p) - |X_n - p|$ we have, in either case, $Y_n \geq k(q-p) - |X_n - p|$. Hence

$$\mathbb{E}[Y_n] \geq (q-p)\mathbb{E}[U_n] - \mathbb{E}[|X_n - p|].$$

Since $(Y_n)_{n=1}^{\infty}$ are the winnings on fair games, $\mathbb{E}[Y_n] = 0$, and hence

$$\mathbb{E}[U_n] \leq \frac{\mathbb{E}[|X_n - p|]}{q-p} \leq \frac{M + |p|}{q-p}$$

where $M = \sup_n \mathbb{E}[|X_n|]$. By the Monotone Convergence Theorem the increasing sequence $(U_n)_{n=1}^{\infty}$ converges almost surely to an integrable random variable U and $\lim_{n\to\infty} U_n(\omega) < \infty$ almost surely. Since the cycles for Stephen are precisely the $[p,q]$ upcrossings for the sequence $(X_n)_{n=1}^{\infty}$, we have shown $P(A_{p,q}) = 0$ and hence $\lim_{n\to\infty} X_n = X$ almost surely. By Fatou's lemma

$$\int_{\Omega} |X| dP = \int_{\Omega} (\liminf_{n\to\infty} |X_n|) dP \leq \liminf_{n\to\infty} \int_{\Omega} |X_n| dP \leq M$$

and X is integrable. This completes the proof.

The boundedness hypothesis in the Proposition cannot be removed without some replacement. Since, suppose $(X_n)_{n=1}^{\infty}$ denote a sequence of independent random variables on $(\Omega, \mathcal{F}, P)$ and suppose $P(X_n = n) = P(X_n = -n) = \frac{1}{2}$ for all n. By example 4.1.1 the sequence $(Y_n)_{n=1}^{\infty}$, $Y_n := \sum_{i=1}^{n} X_n$ is a martingale. But the sequence $(Y_n)_{n=1}^{\infty}$ does not converge almost surely.

To obtain stronger convergence we place extra condition on the martingale sequence.

Lemma 4.2.3 *Let X denote an integrable random variable on $(\Omega, \mathcal{F}, P)$ and let $(A_n)_{n=1}^{\infty} \subset \mathcal{F}$. Then*

(a) $\lim_{m\to\infty}(\int_{\{|X|\geq m\}} |X| dP) = 0$*;*

(b) if $\lim_{n\to\infty} P(A_n) = 0$*, then* $\lim_{n\to\infty}(\int_{A_n} |X| dP) = 0$.

Proof The proof can be easily verified.

The next proposition requires a uniform version of the property established in Lemma 4.2.3(a).

Definition 4.2.4 *A set of integrable random variables* $(X_i)_{i\in I}$ *on the probability space* $(\Omega, \mathcal{F}, P)$ *is uniformly integrable if*

$$\lim_{m\to\infty}(\sup_{i\in I}\int_{\{|X|\geq m\}}|X_i|dP) = 0 \tag{4.2.4.1}$$

Let X denote an integrable random variable on $(\Omega, \mathcal{F}, P)$ and let $(\mathcal{F}_n)_{n=1}^{\infty}$ be a filtration on $(\Omega, \mathcal{F}, P)$. For each positive integer n, let $X_n = \mathbb{E}[X|\mathcal{F}_n]$. Then, $|X_n| \leq \mathbb{E}[|X||\mathcal{F}_n]$ almost surely and for positive integers m and n, $\{\omega \in \Omega : |X_n(\omega)| \geq m\} \in \mathcal{F}_n$. Hence

$$\begin{aligned} mP(|X_n| \geq m) &\leq \int_{\{|X_n|\geq m\}}|X_n|dP && (4.2.4.2)\\ &\leq \int_{\{|X_n|\geq m\}}\mathbb{E}[|X||\mathcal{F}_n]dP \\ &= \int_{\{|X_n|\geq m\}}|X|dP \\ &\leq \mathbb{E}[|X|] && (4.2.4.3) \end{aligned}$$

Let ϵ be arbitrary positive number. By lemma 4.2.3 there exists a positive number δ such that $\int_A|X|dP < \epsilon$ whenever $P(A) < \delta$. By (4.2.4.2) and (4.2.4.3) there exists a positive integer m_0 such that $P(|X_n| \geq m) \leq \mathbb{E}[|X|]/m \leq \delta$ for all $m > m_0$ and all n. Lemma 4.2.3(b) and further applications of (4.2.4.2) and (4.2.4.3) show that

$$\int_{\{|X_n|\geq m\}}|X_n|dP \leq \epsilon$$

for all n and all $m > m_0$. Since ϵ was arbitrary this shows that $(X_n)_{n=1}^{\infty}$ is a uniformly integrable sequence of random variables. The following proposition proves the converse.

Proposition 4.2.5 *If $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$ and $(X_n)_{n=1}^{\infty}$ is uniformly integrable, then there exists an integrable random variable X on $(\Omega, \mathcal{F}, P)$ such that $X_n = \mathbb{E}[X|\mathcal{F}_n]$ almost surely for all n and $X_n \to X$ almost surely and in $\boldsymbol{L}^1(\Omega, \mathcal{F}, P)$ as $n \to \infty$.*

Proof Choose m_1 so that $\int_{\{|X_n \geq m_1|\}} |X_n| dP \leq 1$ for all n. Then

$$\begin{aligned}
\mathbb{E}[|X_n|] &= \int_{\{|X_n \geq m_1|\}} |X_n| dP + \int_{\{|X_n|<m_1\}|X_n|} dP \\
&\leq 1 + \int_{\Omega} m_1 dP \\
&\leq 1 + m_1
\end{aligned}$$

for all n. Hence $(\mathbb{E}[|X_n|])_{n=1}^{\infty}$ is bounded and, by proposition 4.2.2, $(X_n)_{n=1}^{\infty}$ converges almost surely to an integrable random variable X.

Let $Y_n = |X_n - X|$ for all n. Since $|Y_n| \leq |X_n| + |X|$ for all n the sequence $(Y_n)_{n=1}^{\infty}$ is uniformly integrable. Thus we have to show $\mathbb{E}[Y_n] \to 0$ as $n \to \infty$. Let $\epsilon > 0$ be arbitrary. By uniform integrability $\exists\, m_2 > 0$ such that $\int_{\{Y_n \geq m_2\}} Y_n dP \leq \epsilon/2$ for all n. Let $A_n := \{\omega\Omega : Y_n(\omega) < m_2\}$ and let $Z_n = Y_n.1_n$. Then $0 \leq Z_n \leq m_2$ for all n and $(Z_n)_{n=1}^{\infty}$ tends to 0 almost surely as $n \to \infty$. By Dominated Convergence Theorem, $\mathbb{E}[Z_n] \to 0$ as $n \to \infty$. Hence $\exists\, m_3$ such that $\mathbb{E}[Z_n] < \epsilon/2$ for all $n \geq m_3$. If $n \geq m_3$, then

$$0 \leq \mathbb{E}[Y_n] = \int_{\{Y_n \geq m_2\}} Y_n dP + \int_{\Omega} Z_n dP \leq \epsilon.$$

Since $\epsilon > 0$ was arbitrary, this proves $X_n \xrightarrow{L_1(\Omega,\mathcal{F},P)} X$ as $n \to \infty$.

Let $n \geq m$ and $A \in \mathcal{F}_m$. Since $(X_n)_{n=1}^{\infty}$ is a martingale

$$\int_A \mathbb{E}[X_n|\mathcal{F}_m] dP = \int_A X_m dP$$

and, by conditional expectation definition

$$\int_A \mathbb{E}[X_n|\mathcal{F}_m] dP = \int_A X_n dP$$

Hence

$$|\int_A (X_m - X)dP| = |\int_A (X_n - X)dP| \leq \int_\Omega |X_n - X|dP \xrightarrow{n\to\infty} 0$$

this implies

$$\int_A X_m dP = \int_A X dP$$

for all $A \in \mathcal{F}_m$. Hence $X_m = \mathbb{E}[X|\mathcal{F}_m]$ almost surely.

[ch 3][↑]

4.3 Continuous Martingales

Continuous martingales are similar to discrete martingales but indexed by intervals.

Definition 4.3.1 *Let $(\mathcal{F}_t)_{t\in I}$ denote a filtration on $(\Omega, \mathcal{F}, P)$, indexed by an interval I of real numbers, and let $(X_t)_{t\in I}$ denote a set of integrable random variables on $(\Omega, \mathcal{F}, P)$ adapted to filtration; that is, X_t is $\mathcal{F}_t$ measurable for all $t \in I$. Then $(X_t)_{t\in I}$ is a continuous martingale if*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s$$

for all $s, t \in I, s \leq t$.

As in the case of discrete martingales, there is no loss of generality in assuming that $\mathcal{F}_t$ is the σ-field generated by $(X_s)_{s\in I, s\leq t}$. We are interested in two cases $I = [0, T]$, where T is a positive real number, and $I = [0, \infty)$. We can absorb the first case into the second by letting $X_t = X_T$ and $\mathcal{F}_t = \mathcal{F}_T$ for all $t > T$. Thus we will restrict to continuous martingales over $[0, \infty)$ and use the notation $(X_t)_{t\geq 0}$.

It is useful to think of I as an interval of time and $\mathcal{F}_t$ as either the history of the process up to time t or as the information available at time t. If $(X_t)_{t\geq 0}$ is a continuous martingale, then $(X_{t_n})_{n=1}^{\infty}$ is a discrete martingale for any strictly increasing sequence of real numbers tending to infinity.

The following example shows how discrete martingales may be used to interpolate and derive the results for continuous martingales.

Example 4.3.1

Let X denote an integrable random variable on $(\Omega, \mathcal{F}, P)$ and let $(\mathcal{F}_t)_{t\geq 0}$ denote a filtration on the space. Let $X_t = \mathbb{E}[X|\mathcal{F}_t]$ for all t. If $0 < s < t$, the tower law implies

$$\mathbb{E}[X_t|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[X|\mathcal{F}_s] = X_s$$

and $(X_t)_{t\geq 0}$ is a martingale.

Conversely, suppose $(X_t)_{t\geq 0}$ is a martingale adapted to $(\mathcal{F}_t)_{t\geq 0}$ and $\exists$ a strictly increasing sequence of real numbers $(t_n)_{n=1}^{\infty}$, with $t_n \xrightarrow{n\to\infty} \infty$ and $\mathbb{E}[X|\mathcal{F}_{t_n}] = X_{t_n}$ for all n. If $t_n > t \geq 0$, by tower law and martingale property

$$\mathbb{E}[X|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_{t_n}]|\mathcal{F}_t] = \mathbb{E}[X_{t_n}|\mathcal{F}_t] = X_t.$$

Thus

$$\lim_{t\to\infty} \mathbb{E}[|X_t - X|] = 0$$

which implies

$$\lim_{m\to\infty} (\sup_{t\geq 0} \int_{\{X_t \geq m\}} |X_t| dP) = 0$$

i.e. $(X_t)_{t\geq 0}$ is uniformly integrable.

Lemma 4.3.2 *If X is an $N(0,\sigma^2)$ random variable, then*

$$\mathbb{E}[e^X] = e^{\frac{1}{2}\sigma^2}. \tag{4.3.2.1}$$

Proof Here X has density function $f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}$ and by evaluating $\mathbb{E}[e^X]$ we get the lemma.

We denote $\mathbb{E}_P[X]$ and $\mathbb{E}_P[X|\mathcal{F}]$ as the expectation and conditional expectation, respectively, of X with respect to probability measure P.

Proposition 4.3.3 *Let $(W_t)_{t\geq 0}$ denote a collection of random variables on $(\Omega, \mathcal{F}, P)$, and for $t \geq 0$ let $\mathcal{F}_t$ denote the σ-field generated by $(W_s)_{0\leq s\leq t}$. Suppose W_t and $W_t - W_s$ are $N(0,t)$ and $N(0, t-s)$ distributed random variables respectively and $W_t - W_s$ and W_r are independent for all $r, s, t, 0 \leq r \leq s \leq t$. Then the following hold :*

(a) $(W_t)_{t\geq 0}$ *and* $(W_t^2 - t)_{t\geq 0}$ *are martingales.*

(b) if μ and σ are real numbers, then $(e^{\mu t+\sigma W_t})_{t\geq 0}$ is a martingale if and only if $\mu = -\sigma^2/2$;

(c) if $\gamma \in \mathbb{R}$, then $(e^{-\frac{1}{2}\gamma^2 t+\gamma W_t})_{t\geq 0}$ is a uniformly integrable set of random variables;

(d) $X_t := e^{-\frac{1}{2}\gamma^2 t+\gamma W_t}$ for $t \geq 0$, then there exists an integrable random variable X_∞ such that $\lim_{n\to\infty} X_{t_n} = X_\infty$ almost surely for any increasing sequence of real numbers $(t_n)_{n=1}^{\infty}$ which converges to infinity. If $P_\gamma(A) := \int_A X_\infty dP$ when $A \in \mathcal{F}$, then $(\Omega, \mathcal{F}, P_\gamma)$ is a probability space. If $\mathcal{G}$ is a σ-field on Ω with $\mathcal{G} \subset \mathcal{F}$ and Y is an integrable random variable on $(\Omega, \mathcal{F}, P)$, then $\mathbb{E}_{P_\gamma}[Y|\mathcal{G}] = \mathbb{E}_p[X_\infty.Y|\mathcal{G}]$; and if $t \geq s$ and Y is $\mathcal{F}_t$ measurable, then

$$\mathbb{E}_{P_\gamma}[Y|\mathcal{F}_s] = \mathbb{E}_P[X_t.Y|\mathcal{F}_s]. \tag{4.3.3.1}$$

Proof W_t being normal both W_t and W_t^2 are integrable for all t. By our hypothesis $W_t - W_s$ and $(W_t - W_s)^2$ are independent of $\mathcal{F}_s$ for all t and s, $0 \leq s \leq t$.

And for $0 \le s \le t$

$$\begin{aligned}
\mathbb{E}[W_t|\mathcal{F}_s] &= \mathbb{E}[W_t - W_s|\mathcal{F}_s] + \mathbb{E}[W_s|\mathcal{F}_s] \\
&= \mathbb{E}[W_t - W_s|\mathcal{F}_s] + W_s, \text{ taking out what is known,} \\
&= \mathbb{E}[W_t - W_s] + W_s, \text{ independence drops out,} \\
&= W_s, \text{ since } \mathbb{E}[W_t] = \mathbb{E}[W_s] = 0.
\end{aligned}$$

Hence $(W_t)_{t\ge 0}$ is a martingale.

For $0 \le s \le t$, $W_t^2 = (W_t - W_s)^2 + 2(W_t - W_s)W_s + W_s^2$ and

$$\begin{aligned}
\mathbb{E}[W_t^2|\mathcal{F}_s] &= \mathbb{E}[(W_t - W_s)^2|\mathcal{F}_s] + 2\mathbb{E}[(W_t - W_s)W_s|\mathcal{F}_s] + \mathbb{E}[W_s^2|\mathcal{F}_s] \\
&= \mathbb{E}[(W_t - W_s)^2|\mathcal{F}_s] + 2W_s\mathbb{E}[(W_t - W_s)|\mathcal{F}_s] + W_s^2 \text{ taking out what is known} \\
&= \mathbb{E}[(W_t - W_s)^2] + 2W_s\mathbb{E}[W_t - W_s] + W_s^2 \text{ independence drops out} \\
&= t - s + 2W_s.0 + W_s^2 \text{ since } W_t - W_s\ is\ N(0, t-s)
\end{aligned}$$

Hence

$$\mathbb{E}[W_t^2 - t|\mathcal{F}_s] = W_s^2 + t - s - t = W_s^2 - s$$

and $(W_t^2 - t)_{t\ge 0}$ is a martingale, proving (a).

To prove (b), we first suppose that $(e^{\mu t + \sigma W_t})_{t\ge 0}$ is a martingale. From Lemma 4.3.2, $\mathbb{E}[e^{\mu t+\sigma W_t}] = e^{\mu t}.\mathbb{E}[e^{\sigma W_t}] = e^{(\mu+\sigma^2/2)t}$. Proposition 4.1.3 extends to continuous martingales and implies that $e^{(\mu t+\sigma^2/2)t}$ is independent of t. Hence $\mu = -\sigma^2/2$.

Conversely suppose $\mu = \sigma^2/2$. By lemma 4.3.2, $e^{\mu t+\sigma W_t}$ is integrable. If $0 \le s \le t$, then

$$\begin{aligned}
\mathbb{E}[e^{\sigma W_t}|\mathcal{F}_s] &= \mathbb{E}[e^{\sigma(W_t - W_s)}.e^{\sigma W_s}|\mathcal{F}_s] \\
&= e^{\sigma W_s}.\mathbb{E}[e^{\sigma(W_t - W_s)}|\mathcal{F}_s] \\
&= e^{\sigma W_s}.\mathbb{E}[e^{\sigma(W_t - W_s)}] \\
&= e^{\sigma W_s}.e^{\sigma^2(t-s)/2}
\end{aligned}$$

and

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

This shows that $(e^{\mu t+\sigma W_t})_{t\geq 0}$ is a martingale if and only if $\mu = -\frac{\sigma^2}{2}$.

Without loss of generality assume $\gamma > 0$. Let m denote a positive number greater than 1. We use the fact that W_t and $\sqrt{t}W_t$ are identically distributed, which implies

$$\begin{aligned}P(\{e^{-\frac{1}{2}\gamma^2 t+\gamma W_t} \geq m\}) &= P(\{W_t \geq \frac{1}{\gamma}\log m + \frac{1}{2}\gamma t\})\\ &= P(\{W_1 \geq \frac{1}{\gamma\sqrt{t}}\log m + \frac{1}{2}\gamma\sqrt{t}\})\\ &\leq P(\{W_1 \geq \sqrt{2\log m}\}).\end{aligned}$$

By lemma 4.2.3(a)

$$\lim_{m\to\infty}(\sup_{t\geq 0} P(\{e^{-\frac{1}{2}\gamma^2 t-\gamma W_t} \geq m\})) \leq \lim_{m\to\infty} P(\{W_1 \geq \sqrt{2\log m}\}) = 0$$

which proves (c).

By (c), the sequence $(X_n)_{n=1}^\infty$ is uniformly integrable and, proposition 4.2.5 and Example 4.3.1 there exists an itegrable random variable X_∞ such that $X_{t_n} \to X_\infty$ almost surely and $\mathbb{E}_P[|X_{t_n} - X_\infty|] \to 0$ as $n \to \infty$. Since $\mathbb{E}[X_{t_n}] = 1$ and $|\mathbb{E}[X_{t_n}] - \mathbb{E}[X_\infty]| \leq \mathbb{E}[|X_{t_n} - X_\infty|]$ we have $\mathbb{E}_P[X_\infty] = 1$ and, as $X_{t_n}(\omega) \geq 0$ almost surely for each n, $X_\infty \geq 0$ almost surely. Thus, $P_\gamma(A) := \int_A X_\infty dP$ defines a probability measure on $(\Omega, \mathcal{F})$ and

$$\int_A Y dP_\gamma = \int_A X_\infty.Y dP \tag{4.3.3.2}$$

for all $A \in \mathcal{F}$ and all P_γ integrable random variables Y. By Proposition 3.2.4, $\mathbb{E}_{P_\gamma}[Y|\mathcal{G}]$ is almost surely the unique $\mathcal{G}$ measurable random variable on Ω satisfying

$$\int_A \mathbb{E}_{P_\gamma}[Y|\mathcal{G}] dP_\gamma = \int_A Y dP_\gamma \tag{4.3.3.3}$$

and $\mathbb{E}_p[X_\infty.Y|\mathcal{G}]$ is almost surely the unique $\mathcal{G}$ measurable random variable on Ω satisfying

$$\int_A \mathbb{E}_P[X_\infty Y|\mathcal{G}]dP = \int_A X_\infty Y dP \tag{4.3.3.4}$$

for all $A \in \mathcal{G}$. By 4.3.3.2, 4.3.3.3 and 4.3.3.4, $\mathbb{E}_{P_\gamma}[Y|\mathcal{G}] = \mathbb{E}_P[X_\infty.Y|\mathcal{G}]$ almost surely as random variable on $(\Omega, \mathcal{G}, P)$.

If Y is $\mathcal{F}_t$ measurable and $t \geq s$, then

$$\begin{aligned}
\mathbb{E}_{P_\gamma}[Y|\mathcal{F}_s] &= \mathbb{E}_P[X_\infty.Y|\mathcal{F}_s] \\
&= \mathbb{E}_P[\mathbb{E}_P[X_\infty.Y|\mathcal{F}_t]|\mathcal{F}_s] \text{ by tower law} \\
&= \mathbb{E}_P[Y.\mathbb{E}_P[X_\infty|\mathcal{F}_t]|\mathcal{F}_s] \text{ taking out what is known} \\
&= \mathbb{E}_P[Y.X_t|\mathcal{F}_s] \text{ since by Example 4.3.1, } \mathbb{E}_P[X_\infty|\mathcal{F}_t] = X_t
\end{aligned}$$

This proves 4.3.3.1 and hence completes the proof.

[ch 3][↑]

Chapter 5

Brownian Motion

[↑]

5.1 Wiener Process

Processes with stationary independent increments and continuous sample paths were studied by Norbert Wiener and afterwards called *Wiener Processes*. It was later shown that the Brownian motion and Wiener processes are identical, and as a result both terms are used interchangeably today.

5.1.1 Random Walk

Let us consider a random walk starting at 0 with jumps h and $-h$ equally likely at times $\delta, 2\delta, ...$, where h and δ are positive numbers. More precisely, let $\{X_n\}_{n=1}^{\infty}$ be a sequence of independent and identically distributed random variables with

$$P[X_i = h] = P[X_i = -h] = \frac{1}{2}.$$

Now let us extend the random sequence $\{X_n\}_{n=1}^{\infty}$ to a random process $\{Y_t^{\delta,h}\}_{t\geq 0}$ by taking Let $Y_0^{\delta,h} = 0$ and

$$Y_t^{\delta,h} = X_1 + X_2 + ... + X_n, \quad \text{for } t = n\delta$$

and

$$Y_t^{\delta,h} := Y_{[t]}^{\delta,h}, \quad \text{for } t \in [n\delta, (n+1)\delta). \tag{5.1.0.1}$$

Now let us study what will be the limit of the random walk $Y^{\delta,h}$ as $\delta, h \to 0$. Let us first compute the limit of the characteristic function of $Y_t^{\delta,h}$, i.e.

$$\lim_{\delta,h\to 0} \mathbb{E}[e^{i\lambda Y_t^{\delta,h}}]$$

where $\lambda \in \mathbb{R}$ is fixed.

Let $t = n\delta$ and so $n = \frac{t}{\delta}$. Then we have

$$\begin{aligned}
\mathbb{E}[e^{i\lambda Y_t^{\delta,h}}] &= \prod_{j=1}^{n} \mathbb{E}[e^{i\lambda X_j}] \\
&= (\mathbb{E}[e^{i\lambda X_1}])^n \\
&= (\frac{1}{2}e^{i\lambda h} + \frac{1}{2}e^{-i\lambda h}) \\
&= (cos(\lambda h))^n \\
&= (cos(\lambda h))^{\frac{t}{\delta}}
\end{aligned}$$

Thus we must impose a certain relationship between δ and h in order for the limit to exist.

Therefore for small δ and h,

$$(cos(\lambda h))^{\frac{t}{\delta}} \approx e^{-\frac{t\lambda^2 h^2}{2\delta}}.$$

Thus we get

$$\mathbb{E}[e^{i\lambda Y_t^{\delta,h}}] \approx e^{-\frac{t\lambda^2 h^2}{2\delta}}.$$

In particular, if δ and h are related by $h^2 = \delta$, then

$$\lim_{\delta\to 0} \mathbb{E}[e^{i\lambda Y_t^{\delta,h}}] = e^{-\frac{1}{2}t\lambda^2}, \quad \lambda \in \mathbb{R}$$

Thus we have derived the following theorem about the limit of the random walk $Y^{\delta,h}$ as $\delta, h \to 0$ in such a way that $h^2 = \delta$.

Theorem 5.1.1 *Let $Y_t^{\delta,h}$ be the random walk starting at 0 with jumps h and -h equally likely at times $\delta, 2\delta, ...$ Assume that $h^2 = \delta$. Then for each $t \geq 0$, the limit*

$$B_t = \lim_{\delta \to 0} Y_t^{\delta,h}$$

exists in distribution. Moreover, we have

$$\mathbb{E}[e^{i\lambda B_t}] = e^{-\frac{1}{2}t\lambda^2}, \;\; \lambda \in \mathbb{R} \tag{5.1.1.1}$$

On the basis of the above discussion, we would expect the stochastic process B_t to have the following properties:

1. For each t, B_t is a Gaussian random variable with mean 0 and variance t, since from equation 5.1.1.1 the characteristic function is that of a Gaussian variable. *Remark:* Another way to deduce it is by using Central Limit Theorem, described below.

2. The stochastic process B_t has independent increments, i.e. for any $0 \leq t_1 \leq ... \leq t_n$, the random variables

$$B_{t_1}, B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$$

are independent.

Now lets derive the above results from another perspective.

Define a sequence of random processes $\{Y_n^{\delta,h}\}_{n=1}^{\infty}$ as

$$Y_{n,t}^{\delta,h} := \frac{Y_{nt}^{\delta,h}}{\sqrt{n}}.$$

Then by Central Limit Theorem we have

$$Y_{n,t}^{\delta,h} \xrightarrow{w} N(o,t) \text{ as } n \to \infty, \; t \geq 0.$$

Thus, we get that, for each *fixed* $t \geq 0$, the values $Y_{n,t}^{\delta,h}$ of the processes $Y_n^{\delta,h}$ weakly converge to a Gaussian random variable with mean 0 and variance t, as above. However, this does not mean that the sequence $\{Y_n^{\delta,h}\}$ converges, in some sense, to some random process. Equivalently, is it possible to construct, on a single probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random process $B_t,\ t \geq 0$, such that, *for all* $t \geq 0$, we should have

$$Y_{n,t}^{\delta,h} \xrightarrow{w} B_t,\ n \to \infty?$$

The answer to the above question is *Yes.* We will provide a sketch of the construction of the above required space in the next chapter. But for the time being, we state below the existence theorem of Brownian motion.

Theorem 5.1.2 *There exists a random process* $B := \{B_t\}_{t \geq 0}$ *such that*

$$\sum_{i=1}^{k} \alpha_i Y_{n,t_i}^{\delta,h} \xrightarrow{w} \sum_{i=1}^{k} \alpha_i B_{t_i},\ n \to \infty \tag{5.1.2.1}$$

with all collections $0 \leq t_1 < t_2 < ... < t_k,\ \alpha_i \in \mathbb{R},\ i = 1, 2, ..., k \in \mathbb{N}$. *We denote this convergence by* $Y_n^{\delta,h} \xrightarrow{w} B$ *and say that the sequence of the processes* $\{Y_n^{\delta,h}\}$ *weakly converges to* B *in the sense of finite dimensional distributions. The process* B *is called a Brownian motion or Wiener process.*

We know that the increments $B_{t_1}, B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$, for any $0 \leq t_1 \leq ... \leq t_n$ are independent random variables. Now we will see that $B_t - B_s \sim N(0, t-s),\ 0 \leq s \leq t$.

By 5.1.2,

$$Y_{n,t}^{\delta,h} - Y_{n,s}^{\delta,h} \xrightarrow{w} B_t - B_s,\ n \to \infty.$$

On the other hand, by CLT,

$$\begin{aligned}
Y^{\delta,h}_{n,t} - Y^{\delta,h}_{n,s} &= \frac{Y^{\delta,h}_{nt} - Y^{\delta,h}_{ns}}{\sqrt{n}} \\
&= \frac{X_{[ns]+1} + X_{[ns]+2} \cdots + X_{[nt]}}{\sqrt{n}} \\
&\stackrel{d}{=} \frac{X_1 + \cdots + X_{[nt]-[ns]}}{\sqrt{n}} \\
&\stackrel{w}{\longrightarrow} (\sqrt{t-s})N(0,1),\ n \to \infty \\
&\sim N(o, t-s).
\end{aligned}$$

Remark: If for $t > 0$, we define $Y^{\delta,h}_t$ by linearization, instead of taking the greatest integer function of t as in 5.1.0.1, i.e. for $t \in [n\delta, (n+1)\delta)$

$$Y^{\delta,h}_t = \frac{(n+1)\delta - t}{\delta} Y^{\delta,h}_{[t]} + \frac{t - n\delta}{\delta} Y^{\delta,h}_{[t]+\delta}.$$

Then we would get two other important observations:

1. Almost all sample paths of B_t are continuous.

2. The absolute value of the slope of $Y^{\delta,h}$ in each step is

$$\frac{h}{\delta} = \frac{1}{\sqrt{\delta}} \to \infty \ as\ \delta \to 0.$$

   Thus it reflects the possibility that every path B_t can be nowhere differentiable. In fact, if we let $\delta = |t-s|$, then

$$|B_t - B_s| \approx \frac{|t-s|}{\sqrt{\delta}} = |t-s|^{\frac{1}{2}}. \tag{5.1.2.2}$$

   Thus almost all sample paths of B_t have this particular property, which will be referred latter.

5.1.2 Continuity and non-differentiability of Brownian motion

Now returning back to our original definition of the $Y^{\delta,h}_t$, i.e. by the greatest integer function of t, we see that $Y^{\delta,h}_t$ is discontinuous, precisely it has jumps discontinuity.

Still we will see that the limit of these discontinuous jumps leads to a continuous process, i.e. B_t is almost surely everywhere continuous, which is nowhere differentiable. Therefore, it is important to define the continuity of a process and to have a simple condition to check the continuity of a process. Such a rather universal condition was provided by the Kolmogorov. We will state the theorem without the proof.

Definition 5.1.3 *A stochastic process $X := (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is continuous if there exists a $\mathcal{F}-$measurable $A \subset \Omega$ with $P(A) = 1$ such that*

$$t \mapsto X_t(\omega), \quad t \in [0, +\infty) \tag{5.1.3.1}$$

is continuous for all $\omega \in A$.

A path $\omega \in \Omega$ is called $X-$*continuous* if the mapping 5.1.3.1 is continuous

Theorem 5.1.4 (Kolmogorov Continuity Theorem) *If, for a random process $X := (X_t)_{t\geq 0}$, there are constants $C > 0$, $\alpha > 0$, and $\beta > 0$ such that the following condition holds*

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C|t-s|^{\beta+1}$$

then X is continuous almost surely.

Remark: Precisely, the statement of the theorem has small a modification. A random process $\tilde{X}$ is called a *modification* of X if, for all $t \geq 0$, $\tilde{X}_t = X_t$ almost surely. So, in a precise formulation, *"X is continuous almost surely"* must be replaced by *"there exists a continuous modification of X"*.

Theorem 5.1.5 *Brownian Motion is a continuous process.*

Proof Since $B_t - B_s \sim N(0, t-s)$ for $t > s$, we have $\mathbb{E}[|B_t - B_s|^4] = 3|t-s|^2,\ t, s \geq 0$. Therefore, the continuity of the Brownian motion follows from the the Kolmogorov theorem by taking $\alpha = 4,\ \beta = 1,$ and $C = 3$.

Now we want to prove the non-differentiability of the Brownian Motion i.e.

Theorem 5.1.6 *The trajectories of Brownian Motion are nowhere differentiable with probability one.*

Proof It is suffices to prove the non-differentiability of a trajectory of Brownian Motion on any finite time interval $I = [0, T]$.

Suppose it is differentiable at some point $s \in I$, then

$$\lim_{t \to s} \frac{|B_t - B_s|}{|t - s|} = |B_s'| \text{ exists.}$$

Therefore, for any $k > |B_s'|$, there exists $\epsilon > 0$ such that

$$|B_t - B_s| < k(t - s) \text{ for } 0 < t - s < \epsilon.$$

Now for $n\mathbb{N}$ such that $\frac{4}{n} < \epsilon$ and $i = [ns]+1$, we have $0 < \frac{j}{n} - s < \epsilon$, for $j = i, i+1, i = 2, i+3$. Therefore

$$|B_{\frac{j}{n}} - B_s| < k(\frac{j}{n} - s) < k\frac{4}{n}, \text{ for } j = i, i+1, i+2, i+3.$$

Hence,

$$\begin{aligned} |B_{\frac{j}{n}} - B_{\frac{j-1}{n}}| &\leq |B_{\frac{j}{n}} - B_s| + |B_{\frac{j-1}{n}} - B_s| \\ &< k\frac{4}{n} + k\frac{4}{n} = k\frac{8}{n}, \text{ for } j = i+1, i+2, i+3. \end{aligned}$$

Now define the events

$$A_{k,n}^j := \{|B_{\frac{j}{n}} - B_{\frac{j-1}{n}}| < k\frac{8}{n}\}$$

and consider the event

$$A := \bigcup_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \bigcup_{i=1}^{[nT]} \bigcap_{j=i+1}^{i+3} A_{k,n}^{j}.$$

The above says, on taking the assumption that B is differentiable at some point $s \in [0,T]$, we have a $k \in \mathbb{N}^{+}$ such that for all sufficiently large $n \in \mathbb{N}$, or $\exists n_0 \in \mathbb{N}$ such that $\forall n > n_0$ and $n \in \mathbb{N}$, the inequalities $|B_{\frac{j}{n}} - B_{\frac{j-1}{n}}| < k\frac{8}{n}$ holds for $j = i+1, i+2, i+3$. Therefore, A contains the event that a trajectory of the Brownian motion is differentiable at some point $s \in [0,T]$. Thus if $P(A) = 0$, then Brownian motion is not differentiable with probability 1.

Event A is a countable union of the events

$$B_{k,m} := \bigcap_{n=m}^{\infty} \bigcup_{i=1}^{[nT]} \bigcap_{j=i+1}^{i+3} A_{k,n}^{j}, \quad k,m \in \mathbb{N}.$$

Therefore, if $P(B_{k,m}) = 0$ for all $k,m \in \mathbb{N}$, then $P(A) = 0$. Since

$$P(B_{k,m}) \leq P\left(\bigcup_{i=1}^{[nT]} \bigcap_{j=i+1}^{i+3} A_{k,n}^{j}\right) \text{ for } n \geq m.$$

Therefore, it is suffices to show that, for all $k \in \mathbb{N}$,

$$P\left(\bigcup_{i=1}^{[nT]} \bigcap_{j=i+1}^{i+3} A_{k,n}^{j}\right) \to 0, \quad \text{as } n \to \infty.$$

Since the events $A_{k,n}^{j}$, $j = i+1, i+2, i+3$, are independent and

$$|B_{\frac{j}{n}} - B_{\frac{j-1}{n}}| \overset{d}{=} |B_{\frac{1}{n}}| \overset{d}{=} \frac{|B_1|}{\sqrt{n}},$$

for all $i,k,n \in \mathbb{N}$, we have

$$\begin{aligned} P\left(\bigcap_{j=i+1}^{i+3} A_{k,n}^{j}\right) &= \prod_{j=i+1}^{i+3} P(A_{k,n}^{j}) \\ &= \left(P\{\frac{|B_1|}{\sqrt{n}} < k\frac{8}{n}\}\right)^3 \\ &= \left(P\{|B_1| < k\frac{8}{\sqrt{n}}\}\right)^3 \end{aligned}$$

Since

$$P\{|B_1| < x\} = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-\frac{y^2}{2}} dy \leq \frac{2x}{\sqrt{2\pi}}, \ x > 0,$$

we get

$$P\left(\bigcap_{j=i+1}^{i+3} A_{k,n}^j\right) \leq \left(\frac{2 \cdot \frac{8k}{\sqrt{n}}}{\sqrt{2\pi}}\right)^3 = \frac{C_k}{n^{\frac{3}{2}}}$$

with constant C_k depending on k. Finally

$$\begin{aligned} P\left(\bigcap_{j=i+1}^{i+3} A_{k,n}^j\right) &\leq \sum_{i=1}^{[nT]} P\left(\bigcap_{j=i+1}^{i+3} A_{k,n}^j\right) \\ &\leq [nT]\frac{C_k}{n^{\frac{3}{2}}} \\ &\leq \frac{TC_k}{\sqrt{n}} \xrightarrow{n\to\infty} 0. \end{aligned}$$

We finally end this section by combining all the properties discussed above and giving a formal definition of Brownian motion.

Definition 5.1.7 *A stochastic process $(B_t)_{t\geq 0}$ is said to be a Brownian Motion if it satisfies the following properties:*

1. ***Independence of increments:*** *$B_t - B_s$, for $t > s$, is independent of the past, i.e. B_u, $0 \leq u \leq s$, or of $\mathcal{F}_s$, the σ-field generated by $(B_u)_{0\leq u\leq s}$.*
2. ***Normal Stationary increments:*** *$B_t - B_u$, for $t > s$, has Normal distribution with mean 0 and variance $t - s$, i.e $B_t - B_u \sim N(0, t-s)$.*
3. ***Continuity of paths:*** *$(B_t)_{t\geq 0}$ is a continuous process.*

In the next chapter we will see a real example of Brownian motion, which is used in financial market.

Chapter 6

The Black-Scholes Formula

[↑]

6.1 Share Price as a Wiener Process

Till now we have assumed that the share price could achieve only a small finite set of values at a number of specified future times. So to remove this artificial assumption and build a more realistic model, we examine the share price as a random variable.

Consider a company quoted on the stock market with share price X_t at time t, where $t = 0$ is the present, $t < 0$ the past and $t > 0$ the future. We intend to use the history of the company and the share price to make at time t=0 reasonable assumptions about the future. Since the share price is always positive we let $X_t = X_0 e^{A_t}$ where A_t can take any real value. If r is the interest rate and the share price follows this growth rate exactly, then $X_t = X_0 e^{rt}$. However, the value of each company has its own *internal* rate of growth, the *drift* μ, which may be different from the rate of interest. We suppose that the drift is constant[1]. Since the share price itself is a fixed fraction of the perceived value of the company, the drift will reflect itself in the share price, but the share price is also subject to a large number of *independent random changes* brought about by investors. We suppose that these independent random changes are *equally likely* to move the share price in either direction, up or down. Let Z_t be the random variable brought about by these independent random changes. Thus we have two components which affect the share price, $e^{\mu t}$ and e^{Z_t},

[1] in practice it will be a function of time

where $\mathbb{E}[Z_t] = 0$.

Changes in share price are usually considered as a *proportion* of the current price. This implies

$$X_t = X_0\ e^{\mu t} e^{Z_t} = X_0\ e^{\mu t + Z_t}.$$

We suppose that the $Var(Z_t)$, the spread of values taken by Z_t, is finite and strictly positive[2]. As t increases Z_t has more time to wander, and thus it is reasonable to suppose that $Var(Z_t)$ is an increasing function of time t. Observations of different share prices over many periods of time suggest that $Var(Z_t)$ will behave like $tf(t)$ where $f(t)$ is fairly regular, for example continuous and bounded away from both 0 and ∞. For simplicity, we suppose that $f(t)$ is constant and strictly positive. Thus for each stock we have a constant σ such that $Var(Z_t) = \mathbb{E}[Z_t^2] = \sigma^2 t$. We call σ the *volatility* of the stock[3]. The parameters μ and σ are required in our analysis, but the formula for pricing an option will involve only the volatility[4].

The case $\sigma = 1$ may be regarded as the unit measure of volatility. We let $Z_t = \sigma W_t$. This implies $\mathbb{E}[W_t] = 0,\ \mathbb{E}[W_t^2] = t$, and

$$X_t = X_0\ e^{\mu t + \sigma W_t}$$

The random or probabilistic components of the share price, on which we concentrate, is encoded in W_t. We make similar assumptions as were made in earlier examples of call options and follow similar approaches. Suppose investors act independently over disjoint time intervals and that their behaviour over a given time interval depends

[2]Different stocks, however, do have different spreads or variances, with the more speculative stocks having share price with bigger swings.

[3]If $\sigma = 0$, then $X_t = X_0 e^{\mu t}$ and the share price is non-random

[4]There is no universally agreed method for calculating volatility and the topic is currently an active area of research. It is a surprise that the drift does not appear in the Black Scholes formula for pricing an option. This is related to the fact that the expected return in a fair game should not drift up or down. The drift is used in the analysis leading to the Black Scholes formula. In practice the volatility is usually non-constant and may even be a random variable.

only on the length of the interval. We fix $t > 0$ and divide the time interval $[0, t]$ into a large number, n, of sub-intervals of equal length $\Delta t = t/n$. By our assumptions W_t is the limit of equally likely independent up/down jumps of size Δx on each sub-interval as $n \to \infty$. That is in the time sub-interval $[s, s + \Delta t]$, X_s can either take value $X_s e^{\mu\Delta t+\sigma\Delta x}$ or $X_s e^{\mu\Delta t-\sigma\Delta x}$ at time $s + \Delta t$. Let B_j^n denote the random variable which records the movement on the j^{th} sub-interval. This is similar to the situation for symmetric walks[5]. Hence

$$\begin{aligned} P[B_j^n &= \Delta x] = P[B_j^n = -\Delta x] = 1/2, \\ \mathbb{E}[B_j^n] &= 1/2(\Delta x) + 1/2(-\Delta x) = 0, \\ Var(B_j^n) &= \mathbb{E}[(B_j^n)^2] = 1/2(\Delta x)^2 + 1/2(-\Delta x)^2 = (\Delta x)^2. \end{aligned}$$

The cumulative effect of these independent jumps is $\sum_{j=1}^{n} B_j^n$ and

$$\mathbb{E}[\sum_{j=1}^{n} B_j^n] = 0.$$

The sequence $(B_j^n)_{j=1}^n$ consists of identically distributed independent random variables, which implies that

$$Var(\sum_{j=1}^{n}) = nVar(B_j^n) = n(\Delta x)^2.$$

Since $\sum_{j=1}^{n} B_j^n$ tends to W_t as $n \to \infty$ it is reasonable to suppose that

$$n(\Delta x)^2 = Var(\sum_{j=1}^{n} B_j^n) \to Var(W_t) = t \quad as\ n \to \infty.$$

It is thus convenient to let $n(\Delta x)^2 = t \Leftrightarrow (\Delta x)^2 = t/n \equiv \Delta t$. Hence $\sqrt{n}(\Delta x) = \sqrt{t}$, and by Central Limit Theorem,

$$\lim_{n\to\infty} P\left[\frac{\sum_{j=1}^{n} B_j^n}{\sqrt{t}} \leq x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2 dy}$$

[5] as in Example 4.1.4

for any real number x. Hence $W_t/\sqrt{t}$ is an $N(0,1)$ distributed random variable and W_t is $N(0,t)$. The following proposition summarizes above concepts

Proposition 6.1.1 *If $(X_t)_{t\geq 0}$ is the process that gives the share price of a stock at time t, then, under the above assumptions, there exists a constant μ, the drift, and a constant σ, the volatility, such that*

$$X_t = X_0 e^{\mu t+\sigma W_t}$$

where W_t is a normal random variable with mean 0 and variance t.

The above process $(W_t)_{t\geq 0}$ has the following properties:

(a) $W_0 = 0$ almost surely

(b) W_t is $N(0,t)$ distributed for all $t \geq 0$(Gaussian Increments)

(c) for any n and $\{0 = t_0 < t_1 < t_2 ... < t_{n+1}\}$, $(W_{t_i} - W_{t_{i-1}})_{i=1}^n$ is a set of independent random variables(independent increments)

(d) the probability distribution of $W_t - W_s$ depends only on $t-s$ for $0 \leq s \leq t$(stationary increments).

Hence the stochastic process is a *Brownian motion or Wiener processes.*

If $0 \leq s \leq t$, then (a) and (c) implies that $W_t - W_s$, $W_{t-s} - W_0$, W_{t-s} are identically distributed $N(0, t-s)$,by (b), random variable.

We now are in a familiar situation; we have a collection of random variable, in fact an uncountable collection, about which we have certain probabilistic information but we do not have a probability space $(\Omega, \mathcal{F}, P)$ in which to analyze this information. Wiener constructed a suitable measure in 1923. We briefly outline a construction of this measure.

For each positive number t, $W_t(x) = x$ for all $x \in \mathbb{R}$ is an $N(0,t)$ distributed random variable on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_t)$ where $P_t(A) = \frac{1}{\sqrt{2\pi t}} \int_A e^{-\frac{x^2}{2t}} dx$ for any Borel subset $A \subset \mathbb{R}$ and any $t > 0$ and $P_0(0) = 1$. If $(t_i^n)_{i=0}^n$ is a finite increasing sequence of positive real number, with $t_0 = 0$, then $(W_{t_i} - W_{t_{i-1}})_{i=1}^n$ is a set of independent random variables and there exists a product probability measure $\mathbb{P}_t$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ where $t = \{t_0, ..., t_n\}$ such that

$$\mathbb{P}_t(A_1 \times \cdots \times A_n) = P_{t_1 - t_0}(A_1) \times \cdots \times P_{t_n - t_{n-1}}(A_n)$$

for any finite collection, $(A_i)_{i=1}^n$, of Borel subsets of $\mathbb{R}$. If $x = (x_1, ..., x_n) \in \mathbb{R}^n$ and we identify $W_{t_i} - W_{t_{i-1}}$ with the random variable $x \in \mathbb{R}^n \mapsto (W_{t_i} - W_{t_{i-1}})(x_i)$ for all $i,\ 1 \le i \le n$, then $(W_{t_i} - W_{t_{i-1}})_{i=1}^n$ is a set of independent random variables on the product probability space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P}_t)$ and we have the following consistency condition: if $t_{n+1} \ge t_n$ and $t' = (t, t_{n+1}),\ t \in \mathbb{R}^n$, then

$$\lim_{x_{n+1} \to +\infty} \mathbb{P}_{t'}(x \in \mathbb{R}^{n+1}\ :\ W_{t_i} \le x_i,\ i \le n+1) = \mathbb{P}_t(x \in \mathbb{R}^n\ :\ W_{t_i} \le x_i,\ i \le n).$$

The above consistency condition leads to the following question:

How long can the probability measure be extended? Equivalently, can it be extended from countable product space to uncountable product space?

Kolmogorov provided a theorem that suggests its existence. It is stated in the following theorem.

Theorem 6.1.2 (Kolmogorov Extension Theorem) *Suppose that associated with each $0 \le t_1 < t_2 \cdots < t_n,\ n \ge 1$, there is a probability measure $\mu_{t_1, t_2, ..., t_n}$ on $\mathbb{R}^n$. Assume that the family*

$$\{\mu_{t_1, t_2, ..., t_n} : 0 \le t_1 < ... < t_n,\ n = 1, 2, ...\} \tag{6.1.2.1}$$

satisfies the consistency condition

$$\mu_{t_1, ... t_{i-1}, \hat{t}_i, t_{i+1}, ..., t_n}(A_1 \times A_2) = \mu_{t_1, ..., t_n}(A_1 \times \mathbb{R} \times A_2) \tag{6.1.2.2}$$

where $1 \leq i \leq n$ *and* $\hat{t}_i$ *means that* t_i *is deleted. Then there exists a unique probability measure* P *on the space* $(\mathbb{R}^{[0,\infty)}, \mathcal{F})$ *such that*

$$P\{\omega \in \mathbb{R}^{[0,\infty)} : (\omega(t_1), ..., \omega(t_n)) \in A\} = \mu_{t_1,...,t_n}(A) \tag{6.1.2.3}$$

for any $0 \leq t_1 < ... < t_n$, $A \in \mathcal{B}(\mathbb{R}^n)$ *and* $n \geq 1$.

Kolmogorov's Consistency Theorem shows that the required probability space exists. The sample space Ω is the uncountable product space $\mathbb{R}^{[0,\infty)}$. A typical point $\omega \in \Omega$ has the form $(\omega_t)_{t\geq 0}$ and may be identified with the function or path $t \in \mathbb{R} \mapsto \omega_t$. Thus Ω is the space of sample paths. The σ-field $\mathcal{F}_\infty$ is the smallest σ-field on Ω for which the function $\omega \in \Omega \mapsto \omega_t \in \mathbb{R}$ are measurable for all $t \geq 0$. We identify W_t with the function $\omega \mapsto \omega_t$; that is $W_t(\omega) = \omega_t$ for all $\omega \in \Omega$ and $t \geq 0$. We call the measure constructed a *Wiener measure* and denote it by W. If $\mathcal{F}_t$ denote the σ-field generated by $(W_s)_{0\leq s\leq t}$, then $(W_t)_{t\geq 0}$ is a stochastic process on the probability space $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. The following proposition is an upgrading of Proposition 6.1.1 incorporating the results of Wiener.

Proposition 6.1.3 *If* $(W_t)_{t\geq 0}$ *is a Wiener process, then there exists a probability space* $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$*, a filtration on* $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$, $(\mathcal{F}_t)_{t\geq 0}$*, such that* $(W_t)_{t\geq 0}$ *is a stochastic process adapted to the filtration. Moreover, paths of the process are almost surely continuous and almost surely nowhere differentiable with respect to the measure* W.

By proposition 4.3.3, $(W_t)_{t\geq 0}$ and $(W_t^2 - t)_{t\geq 0}$ are martingales. The process $(\mu t + \sigma W_t)_{t\geq 0}$ is called *Brownian motion with drift*, while the process $(Ce^{\mu t+\sigma W_t})_{t\geq 0}$ is called *geometric or exponential Brownian motion.* If interest is a constant rate r, then the discounted share price

$$e^{-rt}X_t = X_0 e^{(\mu-r)t+\sigma W_t} \tag{6.1.3.1}$$

is a random variable on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ and $(e^{-rt}X_t)_{t\geq 0}$ is a stochastic process adapted to the filtration generated by the Wiener process. For a fair price, the buyer requires, by (3.2.5.3), the discounted share price to be a martingale. However, by proposition 4.3.3(b), this will occur only when $\mu - r = -\sigma^2/2$. This shows that the particular Wiener measure we have constructed is not the required risk neutral probability measure. In the latter sections we rectify this matter.

[ch 6][↑]

6.2 Call Option 3

Here we present the Black Scholes Formula for pricing an even more realistic model for option pricing.

We proceed to price a call option so that it represents a fair price for the buyer. Let r denote the interest rate, T denote the maturity date and k denote the strike price of a call option. By, Example 3.2.1, we need to show that there exists a probability measure, say P_N, on the measurable space $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty)$ such that

$$(e^{-rt}X_t)_{t\geq 0} \text{ is a martingale on } (\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty). \tag{6.2.0.2}$$

and afterwards we need to evaluate

$$\mathbb{E}_{P_N}[e^{-rt}(X_T - K)^+|\mathcal{F}_0] = \mathbb{E}_{P_N}[e^{-rt}(X_T - k)^+]. \tag{6.2.0.3}$$

Equation 6.2.0.3 holds because $\mathcal{F}_0 = \mathcal{F}_\phi$ and gives a fair price for the buyer since $(\mathbb{E}_{P_N}[e^{-rt}(X_T - k)^+|\mathcal{F}_t])_{0\leq t\leq T}$ is a martingale.

We assume that the stock has drift μ and strictly positive volatility σ. We partition the interval $[0,t]$ into n adjacent sub-intervals, each of length $\Delta t = t/n$. As noted in the previous section, the share price changes on each sub-intervals by a factor $e^{\mu t \pm \sigma \Delta x}$ where $(\Delta x)^2 = \Delta t$. Thus the risk neutral probability, p, that the discounted share price rises by Δx over a typical sub-interval $[s, s+\Delta t]$ so that a fair price is maintained or equivalently that the martingale property is satisfied, is given by

$$\begin{aligned} p &= \frac{e^{r(\Delta x)^2} - e^{\mu(\Delta x)^2 - \sigma\Delta x}}{e^{\mu(\Delta x)^2+\sigma\Delta x} - e^{\mu(\Delta x)^2-\sigma\Delta x}} \\ &= \frac{e^{(r-\mu)(\Delta x)^2} - e^{-\sigma\Delta x}}{e^{\sigma\Delta x} - e^{-\sigma\Delta x}} \\ &\approx \frac{(r-\mu)(\Delta x)^2 + \sigma\Delta x - \sigma^2(\Delta x)^2/2}{\sigma\Delta x + \sigma^2(\Delta x)^2/2 + \sigma\Delta x - \sigma^2(\Delta x)^2/2} \\ &= \frac{\sigma + (r-\mu-\sigma^2/2)\Delta x}{2\sigma} \\ &= \frac{1}{2}(1 + \frac{(r-\mu-\frac{\sigma^2}{2})}{\sigma}\Delta x) \end{aligned}$$

The above shows that

$$p = \frac{1}{2}(1 + \frac{(r-\mu-\frac{\sigma^2}{2})}{\sigma}\Delta x) + f(\Delta x) \tag{6.2.0.4}$$

where $|f(x)| \leq c|x|^2$ for some constant $c > 0$ and all x close to 0. We use the approximated value of p, neglecting the $f(\Delta x)$ term as its contribution is negligible.

Let $(A_j)^n$ denote the random variable on the j^{th} interval which takes the value $\sigma\Delta x$ with probability p and value $-\sigma\Delta x$ with probability $(1-p)$. Thus the change in the discounted share price over $[0,t]$ is given approximately by the random variable

$$X_0 e^{(\mu-r)t + \sum\limits_{j=1}^{n} A_j^n}.$$

And we have

$$\begin{aligned} \mathbb{E}[A_j^n] &= p(\sigma\Delta x) + (1-p)(-\sigma\Delta x) \\ &= (2p-1)(\sigma\Delta x) \\ &= (r-\mu-\frac{\sigma^2}{2})(\Delta x)^2 \\ \mathbb{E}[(A_j^n)^2] &= p(\sigma\Delta x)^2 + (1-p)(\sigma\Delta x)^2 \\ &= \sigma^2(\Delta x)^2 \end{aligned}$$

Hence

$$Var(A_j^n) = \mathbb{E}[(A_j^n)^2] - \mathbb{E}[A_j^n]^2 \approx \sigma^2(\Delta x)^2.$$

Since the random variables $(A_j^n)_{j=1}^n$ are independent and identically distributed

$$\mathbb{E}[\sum_{j=1}^n A_j^n] = n(r - \mu - \frac{\sigma^2}{2})(\Delta x)^2 = (r - \mu - \frac{\sigma^2}{2})t$$

and

$$Var(\sum_{j=1}^n A_j^n) \approx n\sigma^2(\Delta x)^2 = \sigma^2 t.$$

By Central Limit Theorem

$$\lim_{n\to\infty} P\left[\frac{\sum\limits_{j=1}^n A_j^n - (r - \mu - \frac{\sigma^2}{2})t}{\sigma\sqrt{t}} \leq x\right] = \frac{1}{\sqrt{2\pi}}\int_\infty^x e^{-y^2/2dy} \tag{6.2.0.5}$$

for any $x \in \mathbb{R}$. Hence $\sum\limits_{j=1}^n A_j^n \xrightarrow{D} (r-\mu-\frac{\sigma^2}{2})t+\sigma\tilde{W}_t$[6] as $n \to \infty$ for all $t \geq 0$ where $\tilde{W}_t$ is an $N(0,t)$ distributed random variable. And the properties holding for earlier process $(W_t)_{t\geq 0}$ also holds for $(\tilde{W}_t)_{t\geq 0}$. Thus there exists[7] a probability measure on $\mathbb{R}^{[0,\infty)}$, P_N, such that $(\tilde{W}_t)_{t\geq 0}$ is a Wiener process on the probability space $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, P_N)$. We call P_N the risk neutral probability measure for the share price. Since

$$\mu t + (r - \mu - \frac{\sigma^2}{2})t + \sigma\tilde{W}_t = (r - \frac{1}{2}\sigma^2)t + \sigma\tilde{W}_t$$

we have the following result.

Proposition 6.2.1 *If a stock has drift μ and volatility σ, then there exist two probability measures: W, the Wiener measure, and P_N, the risk neutral probability measure, on the measurable space $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty)$ such that the share price at time t, X_t, has the following properties:*

[6] the convergence is in distribution sense

[7] by Kolmogorov's Consistency Theorem; which we won't be describing here

(a) under W, $X_t = X_0\, e^{\mu t + \sigma W_t}$, and $(W_t)_{t\geq 0}$ is a Wiener process

(b) underP_N, $e^{-rt}X_t = X_0\, e^{-\frac{\sigma^2}{2}t + \sigma \tilde{W}_t}$ and $(\tilde{W}_t)_{t\geq 0}$ is a Wiener process.

The following corollary follows directly from combining proposition 4.3.3(b) and 6.2.1(b)

Corollary 6.2.2 *The discounted share price $(e^{-rt}X_t)_{t\geq 0}$ is a martingale with respect to the risk neutral probability measure P_N.*

[ch 6][↑]

6.3 The Black-Scholes Formula

We let $N(x) = \frac{1}{2\pi}\int_{-\infty}^{x} e^{-y^2/2}dy$ i.e. $N(x) = P[X \leq x]$ where X is and $N(0,1)$ distributed random variable. And from symmetric property $N(-x) = 1 - N(x)$ for all $x \in \mathbb{R}$. The following is Black Scholes formula.

Proposition 6.3.1 *(Black Scholes Formula) Suppose the share price of a stock with volatility σ is X_0 today. For the buyer*

$$X_0 N\left(\frac{log(\frac{X_0}{k}) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) - ke^{-rT}N\left(\frac{log(\frac{X_0}{k}) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \tag{6.3.1.1}$$

is a fair price for a call option with maturity date T and strike price k given that r is the risk-free interest rate.

Proof By proposition 6.2.1(b) and corollary 6.2.2

$$V_0 = \mathbb{E}_{P_N}[e^{-rt}(X_T - k)^+|\mathcal{F}_0] = \mathbb{E}_{P_N}[e^{-rt}(X_T - k)^+]$$

is the buyer's fair price for the option, and it suffices to show that this reduces to 6.3.1.1. By proposition 6.2.1,

$$e^{-rT}(X_T - k)^+ = e^{-rT}(X_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{TY}} - k)^+$$

where Y is an $N(0,1)$ distributed random variable. Thus

$$V_0 = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-rT}(X_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{Tx}} - k)^+ e^{-\frac{1}{2}x^2} dx$$

Since

$$\begin{aligned} X_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{Tx}} - k \geq 0 \quad &\Leftrightarrow \quad e^{\sigma\sqrt{Tx}} \geq (\frac{k}{X_0})e^{-(r-\frac{1}{2}\sigma^2)T} \\ &\Leftrightarrow \quad x \geq \frac{1}{\sigma\sqrt{T}}\left(\log(\frac{k}{X_0}) - (r - \frac{1}{2}\sigma^2)T\right) := T_1 \end{aligned}$$

we have, using the substituting $y = x - \sigma\sqrt{T}$,

$$\begin{aligned} V_0 &= \frac{X_0 e^{-\frac{1}{2}\sigma^2 T}}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{\sigma\sqrt{Tx}-\frac{1}{2}x^2} dx - \frac{ke^{-rt}}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{-\frac{1}{2}x^2} dx \\ &= \frac{X_0}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{-\frac{1}{2}(x-\sigma\sqrt{T})^2} dx - ke^{-rT}(1 - N(T_1)) \\ &= \frac{X_0}{\sqrt{2\pi}} \int_{T_1-\sigma\sqrt{T}}^{\infty} e^{-\frac{1}{2}y^2} dy - ke^{-rT}(1 - N(T_1)) \\ &= X_0(1 - N(T_1 - \sigma\sqrt{T})) - ke^{-rT}(1 - N(T_1)) \end{aligned}$$

Since

$$T_1 - \sigma\sqrt{T} = \frac{1}{\sigma\sqrt{T}}(-\log(\frac{X_0}{k}) - (r + \frac{1}{2}\sigma^2)T)$$

we have

$$\begin{aligned} 1 - N(T_1 - \sigma\sqrt{T}) &= N(-T_1 + \sigma\sqrt{T}) \\ &= N(\frac{\log(\frac{X_0}{k}) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}) \end{aligned}$$

Similarly

$$1 - N(T_1) = N(\frac{\log(\frac{X_0}{k}) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}})$$

Now substituting these two formulas back into the integral representation for V_0, we get the Black Scholes formula.

And to find the value of the option at time $t, 0 \leq t \leq T$, apply the Black Scholes formula 6.3.1.1 with initial price X_t, strike price k and maturity date $T - t$.

Proposition 6.3.2 *(Call-Put Parity) Suppose the share price of a stock with volatility σ is X_0 today. If C_T and P_T denote the fair prices for a call option and a put option, respectively, with maturity date T, strike price k and risk-free interest rate r, then*

$$C_T - P_T = X_0 - ke^{-rt}.$$

Proof Since $C_T = \mathbb{E}[e^{-rT}(X_T - k)^+]$ and $P_T = \mathbb{E}[e^{-rt}(X_T - k)^-]$, proposition 6.2.1 and lemma 4.3.2 imply

$$\begin{aligned} C_T - P_T &= \mathbb{E}[e^{-rT}(X_T - k)^+] - \mathbb{E}[e^{-rT}(X_T - k)^] \\ &= \mathbb{E}[e^{-rt}(X_T - k)] \\ &= X_0 - ke^{-rt} \end{aligned}$$

completing the proof.

6.3.1 Change of Measure Derivation

In the previous section we followed a first principles approach and obtained the Black-Scholes formula by using a limiting process and finite risk neutral probabilities derived from properties of elementary fair games. In this section we take a different approach and use the martingale property 4.3.3.1 in Proposition 4.3.3(d).

We need to solve the following problem. Given a probability space $(\Omega, \mathcal{F}, P)$, a filtration $(\mathcal{F}_t)_{t \geq 0}$, a Wiener process $(W_t)_{t \geq 0}$ adapted to the filtration, μ and r real numbers and $\sigma > 0$, show that there exists a probability measure on $(\Omega, \mathcal{F})$ such that $(e^{(\mu-r)t+\sigma W_t})_{t\geq 0}$ is a martingale under this new measure. Once we have found this measure the derivation of the Black-Scholes formula proceeds as in the final part of the previous section.

Let $Y_t = e^{(\mu-r)t+\sigma W_t}$, and for a fixed $\gamma \in \mathbb{R}$ let $Z_t = e^{-\frac{1}{2}\gamma^2 t+\gamma W_t}$. By Proposition 4.3.3(d), there exists a positive integerable random variable Z_∞ on $(\Omega, \mathcal{F}, P)$ such that $P_\gamma(A) = \int_A Z_\infty dP$ defines a probability measure P_γ on $(\Omega, \mathcal{F})$ and, moreover, if Y is an $\mathcal{F}_t$ measurable P_γ integrable random variable and $0 \leq s \leq t$, then

$$\mathbb{E}_{P_\gamma}[Y|\mathcal{F}_s] = \mathbb{E}_P[Z_t \cdot Y|\mathcal{F}_s].$$

Since $\mathcal{F}_0 = \mathcal{F}_\phi$ this implies $\mathbb{E}_{P_\gamma}[Y] = \mathbb{E}_P[Z_t \cdot Y]$.

If $(Y_t)_{t\geq 0}$ is a martingale on $(\Omega, \mathcal{F}, P_\gamma)$, then the sequence $(\mathbb{E}_{P_\gamma}[Y_t])_{t\geq 0}$ must be independent of t. By lemma 4.3.2,

$$\begin{aligned} \mathbb{E}_{P_\gamma}[Y_t] &= \int_\Omega e^{(\mu-r)t+\sigma W_t} \cdot e^{-\frac{1}{2}\gamma^2 t+\gamma W_t} dP \\ &= e^{(\mu-r-\frac{1}{2}\gamma^2)t+\frac{1}{2}(\sigma+\gamma)^2 t}. \end{aligned}$$

This means we must have

$$\mu - r - \frac{1}{2}\gamma^2 + \frac{1}{2}(\sigma+\gamma)^2 = \mu - r + \frac{1}{2}\sigma^2 + \sigma\gamma = 0,$$

that is

$$\gamma = \frac{-\mu + r - \frac{1}{2}\sigma^2}{\sigma} \tag{6.3.2.1}$$

We now examine $(W_t)_{t\geq 0}$ as a stochastic process on $(\Omega, \mathcal{F}, P_\gamma)$ where γ is given by 6.3.2.1. Fix $x \in \mathbb{R}$ and $t > 0$ and let $f = \mathbb{I}_{(-\infty,x]}$ and $g(y) = e^{-\frac{1}{2}\gamma^2 t+\gamma y}$ for all $y \in \mathbb{R}$.

Then $g(W_t) = Z_t$, and $f(W_t) = \mathbb{I}_{\{W_t \le x\}}$. Since $f(W_t)$ is $\mathcal{F}_t$ measurable

$$
\begin{aligned}
P_\gamma(\{W_t \le x\}) &= \int_{\{W_t \le x\}} dP_\gamma \\
&= \int_\Omega \mathbb{I}_{\{W_t \le x\}} dP_\gamma \\
&= \int_\Omega f(W_t) dP_\gamma \\
&= \int_\Omega f(W_t) \cdot g(W_t) dP \\
&= \mathbb{E}_P[f(W_t) \cdot g(W_t)],
\end{aligned}
$$

and as W_t is $N(0,t)$ distributed $f(W_t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}}$ for all $y \in \mathbb{R}$. Therefore

$$
\begin{aligned}
\mathbb{E}_P[f(W_t) \cdot g(W_t)] &= \int_\mathbb{R} f(y) g(y) f_{W_t}(y) dy \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{1}{2}\gamma^2 t + \gamma y} \cdot e^{-\frac{1}{2t} y^2} dy \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{1}{2t}(y - \gamma t)^2} dy.
\end{aligned}
$$

Hence W_t is $N(\gamma t, t)$ distributed and $\tilde{W}_t := -\gamma t + W_t$ is $N(0,t)$ distributed over $(\Omega, \mathcal{F}, P_\gamma)$. Our choice of γ in 6.3.2.1 implies

$$
\begin{aligned}
Y_t = e^{(\mu - r)t + \sigma(\gamma t + \tilde{W}_t)} &= e^{(\mu r + \gamma\sigma)t + \sigma \tilde{W}_t} \\
&= e^{-\frac{1}{2}\sigma^2 t + \sigma \tilde{W}_t}
\end{aligned}
$$

on $(\Omega, \mathcal{F}, P_\gamma)$. To show that $(\tilde{W}_t)_{t \ge 0}$ is a Wiener process on $(\Omega, \mathcal{F}, P_\gamma)$ we must show that it has independent, stationary and Gaussian increments. Now for $0 \le s \le t$,

$$
\begin{aligned}
P_\gamma(\{W_t - W_s \le x\}) &= \int_\Omega \mathbb{I}_{\{W_t - W_s \le x\}} Z_t dP \\
&= \int_\Omega f(W_t - W_s) e^{-\frac{1}{2}\gamma^2 t + \gamma(W_t - W_s)} e^{\gamma W_s} dP \\
&= \mathbb{E}_P[f(W_t - W_s) g(W_t - W_s) e^{\gamma W_s}] \\
&= \mathbb{E}_P[f(W_t - W_s) g(W_t - W_s)] \cdot \mathbb{E}_P[e^{\gamma W_s}] \\
&= \frac{e^{\frac{1}{2}\gamma^2 s}}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{x} f(y) g(y) e^{-\frac{y^2}{2(t-s)}} dy
\end{aligned}
$$

by lemma 4.3.2 and since $W_t - W_s$ is $N(0, t-s)$. Hence

$$\begin{aligned} P_\gamma(\{W_t - W_s \leq x\}) &= \frac{e^{\frac{1}{2}\gamma^2 s}}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{x} e^{-\frac{1}{2}\gamma^2 t + \gamma y - \frac{y^2}{2(t-s)}} dy \\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{x} e^{-\frac{1}{2(t-s)}(y-(t-s)\gamma)^2} dy. \end{aligned}$$

This shows that $W_t - W_s$ is an $N((t-s)\gamma, t-s)$ distributed random variable on the space $(\Omega, \mathcal{F}, P_\gamma)$ and, as $\tilde{W}_t - \tilde{W}_s = -\gamma(t-s) + (W_t - W_s)$, $\tilde{W}_t - \tilde{W}_s$ has an $N(0, t-s)$ distribution over $\Omega, \mathcal{F}, P_\gamma$. We have shown that the stochastic process $(\tilde{W}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P_\gamma)$ has Gaussian and stationary increments. It remains to show that it has independent increments. It suffices to show that $\tilde{W}_t - \tilde{W}_s$ and $\tilde{W}_s$ are independent when $0 \leq s \leq t$. Let h and l denote real valued bounded Borel measurable functions on $\mathbb{R}$. Since $W_t - W_s$ and W_s are independent $\mathcal{F}_t$ measurable random variables on $(\Omega, \mathcal{F}, P)$, Proposition 4.3.3(d) implies

$$\begin{aligned} \mathbb{E}_{P_\gamma}[h(\tilde{W}_t - \tilde{W}_s) \cdot l(\tilde{W}_s)] &= \mathbb{E}_P[h(W_t - W_s - \gamma(t-s))e^{-\frac{1}{2}\gamma^2(t-s)+\gamma(W_t - W_s)} \\ &\quad \cdot l(W_s - \gamma s) \cdot e^{-\frac{1}{2}\gamma^2 s + \gamma W_s}] \\ &= \mathbb{E}_P[h(W_t - W_s - \gamma(t-s))e^{-\frac{1}{2}\gamma^2(t-s)+\gamma(W_t - W_s)}] \\ &\quad \cdot \mathbb{E}_P[l(W_s - \gamma s) \cdot e^{-\frac{1}{2}\gamma^2 s + W_s}] \\ &= \mathbb{E}_{P_\gamma}[h(\tilde{W}_t - \tilde{W}_s)] \cdot \mathbb{E}_{P_\gamma}[l(\tilde{W}_s)]. \end{aligned}$$

Therefore, $\tilde{W}_t - \tilde{W}_s$ and $\tilde{W}_s$ are independent random variables on $(\Omega, \mathcal{F}, P_\gamma)$. Hence we have proved the following proposition.

Proposition 6.3.3 *Let $(W_t)_{t\geq 0}$ denote a Wiener process on $(\Omega, \mathcal{F}, P)$ and let $Y_t = e^{(\mu - r)t + \sigma W_t}$ for $t \geq 0$ where $\mu, r, \sigma \in \mathbb{R}$ and $\sigma > 0$. If*

$$\gamma = \frac{-\mu + \gamma - \frac{1}{2}\sigma^2}{\sigma},$$

then on the probability space $(\Omega, \mathcal{F}, P_\gamma)$

$$Y_t = e^{-\frac{1}{2}\sigma^2 t + \sigma \tilde{W}_t},$$

where $\tilde{W}_t = -\gamma t + W_t$ *and* $(\tilde{W}_t)_{t\geq 0}$ *is a Wiener process. Moreover,* $(Y_t)_{t\geq 0}$ *is a martingale.*

By Proposition 6.2.1(b) and Proposition 6.3.3 the discounted share price has the same distribution under P_N and P_γ, γ as above, and hence both give the same formula, the Black-Scholes formula, for a call option. We now show that $P_N = P_\gamma$ on $(\Omega, \mathcal{F}_\infty)$ where $\mathcal{F}_\infty$ is the σ-field generated by $(X_t)_{t\geq 0}$.

If A is an $\mathcal{F}_t$ measurable subset of Ω, then, we can have a bounded Borel measurable function f on $\mathbb{R}$ such that $f(\tilde{W}_t) = \mathbb{I}_A$. Since $\tilde{W}_t$ has the same distribution under P_γ and P_N and $(\tilde{W}_s)_{0\leq s\leq t}$ generates the σ-filed $\mathcal{F}_t$, we have

$$P_\gamma(A) = \mathbb{E}_{P_\gamma}[\mathbb{I}_A] = \mathbb{E}_{P_\gamma}[f(\tilde{W}_t)] = P_N(A).$$

Hence $P_\gamma = P_N$ on $\bigcup_{t\geq 0} \mathcal{F}_t$ and hence on $\mathcal{F}_\infty$.

Chapter 7

Ito's Calculus

[↑]

7.1 Convergence of Random Variable

We will see that the Riemann Sums involved in constructing the Itô's integral do not appear to converge almost surely, and hence we are obliged to consider a weaker form of convergence, convergence in measure, which involves almost sure convergence of subsequences. To prepare for this definition lets have another notion of convergent sequences of real numbers.

Lemma 7.1.1 *A sequence of real numbers $(a_n)_{n=1}^{\infty}$ converges to the real number a if and only if every subsequence of $(a_n)_{n=1}^{\infty}$ contains a subsequence which converges to a.*

Now we define the above discussed weaker form of convergence by the following proposition.

Proposition 7.1.2 *Let $(X_n)_{n=1}^{\infty}$ and X denote random variables on a probability space $(\Omega, \mathcal{F}, P)$. The following conditions are equivalent:*

(a) $\lim_{n\to\infty} \int_\Omega \frac{|X_n - X|}{1+|X_n - X|} dP = 0$*;*

(b) $\lim_{n\to\infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\}) = 0$ *for every* $\epsilon > 0$*.*

(c) every subsequence of $(X_n)_{n=1}^{\infty}$ contains a subsequence which converges almost surely to X.

If these equivalent conditions are satisfied, the $(X_n)_{n=1}^{\infty}$ converges in measure (or probability) to X

Proof

- (a)$\Rightarrow$(b) Let $\epsilon > 0$ be arbitrary,

  $$A_n := \{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}$$

  and

  $$B_n := \{\omega \in \Omega : \frac{|X_n(\omega) - X(\omega)|}{1 + |X_n(\omega) - X(\omega)|} \geq \frac{\epsilon}{1+\epsilon}\}.$$

  Since $A_n \subset B_n$

  $$\begin{aligned} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}) &\leq \frac{1+\epsilon}{\epsilon} \int_{B_n} \frac{1+\epsilon}{\epsilon} dP \\ &\leq \frac{1+\epsilon}{\epsilon} \int_{\Omega} \frac{|X_n - X|}{1 + |X_n - X|} dP \end{aligned}$$

  which tends to 0 as $n \to \infty$.

- (b)$\Rightarrow$(c) If $(X_n)_{n=1}^{\infty}$ satisfies (b), then every subsequence of $(X_n)_{n=1}^{\infty}$ satisfies (b) and it is suffices to show that $(X_n)_{n=1}^{\infty}$ contains a subsequence which converges almost surely to X. For each positive integer k, choose a positive integer n_k such that

  $$P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \leq \frac{1}{2^k}\}) \geq 1 - \frac{1}{2^k}$$

  for all $n \geq n_k$. Let $C_k := \{\omega \in \Omega : |X_n(\omega) - X(\omega)| \leq \frac{1}{2^k}\}$ and let $m_k = n_1 + n_2 + \cdots + n_k$ for all k. The sequence $(m_k)_{k=1}^{\infty}$ is strictly increasing. If $\omega \in D_k := \bigcap_{l \geq k} C_l$, then $X_{m_k}(\omega) \to X(\omega)$ as $k \to \infty$. Then

  $$P(D_k^c) = P(\bigcup_{l \geq k} C_l^c) \leq \sum_{l=k}^{\infty} 2^{-l} = 2^{-k+1}.$$

  If $\omega \in D := \bigcup_{k \geq l} D_k$, then $X_{m_k}(\omega) \to X(\omega)$ as $k \to \infty$ and $P(D) \geq P(D_{k+1}) \geq 1 - \frac{1}{2^k}$ for all k. Hence $P(D) = 1$ and $X_{m_k} \to X$ almost surely as $k \to \infty$.

- (c)$\Rightarrow$(a) Let $Y_n = \frac{|X_n - X|}{1+|X_n - X|}$ and $a_n = \int_\Omega Y_n dP$ for all n. Then $|Y_n(\omega)| \leq 1$ for all n and ω. Given any subsequence $(Y_{n_j})_{j=1}^{\infty}$ of $(Y_n)_{n=1}^{\infty}$ we can, by hypothesis, choose a subsequence of $(Y_{n_j})_{j=1}^{\infty}$, $(Y_{n_{j_k}})_{k=1}^{\infty}$, which converges to 0 almost surely. By Dominated Convergence Theorem, $(a_{n_{j_k}})_{k=1}^{\infty}$ converges to 0, and by Lemma 7.1.1, the sequence $(a_n)_{n=1}^{\infty}$ converges to 0.

The following are certain properties of convergence in measure.

(a) If $(X_n)_{n=1}^{\infty}$ converges in measure to both X and Y, then $X = Y$ almost surely.

(b) If $(X_n)_{n=1}^{\infty}$ converges in measure to X, then every subsequence of $(X_n)_{n=1}^{\infty}$ converges in measure to X.

(c) Almost sure convergence implies convergence in measure.

(d) If every subsequence of $(X_n)_{n=1}^{\infty}$ contains a subsequence which converges in measure and any two such limits are almost surely equal, then the sequence $(X_n)_{n=1}^{\infty}$ converges in measure.

(e) If $(X_n)_{n=1}^{\infty}$ and $(Y_n)_{n=1}^{\infty}$ converge in measure to X and Y, respectively, and $a, b \in \mathbb{R}$, then $(aX_n + bY_n)_{n=1}^{\infty}$ converges in measure to $aX + bY$.

Now lets state some of the mathematical concepts that will used in the latter sections.

Proposition 7.1.3 *If $X, Y \in L^2(\Omega, \mathcal{F}, P)$, then*

(a) Cauchy-Schwarz Inequality :

$$||X \cdot Y||_1 \leq ||X||_2 \cdot ||Y||_2 \,,$$

(b) The Triangle Inequality

$$||X + Y||_2 \leq ||X||_2 + ||Y||_2,$$

(c) $||X||_1 \le ||X||_2$.

Proposition 7.1.4 *If* $(X_n)_{n=1}^{\infty}$ *is a Cauchy Sequence in* $L^i(\Omega, \mathcal{F}, P)$*, i=1,2, then there exists* $X \in L^i(\Omega, \mathcal{F}, P)$ *such that* $||X_n - X||_i \to 0$ *for i=1,2.*

i.e $L^i(\Omega, \mathcal{F}, P)$ *is complete.*

Remark: We complete this section by comparing the different ways of convergence of sequence

$$L^2 \; convergence \Rightarrow L^1 \; convergence \Rightarrow convergence \; in \; measure.$$

[ch 7][↑]

7.2 Riemann Stieltjes Integral

7.2.1 Variation and Quadratic Variation of a function

If g is a function of real variable, its variation over the interval $[a, b]$ is defined as

$$V_g([a, b]) = \sup \sum_{i=1}^{n} |g(t_i^n) - g(t_{i-1}^n)|,$$

where the supremum is taken over partitions $a = t_0^n < t_1^n < \cdots < t_n^n = b$.

Thus by triangle inequality the above sum increases on the partitions become finner and finner. Therefore, the variation of g is

$$V_g([a, b]) = \lim_{\delta_n \to 0} \sum_{i=1}^{n} |g(t_i^n) - g(t_{i-1}^n)|,$$

where $\delta_n = \max_{1 \le i \le n}(t_i - t_{i-1})$. Now if g is a function of $t \ge 0$, then the variation of g as a function of t is defined by

$$V_g(t) = V_g([0, t]).$$

Definition 7.2.1 *g is of finite variation if $V_g(t) < \infty$ for all t. And g is of bounded variation is $\sup_t V_g(t) < \infty$.*

Theorem 7.2.2 (Jordan Decomposition) *Any function $g : [0, \infty) \to \mathbb{R}$ of finite variation can be expressed as the difference of two increasing function*

$$g(t) = a(t) - b(t).$$

One such decomposition is given by

$$a(t) = V_g(t),\ b(t) = V_g(t) - g(t).$$

But the representation as the difference of two increasing function may not be unique, as

$$g(t) = \frac{1}{2}(V_g(t) + g(t)) - \frac{1}{2}(V_g(t) - g(t))$$

We now define the Quadratic Variation of a function.

Definition 7.2.3 *If g is a function of real variable, define its quadratic variation over the interval $[0, t]$ as the limit (when it exists)*

$$[g](t) = \lim_{\delta_n \to 0} \sum_{i=1}^{n} (g(t_i^n) - g(t_{i-1}^n))^2$$

where the limit is taken over partitions $0 = t_0^n < t_1^n < \cdots < t_n^n = t$, with $\delta_n = max_{1 \le i \le n}(t_i - t_{i-1})$.

7.2.2 Riemann Stieltjes Integral

The Stieltjes Integral is an integral of the form $\int_a^b f(t)dg(t)$, where g is a function of finite variation. Since a function of finite variation is a difference of two increasing functions, it is sufficient to define the integral with respect to monotone functions.

Definition 7.2.4 *The Stieltjes Integral of f with respect to a monotone function g over an interval $[a, b]$ is defined as*

$$\int_a^b f dg = \int_a^b f(t)dg(t) = \lim_{\delta_n \to 0} \sum_{i=1}^n f(\xi_i^n)(g(t_i^n) - g(t_{i-1}^n)), \tag{7.2.4.1}$$

where t_i^n, s represent partitions of the interval, $a = t_0^n < t_1^n < \cdots < t_n^n = b$, $\delta_n = max_{1 \le i \le n}(t_i - t_{i-1})$ and $\xi_i^n \in [t_{i-1}^n, t_i^n]$.

Now suppose f is monotonically increasing and continuous and g is continuous, then we can use the integration by parts formula to define

$$\int_a^b f(t)dg(t) \equiv [f(t)g(t)]_a^b - \int_a^b g(t)df(t). \tag{7.2.4.2}$$

where the integral in the right-hand side is defined by Riemann Stieltjes Integral. Thus the natural question arises

When can we define the integral $\int_a^b f(t)g(t)$, for any continuous functions f and g ?

The answer to the above question is not trivial but we can have certain conditions to check non-integrability of certain class of functions.

Consider the special case $f = g$, i.e., the integral

$$\int_a^b f(t)df(t).$$

Let $a = t_0^n < t_1^n < \cdots < t_n^n = b$ be a partition of $[a, b]$. Let L_n and R_n denote the lower and upper Riemann Sums respectively, i.e.

$$L_n = \sum_{i=1}^n f(t_{i-1})(f(t_i) - f(t_{i-1})), \tag{7.2.4.3}$$

$$R_n = \sum_{i=1}^n f(t_i)(f(t_i) - f(t_{i-1})). \tag{7.2.4.4}$$

Now for the Riemann Sum to exists $L_n = R_n$, as $\delta_n \to 0$.

$$R_n - L_n = \sum_{i=1}^{n} (f(t_i) - f(t_{i-1}))^2. \tag{7.2.4.5}$$

$$R_n + L_n = \sum_{i=1}^{n} (f(t_i)^2 - f(t_{i-1})^2) = f(b)^2 - f(a)^2. \tag{7.2.4.6}$$

Therefore, R_n L_n are given by

$$R_n = \frac{1}{2}\left(f(b)^2 - f(a)^2 + \sum_{i=1}^{n} (f(t_i) - f(t_{i-1}))^2\right).$$
$$L_n = \frac{1}{2}\left(f(b)^2 - f(a)^2 - \sum_{i=1}^{n} (f(t_i) - f(t_{i-1}))^2\right).$$

The quadratic variation of f over $[a, b]$ is $\lim_{\delta_n \to 0} R_n - L_n$. Obviously, $\lim_{\delta_n \to 0} R_n \neq \lim_{\delta_n \to 0} L_n$ if and only if the quadratic variation of the function f is non-zero. Hence the Riemann Stieltjes Integral is not defined for non-zero quadratic variation functions.

Let us consider the following example. Suppose f is continuous function satisfying the condition

$$|f(t) - f(s)| \approx |t - s|^{\frac{1}{2}}. \tag{7.2.4.7}$$

In this case, we have

$$0 \leq R_n - L_n \approx \sum_{i=1}^{n} (t_i - t_{i-1}) = b - a$$

Hence, the quadratic variation is $b - a \neq 0$ for $a \neq b$. Consequently, the Riemann Stieltjes Integral cannot be defined for $g = f$.

Remark: We can see that from equation 5.1.2.2, R-S Integral cannot be defined for Brownian motion B.

Now we will see Brownian Motion has other two properties that inhibits its Riemann Stieltjes Integration.

Theorem 7.2.5 (Non-zero Quadratic variation) *Let $(B_t)_{t\geq 0}$, be a Brownian motion. And let $\Delta^n := \{s = t_0^n < t_1^n < \cdots < t_{k_n}^n = t\}$, be a sequence of partitions of the interval $[s,t]$ such that $|\Delta^n| :=\ max_i \Delta t_i^n \to 0,\ n \to \infty$. Then*

$$\sum_{i=0}^{k_n-1} (\Delta B_i^n)^2 = \sum_{i=0}^{k_n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2 \xrightarrow{L^2} t-s,\ n \to \infty. \tag{7.2.5.1}$$

Remark: From equation 7.2.5.1, a Brownian motion is said to have quadratic variation $t-s$ in any interval $[s,t]$. Symbolically, represented by $(dB_t)^2 = dt$.

Proof Let $\Delta^n B_i = B_{t_{i+1}^n} - B_{t_i}^n$ and $\Delta^n t_i = t_{i+1}^n - t_i^n$.

Then

$$\begin{aligned} \epsilon_n &:= \mathbb{E}[(\sum_i \Delta^n B_i^2 - (t-s))^2] \\ &= \mathbb{E}[(\sum_i \Delta^n B_i^2 - \sum_i \Delta^n t_i)^2] \\ &= \mathbb{E}[(\sum_i (\Delta^n B_i^2 - \Delta^n t_i))^2]. \end{aligned}$$

Since the increments of Brownian Motion $\Delta^n B_i$ is independent, the random variable $\Delta^n B_i^2 - \Delta^n t_i$, as their increments, are also independent. Using this and the equality $\mathbb{E}[(\Delta^n B_i^2 - \Delta^n t_i)] = 0$, we have

$$\begin{aligned} \epsilon_n &= var\left(\sum_i (\Delta^n B_i^2 - \Delta^n t_i)\right) \\ &= \sum_i var(\Delta^n B_i^2 - \Delta^n t_i) \\ &= \sum_i \mathbb{E}[(\Delta^n B_i^2 - \Delta^n t_i)^2] \\ &= \sum_i (\mathbb{E}[\Delta^n B_i^4] - 2\Delta^n t_i \mathbb{E}[\Delta^n B_i^2] + \Delta^n t_i^2) \\ &= \sum_i (3\Delta^n t_i^2 - 2\Delta^n t_i^2 + \Delta^n t_i^2) = 2\sum_i \Delta^n t_i^2 \\ &\leq 2|\Delta^n| \sum_i \Delta^n t_i \\ &= 2(t-s)|\Delta^n| \to 0. \end{aligned}$$

Theorem 7.2.6 *Brownian Motion has infinite variation in every interval* $[s,t]$ *almost surely.*

Proof Using the same notion, we have

$$\begin{aligned}\sum_i |\Delta^n B_i| &= \sum_i \frac{\Delta^n B_i^2}{|\Delta^n B_i|} \\ &\geq \sum_i \frac{\Delta^n B_i^2}{max_j\ |\Delta^n B_j|} \\ &= \frac{1}{max_j\ |\Delta^n B_j|} \sum_i \Delta^n B_i^2.\end{aligned}$$

By theorem 7.2.5, $\sum_i \Delta^n B_i^2 \xrightarrow{L^2} t-s$, and hence $\sum_i \Delta^n B_i^2 \xrightarrow{P} t-s$ as $n \to \infty$. Therefore, $\sum_i \Delta^n B_i^2 \to t-s,\ n \to \infty$ almost surely. Since the trajectories of Brownian motion are uniformly continuous functions on $[s,t]$, $max_j|\Delta^n B_j| \to 0$ a.s. Therefore,

$$\sum_i |\Delta^n B_i| = \sum_i |B_{t_{i+1}^n} - B_{t_i^n}| \to \infty,\ n \to \infty.$$

and thus the variation of Brownian motion is infinite in any interval $[s,t]$.

[ch 7][↑]

7.3 Stochastic Riemann Integral

Definition 7.3.1 *A stochastic process* $X := (X_t)_{t\geq 0}$ *on* $(\Omega, \mathcal{F}, P)$ *is continuous if there exists a* $\mathcal{F}-measurable$ $A \subset \Omega$ *with* $P(A) = 1$ *such that*

$$t \mapsto X_t(\omega),\ \ t \in [0, +\infty) \tag{7.3.1.1}$$

is continuous for all $\omega \in A$.

A path $\omega \in \Omega$ is called $X-continuous$ if the mapping 7.3.1.1 is continuous.

Remark: Any Wiener Process is a continuous process.

In view of our above discussion, we let $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ denote the probability space for Wiener Process $(W_t)_{t\geq 0}$, where $\mathcal{F}_\infty$ denote the $\sigma - field$ generated by $(W_t)_{t\geq 0}$ and $\mathcal{F}_t$ the $\sigma - field$ generated by $(W_s)_{0\leq s\leq t}$ for all $t \geq 0$. Since $W_t(\omega) = \omega_t$, when $\omega = (\omega_t)_{t\geq 0} \in \mathbb{R}^{[0,\infty)}$, $W - continuity$ coincides with the usual notion of continuity.

Let $B : [0, \infty) \to \mathbb{R}$ be a continuously differentiable function. To define $\int_0^T X_t dB(t)$,$0 < T < \infty$, we consider the Riemann Sum corresponding to the partition of $[0, T]$, given by $\mathcal{P}_n := 0 = t_0, t_1, ..., t_n = T$,

$$\sum_{i=0}^{n-1} X_{t_i}(B(t_{i+1}) - B(t_i)). \tag{7.3.1.2}$$

So by Mean Value Theorem, $\forall i\ \exists\ t_i^*,\ t_i < t_i^* < t_{i+1}$ such that $B(t_{i+1}) - B(t_i) = B'(t_i^*)(t_{i+1} - t_i)$. Hence 7.3.1.2 becomes

$$\sum_{i=0^{n-1}} X_{t_i} B'(t_i^*)(t_{i+1} - t_i).$$

Now if the mapping $t \mapsto X_t(\omega)$ is continuous then

$$\sum_{i=0}^{n-1} X_{t_i}(\omega) B'(t_i^*)(t_{i+1} - t_i) \xrightarrow{mesh\ \mathcal{P}_n \to 0} \int_0^T X_t(\omega) B'(t) dt$$

Hence if $(X_t)_{t\geq 0}$ is a continuous process, then

$$\omega \mapsto \left(\int_0^T X_t dB(t) \right)(\omega) := \int_0^T X_t(\omega) B'(t) dt$$

is the almost sure limit of a sequence of measurable Riemann Sums and may be regarded as a random variable, $\int_0^T X_t dB(t)$, on $(\Omega, \mathcal{F}, P)$. Therefore, $(\int_0^T X_t dB(t))_{T\geq 0}$ can be considered as a stochastic process.

Proposition 7.3.2 *If $X := (X_t)_{t\geq 0}$ is a continuous process on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, then $(\int_0^T X_t dB(t))_{T\geq 0}$ is a continuous process on $(\Omega, \mathcal{F}, P)$ adapted to filtration $(\mathcal{F}_t)_{t\geq 0}$.*

Proof The Riemann sums for $\int_0^T X_t dt$ are $\mathcal{F}_T$ measurable and, as the integral along each $X - continuous$ path is the limit of a sequence of Riemann Sums, $\int_0^T X_t dt$ is $\mathcal{F}_T$ measurable.

If ω is $X - continuous$ and $\epsilon \in \mathbb{R}$ is arbitrary, then

$$\left|\int_0^{T+\epsilon} X_t(\omega)dt - \int_0^T X_t(\omega)dt\right| \leq |\epsilon| \cdot max\{|X_t(\omega)| : T - |\epsilon| \leq t \leq T + |\epsilon|\}.$$

By continuity, $max\{|X_t(\omega)| : T - |\epsilon| \leq t \leq T + |\epsilon|\} \rightarrow |X_T(\omega)|$ as $\epsilon \rightarrow 0$.

In the succeeding section we want to define the integral of a process,$(X_t)_{t\geq 0}$, with respect to another process, i.e. $B(t)$ is a process $(B_t)_{t\geq 0}$, known as *Ito's Integral.* But before that we define *step process* and study calculus on it.

Definition 7.3.3 *A step process on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ is a stochastic process X of the form,*

$$X = \sum_{i=0}^{k-1} X_{t_i} \mathbb{I}_{[t_i, t_{i+1})}$$

where X_{t_i} is $\mathcal{F}_{t_i} - measurable$ random variable for all i and $0 = t_0 < t_1 ... < t_k$.

A step process is not usually a continuous process. However, it is Riemann Integrable over all finite intervals along all paths.

For $T > 0$, we have the product measure space $(\Omega \times [0, T], \mathcal{F} \times \mathcal{B}_T, P \times m)$, where $\mathbb{B}_T$ is the borel field on $[0, T]$.

Suppose $X = X_{t_i} \mathbb{I}_{[t_i, t_{i+1})}$, then

$$X_t(\omega) = \begin{cases} X_{t_i}(\omega), & \text{if } t_i \leq t \leq t_{i+1} \\ 0, & otherwise \end{cases}$$

and if $B \in \mathcal{B}(\mathbb{R})$, then

$$X^{-1}(B) = \begin{cases} X_{t_i}^{-1}(B) \times [t_i, t_{i+1}), & \text{if } 0 \notin B \\ (X_{t_i}^{-1}(B) \times [t_i, t_{i+1})) \cup (\Omega \times ([0, t_i) \cup [t_{i+1}, T))), & \text{if } 0 \in B \end{cases}$$

This shows that X is $\mathcal{F} \times \mathcal{B}_T - measurable$. Hence step process, finite sum of X's, is $\mathcal{F} \times \mathcal{B}_T - measurable \; \forall \; T \geq 0$.

If $(X_t)_{t\geq 0}$ is a stochastic process on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and $\{[t_i^n, t_{i+1}^n)\}_{i=0}^{k_n-1}$ is a partition of $[0, t_{k_n}^n)$, then we let

$$X^n = \sum_{i=0}^{k_n-1} X_{t_i^n} \mathbb{I}_{[t_i^n, t_{i+1}^n)}. \tag{7.3.3.1}$$

Lemma 7.3.4 *If X is a continuous stochastic process on $(\Omega, \mathcal{F}, P)$ adapted to filtration $(\mathcal{F}_t)_{t\geq 0}$, then X is $\mathcal{F} \times \mathcal{B}_T - measurable$ for all $T > 0$.*

Proof We choose $(t_i^n)_{i=0}^{k_n}$ and X_n as above and suppose $t_{k_n}^n = T$ and $\sup\{|t_{i+1}^n - t_i^n| : 1 \leq i \leq k_n\} \leq \frac{1}{n}$. Since each X^n is a step process, the sequence $(X_n)_{n=1}^{\infty}$ consists of $\mathcal{F} \times \mathcal{B}_T$ measurable functions. Suppose all paths in A are $X - continuous$ where $A \in \mathcal{F}$ and $P(A) = 1$. If $\omega \in A$ and $t \in [0, T]$ are fixed, then, for each n, there exists a unique t_i^n, $|t - t_i^n| \leq \frac{1}{n}$, such that $X^n(t, \omega) = X_t^n(\omega) = X_{t_i^n}(\omega)$. By continuity,

$$X_t^n(\omega) = X_{t_i^n}(\omega) \to X_t(\omega), \text{ as } n \to \infty$$

for all $\omega \in A$ and all $t,\ 0 \leq t \leq T$. Since

$$P \times m(A \times [0, T)) = P(A) \times m([0, T)) = T$$

the sequence $(X_n)_{n=1}^{\infty}$ converges almost everywhere to X on the measure space $(\Omega \times [0, T), \mathcal{F} \times \mathcal{B}, P \times m)$ and, as each X^n is $\mathcal{F} \times \mathcal{B}_T$ measurable, this completes the proof.

Now we recall the results of Fubini's Theorem that shall be required the latter phase. If X is a measurable function on $(\Omega \times [0, T), \mathcal{F} \times \mathcal{B}, P \times m)$, then X is integrable if and only if all three of the following integrals are well-defined and finite:

$$\begin{aligned}
\int_{\Omega\times[0,T)} |X| d(P \times m) &= \int_{\Omega} \left(\int_{[0,T)} |X(., \omega)| dm \right) dP \\
&= \int_{[0,T)} \left(\int_{\Omega} |X(t, .)| dP \right) dm.
\end{aligned}$$

And when X is integrable

$$\begin{aligned}\int_{\Omega\times[0,T)} Xd(P\times m) &= \int_{\Omega}\left(\int_{[0,T)} X(.,\omega)dm\right)dP \\ &= \int_{[0,T)}\left(\int_{\Omega} X(t,.)dP\right)dm.\end{aligned}$$

Our final result in this section is a simple criterion which guarantees that the Riemann stochastic integral of a continuous process is an integrable random variable.

Proposition 7.3.5 *If X is a continuous process on $(\Omega, \mathcal{F}, P)$, then $\int_0^T |X_t|dt$ is an integrable random variable if and only if $\int_0^T \mathbb{E}[|X_t|]dt < \infty$. If $\int_0^T |X_t|dt$ is integrable, then $\int_0^T X_t dt$ is integrable and*

$$\mathbb{E}[\int_0^T X_t dt] = \int_0^T \mathbb{E}[X_t]dt.$$

Proof By Fubini's Theorem

$$\int_{\Omega}\left(\int_0^T |X_t|dt\right)dP = \int_0^T\left(\int_{\Omega}|X_t|dP\right)dt = \int_0^T \mathbb{E}[|X_t|]dt$$

The left hand side is finite iff the random variable $\omega \mapsto \int_0^T |X_t(\omega)|dt$ is integrable. Since $|\int_0^T X_t(\omega)dt| \leq \int_0^T |X_t(\omega)|dt$ for all X-continuous paths, implies $\int_0^T X_t dt$ is integrable whenever $\int_0^T |X_t|dt$ is integrable, and a further application of Fubini's Theorem completes the proof.

[ch 7][↑]

7.4 The Ito's Integral

Now we want to define $\int_0^T X_t dY_t$, where $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are processes. Our main concern will be when $Y_t = f(W_t)$, where $(W_t)_{t\geq 0}$ is a Wiener process and $f : \mathbb{R} \mapsto \mathbb{R}$ is a real valued function. In this section we discuss the case $Y_t = W_t$.

We know that $(W_t)_{t\geq 0}$ is a continuous process but nowhere differentiable almost surely, so if we attempt, as when defining Riemann Stochastic Integration, to define $\int_0^T X_t dW_t$ as the limit of Riemann Sums i.e.

$$\int_0^T X_t dW_t = \lim_{n\to\infty} \sum_{i=0}^{k_n-1} X_{t_i^n}(\omega)(W_{t_{i+1}^n}(\omega) - W_{t_i^n}(\omega))$$

along different paths, we will run into difficulties as $(W_t)_{t\geq 0}$ lacks differentiability.

Definition 7.4.1 *Let $X := (X_t)_{t\geq 0}$ and $Y := (Y_t)_{t\geq 0}$ denote two continuous processes on the probability space $(\Omega, \mathcal{F}, P)$ both adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. If $0 \leq S \leq T$ and there exists a random variable that we denote by $\int_S^T X_t dY_t$ such that for any sequence of partitions of $[S,T]$, $\mathcal{P} := (t_i^{k_n})_{i=0}^{k_n}$,*

$$\sum_{i=0}^{k_n-1} X_{t_i^n} \cdot (Y_{t_{i+1}^n} - Y_{t_i^n}) \to \int_S^T X_t dY_t$$

in measure as mesh $(\mathcal{P}) \to 0$, then we say that X is Itô integrable with respect to Y and call $\int_S^T X_t dY_t$ the Itô integral of X with respect to Y.

It is important in the above Riemann Sums that the process being integrated is evaluated at the left endpoint of each subinterval in the partition of $[0,T]$. A different choice would lead to a different type of integral, as explained latter in Example 7.5.1. Since

$$\int_0^S X_t dW_t + \int_S^T X_t dW_t = \int_0^T X_t dW_t$$

we restrict our study to integrals over $[0,T]$.

Lemma 7.4.2 *Let $X := (X_t)_{t\geq 0} = \sum_{i=0}^{k-1} X_{t_i}\mathbb{I}_{[t_i,t_{i+1})}$ be a step process on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ adapted to filtration $(\mathcal{F}_t)_{t\geq 0}$ and let $0 < t_k < T$, then*

(a)

$$\int_0^T X_t dt = \sum_{i=0}^{k-1} X_{t_i}(W_{t_{i+1}} - W_{t_i});$$

(b) if $\mathbb{E}[|X_{t_i}|] < \infty$ *for all* i*, then* $\int_0^T X_t dW_t$ *is integrable and*

$$\mathbb{E}[\int_0^T X_t dW_t] = 0.$$

This property is known as Zero Mean Property.

(c) if $\mathbb{E}[X_{t_i}^2] < \infty$ *for all* i*, then*

$$\mathbb{E}[|\int_0^T X_t dW_t|^2] = \mathbb{E}[\int_0^T X_t^2 dt] = \int_0^T \mathbb{E}[X_t^2] dt.$$

This property is known as Isometry Property.

Proof (a) By linearity we may suppose $X = X_r \mathbb{I}_{[r,s)}$ where $0 \leq r < s \leq T$. Let $\mathcal{P}_n$ denote the partition of $[0,T)$ by $(t_i^n)_{i=0}^{k_n}$ and suppose $t_j^n < r \leq t_{j+1}^n \leq t_l^n < s \leq t_{l+1}^n$. The Riemann Sum for X with respect to $\mathcal{P}_n$ is

$$\sum_{i=j+1}^{l} X_r(W_{t_{i+1}^n} - W_{t_i^n}) = X_r(W_{t_{l+!}^n} - W_{t_{j+1}^n}).$$

Since $|r - t_{j+1}^n| \leq mesh(\mathcal{P}_n)$ and $|s - t_{l+1}^n| \leq mesh(\mathcal{P}_n)$, the Riemann sums converge, as $mesh(\mathcal{P}_n) \to 0$, along all $W - continuous$ paths and hence almost surely, to $X_r(W_s - W_r)$.

(b) Since $X := (X_t)_{t \geq 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$, X_{t_i} is $\mathcal{F}_{t_i} - measurable$. Hence X_{t_i} and $W_{t_{i+1}} - W_{t_i}$ are independent random variables and, hence $\mathbb{E}[X_{t_i} \cdot (W_{t_{i+1}} - W_{t_i})] = \mathbb{E}[X_{t_i}] \cdot \mathbb{E}[W_{t_{i+1}} - W_{t_i}] = 0$. Applying (a) completes the proof.

(c) Since $\mathbb{E}[(W_{t_{i+1}} - W_{t_i})^2] = t_{i+1} - t_i$ and $X_{t_i} X_{t_j} (W_{t_{i+1}} - W_{t_i})$ and $(W_{t_{j+1}} - W_{t_j})$

are independent random variables when $i < j$,

$$\begin{aligned}
\mathbb{E}[|\int_0^T X_t dW_t|^2] &= \sum_{i=0}^{k-1} \mathbb{E}[X_{t_i}^2 (W_{t_{i+1}} - W_{t_i})^2] \\
&+ \sum_{0 \le i \ne j \le k-1} \mathbb{E}[X_{t_i} X_{t_j} (W_{t_{i+1} - W_{t_i}})(W_{t_{j+1}} - W_{t_j})] \\
&= \sum_{i=0}^{k-1} \mathbb{E}[X_{t_i}^2] \cdot \mathbb{E}[(W_{t_{i+1}} - W_{t_i})^2] \\
&+ 2 \sum_{0 \le i < j \le k-1} \mathbb{E}[X_{t_i} X_{t_j} (W_{t_{i+1}} - W_{t_i})] \cdot \mathbb{E}[(W_{t_{j+1}} - W_{t_j})] \\
&= \sum_{i-0}^{k-1} \mathbb{E}[X_{t_i}^2] \cdot (t_{i+1} - t_i) \\
&= \int_0^T \mathbb{E}[X_t^2] dt \\
&= \mathbb{E}[\int_0^T X_t^2 dt], \text{ by Proposition 7.3.5.}
\end{aligned}$$

This complets the proof.

Proposition 7.4.3 *If $X := (X_t)_{t \ge 0}$ is a continuous stochastic process on the probability space $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ adapted to filtration $(\mathcal{F}_t)_{t \ge 0}$, then X is Itô integrable with respect to the Wiener process $W := (W_t)_{t \ge 0}$*

Proof We first suppose that the process is bounded, i.e. $\exists M > 0$ such that $|X_t(\omega) \le M|$ for all t, ω. For each n, let $\mathcal{P}_n$ denote the partition of $[0, T)$ into adjacent subintervals $(t_i^n)_{i=0}^{k_n}$ and suppose the *mesh*$(\mathcal{P}_n) \to 0$ as $n \to \infty$. If $X^n := \sum_{i=0}^{k_n - 1} X_{t_i^n} \cdot \mathbb{I}_{[t_i^n, t_{i+1})}$, then by Lemma 7.4.2(a),

$$\mathcal{R}(X, \mathcal{P}_n) := \sum_{i=0}^{k_n - 1} X_{t_i^n} \cdot (W_{t_{i+1}^n} - W_{t_i^n}) = \int_0^T X_t^n dW_t$$

is the Riemann Sum for X with respect to $\mathcal{P}_n$. By Lemma 7.4.2(c)

$$\left|\left| \int_0^T X_t^n dW_t - \int_o^T X_t^m dW_t \right|\right|_2^2 = \int_{\mathbb{R}^{[0,\infty)}} \int_0^T |X^n - X^m|^2 dt dW. \qquad (7.4.3.1)$$

For any pair (t, ω) and positive integers n and m

$$|X^n - X^m|^2(t, \omega) = |X_t^n - X_t^m|^2(\omega) = |X_{t_i^n}(\omega) - X_{t_j^m}(\omega)|^2 \leq 4M^2.$$

for some i and j where $|t - t_i^n| \leq |t_{i+1}^n - t_i^n|$ and $|t - t_j^m| \leq |t_{j+1}^m - t_j^m|$. Hence, if ω is $X - \textit{continuous}$ path, $|X^n - X^m|^2 \to 0$ almost surely as $n, m \to \infty$ on the product measure space $(R^{[0,\infty)} \times [0,T], \mathcal{F} \times \mathcal{B}_T, W \times m)$. Let $(n_j)_{j=1}^{\infty}$ be a strictly increasing sequence of positive integers. The Dominated Convergence Theorem and 7.4.3.1 imply $||X^{n_{j+1}} - X^{n_j}||_2 \to 0$ as $j \to \infty$ and, hence $(X^n)_{n=1}^{\infty}$ is an L^2 Cauchy, and hence convergent, sequence in the product measure space. By 7.4.3.1 and Proposition 7.4.5, X is Ito's integrable with respect to Wiener process. This completes the proof for any continuous bounded process and shows, moreover, that

$$\left|\left|\int_0^T X_t^n dW_t - \int_0^T X_t dW_t\right|\right|_2 \to 0 \text{ as } n \to \infty. \tag{7.4.3.2}$$

We now remove the boundedness hypothesis. For each positive integer m define a continuous function

$$\phi_m : \mathbb{R} \mapsto \mathbb{R}$$

by

$$\phi_m(x) = \begin{cases} x \text{ for } |x| \leq m \\ 0 \text{ for } |x| \geq m+1 \\ \text{linearly interpolated for } x \in [m, m+1] \end{cases}$$

Since ϕ_m is continuous, $(\phi_m(X_t))_{t\geq 0}$ is a continuous bounded, and hence integrable, process on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. For each m let Y_m denote a random variable on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ such that

$$\mathcal{R}(\phi_m(X), \mathcal{P}_n) \to Y_m$$

in measure as $n \to \infty$. Now choose inductively a sequence of measurable sets $(A_m)_{m=1}^{\infty}$, consisting of $X - \textit{continuous}$ paths such that $W(A_m) = 1$ for all m and a sequence of subsequences of $(\mathcal{P}_n)_{n=1}^{\infty}$, $(\mathcal{P}_{(n,m)})_{n=1}^{\infty}$, $m = 1, 2, \ldots$, such that for all integers m:

(a) $(\mathcal{P}_{(n,m+1)})_{n=1}^{\infty}$ is a subsequence of $(\mathcal{P}_{(n,m)})_{n=1}^{\infty}$,

(b) $\mathcal{R}(\phi_m(X), \mathcal{P}_{(n,m)}) \to Y_m$ pointwise on A_m as $n \to \infty$.

Let $A := \bigcap_{m=1}^{\infty} A_m$. Then

$$W(A^c) = W(\bigcup_{m=1}^{\infty} A_m^c) \leq \sum_{m=1}^{\infty} W(A_m^c) = 0$$

and $W(A) = 1$. Fix $\omega \in A$. By path continuity there exists an integer l such that $|X_t(\omega)| \leq l$ for all $t \in [0, T]$. For $m > k > l$, $\phi_m(X)(\omega) = X(\omega) = \phi_k(X)(\omega)$ on $[0, T]$. Hence

$$\mathcal{R}(\phi_m(X), \mathcal{P}_n)(\omega) = \mathcal{R}(X, \mathcal{P}_n)(\omega) = \mathcal{R}(\phi_k(X), \mathcal{P}_n)(\omega)$$

for all n and

$$\begin{aligned} Y_m(\omega) &= \lim_{n\to\infty} \mathcal{R}(\phi_m(X), \mathcal{P}_{(n,m)})(\omega) \\ &= \lim_{n\to\infty} \mathcal{R}(\phi_k(X), \mathcal{P}_{(n,k)})(\omega) \\ &= Y_k(\omega) \end{aligned}$$

and the sequence $(Y_m)_{m=1}^{\infty}$ converges pointwise to a finite limit on A. Since $(\mathcal{P}_{(n,n)})_{n=m}^{\infty}$ is a subsequence of $(\mathcal{P}_{(n,m)})_{n=1}^{\infty}$ for all m and

$$\lim_{n\to\infty} \mathcal{R}(X, \mathcal{P}_{(n,n)}) = \lim_{m\to\infty} Y_m$$

almost surely, this means that any sequence of Riemann Sums for $(X_t)_{t\geq 0}$ contains a subsequence which converges almost surely to $\lim_{m\to\infty} Y_m$ and, as $\lim_{m\to\infty} Y_m$ is the same for all choices of Riemann Sums, the Riemann Sums converge in measure and $\int_0^T X_t dW_t$ exists, completing the proof.

Remark: The above proof shows that

$$\int_0^T \phi_m(X_t)dW_t = \int_0^T (\phi_m(X))_t d_t \to \int_0^T X_t dW_t \qquad (7.4.3.3)$$

almost surely as $m \to \infty$.

Now lets discuss the properties of the process $(\int_0^T X_t dW_t)_{T\geq 0}$. If $X := (X_t)_{t\geq 0}$ is a continuous process, then it does not follow that the process $(\int_0^T X_t dW_t)_{T\geq 0}$ is continuous. This inconvenience can be dealt by changing the process slightly so that continuity is achieved and all essential features preserved. A *modification* of a stochastic process $X := (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ is a stochastic process adapted to the same filtration, for which

$$P\{\omega \in \Omega : X_t(\omega) = Y_t(\omega)\} = 1$$

for all $t \geq 0$. If $X := (X_t)_{t\geq 0}$ is a continuous process, then $(\int_0^T X_t dW_t)_{T\geq 0}$ admits a continuous modification.

Proposition 7.4.4 *Let* $X := (X_t)_{t\geq 0}$ *be a continuous process on* $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ *adapted to filtration* $(\mathcal{F}_t)_{t\geq 0}$*. If* $\int_0^T \mathbb{E}[|X_t|^2]dt < \infty$*, then*

$$\mathbb{E}[|\int_0^T X_t dW_t|^2] = \int_0^T \mathbb{E}[|X_t|^2]dt = \mathbb{E}[\int_0^T |X_t|^2 dt]$$

and there exists a sequence of step processes $(X_t^n)_{t\geq 0}$ *with* $\mathbb{E}[|X_t^n|^2] < \infty$ *for all* n *and* t *such that*

$$\left|\left|\int_0^T X_t^n dW_t - \int_0^T X_t dW_t\right|\right|_2 \to 0 \text{ as } n \to \infty.$$

Proof By Fubini's Theorem

$$\int_0^T \mathbb{E}[|X_t|^2]dt = \int_{\mathbb{R}^{[0,\infty)}} \int_0^T |X|^2 dt dW = \mathbb{E}[\int_0^T |X_t|^2 dt] \qquad (7.4.4.1)$$

First suppose the process is bounded. By 7.4.3.2 there is a sequence of Riemann Sums $(\int_0^T X_t^{n_j} dW_t)_{j=1}^\infty$ which converges to $\int_0^T X_t dW_t$ in $L^2(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ as $j \to \infty$.

If $|X_t(\omega)| \leq M$ for all t and all paths ω, then $|X_t^{n_j}(\omega)| \leq M$ for all j and all ω. Since $|X_t^{n_j}(\omega) - X_t(\omega)| \to 0$ almost surely as $j \to \infty$ the Dominated Convergence Theorem and Lemma 7.4.2(c) imply

$$\begin{aligned}
\mathbb{E}[|\int_0^T X_t dW_t|^2] &= \lim_{j\to\infty} \mathbb{E}[|\int_0^T X_t^{n_j} dW_t|^2] \\
&= \lim_{j\to\infty} \mathbb{E}[\int_0^T |X_t^{n_j}|^2 dt] \\
&= \lim_{j\to\infty} \int_{\mathbb{R}^{[0,\infty)}} \int_0^T |X^{n_j}|^2 dt dW \\
&= \int_{\mathbb{R}^{[0,\infty)}} \int_0^T |X|^2 dt dW.
\end{aligned}$$

By 7.4.4.1 this establishes the proposition for a continuous bounded process.

For an arbitrary process $(X_t)_{t\geq 0}$ let ϕ_m denote the function previously defined. By 7.4.3.3, $\int_0^T \phi_m(X_t) dW_t \to \int_0^T X_t dW_t$ almost surely as $m \to \infty$. Since $|(\phi_m(X_t))(\omega)| \leq |X_t(\omega)|$ and $|\phi_m(X_t) - \phi_n(X_t)| \to 0$ almost surely as $m, n \to \infty$, 7.4.4.1 and the Dominated Convergence Theorem imply

$$\begin{aligned}
\left|\left|\int_0^T \phi_m(X_t) dW_t - \int_0^T \phi_n(X_t) dW_t\right|\right|_2 &= \int_{\mathbb{R}^{[0,\infty)}} \int_0^T |\phi_m(X) - \phi_n(X)|^2 dt dW \\
&\to 0 \text{ as } n, m \to \infty.
\end{aligned}$$

By Proposition 7.1.4, the sequence $(\int_0^T \phi_m(X_t) dW_t)_{m=1}^\infty$ converges to some $Y \in L^2(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$, and, as it also converges almost surely to $\int_0^T X_t dW_t$, we have $\int_0^T X_t dW_t \in L^2(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$. Hence

$$\begin{aligned}
\mathbb{E}[|\int_0^T X_t dW_t|^2] &= \lim_{m\to\infty} \mathbb{E}[|\int_0^T \phi_m(X_t) dW_t|^2] \qquad (7.4.4.2) \\
&= \lim_{m\to\infty} \int_{\mathbb{R}^{[0,\infty)}} \int_0^T |\phi_m(X)|^2 dt dW \\
&= \int_{\mathbb{R}^{[0,\infty)}} \int_0^T |X|^2 dt dW.
\end{aligned}$$

By combining 7.4.4.1 and 7.4.4.2, the proof is completed.

Proposition 7.4.5 *Let* $X := (X_t)_{t\geq 0}$ *denote a continuous stochastic process on* $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ *adapted to filtration* $(\mathcal{F}_t)_{t\geq 0}$*. If* $\int_0^T \mathbb{E}[|X_t|^2]dt < \infty$ *for all* $T > 0$*, then* $(\int_0^T X_t dW_t)_{T\geq 0}$ *is a martingale.*

Proof Since $\int_0^T X_t dW_t$ is an almost sure limit of Riemann Sums which are $\mathcal{F}_t - measurable$, it is $\mathcal{F}_t - measurable$ and, by Proposition 7.1.3(c) and 7.4.4.1, it is integrable. It remains to show, for $0 \leq S \leq T$, that

$$\mathbb{E}[\int_0^T X_t dW_t|\mathcal{F}_S] = \int_0^S X_t dW_t.$$

By linearity

$$\mathbb{E}[\int_0^T X_t dW_t|\mathcal{F}_S] = \mathbb{E}[\int_0^S X_t dW_t|\mathcal{F}_S] + \mathbb{E}[\int_S^T X_t dW_t|\mathcal{F}_S].$$

Since $\int_0^S X_t dW_t$ is $\mathcal{F}_S - measurable$ we obtain, on taking out what is known, $\mathbb{E}[\int_0^S X_t dW_t|\mathcal{F}_S] = \int_0^S X_t dW_t$.

Let $(X_t^n)_{t\geq 0}$ denote the sequence of step processes given in Proposition 7.4.4. If $t_i^n \geq S$ then

$$\begin{aligned}
\mathbb{E}[X_{t_i^n}(W_{t_{i+1}^n} - W_{t_i^n})|\mathcal{F}_S] &= \mathbb{E}[\mathbb{E}[X_{t_i^n}(W_{t_{i+1}^n} - W_{t_i^n})|\mathcal{F}_{t_i^n}]|\mathcal{F}_S] \\
&= \mathbb{E}[X_{t_i^n}\mathbb{E}[(W_{t_{i+1}^n} - W_{t_i^n})|\mathcal{F}_{t_i^n}]|\mathcal{F}_S] \\
&= \mathbb{E}[X_{t_i^n}\mathbb{E}[W_{t_{i+1}^n} - W_{t_i^n}]|\mathcal{F}_S] \\
&= \mathbb{E}[W_{t_{i+1}^n} - W_{t_i^n}] \cdot \mathbb{E}[X_{t_i^n}|\mathcal{F}_S] \\
&= 0.
\end{aligned}$$

By linearity, $\mathbb{E}[\int_S^T X_t^n dW_t|\mathcal{F}_S] = 0$. If $\int_0^T \mathbb{E}[|X_t|^2]dt < \infty$, Proposition 7.4.4 and

7.1.3(c), imply

$$\begin{aligned}
\mathbb{E}[|\mathbb{E}[\int_S^T X_t dW_t|\mathcal{F}_S]|] &= \mathbb{E}[|\mathbb{E}[\int_S^T (X_t - X_t^n)dW_t|\mathcal{F}_S]|] \\
&\leq \mathbb{E}[|\int_S^T (X_t - X_t^n)dW_t|] \\
&= \left|\left|\int_S^T (X_t - X_t^n)dW_t\right|\right|_1 \\
&\leq \left|\left|\int_S^T (X_t - X_t^n)dW_t\right|\right|_2 \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}$$

Hence $\mathbb{E}[\int_S^T X_t dW_t|\mathcal{F}_S] = 0$ almost surely, and this completes the proof.

[ch 7][↑]

7.5 Itô's Process and Itô's Lemma

Now we have defined the integral $\int_0^T X_t dW_t$ for a continuous process $X := (X_t)_{t\geq 0}$, but our main interest lies in evaluating $\int_0^T X_t df(t, W_t)$ where $f(t,s) = ce^{\mu t + \sigma s}$, since the share price is of this form. But as a first step in analyzing this integral consider $\int_0^T X_t df(W_t)$ where $f : \mathbb{R} \mapsto \mathbb{R}$ is sufficiently regular. If f has continuous first derivative then the non-random analogue considered earlier

$$\int_0^T g(x)df(x) = \int_0^T g(x)f'(x)dx \tag{7.5.0.1}$$

suggest that $\int_0^T X_t f'(W_t)dW_t$ will appear in the answer.

Now suppose $f : \mathbb{R} \mapsto \mathbb{R}$ has continuous first and second derivatives then the non-random analogue transformed to

$$\int_0^T g(x)df(x) = \int_0^T g(x)f'(x)dx + \frac{1}{2}\int_0^T g(x)f''(x)d^2x. \tag{7.5.0.2}$$

Equivalently the differential form is

$$df(x) = f'(x)dx + \frac{1}{2}f''(x)d^2x$$

which can be considered as the second order Taylor expansion. Thus the above suggest

$$\int_0^T X_t df(W_t) = \int_0^T X_t f'(W_t) dW_t + \frac{1}{2}\int_0^T X_t f''(W_t) d^2 W_t.$$

While proving 7.5.0.1 and 7.5.0.2 we have used *Mean Value Theorem* and hence it's obvious that we need an extended version of it. And in 7.5.0.2 we encountered $\int_0^T d^2x$ and hence we need to define $\int_0^T d^2 W_t$.

We maintain our previous notation and for each positive integer n let $(t_i^n)_{i=0}^{k_n}$ denote the set of points which partitions $[0,T)$ into k_n adjacent sub-intervals. The points t_{i*}^n and t_{i**}^n, in the following proof, are arbitrary chosen to satisfy $t_i^n \leq t_{i*}^n$ and $t_{i**}^n < t_{i+1}^n$.

Proposition 7.5.1 *Let* $X := (X_t)_{t\geq 0}$ *and* $Y := (Y_t)_{t\geq 0}$ *denote continuous process on* $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$*, adapted to the filtration* $(\mathcal{F}_t)_{t\geq 0}$*. Then*

$$\sum_{i=0}^{k_n-1} X_{t_{i*}^n} Y_{t_{i**}^n} (W_{t_{i+1}^n} - W_{t_i^n})^2 \to \int_0^T X_t Y_t dt \tag{7.5.1.1}$$

in measure as $\sup_{0\leq i\leq k_n} |t_{i+1} - t_i^n| \to 0$

Proof We know that the quadratic variation of Wiener process in any interval is the difference between the end points of that interval, i.e.

$$S_n := \sum_{i=0}^{k_n-1} |(W_{t_{i+1}^n}(\omega) - W_{t_i^n}(\omega))^2 - (t_{i+1}^n - t_i^n)| \to 0 \tag{7.5.1.2}$$

in measure as $\sup_{0\leq i\leq k_n} |t_{i+1} - t_i^n| \to 0$. Now if ω is a continuous path for both $X := (X_t)_{t\geq 0}$ and $Y := (Y_t)_{t\geq 0}$, then there exists M_ω such that $|X_t(\omega)| \leq M_\omega$ and $|Y_t(\omega) \leq M_\omega|$ for all $t \in [0,T]$. Hence

$$\begin{aligned} T_n(\omega) &:= \left| \sum_{i=0}^{k_n-1} X_{t_{i*}^n}(\omega) Y_{t_{i**}^n}(\omega)((W_{t_{i+1}^n}(\omega) - W_{t_i^n}(\omega))^2 - (t_{i+1}^n - t_i^n)) \right| \\ &\leq M_\omega^2 \sum_{i=0}^{k_n-1} |(W_{t_{i+1}^n}(\omega) - W_{t_i^n}(\omega))^2 - (t_{i+1}^n - t_i^n)| \\ &= M_\omega^2 |S_n(\omega)|. \end{aligned}$$

By 7.5.1.2, every subsequence $(S_{n_j})_{j=1}^{\infty}$ of $(S_n)_{n=1}^{\infty}$ contains a subsequence $(S_{n_{j_k}})_{k=1}^{\infty}$ which converges almost surely to 0. This implies $(T_{n_{j_k}})_{k=1}^{\infty}$ converges almost surely to 0. Hence

$$\sum_{i=0}^{k_n-1} X_{t^n_{i*}} Y_{t^n_{i**}}((W_{t^n_{i+1}}(\omega) - W_{t^n_i}(\omega))^2 - (t^n_{i+1} - t^n_i)) \to 0$$

in measure as $\sup_{0\le i<k_n}|t^n_{i+1} - t^n_i| \to 0$, completing the proof.

Example 7.5.1

Now lets use the above lemma to show the difference between evaluating the process at the left and right end points of intervals in Riemann Sums. When $X := (X_t)_{t\ge 0} = W := (W_t)_{t\ge 0}$ in the notation of Proposition 7.5.1 is

$$\sum_{i=0}^{k_n-1} W_{t^n_{i+1}}(W_{t^n_{i+1}} - W_{t^n_i}) - \sum_{i=0}^{k_n-1} W_{t^n_i}(W_{t^n_{i+1}} - W_{t^n_i}) = \sum_{i=0}^{k_n-1}(W_{t^n_{i+1}} - W_{t^n_i})^2$$

and, by 7.5.1.1, this converges in measure to T. This shows that

$$\sum_{i=0}^{k_n-1} W_{t^n_{i+1}}(W_{t^n_{i+1}} - W_{t^n_i}) \nrightarrow \int_0^T W_t dW_t$$

in measure as $\sup_{0\le i<k_n}|t^n_{i+1} - t^n_i| \to 0$.

Now suppose $f : \mathbb{R} \mapsto \mathbb{R}$ has continuous first and second derivatives. And let ω denote a $W-$*continuous* path and let $0 \le i \le k_n$. Then on applying the extended *Mean-Value Theorem* to the function on the interval $[W_{t^n_i}(\omega), W_{t^n_{i+1}}(\omega)]$ we obtain $\alpha \in [W_{t^n_i}(\omega), W_{t^n_{i+1}}(\omega)]$ such that

$$\begin{aligned} f(W_{t^n_{i+1}}(\omega)) - f(W_{t^n_i}(\omega)) &= f'(W_{t^n_i}(\omega)) \cdot (W_{t^n_{i+1}}(\omega) - W_{t^n_i}(\omega)) \\ &+ \frac{1}{2} f''(\alpha) \cdot (W_{t^n_{i+1}}(\omega) - W_{t^n_i}(\omega))^2. \end{aligned}$$

Since the function $t \to W_t(\omega)$ is continuous on $[t^n_i, t^n_{i+1}]$, the *Intermediate Value Theorem* implies that there exists $t^*_i \in [t^n_i, t^n_{i+1}]$, which depends on ω, i, n, such that $\alpha = W_{t^*_i}(\omega)$.

Let $X := (X_t)_{t\geq 0}$ denote a continuous process on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$, adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and A denote an $\mathcal{F}_\infty - measurable$ subset of $\mathbb{R}^{[0,\infty)}$, with $W(A) = 1$, consisting of paths which are both X and W continuous. For all $\omega \in A$ we have

$$\sum_{i=0}^{k_n-1} X_{t_i^n}(\omega)\cdot(f(W_{t_{i+1}^n}(\omega)) - f(W_{t_i^n}(\omega))) \tag{7.5.1.3}$$

$$= \sum_{i=0}^{k_n-1} X_{t_i^n}(\omega)\cdot f'(W_{t_i^n}(\omega))(W_{t_{i+1}^n}(\omega) - W_{t_i^n}(\omega)) \tag{7.5.1.4}$$

$$+\frac{1}{2}\sum_{i=0}^{k_n-1} X_{t_i^n}(\omega)\cdot f''(W_{t_i^*}(\omega))(W_{t_{i+1}^n}(\omega) - W_{t_i^n}(\omega))^2. \tag{7.5.1.5}$$

By Proposition 7.4.3 and 7.5.1 the series in 7.5.1.4 and 7.5.1.5 converges in measure to $\int_0^T X_t f'(W_t)dW_t$ and $\frac{1}{2}\int_0^T X_t f''(W_t)dt$, respectively, as $\sup_{0\leq i<k_n}|t_{i+1}^n - t_i^n| \to 0$. Hence the sequence in 7.5.1.3 converges in measure as $\sup_{0\leq i<k_n}|t_{i+1}^n - t_i^n| \to 0$ and $\int_0^T X_t df(W_t)$ is well defined. Now we have the following stochastic version of the *Fundamental Theorem of Calculus.*

Proposition 7.5.2 *If $f : \mathbb{R} \to \mathbb{R}$ has continuous first and second derivatives and $X := (X_t)_{t\geq 0}$ is a continuous process on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, then*

$$\int_0^T X_t df(W_t) = \int_0^T X_t f'(W_t)dW_t + \frac{1}{2}\int_0^T X_t f''(W_t)dt \tag{7.5.2.1}$$

The above result is a special case of Itô's Lemma. Now we introduce some *symbolic differential notation* as

$$(dW_t)^2 = dt,\ (dt)^2 = 0 \text{ and } dt\cdot dW_t = dW_t\cdot dt = 0.$$

The first one symbolic represents the quadratic variation of Wiener process and the second and third is the symbolic representation of the limits of the following Riemann sums, respectively

$$\sum_{i=0}^{k_n-1}(t^n_{i+1}-t^n_i)^2 \text{ and } \sum_{i=0}^{k_n-1}(t^n_{i+1}-t^n_i)\cdot(W_{t^n_{i+1}}-W_{t^n_i}).$$

Hence the integral equation in proposition 7.5.2 can be re-written in differential form as

$$df(W_t) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt. \tag{7.5.2.2}$$

We now integrate with respect to the process $(g(t)f(W_t))_{t\geq 0}$ where $f, g : \mathbb{R} \to \mathbb{R}$. Then

$$\begin{aligned}
&g(t^n_{i+1})f(W_{t^n_{i+1}}) - g(t^n_i)f(W_{t^n_i})\\
&= (g(t^n_{i+1})f(W_{t^n_{i+1}}) - g(t^n_i)f(W_{t^n_{i+1}})) + (g(t^n_i)f(W_{t^n_{i+1}}) - g(t^n_i)f(W_{t^n_i}))\\
&= f(W_{t^n_{i+1}})(g(t^n_{i+1}) - g(t^n_i)) + g(t^n_i)(f(W_{t^n_{i+1}}) - f(W_{t^n_i})).
\end{aligned}$$

If g is continuously differentiable and f is twice continuously differentiable then, by Proposition 7.5.2, the Riemann sums for the continuous adapted process $X := (X_t)_{t\geq 0}$ with respect to $(g(t)f(W_t))_{t\geq 0}$ converges in measure as we take finer and finer partitions and we obtain:

$$\begin{aligned}
d(g(t)f(W_t)) &= f(W_t)g'(t)dt + g(t)df(W_t)\\
&= f(W_t)g'(t)dt + g(t)f'(W_t)dW_t + \frac{1}{2}g(t)f''(W_t)dt.
\end{aligned}$$

And hence the following proposition.

Proposition 7.5.3 *If $f, g : \mathbb{R} \to \mathbb{R}$ where g is continuously differentiable and f is twice continuously differentiable, then for any continuous process $X := (X_t)_{t\geq 0}$ on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ adapted to filtration $(\mathcal{F}_t)_{t\geq 0}$*

$$\begin{aligned}
\int_0^T X_t d(g(t)f(W_t)) &= \int_0^T X_t(f(W_t)g'(t) + \frac{1}{2}g(t)f''(W_t))dt\\
&\quad + \int_0^T X_t g(t)f'(W_t)dW_t
\end{aligned}$$

Example 7.5.2

(a) If $X := (X_t)_{t\geq 0}$ is any continuous process, then from the Riemann sums,

$$\int_0^T dX_t = X_T - X_0.$$

(b) Let $f(t) = t^2$ for all $t \in \mathbb{R}$. Since $f'(t) = 2t$ and $f''(t) = 2$, Proposition 7.5.2 and (a) imply

$$W_T^2 = \int_0^T d(W_t^2) = 2\int_0^T W_t dW_t + \frac{1}{2}\int_0^T 2dt = 2\int_0^T W_t dW_t + T.$$

By Proposition 7.4.5, $(W_t^2 - t)_{t\geq 0}$ is a martingale, and we recover the martingale property of Wiener process.

(c) In the previous chapter we showed, under reasonable assumptions, that the share price X_t satisfied

$$X_t = X_0 e^{\mu t + \sigma W_t}$$

for all $t \geq 0$. We now apply Proposition 7.5.3 with $f(t) = X_0 e^{\sigma t}$ and $g(t) = e^{\mu t}$ and obtain, using above symbolic notation,

$$\begin{aligned}
dX_t &= g'(t)f(W_t)dt + g(t)f'(W_t)dW_t + \frac{1}{2}g(t)f''(W_t)dt \\
&= \mu e^{\mu t} X_0 e^{\sigma W_t} dt + e^{\mu t}\sigma X_0 e^{\sigma W_t} dW_t + \frac{1}{2}e^{\mu t}\sigma^2 X_0 e^{\sigma W_t} dt \\
&= (\mu + \frac{1}{2}\sigma^2) X_0 e^{\mu t + \sigma W_t} dt + \sigma X_0 e^{\mu t + \sigma W_t} dW_t \\
&= (\mu + \frac{1}{2}\sigma^2) X_t dt + \sigma X_t dW_t.
\end{aligned}$$

Hence

$$X_T - X_0 = \int_0^T dX_t = \int_0^T (\mu + \frac{1}{2}\sigma^2) X_t dt + \int_0^T \sigma X_t dW_t. \qquad (7.5.3.1)$$

By Proposition 7.4.5, $X := (X_t)_{t\geq 0}$ is a martingale if $\mu = -\frac{\sigma^2}{2}$, same result proved earlier in other way. Hence if $X := (X_t)_{t\geq 0}$ can be turned into a martingale, by change of measure, then by 7.5.3.1 and (a) we obtain

$$\int_0^T dX_t = X_T - X_0 = \int_0^T \sigma X_t d\tilde{W}_t$$

where $(\tilde{W}_t)_{t\geq 0}$ is a Wiener process. Symbolically

$$dX_t = \sigma X_t d\tilde{W}_t.$$

(d) Let $X_t = \sigma e^{-\alpha t}\int_0^t e^{\alpha s} dW_s$. The process $(X_t)_{t\geq 0}$ is called an *Ornstein-Uhlenbeck Process*. If $Y_t := \int_0^t e^{\alpha s} dW_s$; then $dY_t = e^{\alpha t} dW_t$, and if $g(t) = \sigma e^{-\alpha t}$, then $X_t = g(t)Y_t$. Hence

$$\begin{aligned} dX_t &= g'(t)Y_t dt + g(t)dY_t \\ &= -\alpha\sigma e^{-\alpha t} Y_t dt + \sigma e^{-\alpha t} e^{\alpha t} dW_t \\ &= -\alpha X_t dt + \sigma dW_t \end{aligned}$$

and on using (a),

$$X_T = -\alpha \int_0^T X_t dt + \sigma \int_o^T dW_t \qquad (7.5.3.2)$$

$$= -\alpha \int_0^T X_t dt + \sigma W_T \qquad (7.5.3.3)$$

Equations 7.5.3.1 and 7.5.3.2 are stochastic integral equations. In the above examples we started with a process and showed that it satisfied a certain stochastic integral equation. Stochastic process which admit a special integral representation are called Itô's Processes.

Definition 7.5.4 *A stochastic process $X := (X_t)_{t\geq 0}$ is called an Itô's process if there exists continuous process $Y = (Y_t)_{t\geq 0}$ and $Z = (Z_t)_{t\geq 0}$ on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ adapted to* *the filtration* $(\mathcal{F}_t)_{t\geq 0}$ *such that for all* $t \geq 0$, $\int_0^t |Y_s|ds < \infty$, $\int_0^t \mathbb{E}[Z_s^2]ds < \infty$, *and*

$$X_t - X_0 = \int_0^t Y_s ds + \int_0^t Z_s dW_s. \tag{7.5.4.1}$$

Informally written as

$$dX_t = Y_t dt + Z_t dW_t.$$

Itô's processes have a particularly useful *uniqueness* property, i.e. if

$$\int_0^t Y_s ds + \int_0^t Z_s dW_s = \int_0^t U_s ds + \int_0^t V_s dW_s$$

for all t, then $Y_t = U_t$ and $Z_t = V_t$ almost surely as random variables for all t.

The following is the most important version of Itô's lemma which gives an explicit formula for $(u(t, X_t))_{t\geq 0}$, as an Itô's process, when $X := (X_t)_{t\geq 0}$ is an Itô's process and $u : \mathbb{R}^2 \mapsto \mathbb{R}$ has continuous first and second order partial derivatives. If u is a function of (t, s), let $u_1 = \frac{\partial u}{\partial t}$, $u_2 = \frac{\partial u}{\partial s}$, $u_{11} = \frac{\partial^2 u}{\partial t^2}$, $u_{12} = \frac{\partial^2 u}{\partial t \partial s}$, $u_{22} = \frac{\partial^2 u}{\partial s^2}$, etc.

Proposition 7.5.5 *(Itô's Lemma) Let* $u : \mathbb{R}^2 \mapsto \mathbb{R}$ *have first and second order partial derivatives and let* $X := (X_t)_{t\geq 0}$ *denote an Itô's process with representation (7.5.4.1). A modified version of* $(u(t, X_t))_{t\geq 0}$ *is an Itô's process with representation*

$$\begin{aligned} u(t, X_t) - u(0, X_0) &= \int_0^t (u_1(s, X_s) + u_2(s, Y_s)Y_s + \frac{1}{2}u_{22}(s, X_s)Z_s^2)ds \\ &+ \int_0^t u_2(s, X_s)Z_s dW_s. \end{aligned}$$

Proof We derive the formula *symbolically* using Tylor series expansion in two variables:

$$\begin{aligned} u(t + \Delta t, s + \Delta s) - u(t, s) &\approx u_1(t, s)\Delta t + u_2(t, s)\Delta s + \frac{1}{2}u_{11}(t, s)(\Delta t)^2 \\ &+ u_{12}(t, s)\Delta t \Delta s + \frac{1}{2}u_{22}(t, s)(\Delta s)^2. \end{aligned}$$

Since $dX_t = Y_t dt + Z_t W_t$ we have

$$\begin{aligned} dt \cdot X_t &= dt(Y_t dt + Z_t dW_t) \\ &= Y_t (dt)^2 + Z_t (dt \cdot W_t) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} (dX_t)^2 &= (Y_t dt + Z_t W_t)^2 \\ &= Y_t^2 (dt)^2 + 2Y_t Z_t (dt \cdot W_t) + Z_t^2 (dW_t)^2 \\ &= Z_t^2 dt \end{aligned}$$

Using the Tylor series expansion we obtain

$$\begin{aligned} du(t, X_t) &= u_1(t, X_t)dt + u_2(t, X_t)dX_t + \frac{1}{2}u_{11}(t, X_t)(dt)^2 + u_{12}(t, X_t)dt \cdot dX_t \\ &+ \frac{1}{2}u_{22}(t, X_t)(dX_t)^2 \\ &= u_1(t, X_t)dt + u_2(t, X_t)(Y_t dt + Z_t dW_t) + \frac{1}{2}u_{22}(t, X_t)Z_t^2 dt \\ &= \left(u_1(t, X_t) + u_2(t, X_t)Y_t + \frac{1}{2}u_{22}(t, X_t)Z_t^2\right) dt + u_2(t, X_t)Z_t^2 dW_t. \end{aligned}$$

[ch 7][↑]

7.6 Call Options 4

We now return to the problem of hedging any claim on a call option. We recall that

$$X_t := X_0 e^{\mu t + \sigma W_t}$$

denotes the share price at times $t \geq 0$, r is the interest rate, k is the strike price and T is the maturity date on the option. For convenience we let $c := \mu + \frac{1}{2}\sigma^2$. If V_t is the value of the option at time t, then $V_T = (X_T - k)^+$. To hedge any claim on the option, taking for one share, a portfolio of θ_t shares and β_t units of a riskless bond are held at time t. We suppose that one unit of the bond is worth $B(t) = e^{rt}$ at time t. Hence $dB(t) = B'(t)dt = re^{rt}dt$. This provides a hedge if

$$V_t = \theta_t X_t + e^{rt}\beta_t \tag{7.6.0.1}$$

for all t, $0 \leq t \leq T$. By 3.3.0.17

$$\int_0^t \theta_s dX_s + \int_0^t re^{rs}\beta_s ds = \int_0^t dV_s = V_t - V_0 \tag{7.6.0.2}$$

and hence, by Example 7.5.2(c),

$$V_t - V_0 = \int_0^t (c\theta_s X_s + re^{rs}\beta_s)ds + \int_0^t \sigma\theta_s X_s dW_s \tag{7.6.0.3}$$

for $0 \leq t \leq T$. By 7.6.0.1, $\beta_s = e^{-rs}(V_s - \theta_s X_s)$, and substituting this into 7.6.0.3 we obtain

$$\begin{aligned} V_t - V_0 &= \int_0^t (c\theta_s X_s + re^{rs}(e^{-rs}(V_s - \theta_s X_s)))ds + \int_0^t \sigma\theta_s X_s dW_s \\ &= \int_0^t ((c-r)\theta_s X_s + rV_s)ds + \int_0^t \sigma\theta_s X_s dW_s. \end{aligned}$$

Now suppose $V_t = u(t, X_t)$ where $u : \mathbb{R}^2 \mapsto \mathbb{R}$ is twice continuously differentiable. To determine the hedging strategy we must find u. The seller's price will be then $u(0, X_0)$. To complete our programme we must show that this coincides with the price given in Proposition 6.3.1. By Itô's Lemma

$$\begin{aligned} V_t - V_0 &= \int_0^t (u_1(s, X_s) + cX_s u_2(s, X_s) + \frac{\sigma^2}{2} X_s^2 u_{22}(s, X_s))ds \\ &+ \int_0^t \sigma X_s u_2(s, X_s) dW_s. \end{aligned}$$

We now have two Itô's representations for the process $(V_t)_{t\geq 0}$. By uniqueness, $\sigma X_t \theta_t = \sigma X_t u_2(t, X_t)$, and since $X_t > 0$ for all t, this implies $\theta_t = u_2(t, X_t)$. Moreover,

$$\begin{aligned} (c-r)\theta_t X_t + rV_t &= (c-r)X_t u_2(t, X_t) + ru(t, X_t) \\ &= u_1(t, X_t) + cX_t u_2(t, X_t) + \frac{\sigma^2}{2} X_t^2 u_{22}(t, X_t) \end{aligned}$$

and hence

$$ru(t, X_t) = u_1(t, X_t) + rX_t u_2(t, X_t) + \frac{\sigma^2}{2} X_t^2 u_{22}(t, X_t) \tag{7.6.0.4}$$

almost surely as random variables. Given any positive real number x and any $t > 0$ we can choose a path ω such that $X_t(\omega) = x$. Evaluating 7.6.0.4 at ω we get

$$ru(t, x) = u_1(t, x) + rxu_2(t, x) + \frac{\sigma^2}{2} x^2 u_{22}(t, x). \tag{7.6.0.5}$$

This is a *partial differential equation.*[1] The conditions on the option mean that our solution to 7.6.0.5 must satisfy the boundary condition $u(T, x) = (x - k)^+$. It is generally quite difficult to obtain explicit solutions for partial differential equations. In this particular case we know, from Proposition 6.3.1, the only solution that will achieve our final goal of hedging any claim on the option. It is thus a matter of verifying that it satisfies 7.6.0.5. Let $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$ and $n(x) := N'(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$

Proposition 7.6.1 *The portfolio consisting of*

$$\theta_t = N\left(\frac{\log(\frac{X_t}{k}) + (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}\right) \tag{7.6.1.1}$$

shares and

$$\beta_t = -ke^{-rT} N\left(\frac{\log(\frac{X_t}{k}) + (r - \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}\right) \tag{7.6.1.2}$$

riskless bonds at time t, $0 \le t \le T$, hedges any claim on an option for one share with strike price k, maturity date T, given that $X_t = X_0 e^{\mu t + \sigma W_t}$ is the share price at time t, the interest rate r is fixed for the duration of the option and σ, the volatility, is constant. Moreover, the value V_t of the option at time t is given by

$$V_t = \theta_t X_t + \beta_t e^{rt}.$$

[1]Equation 7.6.0.5 does not involve the drift μ and hence the solution will also will also be independent of the drift. If we had known this when we obtained Proposition 6.1.1 and 6.1.3, we could have let $\mu = -\frac{\sigma^2}{2}$ and obtained the required martingale immediately. Equation 7.6.0.4 is sometimes called the *Black-Scholes equation.*

Proof We need to show that $u(t,x)$ given by

$$xN\left(\frac{\log(\frac{x}{k})+(r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)-ke^{-r(T-t)}N\left(\frac{\log(\frac{x}{k})+(r-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)$$

solves 7.6.0.5 and satisfies the above boundary condition. If $x>k$ and $t<T$, then

$$\lim_{t\to T}N\left(\frac{\log(\frac{x}{k})+(r\pm\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)=N(+\infty)=1$$

while if $x<k$ and $t<T$, then

$$\lim_{t\to T}N\left(\frac{\log(\frac{x}{k})+(r\pm\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)=N(-\infty)=0.$$

This implies

$$u(T,x)=\lim_{t\to T}u(t,x)=\begin{cases}x-k, & \text{if } x\geq k\\ 0, & \text{if } x\leq k\end{cases}$$

thus $u(T,x)=(x-k)^+$ and u satisfies the required boundary condition at $t=T$. Moreover, when $t=0$ we recover the Black-Scholes formula.

Now we to verify that u is a solution for 7.6.0.5. It consists of taking the partial derivatives and substituting them into 7.6.0.5. The following substitution can be used:

$$g(t,x):=N\left(\frac{\log(\frac{x}{k})+(r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right),\ h(t,x):=g(t,x)-\sigma\sqrt{T-t}$$

Using this notation we have

$$\begin{aligned}
\sqrt{2\pi n}(g(t,x)) &= e^{-\frac{1}{2}(h(t,x)+\sigma\sqrt{T-t})^2}\\
&= e^{-\frac{1}{2}h(t,x)^2}\cdot e^{-\sigma h(t,x)\sqrt{T-t}-\frac{1}{2}\sigma^2(T-t)}\\
&= \sqrt{2\pi n}(h(t,x))e^{\log(\frac{k}{x})-(r-\frac{\sigma^2}{2})(T-t)-\frac{\sigma^2}{2}(T-t)}\\
&= \sqrt{2\pi n}(h(t,x))\frac{k}{x}e^{-r(T-t)}
\end{aligned}$$

and hence

$$xn(g(t,x))=ke^{-r(T-t)}n(h(t,x)). \tag{7.6.1.3}$$

On taking partial derivatives and substituting 7.6.1.3 we obtain

$$
\begin{aligned}
g_1(t,,x) &= \frac{\log(\frac{x}{k})}{2\sigma(T-t)^{3/2}} - \frac{(r+\frac{1}{2}\sigma^2)}{2\sigma\sqrt{T-t}}, \\
h_1(t,x) &= g_1(t,x) + \frac{\sigma}{2\sqrt{T-t}}, \\
g_2(t,x) &= h_2(t,x) = \frac{1}{x\sigma\sqrt{T-t}}, \\
u(t,x) &= xN(g(t,x)) - ke^{-r(T-t)}N(h(t,x)), \\
u_1(t,x) &= -\frac{x\sigma n(g(t,x))}{2\sqrt{T-t}} - rke^{-r(T-t)}N(h(t,x)), \\
u_2(t,x) &= N(g(t,x)) \\
u_{22}(t,x) &= \frac{n(g(t,x))}{x\sigma\sqrt{T-t}}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
u_1(t,x) + rxu_2(t,x) + \frac{1}{2}\sigma^2x^2u_{22}(t,x) &= rxN(g(t,x)) - rke^{-r(T-t)}N(h(t,x)) \\
&= ru(t,x)
\end{aligned}
$$

as required. Since we have already observed that $\theta_t = u_2(t,X_t) = N(g(t,X_t))$, and as

$$
\begin{aligned}
\beta_t &= e^{-rt}(V_t - \theta_t X_t) \\
&= e^{-rt}(u(t,X_t) - u_2(t,X_t)X_t) \\
&= -ke^{-rT}N(h(t,X_t)),
\end{aligned}
$$

the formula for a hedging portfolio are now readily available. On substituting we obtain 7.6.1.1 and 7.6.1.2, completing the proof.

Since $0 \leq N(x) \leq 1$ the amount of shares in the portfolio will always be less than 1 and the borrowings will never exceed the strike price k. It can be shown that the agreed fair price for a call option, obtained in Proposition 6.3.1 and 7.5.2.1, is an *arbitrage free price.* [ch 7][↑]

Chapter 8

Summary and Conclusion

[↑]

A number of questions, mathematical and financial, can arise such as

How realistic is the Black-Scholes Formula? What role does the Black Scholes Formula play today in pricing options? etc.

The model we have developed is, perhaps, the simplest available showing the fundamental ideas, financial and mathematical, which arises in this complex area of applied mathematics. It is only a model and no model of is the perfect representation of the real world. A model can be judged only in comparison with the alternatives. The Black Scholes Model was a big improvement on what preceded it. Since 1973 a large amount of research has appeared, devoted to refining this model, and more realistic models are now available. All, however, involve the probabilistic methods of stochastic calculus and it could not be otherwise, as any attempt to predict future events must involve a random component. Consider the Black Scholes model a first, rather than a final, as approximation to the real world of option pricing.

We took some key assumptions in the model such as share prices change continuously, that re-balancing the hedging portfolio is continuous and costless, there are no transaction costs, the interest rate is constant, that all information on the history of the stock up to time t is reflected in the share price X_t and that the discrete approximation converges to the correct limit, i.e normal. Traders freely admit that the overall combination is unrealistic. On the other hand, it is possible to use market prices for options and the Black-Scholes formula to estimate *volatility*, and the formula is used by some companies to calculate emplyoee options as operating expenses. It is quite likely that many further uses will be found for this remarkable model.

We conclude by mentioning two particularly important results. In Example 7.5.2(c) we saw that the share price satisfied the stochastic differential equation

$$dX_t = (\mu + \frac{1}{2}\sigma^2)X_t dt + \sigma X_t dW_t$$

and that, be an appropriate change of measure, it satisfied

$$dX_t = \sigma X_t d\tilde{W}_t$$

where $(W_t)_{t\geq 0}$ and $(\tilde{W}_t)$ are Wiener processes. By Proposition 7.4.5 this change of measure turned the share into martingale. More generally, given a stochastic differential equation

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t$$

and a function g, then under fairly general conditions on f, g and σ, the *Girsanov-Theorem* gives an explicit change of measure formula which shows that $(X_t)_{t\geq 0}$ satisfies the stochastic differential equation

$$dX_t = g(X_t)dt + \sigma(X_t)d\tilde{W}_t.$$

Proposition 7.6.1 was obtained by solving the partial differential equation 7.6.0.5 which we derived from the stochastic equation 7.6.0.4. This a special case of the following proposition which connects partial and stochastic differential equations.

Proposition 8.0.2 (Feymann-Kac Formula) *If $r \in \mathbb{R}$, μ, σ and ϕ are sufficiently smooth functions of one real variable, then the solution of the partial differential equation*

$$ru(t,x) = u_1(t,x) + \mu(x)u_2(t,x) + \frac{\sigma^2(x)}{2}u_{22}(t,x)$$

with boundary condition $u(T,x) = \phi(x)$ *is given by*

$$u(t,x) = e^{-r(T-t)}\mathbb{E}[\phi(X_T)|X_t = x]$$

where $(X_t)_{t\geq 0}$ *satisfies the stochastic differential equation*

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

and $(W_t)_{t\geq 0}$ *is a Wiener process adapted to the filtration* $(\mathcal{F}_t)_{t\geq 0}$.

References

[1] Sean Dineen, *Probability Theory in Finance; A Mathematical guide to the Black Scholes Formula*, Graduate Studies in Mathematics (Vol 70), American Mathematical Society

[2] Vigirdas Mackevicius, *Introduction to Stochastic Analysis*, Integrals and Differential Equation, Wiley-ISTE; 1 edition (February 7, 2013)

[3] Hui-Hsiung Kuo, *Introduction to Stochastic Integration*, Universitext, Springer; 2006 edition (November 15, 2005)

[4] Fima C Klebaner, *Introduction to Stochastic Calculus With Applications*, Imperial College Press; 2nd Revised edition edition (June 20, 2005)

[5] R.J.Williams, *Introduction to the Mathematics of Finance*, Graduate Studies in Mathematics (Vol 72), American Mathematical Society