Essentials for Financial Derivatives

YIN JUNQI

National University of Singapore

February 13, 2025

Contents

1 Introduction **1**
- 1.1 Overview 1
- 1.2 Forward 2
- 1.3 European Vanilla Options 2
- 1.4 Other Types of Options 4
- 1.5 Leverage 5
- 1.6 Pricing Option 5

2 Discrete Time Models **6**
- 2.1 1 Period Binomial Model 6
- 2.2 Multi-Period Binomial Model 8
- 2.3 Discrete Model for American Options 13
- 2.4 The Trinomial Model 15

3 Math Preliminaries **17**
- 3.1 Stochastic Processes 17
- 3.2 Brownian Motion(Wiener Process) 18
- 3.3 Quadratic Variation 20
- 3.4 Itô's Integral 23
- 3.5 Itô's Lemma 29
- 3.6 Geometric Brownian Motion 31
- 3.7 Stochastic Differential Equations 32

4 Black-Scholes Model **35**
- 4.1 Self-Financing Portfolio and Black-Scholes PDE 36
- 4.2 Numerical Computation and Stablility 38
- 4.3 Risk-Neutral Framework 41
- 4.4 Black-Scholes Formula and Greeks 47
- 4.5 Dividend-Paying Stock 50
- 4.6 Change of Numeraire 51
- 4.7 Multi-Dimensional Black-Scholes Model 53
  - 4.7.1 PDE Framework 53
  - 4.7.2 Exchange Options 57

5 Implied Volatility Skew(Smile) **59**
- 5.1 Implied Volatility 59
- 5.2 Volatility Smile 60
- 5.3 Market Implied Density 61
- 5.4 Local Volatility Model: Dupire's Equation 62
- 5.5 Stochastic Volatility Model 63

6 Exotic Options **67**
- 6.1 Barrier Options 67
  - 6.1.1 Down-and-In Call Options, Reflection Principle 68
  - 6.1.2 Up-and-Out Call Options 73
  - 6.1.3 PDE Framework 75
- 6.2 Asian Options 76
  - 6.2.1 Geometric Asian Options 77
  - 6.2.2 PDE Framework: For Path-Dependent Options 80
  - 6.2.3 Rogers-Shi Method 80
- 6.3 Lookback Options 82
- 6.4 Foreign Exchange (FX) Options 84

7 American Options · · · · · · · · · · **88**
7.1 Early Exercise and Optimal Stopping · · · · · · · · · · 88
7.2 American Put Options · · · · · · · · · · 91
7.3 Perpetual American Options · · · · · · · · · · 93
8 Interest Rate Modelling(An Introduction) · · · · · · · · · · **95**
8.1 The Bank Account Revisited · · · · · · · · · · 95
8.2 Tradeable Interest Rate (Bonds) · · · · · · · · · · 96
8.3 Interest Rate Derivatives Modelling · · · · · · · · · · 97
8.4 Forward · · · · · · · · · · 99
8.4.1 Forward Pricing · · · · · · · · · · 99
8.4.2 Forward Measure · · · · · · · · · · 100
8.5 Swaps · · · · · · · · · · 102
8.6 Zero Curve Bootstrapping · · · · · · · · · · 104
8.7 Introduction to Credit Risk · · · · · · · · · · 105
9 Problem Set · · · · · · · · · · **108**
9.1 Set 1 · · · · · · · · · · 108
9.2 Set 2 · · · · · · · · · · 109
9.3 Set 3 · · · · · · · · · · 113
A Expectation, Information and Conditioning · · · · · · · · · · **118**
A.1 Expectation · · · · · · · · · · 118
A.2 Convergence of Integrals · · · · · · · · · · 120
A.3 Information and Conditional Expectation · · · · · · · · · · 121
B Monte Carlo Simulation · · · · · · · · · · **122**
C Another Way to Derive Black-Scholes PDE · · · · · · · · · · **123**
D An Abbreviated Proof of Girsanov's Theorem · · · · · · · · · · **125**
E Bilingual Table · · · · · · · · · · **127**

List of Tables

1.1 Payment Streams of Forward Contract . . . 2

3.1 Multiplication Table . . . 22

4.1 Discretization Error and Stability of Schemes . . . 40

5.1 Volatility Smile . . . 60

6.1 Types of Barrier Options . . . 68
6.2 Conditions for Barrier Options . . . 76

7.1 American v.s. European . . . 92
7.2 American v.s. European . . . 92

E.1 Bilingual Table . . . 127

List of Figures

1.1 Derivative . . . 1
1.2 Payoff of a Call Option . . . 3
1.3 Payoff of a Put Option . . . 3
1.4 Barriers . . . 4

2.1 Binomial Tree . . . 6
2.2 3-Period Binomial Tree . . . 9
2.3 Binomial Model(European Option) . . . 12
2.4 Binomial Tree for Asset Prices . . . 13
2.5 American Option . . . 14
2.6 Trinomial Model . . . 16

3.1 Martingales . . . 18
3.2 Wiener Process . . . 19
3.3 Quadratic Variation Comparison . . . 23
3.4 Simple Process . . . 24
3.5 Simple Process to Brownian Motion . . . 27
3.6 Geometric Brownian Motion . . . 32

4.1 Black-Scholes Model . . . 35
4.2 Finite Difference Grid . . . 38
4.3 Stability of Different Schemes . . . 41
4.4 Price of a European call option . . . 48
4.5 Greeks of a European call option . . . 49
4.6 Price of a European call option with two underlying assets . . . 55
4.7 Price of a Basket option with two underlying assets . . . 56

5.1 Volatility Smile . . . 61
5.2 Heston Model . . . 66

6.1 Reflection Principle . . . 70
6.2 Down-and-In v.s. Down-and-Out Call Option . . . 73
6.3 Price of Geometric Asian Call Option . . . 80

7.1 Continuation Region and Exercise Boundary for American Put Option . . . 93

8.1 Payments dates presence . . . 104

A.1 Riemann Integral . . . 119
A.2 Lebesgue Integral . . . 119

1 Introduction

1.1 Overview

This course covers one of the major questions in Mathematical Finance: How do we price financial products? To answer this question, we will see:

1. Several mathematical tools from stochastic calculus(continuous time):
   (a) Brownian motion, Quadratic variation
   (b) Itô's Integral and Itô's Lemma
   (c) Stochastic differential equations
2. Discrete-time Models;
3. Important concepts for the valuation of derivatives:
   (a) Arbitrage-free
   (b) Self-financing and Delta-hedging
   (c) Pricing via Replication and PDEs
   (d) Risk-neutral measure
4. The Black-Scholes model and its generalizations. How to apply these concepts and tools to price specific options: Call/Put; Barrier; Asian;etc.

So first, what is a financial derivative?

Definition 1.1. *A financial derivative is a contract between two parties that specifies conditions under which payments are to be made between the parties. The value of the derivative is determined by the price of an* ***underlying asset****. The most common types of derivatives are futures, forwards, options, and swaps.*

Financial derivatives play a crucial role in the financial markets. They are used for **Risk management, Operation, and Speculation**.

Any asset can be the underlying of an option, like stocks, bonds, currencies, oil, soybeans, etc. The most common underlying assets are stocks.

Figure 1.1: Derivative

A very important assumption is that no profit can be made without taking risk. This is called the **no-arbitrage** assumption. If there were an arbitrage opportunity, then the market would react immediately to eliminate it. Any arbitrage opportunity is limited by transaction costs related to the liquidities of the market.

Definition 1.2. ***Arbitrage*** *is a transaction that guarantees profit without any risk. Also called a* ***free lunch****. For $t_0 < T$ an arbitrage is the existence of any asset with value $V(t)$ at time t where either of the below hold:*

$$V(t_0) < 0 \text{ and } \mathbb{P}(V(T) \geq 0) = 1$$
$$V(t_0) = 0 \text{ and } \mathbb{P}(V(T) > 0) = 1$$

Now we go through basic derivatives.

1.2 Forward

Definition 1.3. ***Forward contract****: Agreement to buy an asset(say,a share of Apple) at a future date T at a price K which is* ***agreed today****. To an investor, the value of a forward contract on maturity date is:*

$$V(T) = S(T) - K$$

Consider 2 strategies:

1. I enter the forward contract today and buy the share of Apple at time T at price K.
2. I borrow S_0 from the bank and buy a share of Apple, then hold it until time T; pay back with interest S_0e^{rT} to the bank.

This leads to the following payment streams:

Table 1.1: Payment Streams of Forward Contract

	$t = 0$	$t = T$
Str.A	0	$S_T - K$
Str.B	S_0	$S_T - S_0e^{rT}$
	Flow	**Value**

By no arbitrage, the return of the two strategies should be the same. So we have:

$$K = e^{rT}S_0$$

For example, if $K < S_0e^{rT}$, one can long Str. A and short Str. B. Enters the forward contract, shorts the stock, invests the cash at a riskless rate → earns riskless profit ($S_0e^{rT} - K$).

Using the following Python code to calculate the forward price:

```python
import numpy as np
def forward_price(S0, r, T):
    return S0*np.exp(r*T)
```

1.3 European Vanilla Options

Definition 1.4. *A European* ***call option*** *is a contract which gives the investor the option to buy the underlying asset at maturity date T, for a strike price K.*

Obviously, the payoff is $V(T) = \max(S(T) - K, 0)$. We write this as

$$C(T) = \max(S(T) - K, 0) := (S(T) - K)^+$$

Unlike the forward contract, which is a commitment to buy the underlying asset, the call option gives the investor all of the upside and none of the downside. An oil consumer can buy a call option on oil; if oil prices rise above K then the oil consumer can **exercise the option** to buy the oil at the lower price K. The option is then said to be **in the money**. On the other hand if the oil price falls below the strike K then the option is **out of the money**.

Figure 1.2: Payoff of a Call Option

Definition 1.5. *A **put option** is the downside analogue to the call option. It is a contract which gives the investor the option to **sell** the underlying asset at maturity date T , for a strike price K.*

If prices fall below K, then the option is in the money and the option buyer can exercise it, which **in essence forces the option seller to buy** the oil at strike K even though oil prices in the market are lower. To an option buyer, the value of a put option contract on maturity date is

$$V(T) = \max(K - S(T), 0) := (K - S(T))^+$$

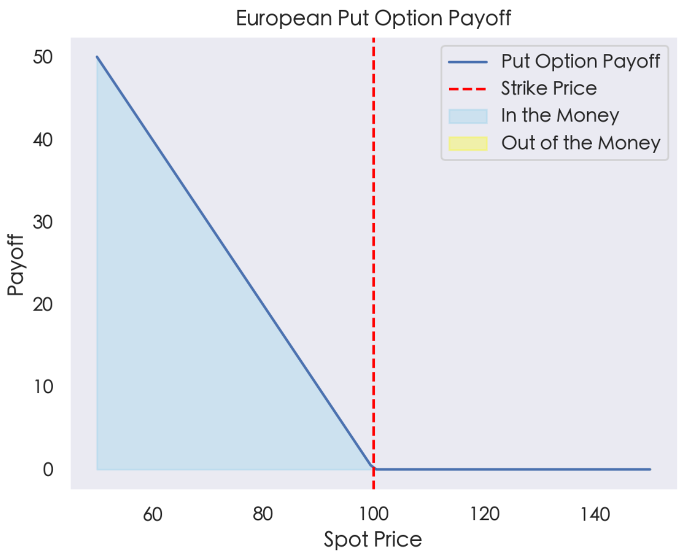

Figure 1.3: Payoff of a Put Option

The formulations above imply the below model independent result, famously known as:

Theorem 1.1. ***Put-Call Parity**: For European vanilla options, the following holds:*

$$C_T - P_T = (S_T - K)^+ - (K - S_T)^+ = S_T - K$$

Since this relationship is valid at time T it must also be valid at any prior time t. If not, there exists and arbitrage. For example, if $C_T - P_T < S_T - K$, then at time T, one can immediately buy a call option, sell a put option, and use the contract to buy the stock at K. This will yield a riskless profit of $S_T - K - (C_T - P_T) > 0$.

1.4 Other Types of Options

Binary Options

Definition 1.6. *A **binary option** is a type of option with a **fixed payout** in which you predict the outcome from two possible results. If your prediction is correct, you receive the agreed payout. If not, you lose your initial stake, and nothing more.*

To an option buyer, the value of a binary call option contract on maturity date is:

$$V_T = \mathbb{1}_{S_T > K}$$

where $\mathbb{1}_{S_T > K}$ is the indicator function.

The formula for the binary put is analogous, and the two formulations imply the parity relationship below.

$$\mathbb{1}_{S_T < K} + \mathbb{1}_{S_T > K} = 1$$

Barrier Options

Definition 1.7. *A **barrier option** is a type of option whose payoff depends on whether the underlying asset's price **reaches a certain level** during a certain period of time.*

- The option contract might specify that the payoff becomes void if the underlying *breaches* a certain level. The option then **knocks out**.
- Analogously the contract terms could specify the converse; that the payoff is only awarded if that barrier is *reached* prior to maturity date. Then it **knocks in**.

There are four interesting cases for the call and put cases. Using notation $M_T = \max_{0 \le t \le T} S_t$ which denotes the maximum value that the underlying asset attains during the period $[0, T]$ and m_t to denote the lifetime minimum, and B to be the contracted barrier level.

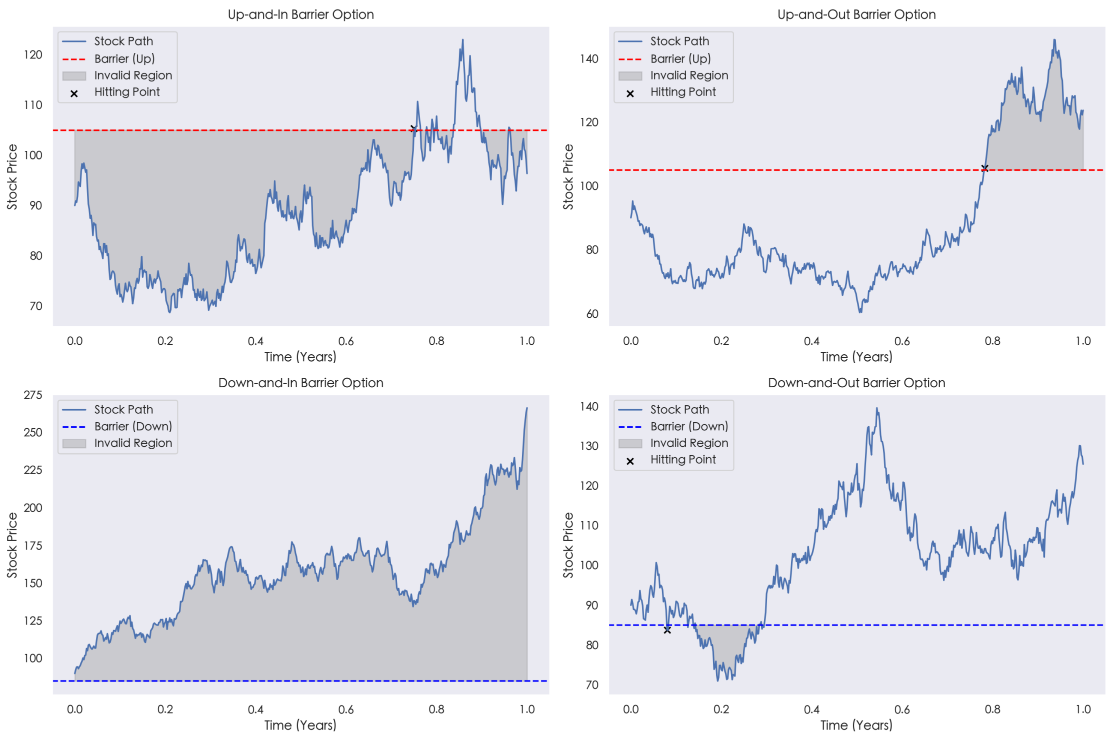

Figure 1.4: Barriers

1. **Up-and-out**: Spot price starts below the barrier level and has to move up for the option to be knocked out.

$$\begin{aligned} V_T^c &= \max\left(S_T - K\right)^+ \mathbb{1}_{M_T < B} \quad && \text{call option} \\ V_T^p &= \max\left(K - S_T\right)^+ \mathbb{1}_{M_T < B} \quad && \text{put option} \end{aligned}$$

2. **Down-and-out**: Spot price starts above the barrier level and has to move down for the option to become null and void.
$$\begin{aligned} V_T^c &= \max\left(S_T - K\right)^+ \mathbb{1}_{m_T > B} && \text{call option} \\ V_T^p &= \max\left(K - S_T\right)^+ \mathbb{1}_{m_T > B} && \text{put option} \end{aligned}$$

3. **Up-and-in**: Spot price starts below the barrier level and has to move up for the option to be activated.
$$\begin{aligned} V_T^c &= \max\left(S_T - K\right)^+ \mathbb{1}_{M_T \geq B} && \text{call option} \\ V_T^p &= \max\left(K - S_T\right)^+ \mathbb{1}_{M_T \geq B} && \text{put option} \end{aligned}$$

4. **Down-and-in**: Spot price starts above the barrier level and has to move down for the option to be activated.
$$\begin{aligned} V_T^c &= \max\left(S_T - K\right)^+ \mathbb{1}_{m_T \leq B} && \text{call option} \\ V_T^p &= \max\left(K - S_T\right)^+ \mathbb{1}_{m_T \leq B} && \text{put option} \end{aligned}$$

1.5 Leverage

Definition 1.8. ***Leverage*** *is the use of borrowed money to increase the potential return of an investment. It has the potential to magnify gains, but also magnify losses.*

Example 1. *Financial derivatives help investors adjust their financial exposures in line with their risk/rewards appetites. A stock might cost* \$100 *to buy outright, and would give an investor a* 10% *return if it appreciates* 10%.

However, a 1 year call option with strike 100 on the stock might only cost \$10*. Buying 10 call options would require the same level of initial investment as buying the stock outright but in the latter, a* 10% *appreciation gives the investor a* 100% *return.*

But the higher reward comes at higher risk; if the call options expires with the stock price at 99 , then the investor's initial \$100 *investment is completely wiped out. An investor who chose to buy the stock would only experience* 1% *loss.* 1% *loss versus* 100% *loss - that's what leverage can do.*

1.6 Pricing Option

Recall that a financial derivative is a contract between two people (a buyer and a seller).

- At time t, the buyer of the contract pays to the seller a certain price.
- At time T, the seller of the contract pays to the buyer a certain payoff $\varphi\left(S_T\right)$, where S_T is the stock price at time T.

The question is How do we calculate the price of the contract. The answer is to use the concept of **replication**. We will see this in the next chapter.

Proposition 1.1. *Arbitrage limits for call and put options:*

$$\begin{gathered} \textit{Call Options:} \\ C(0) > 0 \\ C(0) < S(0) \\ \textit{Put Options:} \\ P(0) > 0 \\ P(0) < Ke^{-rT} < K \end{gathered}$$

For call options, if $V(0) > S(0)$, then one can sell the call option, buy the stock. This will yield a riskless profit of $V(0) - S(0) > 0$ cause at time T, one will either handover the stock and receive K or not. One will have a cash flow $S(T) - K > 0$, or 0 if $S(T) < K$.

For put options, if $V(0) > Ke^{-rT}$, then one can sell the put option, invest Ke^{-rT} in the money market. At time T, one will either buy the stock paying K or not. One will have a cash flow $K > 0$ if not exercised, or $S(T) > 0$ if $S(T) > K$.

2 Discrete Time Models

In this chapter, we will see how to price options in a discrete-time model. We will see the following concepts:

1. Binomial model
2. Multi-period binomial model
3. American options with discrete model
4. Trinomial model introduction.

2.1 1 Period Binomial Model

Definition 2.1. *A **portfolio** is a combination of positions in different financial assets. Assume you have a position of α in the money market account and β in the stock. The portfolio is the vector $(\alpha, \beta) \in \mathbb{R}^2$, the value of the portfolio is $\alpha B_t + \beta S_t$.*

There are two points in time: $t = 0$ (today) and $t = T$ (*maturity*). Money can be either put in/borrowed from the bank or invested in a stock. Putting y dollars in the bank at time 0 earns interest. At time T, it becomes ye^{rT}.

Assume the price of a unit of the money market account at time 0 is $B_0 = 1$. The price of a unit of the money market account at time T is $B_T = e^{rT}$.

The price of the stock at time 0 is $S_0 = s$. The price of the stock at time T is $S_T = s \cdot X$, where

$$X = \begin{cases} u & \text{with probability } p_u \in [0,1] \\ d & \text{with probability } p_d \in [0,1] \end{cases}$$

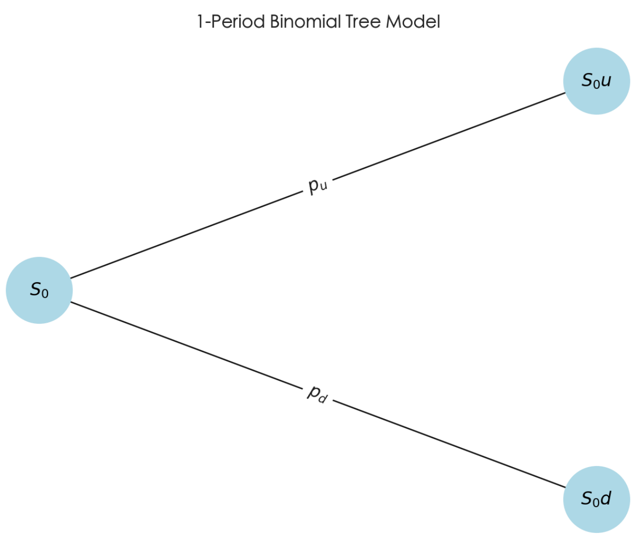

Figure 2.1: Binomial Tree

Frist we need to include important assumptions. An arbitrage portfolio is a portfolio h such that

1. $V_0^h = 0$,
2. $V_T^h \geq 0$,

3. $V_T^h > 0$ with positive probability.

We assume that the market is arbitrage-free. This means that there is no arbitrage portfolio.

Proposition 2.1. *Assume that $V_T^{h*} \geq V_T^h$. Then, in the absence of arbitrage $V_0^{h*} \geq V_0^h$.*

Proof. Consider the portfolio $\tilde{h} = h^* - h - \left(V_0^{h^*} - V_0^h\right)$, then

$$V_0^{\tilde{h}} = 0; \quad V_T^{\tilde{h}} = V_T^{h^*} - V_T^h - \left(V_0^{h^*} - V_0^h\right) e^{rT}$$

If $V_0^{h^*} < V_0^h$, then $V_T^{\tilde{h}} > 0 \rightarrow \tilde{h}$ is an arbitrage portfolio, contradicting the absence of arbitrage, thus $V_0^{h^*} \geq V_0^h$. □

Corollary 2.1. ***Law of One Price**: If two portfolios have the same payoff at time T, then they must have the same price at time 0. Assume that $V_T^{h*} = V_T^h$. Then, in the absence of arbitrage $V_0^{h*} = V_0^h$.*

Proof.

$$V_T^{h^*} = V_T^h \rightarrow V_T^{h^*} \geq V_T^h \rightarrow V_0^{h^*} \geq V_0^h$$
$$V_T^{h^*} = V_T^h \rightarrow V_T^h \geq V_T^{h^*} \rightarrow V_0^h \geq V_0^{h^*}$$

Thus

$$V_0^h = V_0^{h^*}$$

□

Consider a portfolio that is long 1 call option, short 1 stock and long Ke^{-rT} in the risk-free asset. Let $c(K, T)$ be the price of the call option. The value of the portfolio today is $c(K, T) - S_0 + Ke^{-rT}$. The value of the portfolio at T is $(S_T - K)^+ - S_T + K = (K - S_T)^+$. By the law of one price, we have:

Theorem 2.1. ***Put-Call Parity**: For European vanilla options, the following holds:*

$$C(S_t, t) - P(S_t, t) = S_t - Ke^{-r(T-t)}$$

Theorem 2.2. *The binomial model is arbitrage free if and only if $d < e^{rT} < u$.*

Proof. " ⇒ "

Let a portfolio $h = (-S_0, 1)$ Then, we have $V_0^h = -1 \cdot S_0 + 1 \cdot S_0 = 0$. At time T we compute

$$V_T^h = \begin{cases} S_0\left(u - e^{rT}\right) & \text{with probability } p_u \\ S_0\left(d - e^{rT}\right) & \text{with probability } p_d \end{cases}$$

The case $e^{rT} \leq d < u$ can be excluded because h is, by assumption, not an arbitrage strategy. The case $d < u \leq e^{rT}$ can be excluded because $-h$ is, by assumption, not an arbitrage strategy.

" ⇐ "

Consider a potential arbitrage portfolio $h = (\alpha, \beta)$ with $V_0^h = \alpha + \beta S_0 = 0$. Then, we use $\alpha = -\beta S_0$ and see that

$$V_T^h = \begin{cases} \alpha e^{rT} + \beta S_0 u = \beta S_0\left(u - e^{rT}\right) & \text{with probability } p_u \\ \alpha e^{rT} + \beta S_0 d = \beta S_0\left(d - e^{rT}\right) & \text{with probability } p_d \end{cases}$$

By Assumption $\left(u - e^{rT}\right) > 0$ and $\left(d - e^{rT}\right) < 0$. Hence, h cannot be an arbitrage portfolio. □

Note 1. *If $d > e^{rT}$, one can borrow S_0 from bank at time 0. And time T, at worst case one get dS_0 and need to pay back $S_0 e^{rT}$ to the bank, because $d > e^{rT}$, one can make a riskless profit.*

If $u < e^{rT}$, one can short the stock and invest it to the bank at time 0. At time T, at worst case one get $S_0 e^{rT}$ and need to buy back $S_0 u$, because $u < e^{rT}$, one can make a riskless profit.

Consider the portfolio $h = (\alpha, \beta)$. The portfolio h **replicates** the derivative φ if

$$\begin{aligned} \alpha \cdot e^{rT} + \beta \cdot S_0 \cdot u &= \varphi(S_0 \cdot u) \\ \alpha \cdot e^{rT} + \beta \cdot S_0 \cdot d &= \varphi(S_0 \cdot d). \end{aligned}$$

This is a trivial system of linear equations:

$$\begin{aligned} \alpha &= e^{-rT} \cdot \frac{u \cdot \varphi(S_0 \cdot d) - d \cdot \varphi(S_0 \cdot u)}{u-d}, \\ \beta &= \frac{1}{S_0} \cdot \frac{\varphi(S_0 \cdot u) - \varphi(S_0 \cdot d)}{u-d}. \end{aligned}$$

By no arbitrage, the price at time $t = 0$ of the derivative is the value at $t = 0$ of the replicating portfolio. The price is

$$V_0^h = \alpha + \beta \cdot S_0 = e^{-rT}\left[\frac{e^{rT}-d}{u-d}\varphi(S_0 \cdot u) + \frac{u-e^{rT}}{u-d}\varphi(S_0 \cdot d)\right]$$

Define

$$q_u := \frac{e^{rT}-d}{u-d}, \quad q_d := \frac{u-e^{rT}}{u-d}$$

If the market is arbitrage-free, q_u and q_d are **probabilities**. Define $\mathbb{Q}$ as the probability measure such that $\mathbb{Q}(S_T = S_0 \cdot u) = q_u$ and $\mathbb{Q}(S_T = S_0 \cdot d) = q_d$.

Write this into a theorem:

Theorem 2.3. *Assume the bionomial model market is arbitrage-free. Then the price at time $t = 0$ of the derivative that at $t = T$ has payoff $\varphi(S_T)$ is*

$$e^{-rT}\mathbb{E}^{\mathbb{Q}}[\varphi(S_T)].$$

- Note that The **real-world probabilities** p_u and p_d do not matter. Only the probability measure $\mathbb{Q}$ matters.
- Notice that for stock, $s = e^{-rT}E^{\mathbb{Q}}[S_T]$. A probability measure with this property is called a **risk-neutral probability** (in the binomial model, it is unique).
- The price of the derivative is the discounted expectation of the payoff under the risk-neutral probability measure.

2.2 Multi-Period Binomial Model

A multi period model is built by stacking together single period models; this creates what is called a binomial tree or lattice.

There are finite time steps $t = 0, \cdots, T$. The risk-free asset has price dynamics

$$\begin{aligned} B_t &= e^{rt}B_0 \\ B_0 &= 1. \end{aligned}$$

The stock has price dynamics

$$S_{t+1} = S_t \cdot X_t = \begin{cases} S_t \cdot u & \text{with probability } p_u \\ S_t \cdot d & \text{with probability } p_d \end{cases} \quad S_0 = s,$$

where $X_0, \cdots, X_{T-1}$ are i.i.d. random variables with $X_t \in \{u, d\}$.

A 3-period binomial tree is shown as 2.2.

For an n peroid model, the states can then be stored in $(n+1) \times (n+1)$ matrices. **S** holds the asset prices at time step j and level node i where i counts rows and j counts columns. **V** similarly stores the derivative prices. The result of the numerical computation, will eventually come out at grid node $\mathbf{V}[0,0]$.

The matrix **S** is filled out as follows, this is a trivial exercise since all the parameters required are model inputs.

$$\mathbf{S} = \begin{pmatrix} S_0 & u \cdot S_0 & u^2 \cdot S_0 & u^3 \cdot S_0 & u^4 \cdot S_0 & \dots & u^n \cdot S_0 \\ 0 & d \cdot S_0 & ud \cdot S_0 & u^2d \cdot S_0 & u^3d \cdot S_0 & \dots & u^{n-1}d \cdot S_0 \\ 0 & 0 & d^2 \cdot S_0 & ud^2 \cdot S_0 & u^2d^2 \cdot S_0 & \dots & u^{n-2}d^2 \cdot S_0 \\ 0 & 0 & 0 & d^3 \cdot S_0 & ud^3 \cdot S_0 & \dots & u^{n-3}d^3 \cdot S_0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & d^{n-1} \cdot S_0 & ud^{n-1} \cdot S_0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & d^n \cdot S_0 \end{pmatrix}$$

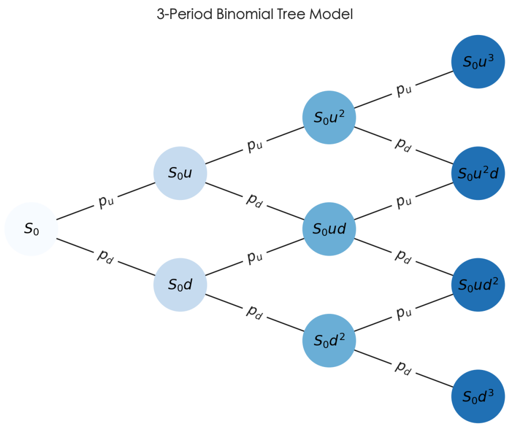

Figure 2.2: 3-Period Binomial Tree

The matrix $\mathbf{V}$ is filled out by induction. We start with the last column which contains the payoff of the options. For example the matrix below contains the terminal payoffs for European Put options with strike K.

$$\mathbf{V}=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & (K-\mathbf{S}[0,n])^+ \\ 0 & 0 & 0 & 0 & 0 & \cdots & (K-\mathbf{S}[1,n])^+ \\ 0 & 0 & 0 & 0 & 0 & \cdots & (K-\mathbf{S}[2,n])^+ \\ 0 & 0 & 0 & 0 & 0 & \cdots & (K-\mathbf{S}[3,n])^+ \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & (K-\mathbf{S}[n-1,n])^+ \\ 0 & 0 & 0 & 0 & 0 & \cdots & (K-\mathbf{S}[n,n])^+ \end{pmatrix}$$

The remaining columns for $\mathbf{V}$ are filled out according to the single period replication formula:

$$\mathbf{V}[i,j]=\frac{q_i^u\mathbf{V}[i,j+1]+\left(1-q_i^u\right)\mathbf{V}[i+1,j+1]}{e^{r\Delta t}} \qquad \forall 0\leq j<n, 0<i\leq j$$

where q_i^u is the risk-neutral probability of the stock price going up from node i to node $i+1$.

The induction must happen **backwards** going from $j=(n-1)\rightarrow 0$. This is no different to the single period model, except that we now have many of them stacked together, with nodes at later time nodes feeding the nodes at carlier time nodes, as we move sweep back through the tree. Note that q_i^u can be a function of j if r is a function of time. But it typically does not vary across nodes in a particular column.

Example 2. *Value a European Call option in a 2 period model:*

1. *Value at time* $T/2: C_u=e^{-rT/2}\left(q_uC_{uu}+q_dC_{ud}\right)$ *(if price went up),*
2. *Value at time* $T/2: C_d=e^{-rT/2}\left(q_uC_{ud}+q_dC_{dd}\right)$ *(if price went down),*
3. *Value today:*
$$C=e^{-rT/2}\left(q_uC_u+q_dC_d\right)=e^{-rT}\left(q_u^2C_{uu}+2q_uq_dC_{ud}+q_d^2C_{dd}\right).$$

For the generalized case, we have the following theorem:

Theorem 2.4. *The price at time* $t=0$ *of the derivative that at* $t=T$ *has payoff* $\varphi\left(S_T\right)$ *is*

$$V(0)=e^{-rT}\sum_{i=0}^{T}\binom{T}{i}q_u^{T-i}q_d^i\varphi\left(S_0\cdot u^{T-i}\cdot d^i\right)=e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\varphi\left(S_T\right)\right].$$

Example 3. *Consider a 1-peroid binomimal model with parameters*

$$p_u=0.4, p_d=0.6, r=\ln(1.1), T=1, d=0.9, u=1.05, S_0=10$$

Due to the fact that $e^r=1.1>u=1.05$, *an arbitrage opportunity exists. We can short the stock and invest the proceeds in the risk-free asset. Even under the case that the stock price goes up to 10.5 , we can still cover the debt and have a profit of 0.5 , because the risk-free asset will have a value of 11 at time 1.*

Change $u = 1.1$ as well as $r = 0$. Moreover, assume there's a second intermediate peroid(the market can change by the factor d or u between times 0 and 0.5 and between times 0.5 and 1). In this 2 peroid binomial model, compute the price of an at-the-money ***LOOKBACK OPTION*** *with payoff*

$$\varphi(S_{0.5}, S_1) = \max(\{S_0, S_{0.5}, S_1\} - 10)^+$$

in 2 ways: Via replication and Via risk-neutral valuation.

1. *The binomial tree is :*

$$\begin{pmatrix} 10 & 11 & 12.1 \\ 0 & 9 & 9.9 \\ 0 & 0 & 8.1 \end{pmatrix}$$

   Let α, β denote the number of risk-free asset and stock, respectively.
   By definition of the option, we have: at node $S_{0.5} = 11$,

$$\begin{cases} \alpha_1 + \beta_1 \times 12.1 = 2.1 \\ \alpha_1 + \beta_1 \times 9.9 = 1 \end{cases}$$

   yielding $\alpha_1 = -3.95, \beta_1 = 0.5$, and option value $V_u = 0.5 \times 11 - 3.95 = 1.55$
   At node $S_{0.5} = 9$, we have $\alpha_2 = \beta_2 = 0$, and option value $V_d = 0$
   At node $S_0 = 10$, solve

$$\begin{cases} \alpha_0 + \beta_0 \times 11 = 1.55 \\ \alpha_0 + \beta_0 \times 9 = 0 \end{cases}$$

   we get $\alpha_0 = -6.975, \beta_0 = 0.775$, and the option price is $V_0 = 0.775 \times 10 - 6.975 = 0.775$

2. *Via risk-neutral valuation: The risk-neutral probability is*

$$q_u = \frac{e^r - d}{u - d} = 0.5, \quad q_d = \frac{u - e^r}{u - d} = 0.5$$

   The option price is

$$V_0 = e^{-r}\left(q_u^2 V_{uu} + 2q_u q_d V_{ud} + q_d^2 V_{dd}\right) = 0.25 \times 3.1 = 0.775$$

The Limit of Multi-Period Binomial Model: Cox-Ross-Rubinstein

Approximate the continuous-time process from 0 to T by discretetime process of N time steps: $0, \Delta t, 2\Delta t, \ldots, N\Delta t; \Delta t = T/N$.

The N-period binomial model is used to model the stock price movement. It is constructed by setting

$$u = e^{\sigma\sqrt{\Delta t}}, d = e^{-\sigma\sqrt{\Delta t}}$$

The market-neutral probabilities, $p_u = \dfrac{e^{r\Delta t} - d}{u - d}; p_d = 1 - p_u$

The CRR is a discrete-time approximation of the continuous-time geometric Brownian motion used in the Black-Scholes model.

$$\begin{aligned} \mathbb{E}(\Delta S) &= (p_u u + p_d d)\, S - S = \left(e^{r\Delta t} - 1\right) S \sim rS\Delta t \\ \mathrm{Var}(\Delta S) = \mathbb{E}\left((\Delta S)^2\right) - (\mathbb{E}(\Delta S))^2 &= \left(p_u(u-1)^2 S^2 + p_d(1-d)^2 S^2\right) - (\mathbb{E}(\Delta S))^2 \\ &\sim \left(\frac{1}{2}\sigma^2\Delta t + \frac{1}{2}\sigma^2\Delta t\right) S^2 = \sigma^2 S^2 \Delta t \end{aligned}$$

This is an equivalent discrete approximation of geometric Brownian motion

$$\Delta S(t) = rS(t)\Delta t + \sigma S z(t)\sqrt{\Delta t}, \quad z(t) \sim \mathcal{N}(0,1)$$

Using the following Python code to build a multi-period binomial model, which can approximate the Black-Scholes model.

```python
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

def binomial_tree(S, K, T, r, sigma, option_type="call", American = False, steps=100):
```

```python
    dt = T / steps
    u = np.exp(sigma * np.sqrt(dt))  # Up factor
    d = 1 / u  # Down factor
    p = (np.exp(r * dt) - d) / (u - d)  # Risk-neutral probability
    discount = np.exp(-r * dt)

    # Initialize asset prices and option values
    asset_prices = np.zeros((steps + 1, steps + 1))
    option_values = np.zeros((steps + 1, steps + 1))

    for i in range(steps + 1):
        for j in range(i + 1):
            asset_prices[j, i] = S * (u ** (i - j)) * (d ** j)

    # Compute option values at expiration
    if option_type == "call":
        option_values[:, steps] = np.maximum(0, asset_prices[:, steps] - K)
    elif option_type == "put":
        option_values[:, steps] = np.maximum(0, K - asset_prices[:, steps])

    # Backward induction
    for i in range(steps - 1, -1, -1):
        for j in range(i + 1):
            option_values[j, i] = discount * (p * option_values[j, i + 1] + (1 - p) * option_values[j + 1, i + 1])
            if American:
                # American option can be exercised at earlier time
                if option_type == "call":
                    option_values[j, i] = np.maximum(option_values[j, i], asset_prices[j, i] - K)
                elif option_type == "put":
                    option_values[j, i] = np.maximum(option_values[j, i], K - asset_prices[j, i])

    return asset_prices, option_values[0,0]

# Payoff functions
def payoff(S,K,option_type="call"):
    if option_type == "call":
        return np.maximum(S - K, 0)
    elif option_type == "put":
        return np.maximum(K - S, 0)

# Validate the binomial tree model
S = np.linspace(50, 150, 100)
K = 100
T = 1
r = 0.05
sigma = 0.2
steps = 100

# Calculate option prices and payoffs
call_prices = np.zeros(len(S))
put_prices = np.zeros(len(S))
for S0 in S:
    _, call_price = binomial_tree(S0, K, T, r, sigma, option_type="call", steps=steps)
    call_prices[np.where(S == S0)] = call_price

    _, put_price = binomial_tree(S0, K, T, r, sigma, option_type="put", steps=steps)
    put_prices[np.where(S == S0)] = put_price
```

```python
call_payoffs = payoff(S, K)
put_payoffs = payoff(S, K, option_type="put")

plt.figure(figsize=(10, 6))
plt.subplot(1, 2, 1)
sns.lineplot(x=S, y=call_prices, label="Call Price", color="green")
sns.lineplot(x=S, y=call_payoffs, label="Call Payoff", linestyle="--", color="blue")
plt.title("Call Option Payoff and Pricing")

plt.subplot(1, 2, 2)
sns.lineplot(x=S, y=put_prices, label="Put Price", color="orange")
sns.lineplot(x=S, y=put_payoffs, label="Put Payoff", linestyle="--", color="blue")

plt.title("Put Option Payoff and Pricing")
plt.xlabel("Stock Price")
plt.ylabel("Payoff/Price")
plt.legend()

plt.show()
```

The output figure is shown as 2.3.

Figure 2.3: Binomial Model(European Option)

Using the following Python code to visualize the binomial tree(asset prices):

```python
import matplotlib.cm as cm
def plot_binomial_tree(tree, title):
    T = tree.shape[1] - 1
    fig, ax = plt.subplots()
    colors = cm.Blues(np.linspace(0.3, 1, T + 1))
    for t in range(T + 1):
        for i in range(t + 1):
            ax.plot(t, tree[i, t], 'o', color=colors[t], markersize=4)
            if t > 0:
                if i < t:
                    ax.plot([t - 1, t], [tree[i, t - 1], tree[i, t]], color=colors[t],
                    ↪ linewidth=1)
                if i > 0:
                    ax.plot([t - 1, t], [tree[i - 1, t - 1], tree[i, t]],
                    ↪ color=colors[t], linewidth=1)
    ax.set_xlabel('Time')
    ax.set_ylabel('Price of the underlying asset')
```

```
    ax.set_title(title)
    plt.show()

S0 = 100
asset_tree,_ = binomial_tree(S0, K, T, r, sigma, option_type="call", steps=20)
plot_binomial_tree(asset_tree, "Binomial Tree for Asset Prices")
```

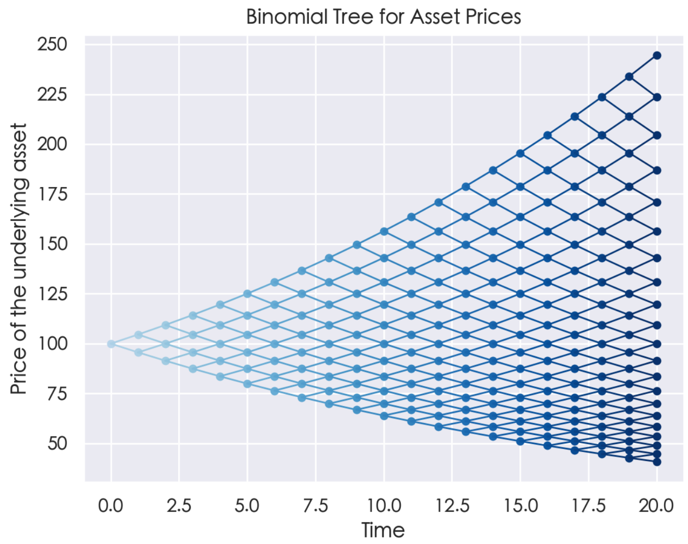

Figure 2.4: Binomial Tree for Asset Prices

2.3 Discrete Model for American Options

The binomial model can be used to price American options. The only difference is that the option can be exercised at any time before maturity. The price of an American option is the maximum of the price of the option at that time and the payoff of the option at that time. We could have seen the difference in the code above.

If exercised at time $t < T$,

- an American call option has payoff $(S_t - K)^+$,
- an American put option has payoff $(K - S_t)^+$.

A rational investor will always exercise an option **if the value of doing so (given by the intrinsic value) is greater than the value of holding onto it (the value of replicating the one step option)**. That is, at each node compare exercise value of option and market value (continuation value) of option. For Amrican Put, that is:

$$\mathbf{V}[i,n] = \max\{K - \mathbf{S}[i,n], 0\}$$

and for $0 \leq j < n$,

$$\mathbf{V}[i,j] = \max\left[\frac{q_i^u \mathbf{V}[i,j+1] + \left(1 - q_i^u\right)\mathbf{V}[i+1,j+1]}{e^{r\Delta t}}, (K - \mathbf{S}[i,j])^+\right]$$

Example 4. *Set discrete rate $r = 0.25$, and risk-neutral probabilities $q_u = q_d = 0.5$. Consider a 2-period binomial model with $S_0 = 4, K = 5$ for an American Put option. The binomial tree is:*

$$\begin{pmatrix} 4 & 8 & 16 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

The option value satisfies:

$$v_2(s) = \max\{5 - s, 0\}$$

$$v_n(s) = \max\left\{\frac{2}{5}\left[v_{n+1}\left(\frac{s}{2}\right) + v_{n+1}(2s)\right], 5 - s\right\}, \quad n = 0, 1$$

Hence the corresponding option values are:

$$\begin{pmatrix} 1.36 & 0.4 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{pmatrix}$$

Call the function we defined above, setting the American flag to True.

```
_, call_price = binomial_tree(S0, K, T, r, sigma, option_type="call", steps=steps,
    American=True)
_, put_price = binomial_tree(S0, K, T, r, sigma, option_type="put", steps=steps,
    American=True)
```

Figure 2.5: American Option

We can see that **An American option is at least as valuable as a European option**. Let C_E and C_A be the prices of European and American call options, respectively. Then $C_A \geq C_E$. The same holds for put options.

Lemma 2.1. ***Bounds for Call Options***:*Suppose the market is arbitrage-free. Then, we have*

$$\left(S_t - Ke^{-r(T-t)}\right)^+ \overset{(1)}{\leq} C_E\left(S_t, t\right) \leq C_A\left(S_t, t\right) \overset{(2)}{\leq} S_t$$

Proof.

1. At time t, buy C_E, sell S_t, invest $Ke^{-r(T-t)}$ at the riskless rate in the currency market. This portfolio gives $V_t = C_E\left(S_t, t\right) - S_t + Ke^{-r(T-t)}$. The payoff at time T is $V_T = C_E\left(S_T, T\right) - S_T + K$. Since $V_T \geq 0$, due to Proposition 2, we have $C_E\left(S_t, t\right) \geq \left(S_t - Ke^{-r(T-t)}\right) \geq \max(0, S_t - Ke^{-r(T-t)})$.

2. Suppose $C_A(S_t, t) > S_t$. Then, at time t, buy S_t, sell C_A, we get a positive cash flow $C_A(S_t, t) - S_t > 0$. At time $\tau \leq T$, handover S_τ and receive K, we have

$$S_\tau - (S_\tau - K)^+ > 0$$

This is an arbitrage portfolio, which contradicts the assumption that the market is arbitrage-free. Thus, $C_A(S_t, t) \leq S_t$.

□

For American Call, we have the following theorem:

Theorem 2.5. *The option to exercise the call option early is worthless. If stock pays no dividends,then*

$$C_A\left(S_t, t\right) = C_E\left(S_t, t\right)$$

Proof. If exercised early, we have

$$(S_\tau - K)^+ \leq (S_\tau - Ke^{-r(T-\tau)})^+ \leq C_E(S_\tau, \tau)$$

It means the exercise value is less than the continuation value. Thus, the option to exercise the call option early is worthless. □

Lemma 2.2. ***Bounds for Put Options***:*Suppose the market is arbitrage-free. Then, we have*

$$\left(Ke^{-r(T-t)} - S_t\right)^+ \overset{(1)}{\leq} P_E\left(S_t, t\right) \leq P_A\left(S_t, t\right) \overset{(2)}{\leq} K.$$

Proof. 1. By put-call parity, we have

$$C_E\left(S_t, t\right) = P_E\left(S_t, t\right) + S_t - Ke^{-r(T-t)} \geq 0$$

Thus, $P_E\left(S_t, t\right) \geq \left(Ke^{-r(T-t)} - S_t\right)^+$.

2. Suppose $P_A(S_t, t) > K$, then at time t, sell P_A, at time $\tau \leq T$, buy S_τ, pay K, gaining $P_A - K > 0$. This forms an arbitrage, contradicting the no-arbitrage assumption. Thus, $P_A(S_t, t) \leq K$.

□

Strategy for American Put Options

Consider the random variable

$$\tau = \min\left\{t \geq 0 : S_t \leq K - Ke^{-r(T-t)}\right\}.$$

Strategy is to exercise at time $\min\{\tau, T\}$.

Recap that $P_E\left(S_t, t\right) < Ke^{-r(T-t)}$.Then,

1. If $\tau < T$, it is better to exercise the option early, cause by definition,

$$K - S_t \geq K - Ke^{-r(T-t)} > P_E\left(S_t, t\right)$$

2. If $\tau \geq T$, the payoff is equal to that of the European option.

2.4 The Trinomial Model

The trinomial model is an extension of the binomial model. The stock price can move up, down, or stay the same. The model is constructed as follows:

The price of the stock at time 0 is $S_0 = s$. The price of the stock at time T is $S_T = s \cdot X$, where

$$X = \begin{cases} u & \text{with probability } p_u \\ 1 & \text{with probability } p_e \\ d & \text{with probability } p_d \end{cases}$$

Also, $u > e^{rT} > d$ and $p_u + p_e + p_d = 1$.

Example 5. *Assume $r = 0, u = 1.1, d = 0.9$ and $s = 100$. The price of the stock today is 100 . The price at $t = T$ can be 90,100 or 110 . What is the price of a call with strike 100?*

By replicating the option, we have

$$\begin{aligned} \alpha + \beta \cdot 90 &= 0 \\ \alpha + \beta \cdot 100 &= 0 \\ \alpha + \beta \cdot 110 &= 10 \end{aligned}$$

The system has no solution There is no replicating portfolio.

Proposition 2.2. *The trinomial model is incomplete.*

In spite of this fact, we can still find some price bounds for the option.

Definition 2.2. *A portfolio h is said to* ***super-replicate*** *the financial derivative with payoff $\varphi\left(S_T\right)$ if $V_T^h \geq \varphi\left(S_T\right)$.*

A portfolio h is said to ***sub-replicate*** *the financial derivative with payoff $\varphi\left(S_T\right)$ if $V_T^h \leq \varphi\left(S_T\right)$.*

These two portfolios give upper and lower bond for the price of the derivative.

Figure 2.6: Trinomial Model

In the previous example: as shown in 2.6, The three blue dots represent the payoff of the option for the three possible outcomes.

The red line represents the cheapest super-replicating portfolio, $h = (-45, 1/2)$ with $V_0^h = -45 + 100/2 = 5$.

The green lines represent the most expensive sub-replicating portfolios, $h^* = (0, 0)$ and $h = (-100, 1)$ with $V_0^{h*} = V_0^h = 0$.

Any price between 0 and 5 is no-arbitrage price.

For any $q \in (0, 1)$, the probability measure $\mathbb{Q}$ s.t. $\mathbb{Q}\left(S_T = 90\right) = \mathbb{Q}\left(S_T = 110\right) = (1-q)/2$ and $\mathbb{Q}\left(S_T = 100\right) = q$ is a risk-neutral measure, as $\mathbb{E}^{\mathbb{Q}}\left[S_T\right] = S_0$

For any risk-neutral measure $\mathbb{Q}, e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\varphi\left(S_T\right)\right]$ is a no-arbitrage price.

$$\mathbb{E}^{\mathbb{Q}}\left[\left(S_T - 100\right)^+\right] = 10\frac{1-q}{2} \in [0, 5]$$

3 Math Preliminaries

For continous-time models, some math preliminaries are needed. We will introduce the following concepts:

1. Stochastic Processes
2. Brownian Motion
3. Itô's Lemma
4. Stochastic Differential Equations.

3.1 Stochastic Processes

Definition 3.1. *A **stochastic process** is a collection of random variables $\{X_t : t \in T\}$, where T is the index set. The index set can be discrete or continuous.*

Definition 3.2. *Given a random variable Y with $E[|Y|] < \infty$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and some sub- σ-field $\mathcal{G} \subset \mathcal{F}$, we will define the **conditional expectation** as the almost surely unique **random variable** $\mathbb{E}[Y \mid \mathcal{G}]$ which satisfies the following two conditions:*

1. *$\mathbb{E}[Y \mid \mathcal{G}]$ is $\mathcal{G}$ -measurable;*
2. *$\mathbb{E}[YZ] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]Z]$ for all Z which are bounded and $\mathcal{G}$-measurable.*

$\mathbb{E}[Y \mid \mathcal{G}]$ is an averaging of Y that erases information finer than the information contained in $\mathcal{G}$.

Consider a continuous time process $\{X_s\}$. Information about the past history of X_s up to time t is encoded in the σ-algebra F_t.

Definition 3.3. *The expression $\{\mathcal{F}_t, t \geq 0\}$ is called the **filtration** generated by $\{X_t\}$. More rigorously,*

$$\mathcal{F}_t = \sigma\left(X_s, 0 \leq s \leq t\right)$$

is the σ-algebra generated by $\{X_s, 0 \leq s \leq t\}$.

Notes: a σ-algebra on a set X is a nonempty collection Σ of subsets of X closed under complement, countable unions, and countable intersections. And filtration: ordered (in time) collections of subsets of events.

Martingales

Definition 3.4. *A stochastic process $\{X_t\}$ is a **martingale** with respect to the filtration $\{\mathcal{F}_t\}$ if*

1. *$\mathbb{E}\left[|X_t|\right] < \infty$ for all t,*
2. *$\mathbb{E}\left[X_t \mid \mathcal{F}_s\right] = X_s$ for all $s \leq t$.*

If the equality in 2. is replaced by $\leq$, then $\{X_t\}$ is a **supermartingale**. If the equality in 2. is replaced by $\geq$, then $\{X_t\}$ is a **submartingale**.

Simulations of martingales are shown as 3.1.

We can see that, supermartingales have a downward trend, while submartingales have an upward trend. We have the following theorem:

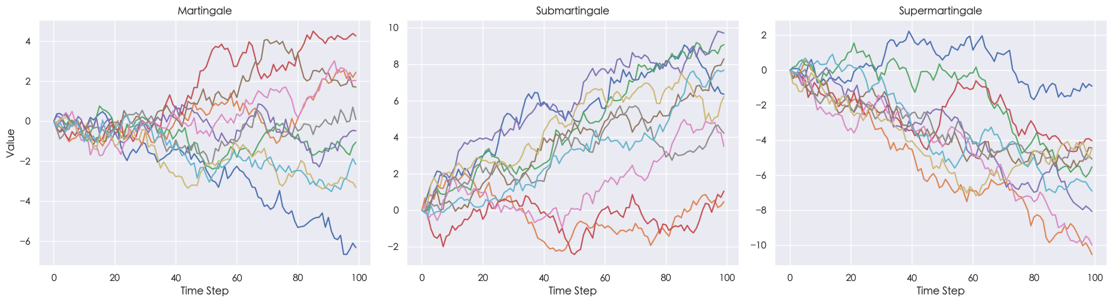

Figure 3.1: Martingales

Theorem 3.1. *If $\{X_t\}$ is a martingale, then $\mathbb{E}[X_t] = \mathbb{E}[X_0]$ for all t.*

Example 6. (***Doob martingale***). *If Z is a random variable with $\mathbb{E}[|Z|] < \infty$. Then we can easily define a martingale by letting $X_t = \mathbb{E}[Z \mid \mathcal{F}_t]$. Indeed, we have for every $0 < s < t < \infty$, almost surely,*

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = \mathbb{E}[\mathbb{E}[Z \mid \mathcal{F}_t] \mid \mathcal{F}_s] = \mathbb{E}[Z \mid \mathcal{F}_s] = X_s.$$

Example 7. *Consider a binomial model. Set the risk-free rate $r > 0$. The stock price can be written as*

$$S_{n+1} = XS_n = \begin{cases} uS_n & \text{with probability } p \\ dS_n & \text{with probability } 1-p \end{cases}, \quad n = 0, 1, 2, \ldots$$

where $0 < d < e^{r\Delta t} < u$. Consider the process

$$M_n = e^{-rn} S_n, \quad n \geq 1$$

Let $\mathbb{E}(X) = \mu$, Compute $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]$:

$$\begin{aligned} \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}\left[e^{-r(n+1)} S_{n+1} \mid \mathcal{F}_n\right] \\ &= e^{-r(n+1)} \mathbb{E}[XS_n \mid \mathcal{F}_n] \\ &= e^{-r(n+1)} S_n \mu \\ &\overset{\text{if it's martingale}}{=\!=\!=\!=} e^{-rn} S_n = M_n \end{aligned}$$

Hence we can solve out $\mu = e^r$, and thus

$$p = \frac{e^r - d}{u - d}$$

So, under risk-neutral probabilities, the discounting process of stocks is a martingale.

3.2 Brownian Motion (Wiener Process)

Definition 3.5. *A stochastic process $W = (W_t)_{t \geq 0}$ is called a **Wiener process** if*

1. *$W_0 = 0$;*
2. *W has independent increments, i.e. if $r < s \leq t < u$ the $W_u - W_t$ and $W_s - W_r$ are independent;*
3. *For $s < t$, $W_t - W_s$ has distribution $\mathcal{N}(0, t-s)$;*
4. *Every path $t \to W_t$ is continuous (with probability 1).*

- Increments are stationary: $W_{t+\Delta t} - W_t$ and $W_{\Delta t} - W_0 = W_{\Delta t}$ have the same distribution $\mathcal{N}(0, \Delta t)$.
- Increments are independent: For $0 < t_1 < t_2 < \cdots < t_m < \infty$, $W_{t_1} - W_{t_0} W_{t_2} - W_{t_1}, \cdots, W_{t_m} - W_{t_{m-1}}$ are independent.

The following theorem states that the Wiener process is the continuous time version of a Gaussian random walk.

Theorem 3.2. *Let $X_1, X_2, \cdots$ be i.i.d random Gaussian variables with mean 0 and variance* 1*. Then, the process*

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

is called a ***Gaussian random walk****.* ***Donsker's Theorem*** *states that the process*

$$W_t = \lim_{n\to\infty} \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}$$

is a Wiener process.

Using this theorem, we can simulate a Wiener process. The following Python code simulates a Wiener process.

```python
import numpy as np
def wiener_process(T, N):
    dt = T / N
    dW = np.sqrt(dt) * np.random.randn(N)
    W = np.cumsum(dW)
    return W
```

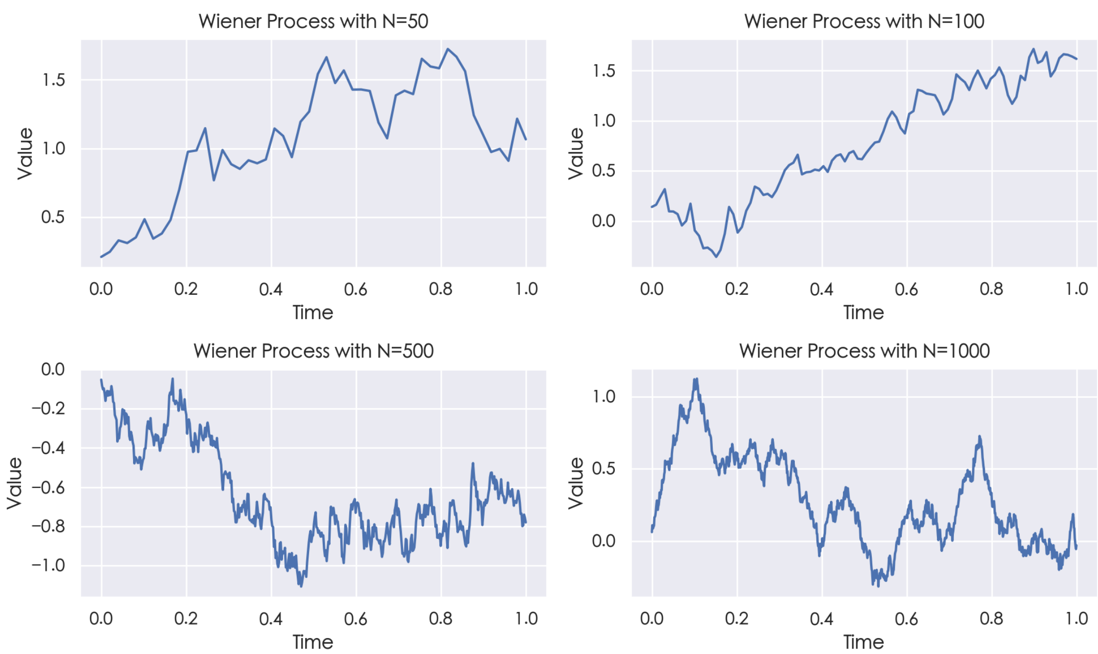

Figure 3.2: Wiener Process

Proposition 3.1. *Some basic properties of Wiener process:*

- *The trajectory of a Wiener process is continuous but nowhere differentiable.*
- *If $s < t$, then $\mathrm{Cov}(W_s, W_t) = \min(s,t) = s$.*
- *W_t is a martingale.*

Proof. 1. For Brownian motion,

$$\left\| \frac{W_{t+h} - W_t}{h} \right\| \approx \frac{\sqrt{h}}{h} = h^{-1/2} \to \infty \text{ as } h \to 0$$

Hence, the trajectory is nowhere differentiable.

2. For the covariance,

$$\mathrm{Cov}\left(W_s, W_t\right) = \mathbb{E}\left[W_s W_t\right] = \mathbb{E}\left[W_s\left(W_t - W_s + W_s\right)\right] = 0 + \mathbb{E}\left[W_s^2\right] = s$$

3. For the martingale property,

$$\mathbb{E}[W_t \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s + W_s \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s \mid \mathcal{F}_s] + \mathbb{E}[W_s \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s] + W_s = 0 + W_s = W_s$$

□

Note 2. *The covariance matrix of a Wiener process is given by*

$$\mathbf{Var}(W(t_1), W(t_2), \cdots, W(t_n)) = \begin{pmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & \cdots & t_2 \\ \vdots & \vdots & \ddots & \vdots \\ t_1 & t_2 & \cdots & t_n \end{pmatrix}$$

Theorem 3.3. ***Scaling Variance Property****: For any $a > 0$, the process $X_t = \frac{1}{a} W_{a^2 t}$ is a Wiener process.*

Proof. We verify the four characteristic properties.

- $X_0 = \frac{1}{a} W_0 = 0$.
- Take $r < s \leq t < u$. Then $a^2 r < a^2 s \leq a^2 t < a^2 u$. Therefore, $W_{a^2 u} - W_{a^2 t}$ and $W_{a^2 s} - W_{a^2 r}$ are independent, and also $X_u - X_t$ and $X_s - X_r$ are independent.
- $W_{a^2 t} - W_{a^2 s}$ has distribution $\mathcal{N}\left(0, a^2 t - a^2 s\right)$. Therefore, $X_t - X_s$ has distribution $\mathcal{N}(0, t - s)$.
- $t \to X_t$ can be written as composition of continuous functions: $t \to a^2 t \to W_{a^2 t} \to \frac{1}{a} W_{a^2 t}$.

□

Theorem 3.4. ***Time Reversal Property****: For any t_0, the process $X_t = W_{t_0 + t} - W_{t_0}$ is a Wiener process.*

Proof.

- $X_0 = 0$.
- Take $r < s \leq t < u$. Then $X_u - X_t = W_{t_0+u} - W_{t_0} - W_{t_0+t} + W_{t_0} = W_{t_0+u} - W_{t_0+t}$ and $X_s - X_r = W_{t_0+s} - W_{t_0} - W_{t_0+r} + W_{t_0} = W_{t_0+s} - W_{t_0+r}$ are independent.
- $X_t - X_s = W_{t_0+t} - W_{t_0+s}$ has distribution $\mathcal{N}(0, t - s)$.
- $t \to X_t$ can be written as composition of continuous functions: $t \to t_0 + t \to W_{t_0+t} \to W_{t_0+t} - W_{t_0}$.

□

Theorem 3.5. ***Markov Property****: Weiner process has the Markov property. That is, for any $0 \leq s < t$, the conditional distribution of W_t given $\mathcal{F}_s$ is the same as the conditional distribution of W_t given W_s.*

Proof. It is obvious, as $W_t - W_s$ is independent of $\mathcal{F}_s$ and has distribution $\mathcal{N}(0, t - s)$. So, given $W_s = w_s$, W_t has distribution $w_s + \mathcal{N}(0, t - s)$, only depending on w_s and $t - s$. □

Definition 3.6. ***Transition Density****: Given $W_s = w_s$, W_t has density function*

$$p(x|w_s) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(x - w_s)^2}{2(t-s)}\right)$$

3.3 Quadratic Variation

Definition 3.7. *The* ***quadratic variation*** *of a stochastic process $\{X_t\}$ is defined as*

$$\langle X, X \rangle[T] = \lim_{\|\Pi\| \to 0} \sum_{i=1}^{n} \left(X_{t_i} - X_{t_{i-1}}\right)^2$$

where $t_0 = 0 < t_1 < \cdots < t_n = T$ is a partition Π of $[0,T]$ and $\|\Pi\| = \max_{i=1}^{n} |t_i - t_{i-1}|$.

Note 3. *Similarly, we define variation and covariation:*

$$\langle X, Y\rangle[T] = \lim_{\|\Pi\|\to 0} \sum_{i=1}^{n} \left(X_{t_i} - X_{t_{i-1}}\right)\left(Y_{t_i} - Y_{t_{i-1}}\right)$$
$$|X|\,[T] = \lim_{\|\Pi\|\to 0} \sum_{i=1}^{n} \left|X_{t_i} - X_{t_{i-1}}\right|$$

Proposition 3.2. *For any function f with continuous first derivative, i.e. $f'(t)$ is continuous on the interval $[0,T]$, the quadratic variation of f is always zero.*

Proof.

$$\begin{aligned}\langle f, f\rangle[T] &= \lim_{\|\Pi\|\to 0} \sum_{i=1}^{n} \left(f\left(t_i\right) - f\left(t_{i-1}\right)\right)^2 \\ &= \lim_{\|\Pi\|\to 0} \sum_{i=1}^{n} f'^2\left(\xi_i\right)\left(t_i - t_{i-1}\right)^2 \\ &\le \|\Pi\| \sum_{i=1}^{n} (f'(\xi_i))^2 (t_i - t_{i-1}) \to 0 \cdot \int_0^T |f'(t)|^2 dt = 0\end{aligned}$$

□

Since a Brownian motion does not even have derivatives, the quadratic variation may not be zero.

The quadratic variation will play an important role when we discuss the Itô integral. The following result shows that indeed the quadratic variation of a Brownian motion is not zero.

Theorem 3.6. ***Multiplication table for Brownian motion**: We have the following quadratic variations / covariations:*

1. *$\langle W, W\rangle[T] = T$, i.e., $dW_t dW_t = dt$.*
2. *$\langle W_t, t\rangle[T] = 0$, i.e., $dW_t dt = 0$.*
3. *$dtdt = 0$.*

To show this we need a concept of convergenve in L^2.

Definition 3.8. *A sequence of random variables $\{X_n\}$ is said to **converge in** L^2 to a if*

$$\mathbb{E}(X_n) = a, \quad \mathrm{Var}(X_n) \to 0$$

Then we write $X_n \xlongequal{L^2} a$.

Now we can show

Proof. 1. Let $Y_n = \sum_{i=0}^{n-1} \left(W_{t_{i+1}} - W_{t_i}\right)^2$. Then

$$\mathbb{E}\left(Y_n\right) = \sum_{i=0}^{n-1} \mathbb{E}\left[\left(W_{t_{i+1}} - W_{t_i}\right)^2\right] = \sum_{i=0}^{n-1} \left(t_{i+1} - t_i\right) = T$$

$$\begin{aligned}\mathrm{Var}\left(Y_n\right) &= \sum_{i=0}^{n-1} \mathrm{Var}\left[\left(W_{t_{i+1}} - W_{t_i}\right)^2\right] \\ &= \sum_{i=0}^{n-1} \mathbb{E}\left[\left(W_{t_{i+1}} - W_{t_i}\right)^4\right] - \sum_{i=0}^{n-1} \left[\mathbb{E}\left[\left(W_{t_{i+1}} - W_{t_i}\right)^2\right]\right]^2\end{aligned}$$

To continue we can use a trick: if $X \sim \mathcal{N}(0, \sigma)$, then $X^2/\sigma \sim \chi^2(1)$, and $\mathbb{E}\left[X^4\right] = 3\sigma^2$. Therefore,

$$\begin{aligned}\mathbb{E}\left[\left(W_{t_{i+1}} - W_{t_i}\right)^4\right] &= 3\left(t_{i+1} - t_i\right)^2 \\ \operatorname{Var}\left(Y_n\right) &= 3\sum_{i=0}^{n-1}\left(t_{i+1} - t_i\right)^2 - \sum_{i=0}^{n-1}\left(t_{i+1} - t_i\right)^2 \\ &= 2\sum_{i=0}^{n-1}\left(t_{i+1} - t_i\right)^2 \to 0\end{aligned}$$

Hence we obtian $Y_n \stackrel{L^2}{=\!=} T$ which implies $dW_t dW_t = dt$.

2. This follows as

$$\sum_{i=0}^{n-1}\mathbb{E}\left[\left(W_{t_{i+1}} - W_{t_i}\right)\left(t_{i+1} - t_i\right)\right] = 0$$

$$\text{and } \operatorname{Var}\underbrace{\left[\sum_{i=0}^{n-1}\left(W_{t_{i+1}} - W_{t_i}\right)\left(t_{i+1} - t_i\right)\right]}_{\text{The terms inside the summation are independent}} = \sum_{i=0}^{n-1}\operatorname{Var}\left[\left(W_{t_{i+1}} - W_{t_i}\right)\left(t_{i+1} - t_i\right)\right]$$

$$= \sum_{i=0}^{n-1}\left(t_{i+1} - t_i\right)^2\operatorname{Var}\left[\left(W_{t_{i+1}} - W_{t_i}\right)\right] = \sum_{i=0}^{n-1}\left(t_{i+1} - t_i\right)^3$$

$$\leq \|\Pi\|^2\sum_{i=0}^{n-1}\left(t_{i+1} - t_i\right) = \|\Pi\|^2 T \to 0$$

implying $dW_T dT = 0$.

3. This is trivial.

□

Consider two independent Brownian motions W_t and $\tilde{W}_t$. Let V_t be a process with finite variation.Then the mutiplication table is given by:

Table 3.1: Multiplication Table

	dt	dW_t	$d\tilde{W}_t$	dV_t
dt	0	0	0	0
dW_t	0	dt	0	0
$d\tilde{W}_t$	0	0	dt	0
dV_t	0	0	0	0

Use the following Python code to simulate quadratic variation.

```python
def quadratic_variation(path):
    return np.cumsum(np.diff(path, prepend=0)**2)
    # prepend=0 is used to add a 0 at the beginning of the array

W = wiener_process(1, n) # Use the function defined above
QV_W = quadratic_variation(W)

def quadratic_variation_smooth(f, time):
    f_diff = np.gradient(f)
    return np.cumsum(f_diff**2 * (time[1] - time[0]))

def a_smooth_function(t):
    return np.sin(2 * np.pi * t)+np.exp(-t*np.cos(2*np.pi*t))

time = np.linspace(0, 1, n)
f = a_smooth_function(time)
QV_f = quadratic_variation_smooth(f, time)
```

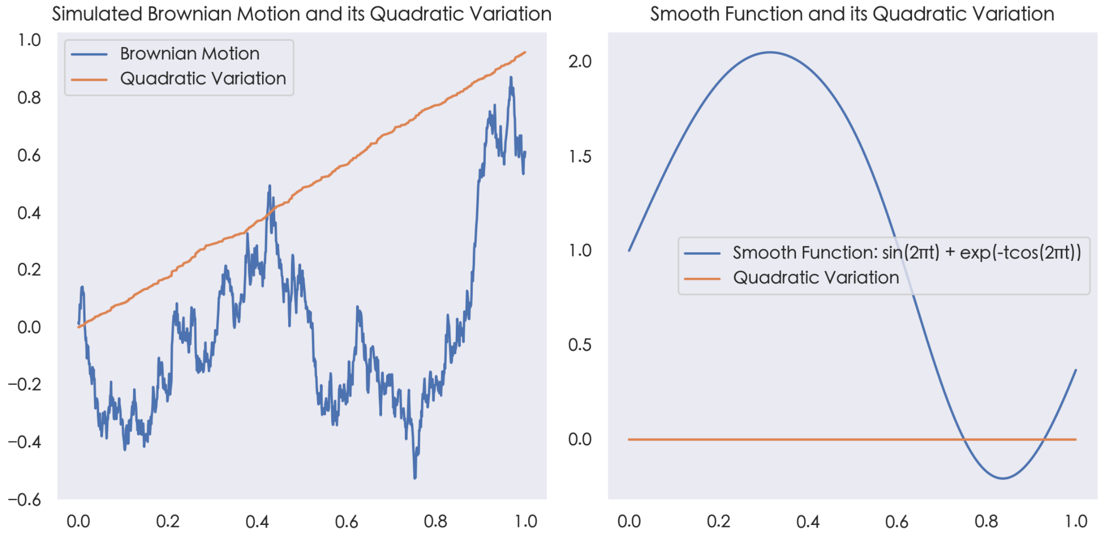

Figure 3.3: Quadratic Variation Comparison

3.4 Itô's Integral

Assume that from time t to time $t+\Delta t$ the price of a stock evolves as

$$S_{t+\Delta t} = S_t + \mu(t, S_t)\Delta t + \sigma(t, S_t)\Delta W_t.$$

The price of the stock at time T is

$$\begin{aligned} S_T &= S_0 + \sum_{i=0}^{k-1} \mu\left(t_i, S_{t_i}\right)\Delta t_i + \sum_{i=0}^{k-1} \sigma\left(t_i, S_{t_i}\right)\Delta W_{t_i} \\ &\approx S_0 + \int_0^T \mu(t, S_t)\, dt + \int_0^T \sigma(t, S_t)\, dW_t \end{aligned}$$

The goal is to study the integral $\int_0^T \sigma(t, S_t)\, dW_t$.
Such a process will be called an **Itô process**. Equivalently, we write

$$dS_t = \mu(t, S_t)\, dt + \sigma(t, S_t)\, dW_t,$$

which is called a **stochastic differential equation**.

The Brownian motion is almost surely nowhere differentiable. Hence, the meaning of dWt is not clear. Let's consider Riemann-Stieltjes integrals first:

$$\int_0^T X(t)dG(t) = \lim_{n\to\infty} \sum_{i=0}^{n-1} X(\zeta_i)\left(G(t_{i+1}) - G(t_i)\right)$$

In calculus, the integral can be defined if the limit exists for **any** sequence ζ_i as long as the norm of the partition Π, i.e. $\|\Pi\| = \max_{0\le i\le n-1} |t_{i+1} - t_i|$, goes to zero. However, this is not the case for Brownian motion.

Consider for example $\int_0^T W(t)dW_t$.Suppose the limit exists. Then as $\|\Pi\| \to 0$, we must have

$$\lim \sum_{i=1}^{n} W(t_i)\left(W(t_{i+1}) - W(t_i)\right) = \lim \sum_{i=1}^{n} W(t_{i+1})\left(W(t_{i+1}) - W(t_i)\right)$$

as we can choose $\zeta_i \in [t_i, t_{i+1}]$ arbitrarily, first $\zeta_i = t_i$ and then $\zeta_i = t_{i+1}$. This implies that

$$\lim \sum_{i=1}^{n} \left(W(t_{i+1}) - W(t_i)\right)\left(W(t_{i+1}) - W(t_i)\right) = 0$$

But we have shown that the quadratic variation of the Brownian motion is T, i.e.

$$\lim \sum_{i=1}^{n} \left(W(t_{i+1}) - W(t_i)\right)^2 = T$$

in probability, yielding a contradiction.

In summary, in general we **CANNOT** interpret $\int_0^T X(t)dW_t$ as a Riemann-Stieltjes integral from calculus, and a different approach is needed.

Define Itô's Integral of Simple Processes

Definition 3.9. *A process X_t is called **simple** if it can be written as*

$$X_t = \sum_{i=0}^{n-1} \xi_i \mathbb{1}_{[t_i, t_{i+1})}(t)$$

where ξ_i are $\mathcal{F}_{t_i}$-measurable random variables.

A simple process is a step function, which is constant on each interval $[t_i, t_{i+1})$. Shown in 3.4 is an example of a simple process.

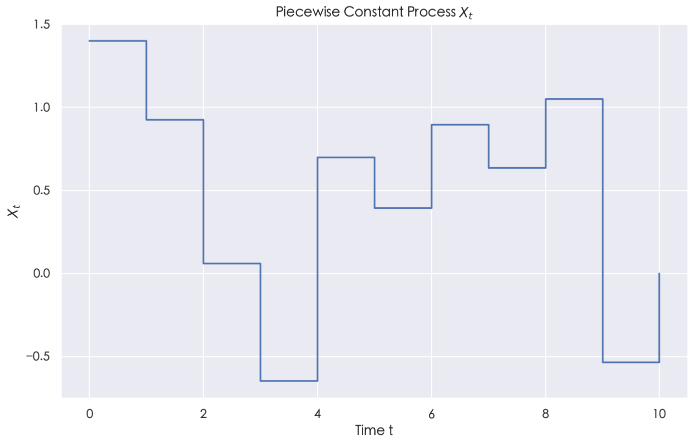

Figure 3.4: Simple Process

The stochastic integral of a simple process is defined as

$$\begin{aligned}\int_0^t X_s dW_s &= \sum_{j=0}^{m-1} \xi_j \left(W_{t_{j+1}} - W_{t_j}\right) + \xi_m \left(W_t - W_{t_m}\right), \text{ if } t_m < t \leq t_{m+1} \\ &= \sum_{j=0}^{n-1} \xi_j \left(W_{t \wedge t_{j+1}} - W_{t \wedge t_j}\right), \text{ where } a \wedge b := \min(a, b)\end{aligned}$$

We define the stochastic integral by using ξ_j, which is adapted to $\mathcal{F}_{t_j}$ (information known up to t_j) at the **left-end point** of the interval $\left[t_j, t_{j+1}\right]$. This is an important feature of Itô integral for two reasons.

1. This is consistent with our requirement that one cannot use any information from the future, Only information from the past and the present.

2. The integral so defined will be a (local) martingale.

Proposition 3.3. *Such defined Itô integral has the following properties:*

1. *Linearity:* $\int_0^t (aX_s + bY_s)\, dW_s = a \int_0^t X_s dW_s + b \int_0^t Y_s dW_s$.

2. ***Martingale property**:* $\int_0^t X_s dW_s$ *is a martingale, and* $\mathbb{E}\left[\int_0^t X_s dW_s\right] = 0$.

3. ***Itô isometry**:* $\mathbb{E}\left[\left(\int_0^t X_s dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t X_s^2 ds\right]$.

Proof. 1. Denote $\Delta W_j = W_{t_{j+1}} - W_{t_j}$. Suppose that $t_n < s \le t_{n+1}$ and $t_m < t \le t_{m+1}, s \le t$ (whence $n \le m$).

$$\begin{aligned}\mathbb{E}\left[\int_0^t X_u dW_u \mid \mathcal{F}_s\right] &= \mathbb{E}\left[\sum_{i=0}^{m-1} \xi_i \Delta W_i + \xi_m \left(W_t - W_{t_m}\right) \mid \mathcal{F}_s\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{n-1} \xi_i \Delta W_i + \xi_n \left(W_s - W_{t_n}\right) \mid \mathcal{F}_s\right]\end{aligned}$$

no matter whether $n < m$ or $n = m$. Therefore,

$$\mathbb{E}\left[\int_0^t X_u dW_u \mid \mathcal{F}_s\right] = \int_0^s X_u dW_u$$

and the proof is completed.

2. Suppose $i \neq j$. Without loss of generality, let us assume that $i < j$. Then

$$\mathbb{E}\left[\xi_i \xi_j \Delta W_j \Delta W_i\right] = \mathbb{E}\left[\xi_i \xi_j \Delta W_i \mathbb{E}\left[\Delta W_j \mid \mathcal{F}_{t_j}\right]\right] = \mathbb{E}\left[\xi_i \xi_j \Delta W_i \mathbb{E}\left[\Delta W_j\right]\right] = 0,$$

because ξ_i, ξ_j and ΔW_i is $\mathcal{F}_{t_j}-$ measurable as $j \ge i + 1$. If $i = j$, then

$$\mathbb{E}\left[\xi_i \xi_j \Delta W_j \Delta W_i\right] = \mathbb{E}\left[\xi_i^2 \mathbb{E}\left[(\Delta W_i)^2 \mid \mathcal{F}_{t_i}\right]\right] = \mathbb{E}\left[\xi_i^2\right](t_{i+1} - t_i),$$

as ξ_i^2 is $\mathcal{F}_{t_i}$-measurable. In summary,

$$\mathbb{E}\left[\xi_j \xi_i \Delta W_j \Delta W_i\right] = \begin{cases} 0, & \text{if } i \neq j; \\ \mathbb{E}\left[\xi_i\right](t_{i+1} - t_i), & \text{if } i = j. \end{cases}$$

Suppose that $t_m < t \le t_{m+1}$. Then

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^t X_s dW_s\right)^2\right] =& \mathbb{E}\left[\sum_{j=0}^{m-1}\sum_{i=0}^{m-1} \xi_i \xi_j \Delta W_j \Delta W_i\right] + \mathbb{E}\left[\xi_m^2 \left(W_t - W_{t_m}\right)^2\right] \\ &+ 2\mathbb{E}\left[\xi_m \left(W_t - W_{t_m}\right) \sum_{i=0}^{m-1} \xi_j \Delta W_j\right] \\ =& \sum_{i=0}^{m-1} \mathbb{E}\left[\xi_i^2\right](t_{i+1} - t_i) + \mathbb{E}\left[\xi_m^2\right](t - t_m) \\ =& \mathbb{E}\left[\sum_{i=0}^{m-1} \xi_i^2 (t_{i+1} - t_i) + \xi_m^2 (t - t_m)\right] \\ =& \mathbb{E}\left[\left(\int_0^t X_s^2 ds\right)\right]\end{aligned}$$

□

Theorem 3.7. ***Quadratic Variation of the Itô Integral**: The quadratic variation of the stochastic integral with respect to a simple process $I(t) = \int_0^t X_S dW_s$ equals*

$$\langle I, I\rangle[T] = \int_0^T X_s^2 ds$$

Proof. We first compute the quadratic variation of the stochastic integral on one of the subintervals $\left[t_j, t_{j+1}\right]$ on which X_s is equal to ξ_j. We can choose partition points

$$t_j = s_0 < s_1 < \cdots < s_k = t_{j+1}$$

and consider

$$\begin{aligned}\sum_{i=0}^{k-1} \left[I(s_{i+1}) - I(s_i)\right]^2 &= \sum_{i=0}^{k-1} \xi_j^2 \left[W_{s_i+1} - W_{s_i}\right]^2 \\ &= \xi_j^2 \sum_{i=0}^{k-1} \left[W_{s_i+1} - W_{s_i}\right]^2\end{aligned}$$

We let $\|\Pi\| = \max_{i=0,1,\cdots,k-1}(s_{i+1} - s_i) \to 0$; Then $\lim_{\|\Pi\|\to 0} \sum_{i=0}^{k-1} \left[W_{s_i+1} - W_{s_i}\right]^2 = t_{j+1} - t_j$.

Therefore, we have

$$\lim_k \sum_{i=0}^{k-1} \left[I(s_{i+1}) - I(s_i)\right]^2 = \xi_j^2 \left(t_{j+1} - t_j\right) = \int_{t_j}^{t_{j+1}} X_s^2 ds$$

where again we have used the fact that X_s is equal to ξ_j for $t_j < s \leq t_{j+1}$. Analogously, the quadratic variation of the stochastic integral I between times t_m and t is $\int_{t_m}^t X_S^2 ds$. Adding up all these pieces, we obtain it. □

Extend to General Processes

The definition of the Itô integral introduced for the simple processes can be extended to a larger class of processes that can be approximated by the simple processes.

Definition 3.10. *Denote $\mathcal{L}^2[0,T]$ to be the set of all $\{\mathcal{F}_t\}$-adapted processes X, for which the mean square for all $T > 0$.*

$$\mathbb{E}\left[\int_0^t X_s^2 ds\right] < \infty$$

Denote the set of all simple processes to be $\mathcal{D}$. We have the very important theorem:

Theorem 3.8. *For any $X \in \mathcal{L}^2[0,T]$, there exists a sequence of simple processes $\{X^n\} \subset \mathcal{D}$ such that*

$$\mathbb{E}\left[\int_0^T |X_s^n - X_s|^2\right] ds \to 0$$

as $n \to \infty$.

Use the following Python code, we can see how simple processes approximate a general process.

```python
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
BM = wiener_process(1, 1000)
def find_simple_process(process,n):
    t_points = np.linspace(0, 1, n)
    simple_process = np.zeros(n)
    for i in range(n):
        simple_process[i] = process[int(i*process.shape[0]/n)]
    return simple_process,t_points

steps=[10,25,50,100]
plt.figure(figsize=(10, 6))
sns.lineplot(x=np.linspace(0, 1, 1000), y=BM, label='Brownian
    Motion',color='skyblue',alpha=0.7)
for n in steps:
    simple_process,t_points = find_simple_process(BM,n)
    plt.step(t_points, simple_process, where='post',label=f'Simple Process with n={n}')
plt.xlabel('Time t')
plt.ylabel('$X_t$')
plt.title('Simple Process to Brownian Motion')
plt.legend()
plt.grid(True)
plt.show()
```

Now we can approximate $X \in \mathcal{L}^2[0,T]$ by simple processes $X^{(n)}$.

Note 4.

- *The stochastic integral $\int_0^T X_s^{(n)} dW_s$ is well-defined for $X^{(n)} \in D$.*
- *Therefore, it is natural to define $\int_0^T X_s dW_s$ as the limit of*

$$I_n(T) = \int_0^T X_s^{(n)} dW_s$$

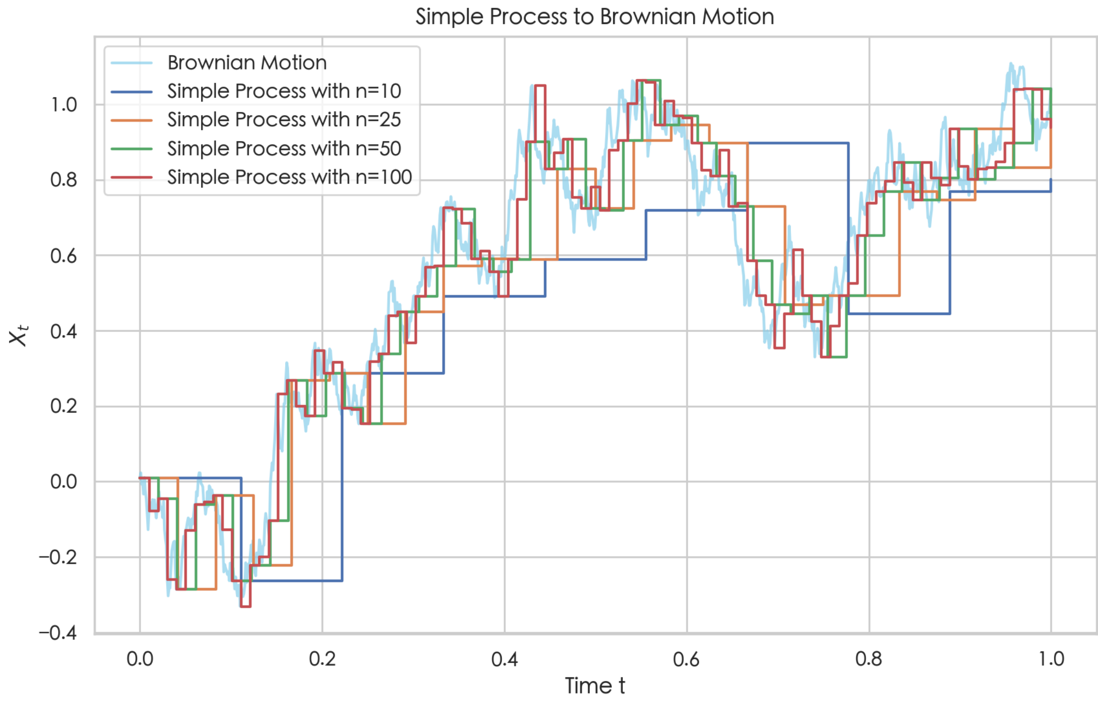

Figure 3.5: Simple Process to Brownian Motion

But we need to show that the ***limit exists and is unique****.*

First, we shall show that the sequence $\{I_n(T), n \geq 1\}$ forms a **Cauchy sequence** in $L^2(\mathbb{P})$, i.e.

$$\mathbb{E}\left[(I_n(T) - I_m(T))^2\right] \to 0, \quad n, m \to \infty.$$

Note that every Cauchy sequence in $L^2(\mathbb{P})$ converges. (Mathematically: The space is complete) This implies that there is a random variable, denoted by $I(T)$, such that $\mathbb{E}\left[(I(T))^2\right] < \infty$ and

$$\lim_{n\to\infty} \mathbb{E}\left[(I_n(T) - I(T))^2\right] = 0.$$

To show this, we have by the linearity of the stochastic integral

$$\begin{aligned}\mathbb{E}\left[(I_n(T) - I_m(T))^2\right] &= \mathbb{E}\left[\left(\int_0^T \left(X_s^{(n)} - X_s^{(m)}\right) dW_s\right)^2\right] \\ &= \mathrm{Var}\left(\int_0^T \left(X_s^{(n)} - X_s^{(m)}\right) dW_s\right)\end{aligned}$$

Therefore, by the variance formula for simple processes(Itô Isometry) we have

$$\mathbb{E}\left[(I_n(T) - I_m(T))^2\right] = \mathbb{E}\left[\int_0^T \left|X_s^{(n)} - X_s^{(m)}\right|^2 ds\right]$$

Using the fact that $(a+b)^2 \leq 2a^2 + 2b^2$, we have

$$\begin{aligned}&\mathbb{E}\left[\int_0^T \left|X_s^{(n)} - X_s^{(m)}\right|^2 ds\right] \\ =&\mathbb{E}\left[\int_0^T \left|X_s^{(n)} - X_s + X_s - X_s^{(m)}\right|^2 ds\right] \\ \leq& 2\mathbb{E}\left[\int_0^T \left|X_s^{(n)} - X_s\right|^2 ds\right] + 2\mathbb{E}\left[\int_0^T X_s - \left|X_s^{(m)}\right|^2 ds\right] \\ \to& 0.\end{aligned}$$

Thus the existence of the limit is shown.

Suppose we have two approximations by simple processes $X_s^{(n)}$ and $\tilde{X}_s^{(n)}$, such that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \left|X_s^{(n)} - X_s\right|^2 ds\right] = 0, \quad \lim_{n\to\infty} \mathbb{E}\left[\int_0^T \left|\tilde{X}_s^{(n)} - X_s\right|^2 ds\right] = 0,$$

Define

$$I_n(T) = \int_0^T X_s^{(n)} dW_s, \quad \tilde{I}_n(T) = \int_0^T \tilde{X}_s^{(n)} dW_s$$

Then we know from above there exist $I(T)$ and $\tilde{I}(T)$ such that

$$\lim_{n\to\infty} \mathbb{E}\left[\left(I_n(T) - I(T)\right)^2\right] = 0, \quad \lim_{n\to\infty} \mathbb{E}\left[\left(\tilde{I}_n(T) - \tilde{I}(T)\right)^2\right] = 0$$

Goal is to show

$$\mathbb{E}\left[(\tilde{I}(T) - I(T))^2\right] = 0$$

and, thus

$$\mathbb{P}(\tilde{I}(T) = I(T)) = 1$$

Since $\tilde{X}_s^{(n)} - X_s^{(n)}$ is still a simple process, we have by the variance formula of the Itô integral that

$$\begin{aligned}
&\mathbb{E}\left[\left(\tilde{I}_n(T) - I_n(T)\right)^2\right] \\
=&\mathbb{E}\left[\left(\int_0^T \tilde{X}_s^{(n)} dW_s - \int_0^T X_s^{(n)} dW_s\right)^2\right] \\
=&\mathbb{E}\left[\left(\int_0^T \left(\tilde{X}_s^{(n)} - X_s^{(n)}\right) dW_s\right)^2\right] \quad \text{(Itô Isometry)} \\
=&\mathbb{E}\left[\left(\int_0^T \left(\tilde{X}_s^{(n)} - X_s^{(n)}\right)^2 ds\right)\right] \quad \text{By } \left((a+b)^2 \le 2a^2 + 2b^2\right) \\
\le&2\mathbb{E}\left[\left(\int_0^T \left(\tilde{X}_s^{(n)} - X_s\right)^2 ds\right)\right] + 2\mathbb{E}\left[\left(\int_0^T \left(X_s - X_s^{(n)}\right)^2 ds\right)\right] \\
\to&0.
\end{aligned}$$

Since $(a+b+c)^2 \le 3\left(a^2 + b^2 + c^2\right)$, we have

$$\begin{aligned}
&\mathbb{E}\left[(\tilde{I}(T) - I(T))^2\right] \\
=&\mathbb{E}\left[\left(\tilde{I}(T) - \tilde{I}_n(T) + \tilde{I}_n(T) - I_n(T) + I_n(T) - I(T)\right)^2\right] \\
\le&3\mathbb{E}\left[\left(\tilde{I}(T) - \tilde{I}_n(T)\right)^2\right] + 3\mathbb{E}\left[\left(\tilde{I}_n(T) - I_n(T)\right)^2\right] + 3\mathbb{E}\left[\left(I_n(T) - I(T)\right)^2\right] \\
\to&0.
\end{aligned}$$

This completes the proof.

Now we can frankly define the Itô integral for general processes.

Definition 3.11. *For any $X \in \mathcal{L}^2[0,T]$, the* ***Itô integral*** *of X with respect to the Brownian motion W is defined as*

$$\int_0^T X_s dW_s := I(T)$$

where $I(T)$ is the limit of the sequence of stochastic integrals $I_n(T)$ with simple processes $X_s^{(n)}$.

$$\int_0^t X_s dW_s := I(T) = \lim_{n\to\infty} I_n(T) = \lim_{n\to\infty} \int_0^T X_s^{(n)} dW_s,$$

where $X_s^{(n)}$ is a sequence of simple processes converging to X in $L^2(\mathbb{P})$.

And properties of the Itô integral are inherited from the simple processes.

Proposition 3.4. *For $X_t, Y_t \in \mathcal{L}^2[0,T]$, the Itô integral has the following properties:*

1. *Adaptivity and Continuity: For each t, $I(T) = \int_0^t X_s dW_s$ is adapted to the past history, i.e. is $\mathcal{F}_t$-measurable, and $I(t)$ is a continuous function of t.*
2. *Linearity: $\int_0^t (aX_s + bY_s)\, dW_s = a\int_0^t X_s dW_s + b\int_0^t Y_s dW_s$.*
3. *Martingility: $\int_0^t X_s dW_s$ is an $\mathcal{F}_t$-martingale, and $\mathbb{E}\left[\int_0^t X_s dW_s\right] = 0$.*
4. *Itô Isometry: $\mathbb{E}\left[\left(\int_0^t X_s dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t X_s^2 ds\right]$.*
5. *Quadratic Variation: $\langle I, I\rangle[T] = \int_0^T X_s^2 ds$.*

Example 8. *Compute the variance of $\frac{1}{2}\int_0^T \sqrt{t} dW_t$. By Itô's isometry, we have*

$$\operatorname{Var}\left(\frac{1}{2}\int_0^T \sqrt{t} dW_t\right) = \frac{1}{4}\int_0^T t dt = \frac{1}{4}\frac{T^2}{2} = \frac{T^2}{8}$$

Corollary 3.1. *Let X be Riemann-integrable, left-continuous and adapted. Then*

$$\int_0^T X_s dW_s = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} X_{s_i}\left(W_{s_{i+1}} - W_{s_i}\right)$$

with $\Pi = \{s_0 = 0 < s_1 < \cdots < s_n = T\}$.

Proof. Consider the simple process $\eta_t^{(n)} = \sum_{j=0}^{n-1} X_{s_j} \cdot \mathbb{1}_{\{s_j < t \le s_{j+1}\}}$. Then, by definition of the Riemann-Integral:

$$\int_0^T \left(\eta_s^{(n)} - X_s\right)^2 ds \to 0 \text{ as } n \to \infty$$

Hence, by the uniqueness of the stochastic integral the assertion follows. □

3.5 Itô's Lemma

Itô's Lemma is the stochastic calculus version of the chain rule: If W_t were differentiable, then the ordinary chain rule would give:

$$\frac{d}{dt}V(W_t) = V'(W_t) \cdot W_t'$$

or in differential notation:

$$dV(W_t) = V'(W_t)\, W_t' dt = V'(W_t)\, dW_t.$$

However, W_t is nowhere differentiable, and the meaning of dW_t is not clear. We have seen that due to the non-zero quadratic variation of Brownian motion, the definition of integrals with respect to Brownian motion is not straightforward like in calculus. Different selections of partition points can lead to different limits, and the limit may not exist. Then the derivative of a function of Brownian motion will be affected too.

The question in general is: Given an SDE: $dS_t = \mu(t, S_t)\, dt + o(t, S_t)\, dW_t$, and a function $V(t, S_t)$, can we write an SDE for V ?

$$dV(t, S_t) = (\quad)dt + (\quad)dW_t$$

The answer is yes, and this is given by Itô's Lemma.

Theorem 3.9. *Let $V(t,S)$ be a $\mathbf{C}^{1,2}$ function and let S_t satisfy the SDE*

$$dS_t = \mu(t, S_t)\, dt + \sigma(t, S_t)\, dW_t.$$

Then

$$dV(t, S_t) = \left(\frac{\partial V}{\partial t} + \mu(t, S_t)\frac{\partial V}{\partial S} + \frac{1}{2}\sigma(t, S_t)^2\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma(t, S_t)\frac{\partial V}{\partial S}dW_t$$

Note 5. *Actually, we write*

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS_t)^2,$$

and use the multiplication table of the Brownian motion to simplify the term $(dS_t)^2$.

Now we need to prove the Itô's Lemma heuristically and non-technically.

Proof. Consider a partition Π with $0 = t_0 < t_1 < \cdots < t_n = T$ and make a Taylor expansion of $V(t, S_t)$.

Assume that all derivatives of V as well as μ and σ are bounded. Moreover μ and σ are simple processes. The general case can then be obtained by approximation. We have

$$\begin{aligned}
V\left(t_{i+1}, S_{t_{i+1}}\right) - V\left(t_i, S_{t_i}\right) &= \frac{\partial V}{\partial t}\left(t_i, S_{t_i}\right)(t_{i+1} - t_i) + \frac{\partial V}{\partial S}\left(t_i, S_{t_i}\right)\left(S_{t_{i+1}} - S_{t_i}\right) \\
&+ \frac{1}{2}\frac{\partial^2 V}{\partial t^2}\left(t_i, S_{t_i}\right)(t_{i+1} - t_i)^2 + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\left(t_i, S_{t_i}\right)\left(S_{t_{i+1}} - S_{t_i}\right)^2 \\
&+ \frac{\partial^2 V}{\partial S \partial t}\left(t_i, S_{t_i}\right)\left(S_{t_{i+1}} - S_{t_i}\right)(t_{i+1} - t_i) + R_i
\end{aligned}$$

with R_i being the remainder.

We write $\Delta t_i = (t_{i+1} - t_i)$, $\Delta S_{t_i} = \left(S_{t_{i+1}} - S_{t_i}\right)$ and $\Delta W_{t_i} = \left(W_{t_{i+1}} - W_{t_i}\right)$, and by summing over all i, we have:

$$\begin{aligned}
V(T, S_T) - V(0, S_0) &= \sum_{i=0}^{n-1}\frac{\partial V}{\partial t}\left(t_i, S_{t_i}\right)\Delta t_i + \sum_{i=0}^{n-1}\frac{\partial V}{\partial S}\left(t_i, S_{t_i}\right)\Delta S_{t_i} \\
&+ \frac{1}{2}\sum_{i=0}^{n-1}\frac{\partial^2 V}{\partial t^2}\left(t_i, S_{t_i}\right)(\Delta t_i)^2 + \frac{1}{2}\sum_{i=0}^{n-1}\frac{\partial^2 V}{\partial S^2}\left(t_i, S_{t_i}\right)\left(\Delta S_{t_i}\right)^2 \\
&+ \sum_{i=0}^{n-1}\frac{\partial^2 V}{\partial S \partial t}\left(t_i, S_{t_i}\right)\Delta S_{t_i}\Delta t_i + \sum_{i=0}^{n-1} R_i
\end{aligned}$$

Since μ and σ are simple processes: We have $\Delta S_{t_i} = \mu\Delta t_i + \sigma\Delta W_{t_i}$ and $\left(\Delta S_{t_i}\right)^2 = \mu^2(\Delta t_i)^2 + 2\mu\sigma\Delta t_i\Delta W_{t_i} + \sigma^2\left(\Delta W_{t_i}\right)^2$.

We consider the limit $n \to \infty$ (i.e. $\|\Pi\| \downarrow 0$) and get

$$\begin{aligned}
&\sum_{i=0}^{n-1}\frac{\partial V}{\partial t}\left(t_i, S_{t_i}\right)\Delta t_i + \sum_{i=0}^{n-1}\frac{\partial V}{\partial S}\left(t_i, S_{t_i}\right)\left(\mu\Delta t_i + \sigma\Delta W_{t_i}\right) \\
&\to \int_0^T\frac{\partial V}{\partial t}dt + \int_0^T\frac{\partial V}{\partial S}\mu dt + \int_0^T\frac{\partial V}{\partial S}\sigma dW_t
\end{aligned}$$

as well as

$$\frac{1}{2}\sum_{i=0}^{n-1}\frac{\partial^2 V}{\partial t^2}\left(t_i, S_{t_i}\right)(\Delta t_i)^2 \to 0$$

Similarly,

$$\frac{1}{2}\sum_{i=0}^{n-1}\frac{\partial^2 V}{\partial S^2}\left(t_i, S_{t_i}\right)\left(\mu^2\Delta t_i^2 + 2\mu\sigma\Delta W_{t_i}\Delta t_i + \sigma^2\left(\Delta W_{t_i}\right)^2\right) \to \frac{1}{2}\int_0^T\sigma^2\frac{\partial V^2}{\partial S^2}dt$$

where we used $\sum_{i=0}^{n-1}\frac{\partial^2 V}{\partial S^2}\left(t_i, S_{t_i}\right)\sigma^2\left(\Delta W_{t_i}^2 - \Delta t_i\right) \to 0$ (the expected value and the variance is 0) essentially for the limit we can replace $\left(\Delta W_{t_i}\right)^2$ by (Δt_i).

Finally, the higher order terms converge to 0 with similar arguments and hence since the derivatives are bounded:

$$\sum_{i=0}^{n-1}\frac{\partial^2 V}{\partial S \partial t}\left(t_i, S_{t_i}\right)\left(\mu\Delta t_i + \sigma\Delta W_{t_i}\right)\Delta t_i + \sum_{i=0}^{n-1}R_i \to 0$$

Together, this implies

$$V(T, S_T) - V(0, S_0) = \int_0^T\left(\frac{\partial V}{\partial t} + \mu\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2\frac{\partial V^2}{\partial S^2}\right)dt + \int_0^T\frac{\partial V}{\partial S}\sigma dW_t$$

which is exactly Itô's Lemma in integral notation.

$$dV = \left(\frac{\partial V}{\partial t} + \mu\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2\frac{\partial V^2}{\partial S^2}\right)dt + \frac{\partial V}{\partial S}\sigma dW_t$$

□

Corollary 3.2. *A reduction form for Itô's Lemma: if $V(t, W_t)$ is a $\mathbf{C}^{1,2}$ function, then*

$$dV = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial W^2}\right)dt + \frac{\partial V}{\partial W}dW$$

Example 9. 1. *Let $V(t, W_t) = W_t^2$. Then*

$$dV = 2W_t dW_t + dt$$

Thus $\int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T$.

2. *Let $V(t, W_t) = W_T^3$. Then*

$$dV = 3W_t^2 dW_t + 3W_t dt$$

Thus $\int_0^T W_t^2 dW_t = \frac{1}{3}W_T^2 - \int_0^T W_t dt$.

3.6 Geometric Brownian Motion

Definition 3.12. *The **Geometric Brownian Motion** is defined as the solution to the SDE*

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

with $S_0 = s_0 > 0$.

Note 6. *The discrete form is written as*

$$\frac{S_{t+\Delta t} - S_t}{S_t} = \mu\Delta t + \sigma\Delta W_t$$

The financial meaning is direct: the return of the stock is a sum of a deterministic part and a random part. Geometric Brownian Motion is the continuous version of ARIMA model.

So what is the solution to this SDE? We first guess the solution $S(t, W_t)$ is variable separable, i.e. $S(t, W_t) = f(t)g(W_t)$. Then we use Itô's Lemma:

$$dS = \left(g\frac{\partial f}{\partial t} + \frac{1}{2}f^2\frac{\partial^2 g}{\partial W^2}\right)dt + f\frac{\partial g}{\partial W}dW$$

Compare it to the SDE, we have

$$\begin{cases} g\dfrac{\partial f}{\partial t} + \dfrac{1}{2}f^2\dfrac{\partial^2 g}{\partial W^2} = \mu S_t \\ \dfrac{\partial g}{\partial W} = \sigma g \end{cases}$$

The second equation is a simple ODE, and the solution is $g(W_t) = Ce^{\sigma W_t}$. Substitute it back to the first equation, note that $g''(W_t) = \sigma^2 g(W_t)$, we have

$$\frac{\partial f}{\partial t} + \frac{1}{2}f^2\sigma^2 = \mu f$$

Hence

$$f(t) = Ae^{(\mu - \frac{1}{2}\sigma^2)t}$$

Therefore, the solution to the SDE is

$$S(t, W_t) = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}$$

We can see that

$$\begin{cases} S_T \to +\infty & \text{if } \mu > \frac{1}{2}\sigma^2 \\ S_T \to 0 & \text{if } \mu < \frac{1}{2}\sigma^2. \end{cases}$$

the mean return should be large enough to offset the risk for a long steady growth of stock price.

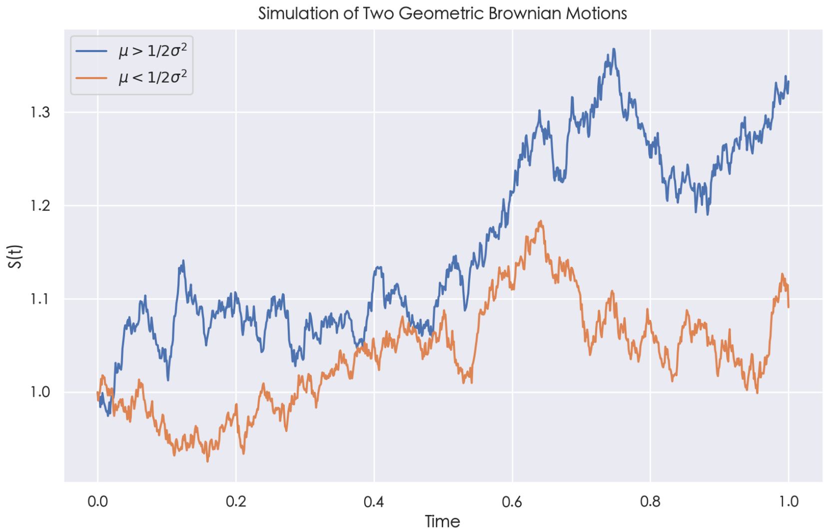

Figure 3.6: Geometric Brownian Motion

Proposition 3.5. *The mean of S_T is*

$$\mathbb{E}\left[S_T\right] = e^{(\mu-\sigma^2/2)T}\mathbb{E}\left[e^{\sigma W_T}\right] = S_0 e^{\mu t}$$

which intuitively makes sense: the mean stock price is nothing but the exponential of the mean rate of return for the stock.

Proof. It is clear that for any constant θ

$$d\left(e^{\theta W_t}\right) = \theta e^{\theta W_t} dW_t + \frac{1}{2}\theta^2 e^{\theta W_t} dt;$$

therefore, $e^{\theta W_t} = 1 + \int_0^t \theta e^{\theta W_s} dW_s + \frac{1}{2}\int_0^t \theta^2 e^{\theta W_s} ds$ and taking expectation,

$$\mathbb{E}\left[e^{\theta W_t}\right] = 1 + \frac{1}{2}\int_0^t \theta^2 E\left[e^{\theta W_s}\right] ds$$

which is equivalent to a simple ordinary differential equation

$$g(t) = \mathbb{E}\left[e^{\theta W_t}\right] \quad \frac{dg(t)}{dt} = \frac{1}{2}\theta^2 g(t), \quad g(0) = 1$$

whose solution is

$$g(t) = \mathbb{E}\left[e^{\theta W_t}\right] = e^{\frac{1}{2}\theta^2 t}$$

Therefore, the mean of S_T is

$$\mathbb{E}\left[S_T\right] = e^{(\mu-\sigma^2/2)T}\mathbb{E}\left[e^{\sigma W_T}\right] = S_0 e^{\mu t}$$

□

3.7 Stochastic Differential Equations

Definition 3.13. *A **Stochastic Differential Equation** (SDE) is an equation that describes the evolution of a random variable X_t with respect to time t and a Brownian motion W_t.*

$$dX_t = \mu\left(t, X_t\right) dt + \sigma\left(t, X_t\right) dW_t$$

where μ and σ are deterministic functions.

If μ is not a function of X_t, then the SDE is called **Linear Gaussian Model**.

$$dX_t = \mu dt + \sigma dW_t$$
$$\int_0^T dX_t = \int_0^T \mu dt + \int_0^T \sigma dW_t$$
$$X_T = X_0 + \mu T + \sigma W_T = X_0 + \mu T + \sigma\sqrt{T}Z, \quad Z \sim \mathcal{N}(0,1)$$

And we have $\mathbb{E}\left[X_T\right] = X_0 + \mu T$ and $\mathrm{Var}\left[X_T\right] = \sigma^2 T$.

Otherwise, if μ is variable seperately formed, and $\sigma = 0$, then the SDE is called **Log-normal Model**.

$$dX_t = \mu X_t dt \quad X_0 \text{ given}$$
$$\int_0^T \frac{dX_t}{X_t} = \int_0^T \mu dt$$
$$X_T = X_0 e^{\mu T}$$

Example 10. ***Ornstein-Uhlenbeck (OU) Process***: *Ornstein-Uhlenbeck process is the solution of the SDE|*

$$dX_t = -a\left(X_t - C\right)dt + \sigma dW_t,$$

where $a > 0, C \in \mathbb{R}$ and $\sigma > 0$ are constants.

The process is mean-reverting, cuz the drift term is proportional to the difference between the current value and the mean value C.

Applying Itô's formula to $e^{at}X_t$ can show that the solution is given by

$$X_t = C\left(1 - e^{-\alpha t}\right) + X_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s.$$

We have seen some simple examples of SDEs, and we can solve them by Itô's Lemma. However, in general, SDEs are not easy to solve. Here we state the strong solution an its conditions for uniqueness and existence.

Definition 3.14. *A strong solution of a SDE is a process X_t with continous sample paths and the following properties:*

1. *X_t is adapted to the filtration $\mathcal{F}_t$.*
2. *$\mathbb{P}(X_0 = \xi) = 1$, where ξ is a given initial value.*
3. *For every $t \geq 0$,*
$$\mathbb{P}\left(\int_0^t |\mu(s, X_s)| + \sigma^2(s, X_s) ds < \infty\right) = 1$$
4. $$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s$$
holds almost surely.

The coefficients $\mu(t,x)$ and $\sigma(t,x)$ are said to satisfy the Lipschitz condition if for every $0 \leq t < \infty$ and every x and y,

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq C|x-y|,$$

and are said to satisfy the linear growth condition if for every $0 \leq t < \infty$ and every x

$$|\mu(t,x)| + |\sigma(t,x)| \leq D(1 + |x|),$$

where C and D are positive constants.

Theorem 3.10. *Let $\mu(t,x)$ and $\sigma(t,x)$ satisfy the Lipschitz and linear growth conditions. Then there exists a unique strong solution to the SDE*
$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

Example 11. ***Constant Elasticity of Variance process (CEV)***: *Consider the SDE*

$$dX_t = \mu X_t dt + \sigma X_t^\beta dW_t$$

where μ, σ, β are constants, $0 < \beta < 1$.

For this SDE, the Lipschitz condition reduces to

$$|\mu x - \mu y| + \left|\sigma x^\beta - \sigma y^\beta\right| = |\mu| \cdot |x - y| + \sigma\left|x^\beta - y^\beta\right| \leq C|x-y|,$$

which implies that

$$\frac{\left|x^\beta - y^\beta\right|}{|x-y|} \leq C' < \infty,$$

for some fixed constant $C' < \infty$. *However, this is violated at* $x = 0$ *when* $0 < y < 1$, *because*

$$\lim_{y \to x} \frac{\left|x^\beta - y^\beta\right|}{|x - y|} = \lim_{y \to 0} \frac{dy^\beta}{dy} = \lim_{y \to 0} ry^{\beta - 1} = \lim_{y \to 0} \frac{\beta}{y^{1-\beta}} = \infty, \text{ for } 0 < \beta < 1,$$

and we CANNOT find a finite constant $C' < \infty$.

Example 12. ***Cox-Ingersoll-Ross (CIR) process****: Consider the SDE*

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$$

where κ, θ, σ *are positive constants.*

The Lipschitz condition will violate at $x = 0$ *by the same reason.*

Finally, note that if a stochastic process has dynamic

$$dX_t = \sigma(t, X_t)dW_t$$

where $\mu = 0$, i.e. **there's no drift term, then the process is a martingale**.

Always make ourselves remember, the process

$$\exp\left(-\frac{\sigma^2}{2}t + \sigma W_t\right)$$

is a martingale. Actually we will see this again in Black-Scholes Model, this process is actually S_t/B_t under the risk-neutral measure, where S_t is the stock price and B_t is the money market account with interest rate r.

4 Black-Scholes Model

Pricing of financial derivatives is a central problem. After preparing basic math tools, we can now introduce the Black-Scholes model, which is the first and most famous model for option pricing. A brief view of the Black-Scholes model is as 4.1.

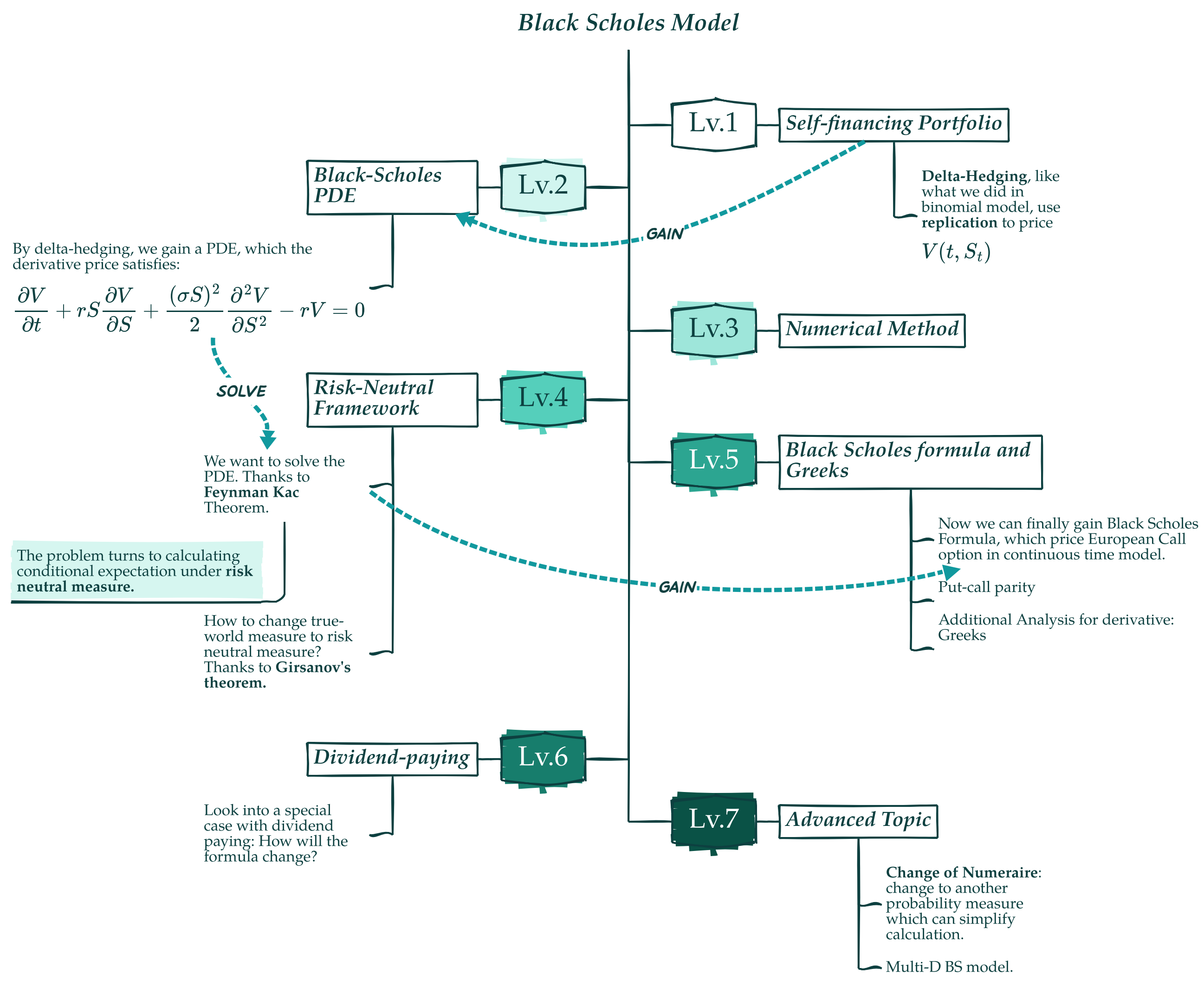

Figure 4.1: Black-Scholes Model

This part includes:

1. Self-Financing Portfolio and Black-Scholes PDE.
2. Numerical Computation and Stablility.
3. Risk-Neutral Framework including Dynkin's Formula, Feynman-Kac Theorem and Girsanov Theorem.

4. Black-Scholes Formula and Greeks.
5. Dividend-Paying Stock.
6. Change of Numeraire.
7. Multi-Dimensional Black-Scholes Model.

4.1 Self-Financing Portfolio and Black-Scholes PDE

Model Bases and Assumptions

The model works by **replicating the derivative's payoff at maturity** using only the underlying asset and a money market account. The assumptions are as follows:

- The **underlying asset** has log-normal distributed and follows the SDE:
$$dS_t = \mu S_t dt + \sigma S_t dW_t$$
μ, σ are deterministic constant for the model (although they can be made deterministic functions of time with little effort) and W_t is standard Brownian motion.
- Investors have access to **money market account** which accrues continuously with continuously compounded rate r.
$$dB_t = rB_t dt$$
- The classic model only considered derivatives of European Style, meaning that there is a clear and defined maturity date.
- Transaction costs and taxes are ignored.
- No restrictions on short selling and liquidity.

Delta-Hedging and Self-Financing Portfolio

Abbreviate the value of the derivative at time t, when the price of the underlying asset is S_t is $V(t, S_t)$.

Set up a portfolio Π_t with one unit of derivative (say, a call option), α_t units of the underlying asset and β_t holding in the money market account. The subscripts in t highlight that these are all some functions of time, as yet undefined.

$$\Pi_t = V_t + \alpha_t S_t + \beta_t B_t$$

The idea of replication in the Black Scholes model is to try and identify α_t and β_t so that they hedge the derivative perfectly. In other words, if we can ensure that $\alpha_t S_t + \beta_t B_t = -V_t, \forall t \in [t_0, T]$ then we can replicate V_t using the simple instruments S_t and B_t. Let's assume for a moment that we can and proceed under this assumption. In that case $\Pi_t = 0, \forall t \in [t_0, T]$, and so $d\Pi_t = 0$. It also means that the product $\beta_t B_t$ (which represents the balance in the money market account) is given in terms of α_t :

$$\beta_t B_t = -(V_t + \alpha_t S_t)$$

To decide on what to choose for α_t we first compare the dynamics of the derivative V_t and try to match it up with the dynamics of $\alpha_t S_t + \beta_t B_t$; after all, this is the meaning of replication. Recall that

$$dV_t = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 V}{\partial S_t^2} (dS_t)^2$$

and

$$d\Pi_t = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 V}{\partial S_t^2} (dS_t)^2 + \alpha_t dS_t + S_t d\alpha_t + \beta_t dB_t + B_t d\beta_t = 0$$

Always remember our goal is to replicate the derivative payoff, so we must have $d\Pi_t = 0$. But this is impossible when Π_t is subjected to randomness (and the presence of dW_t via the presence of dS_t in the equations signifies the presence of randomness). But if we choose $\alpha_t = -\frac{\partial V}{\partial S_t}$ and maintain this $\forall t \in [t_0, T]$ then this hedges out the stochastic element from the portfolio Π_t and below is what is left:

$$d\Pi_t = \frac{\partial V}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V}{\partial S_t^2} (dS_t)^2 + S_t d\alpha_t + \beta_t dB_t + B_t d\beta_t = 0$$

This is called **Delta-Hedging**, which means that we hedge out the risk from value variation due to the underlying asset. We will formally define Greeks including Delta later.

Definition 4.1. *The **Delta** of a derivative is the rate of change of the derivative's price with respect to the price of the underlying asset.*

$$\Delta_t = \frac{\partial V}{\partial S_t}$$

Now let's take a look at $S_t d\alpha_t + B_t d\beta_t$. We neet to ensure that the portfolio does not evolve randomly, so set $S_t d\alpha_t + B_t d\beta_t = 0$, which is called the **self-financing condition**.

Definition 4.2. *The portfolio strategy $h_t = (\alpha_t, \beta_t)$ is **self-financing** if there is no external infusion or withdrawal of money.*

If I want to buy more stocks, I need to sell some risk-free assets (and vice versa).

Discretely, At time $t + \Delta t$, one can change the portfolio $(\alpha_t, \beta_t) \to (\alpha_{t+\Delta t}, \beta_{t+\Delta t})$, but the value of the portfolio must stay the same:

$$\alpha_t B_t + \beta_t S_t \to \alpha_t B_{t+\Delta t} + \beta_t S_{t+\Delta t} = \alpha_{t+\Delta t} B_{t+\Delta t} + \beta_{t+\Delta t} S_{t+\Delta t} \to \alpha_{t+\Delta t} B_{t+2\Delta t} + \beta_{t+\Delta t} S_{t+2\Delta t}$$

Continously, the value change of a portfolio can only come from the change of the underlying asset and the risk-free asset, but not holdings change:

$$dV_t^h = d\left(\alpha_t S_t + \beta_t B_t\right) = \alpha_t dS_t + \beta_t dB_t$$

Now, we have the following equations:

$$d\Pi_t = \frac{\partial V}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V}{\partial S_t^2} \left(\sigma S_t\right)^2 dt + \beta_t r B_t dt = 0$$

Last, plus the replicating condition $\beta_t B_t = -\left(V_t - \frac{\partial V}{\partial S_t} S_t\right)$, we have

Theorem 4.1. ***Black-Scholes PDE**:*

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

Boundary Conditions

The solution to the PDE yields the function $V(t, S_t)$ from which the initial time t_0 price of the derivative can be computed as the value $V(t_0, S_{t_0})$. The Black Scholes PDE is powerful because it applies to any payoff. To solve it for a specific payoff then you need to specify the **boundary conditions that are specific to that payoff**.

The PDE is first order in t and second order in S and so one temporal condition and two spacial conditions must be specified.

- **Temporal Condition**: The derivative has a payoff at maturity, so the value of the derivative at maturity is the payoff. This is the terminal condition.

$$V(T, S_T) = \varphi(S_T) \xlongequal{\text{European Call}} (S_T - K)^+$$

- **Spacial Conditions**:
  - Dirichlet:

$$\lim_{S\to 0} V(t, S) = 0$$
$$\lim_{S\to\infty} V(t, S) = (S - K)^+$$

  In essence Dirichlet explicitly specifies the value of $V(t, S^*)$ at some boundary S^*.

  - Neumann:

$$\lim_{S\to 0} \frac{\partial V}{\partial S}(t, S) = 0$$
$$\lim_{S\to\infty} \frac{\partial V}{\partial S}(t, S) = 1$$

For convenience, we can use the following transformation to simplify the PDE: set $\tau = T - t$,

$$\frac{\partial V}{\partial \tau} = rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV$$

4.2 Numerical Computation and Stablility

The numerical solution to the PDE involves using finite differences to discretize the PDE onto a grid. Where a closed form solution would give us $V(\tau, S)$ for any chosen value of τ and S the numerical solution is an algorithm which will attempt to compute the values at discrete points on the grid, which is a **2D rectangular domain** indexed by τ and S.

The range of S is chosen as the interval $[S_{\min}, S_{\max}]$, and if this is divided up into N_S intervals then there are a set of $(N_S + 1)$ discrete points for S, which are indexed by j :

$$S^j = S_{\min} + \frac{(S_{\max} - S_{\min})}{N_S} j \quad j = 0, 1, 2 \ldots N_S$$

Similarly, τ is discretized into N_T intervals, giving $(N_T + 1)$ discrete points for time, indexed by k :

$$\tau^k = T\left(1 - \frac{k}{N_T}\right) \quad k = 0, 1, 2 \ldots N_T$$

To lighten the notation we define $\Delta S = \dfrac{(S_{\max} - S_{\min})}{N_S}$ and $\Delta t = \dfrac{T}{N_T}$. The grid point (k, j) corresponds to the pair $\left(\tau^k, S^j\right)$ and V_j^k is the value of the derivative at this grid point.

- There are $(N_T + 1) \times (N_S + 1)$ points on the grid;
- $(2N_T + N_S + 1)$ of these points are determined by initial/boundary conditions.
- Leaving $N_T\,(N_S - 1)$ **internal grid points** which must be solved for using equations defined by the PDE.These internal points are V_j^k for $k \in [1, 2, \ldots N_T]$ and $j \in [1, 2 \ldots (N_S - 1)]$.

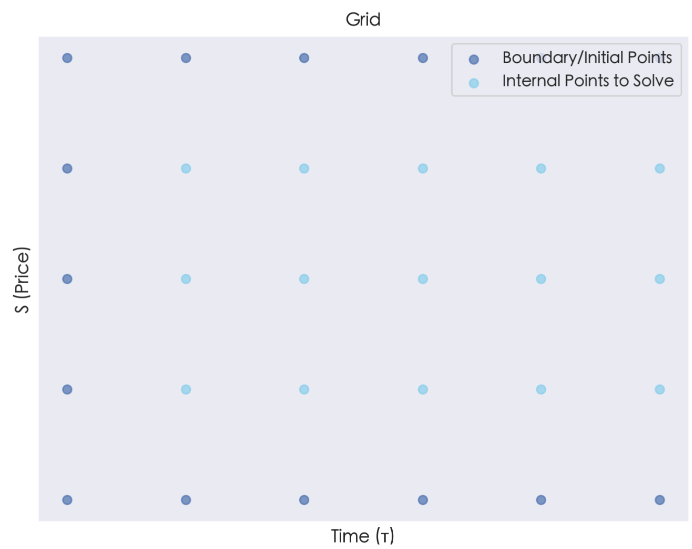

Figure 4.2: Finite Difference Grid

Explicit Scheme(Forward Euler)

The goal is to fill out the inner grid of values for V_j^k. In the explicit scheme V_j^{k+1} is sequentially computed from V_j^k for all j simultaneously. The algorithm naturally steps forward in time (increasing k) using only information **from previous time steps**. This is possible because $V_j^0, \forall j \in [0, N_S]$ is given by the initial conditions. Informally the explicit scheme is a difference equation of the form $V_j^{k+1} = f\left(V_j^k\right) \quad \forall k = 0, 1, 2 \ldots (N_T - 1)$. Applying the discretized PDE at the spacial point j produces the following system of equations.

$$\frac{V_j^{k+1} - V_j^k}{\Delta t} = rS^j\left(\frac{V_{j+1}^k - V_{j-1}^k}{2\Delta S}\right) + \frac{\left(\sigma S^j\right)^2}{2}\left(\frac{V_{j+1}^k + V_{j-1}^k - 2V_j^k}{\Delta S^2}\right) - rV_j^k$$

Rewite it as

$$
\begin{aligned}
V_j^{k+1} &= aV_{j-1}^k + bV_j^k + cV_{j+1}^k \\
a &= \left[\frac{\Delta t}{2\Delta S^2}\left(\sigma S^j\right)^2 - rS^j\frac{\Delta t}{2\Delta S}\right] \\
b &= \left[1 - r\Delta t - \frac{\Delta t}{\Delta S^2}\left(\sigma S^j\right)^2\right] \\
c &= \left[\frac{\Delta t}{2\Delta S^2}\left(\sigma S^j\right)^2 + rS^j\frac{\Delta t}{2\Delta S}\right]
\end{aligned}
$$

This is something like a trinomial tree? Since the next-time value comes from a linear combination of the current-time values of the three adjacent points.

We also have a vector form of the equation:

$$\mathbf{V}^{k+1} = \mathbf{M}\cdot\mathbf{V}^k + \mathbf{P} \quad \forall k = 0, 1, 2\ldots(N_T - 1)$$

where $\mathbf{M}$ is a $(N_S + 1) \times (N_S + 1)$ matrix, $\mathbf{V}^k$ and $\mathbf{P}$ is a $(N_S + 1) \times 1$ vector. $\mathbf{M}$ holds the coefficients as depicted by a, b, c above, and $\mathbf{P}$ is a vector which helps apply the boundary conditions.

Implicit Scheme(Backward Euler)

In the implicit scheme the spacial derivatives are evaluated not at the previous time step, but rather **the current time step**.

$$\frac{V_j^{k+1} - V_j^k}{\Delta t} = rS^j\left(\frac{V_{j+1}^{k+1} - V_{j-1}^{k+1}}{2\Delta S}\right) + \frac{\left(\sigma S^j\right)^2}{2}\left(\frac{V_{j+1}^{k+1} + V_{j-1}^{k+1} - 2V_j^{k+1}}{\Delta S^2}\right) - rV_j^{k+1}$$

The linear system is

$$a'V_{j-1}^{k+1} + b'V_j^{k+1} + c'V_{j+1}^{k+1} = V_j^k$$

where

$$
\begin{aligned}
a' &= \left[rS^j\frac{\Delta t}{2\Delta S} - \frac{\Delta t}{2\Delta S^2}\left(\sigma S^j\right)^2\right] \\
b' &= \left[1 + r\Delta t + \frac{\Delta t}{\Delta S^2}\left(\sigma S^j\right)^2\right] \\
c' &= \left[-rS^j\frac{\Delta t}{2\Delta S} - \frac{\Delta t}{2\Delta S^2}\left(\sigma S^j\right)^2\right]
\end{aligned}
$$

The vector form is

$$\mathbf{M}\cdot\mathbf{V}^{k+1} + \mathbf{P} = \mathbf{V}^k \quad \forall k = 0, 1, 2\ldots(N_T - 1)$$

Here $\mathbf{V}, \mathbf{P} \in \mathbb{R}^{N_S - 1}$ and $\mathbf{M}$ is a square $(N_S - 1) \times (N_S - 1)$ matrix.

Crank-Nicolson Scheme

The Crank-Nicolson scheme is a compromise between the explicit and implicit schemes. It is a second-order accurate scheme and is unconditionally stable. The Crank-Nicolson scheme is

$$
\begin{aligned}
\frac{V_j^{k+1} - V_j^k}{\Delta t} =& \frac{1}{2}\left[rS^j\left(\frac{V_{j+1}^{k+1} - V_{j-1}^{k+1}}{2\Delta S}\right) + \frac{\left(\sigma S^j\right)^2}{2}\left(\frac{V_{j+1}^{k+1} + V_{j-1}^{k+1} - 2V_j^{k+1}}{\Delta S^2}\right) - rV_j^{k+1}\right] \\
&+ \frac{1}{2}\left[rS^j\left(\frac{V_{j+1}^k - V_{j-1}^k}{2\Delta S}\right) + \frac{\left(\sigma S^j\right)^2}{2}\left(\frac{V_{j+1}^k + V_{j-1}^k - 2V_j^k}{\Delta S^2}\right) - rV_j^k\right]
\end{aligned}
$$

The Crank Nicholson scheme is thus the most accurate of the three schemes when it comes to convergence even though it is not much more complicated to implement compared to the Implicit scheme.

Discretization Error and Stability

Table 4.1: Discretization Error and Stability of Schemes

Scheme	Order of Discretization Error	Stability
Explicit	$O\left(\Delta t, \Delta S^2\right)$	Conditionally Stable
Implicit	$O\left(\Delta t, \Delta S^2\right)$	Stable
Crank-Nicolson	$O\left(\Delta t^2, \Delta S^2\right)$	Unconditionally Stable

Definition 4.3. *The **discretization error** is the difference between the true solution and the numerical solution. The error is a function of the grid size* $\Delta S, \Delta t$.

We will use simple ODE to illustrate the stability of the schemes. Consider the ODE

$$\frac{dU_t}{dt} = \beta U_t$$

The solution is $U_t = U_0 e^{\beta t}$. The explicit scheme is

$$\frac{U_{k+1} - U_k}{\Delta t} = \beta U_k + \mathrm{DE}_{k+1}, \quad \mathrm{DE}_k \sim O(\Delta t)$$

where DE_{k+1} is the discretization error and U is the true solution. Let the numerical solution is $\widetilde{U}_k = U_k - \varepsilon_k$, then

$$\varepsilon_{k+1} = (1 + \beta \Delta t)\varepsilon_k + \mathrm{DE}_{k+1}\, \Delta t$$

Call $\lambda = 1 + \beta\Delta t$ is the **Growth Factor**, then the stability condition is $|\lambda| \leq 1$. In Black-Scholes model, the stability condition is $\Delta t \leq C \cdot \Delta S^2$ where C depends on parameters of the model.

For implicit scheme, the growth factor is $\lambda = \dfrac{1}{1 - \beta\Delta t}$, which is always less than 1 unless $\Delta t \sim 1/\beta$. So the implicit scheme is more stable than the explicit scheme. In BS model, we can avoid the stability issue that is inherent in the explicit scheme.

Under Crank Nicholson,

$$\begin{aligned}
\frac{U_{k+1} - U_k}{\Delta t} &= \frac{\beta}{2}\left[U_k + U_{k+1}\right] + \mathrm{DE}_{k+1} \\
U_{k+1}\left[1 - \frac{\beta}{2}\Delta t\right] &= U_k\left[1 + \frac{\beta}{2}\Delta t\right] + \mathrm{DE}_{k+1} \cdot \Delta t \\
U_{k+1} &= \left[\frac{1 + \frac{\beta}{2}\Delta t}{1 - \frac{\beta}{2}\Delta t}\right] U_k + \mathrm{DE}_{k+1} \cdot \Delta t
\end{aligned}$$

The growth factor is $\lambda = \dfrac{1 + \frac{\beta}{2}\Delta t}{1 - \frac{\beta}{2}\Delta t}$, $|\lambda|$ is always less than 1. So the Crank-Nicolson scheme is unconditionally stable.

Use the following Python code to observe the stability of the schemes.

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.interpolate import make_interp_spline

class NumericMethod:
    def __init__(self, T, N, beta, U0=1.0):
        self.T = T
        self.N = N
        self.beta = beta
        self.U0 = U0
        self.dt = self.T / self.N

    def explicit_euler(self):
        t_values = np.arange(0, self.T + self.dt, self.dt)
        U_values = [self.U0]
        for i in range(1, len(t_values)):
            U_values.append(U_values[-1] + self.dt* self.beta * U_values[-1])
        return t_values, U_values
```

```python

    def implicit_euler(self):
        t_values = np.arange(0, self.T + self.dt, self.dt)
        U_values = [self.U0]
        for i in range(1, len(t_values)):
            U_values.append(U_values[-1] / (1 - self.beta * self.dt))
        return t_values, U_values

    def crank_nicolson(self):
        t_values = np.arange(0, self.T + self.dt, self.dt)
        U_values = [self.U0]
        for i in range(1, len(t_values)):
            U_values.append(U_values[-1]* (1 + 0.5 * self.beta * self.dt) / (1 - 0.5 *
             ↪  self.beta * self.dt))
        return t_values, U_values

def plot_results(beta, Time, Num, method, ax, title):
    for i in range(len(Num)):
        problem = NumericMethod(Time, Num[i],beta=beta)
        if method == 'explicit':
            t_values, U_values = problem.explicit_euler()
        elif method == 'implicit':
            t_values, U_values = problem.implicit_euler()
        elif method == 'crank-nicolson':
            t_values, U_values = problem.crank_nicolson()
        # Smooth the curve
        t_smooth = np.linspace(t_values.min(), t_values.max(), 500)
        U_smooth = make_interp_spline(t_values, U_values)(t_smooth)
        ax.plot(t_smooth, U_smooth, label=f'N={Num[i]}')
        # Other settings
```

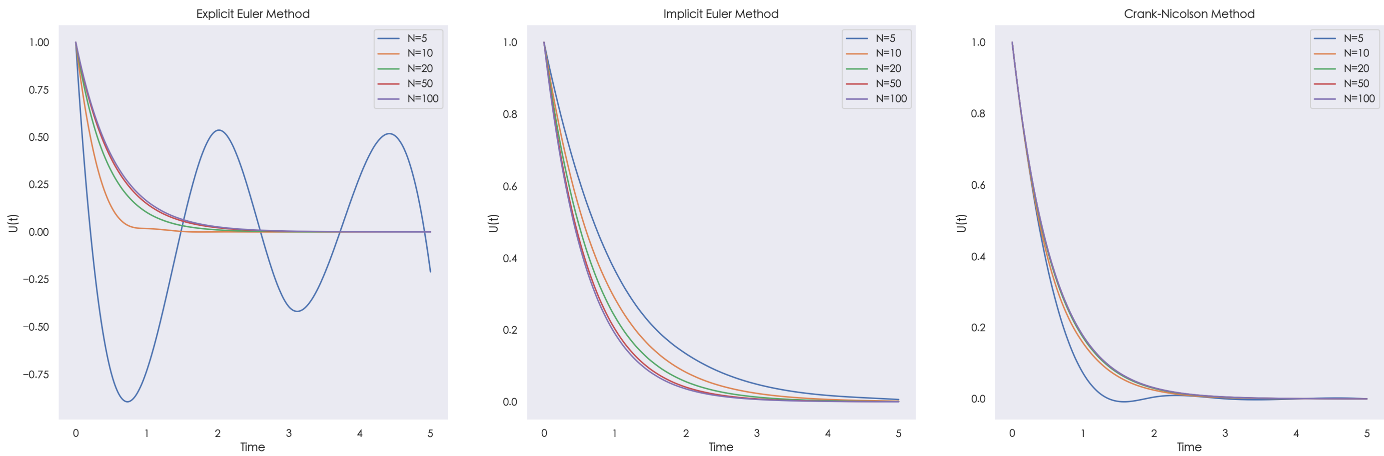

Figure 4.3: Stability of Different Schemes

4.3 Risk-Neutral Framework

Theoredically and mathematically, the Black-Scholes model is based on the risk-neutral framework. Recall that we've already gained the BS PDE, and we can see that **there's no μ which is the expected return of the underlying asset in the PDE**. Then the solution to the PDE will not contain μ.

Dynkin's Formula

Consider a Itô's Lemma of $f(t, x_t)$ where x_t is a stochastic process $dx_t = \mu(t, x_t)dt + \sigma(t, x_t)dW_t$:

$$f(t + dt, x_{t+dt}) = f(t, x_t) + \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dx_t)^2$$

Plus the dynamics of x_t into it,

$$f(t + dt, x_{t+dt}) = f(t, x_t) + \frac{\partial f}{\partial x}\sigma(t, x_t)\, dW_t + \left(\frac{\partial f}{\partial t} + \mu(t, x_t)\frac{\partial f}{\partial x_t} + \frac{1}{2}\sigma(t, x_t)^2\frac{\partial^2 f}{\partial x^2}\right)dt$$

This can be integrated with respect to the time variable:

$$f(T, x_T) = f(t_0, x_0) + \int_{t_0}^{T}\frac{\partial f}{\partial x}\sigma(t, x_t)\, dW_t + \int_{t_0}^{T}\left(\frac{\partial f}{\partial t} + \mu(t, x_t)\frac{\partial f}{\partial x_t} + \frac{1}{2}\sigma(t, x_t)^2\frac{\partial^2 f}{\partial x^2}\right)dt$$

Now taking conditional expectations and remembering that the Itô integral has the property that it's expectation is zero:

$$\mathbb{E}[f(T, x_T) \mid x_0] = f(t_0, x_0) + \mathbb{E}\left[\int_{t_0}^{T}\left(\frac{\partial f}{\partial t} + \mu(t, x_t)\frac{\partial f}{\partial x_t} + \frac{1}{2}\sigma(t, x_t)^2\frac{\partial^2 f}{\partial x^2}\right)dt \,\middle|\, x_0\right]$$

So we have the **Dynkin's Formula**:

Theorem 4.2. *If $f(t, x_t)$ satisfies a PDE*

$$\frac{\partial f}{\partial t} + \mu(t, x_t)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(t, x_t)\frac{\partial^2 f}{\partial x^2} = 0$$

then we would have a elegant result:

$$f(t_0, x_0) = \mathbb{E}[f(T, x_T) \mid x_0]$$

where x_t is a stochastic process $dx_t = \mu(t, x_t)dt + \sigma(t, x_t)dW_t$.

Feynman-Kac Theorem

Let $t_0 = 0$, and $f(t, S_t) = e^{-rt}V(t, S_t)$, using calculus we get:

$$\frac{\partial f}{\partial t} = -re^{-rt}V + e^{-rt}\frac{\partial V}{\partial t}$$
$$\frac{\partial f}{\partial S} = e^{-rt}\frac{\partial V}{\partial S}$$
$$\frac{\partial^2 f}{\partial S^2} = e^{-rt}\frac{\partial^2 V}{\partial S^2}$$

Put all these into Dynkin's Formula, we get:

$$\mathbb{E}\left[e^{-rT}V(T, S_T) \mid S_0\right] = V(0, S_0) + \mathbb{E}\left[\int_0^T e^{-rt}\left(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S_t} + \frac{1}{2}(\sigma S_t)^2\frac{\partial^2 V}{\partial S_t^2} - rV\right)dt \,\middle|\, S_0\right]$$

The term in red is exactly the Black-Scholes PDE, so we have:

Theorem 4.3. ***Feynman-Kac Theorem**: If $V(t, S_t)$ is a function of t and S_t, where S_t has dynamics:*

$$dS_t = rS_t dt + \sigma S_t dW_t$$

*and $V(t, S_t)$ **satisfies the Black-Scholes PDE**, then we have the following pricing formula:*

$$V(0, S_0) = \mathbb{E}\left[e^{-rT}V(T, S_T) \mid S_0\right]$$

or more generally, for any t:

$$V(t, S_t) = \mathbb{E}_{t,s_t}\left[e^{-r(T-t)}V(T, S_T)\right]$$

where $\mathbb{E}_{t,s_t}$ is the conditional expectation given S_t at time t.

Note that Feynman-Kac only works when the underlying asset has drift r which is the risk-free rate. But the model assumption tells us that the underlying asset has drift μ which is the expected return of the asset.

The dynamics in Feynman-Kac are NOT the (real-world) dynamics of the price of the stock. The actual drift of the stock μ does not appear anywhere when calculating the price of a financial derivative.

Like in the binomial model, what matters to compute the price of a derivative is the risk-neutral dynamics of the stock, not the real-world dynamics.

We need to clarify that: The dynamics of the stock price S_t under the **risk-neutral measure** $\mathbb{Q}$ are

$$dS_t = S_t \left(r dt + \sigma dW_t^{\mathbb{Q}} \right),$$

where $W_t^{\mathbb{Q}}$ is a Wiener process under the measure $\mathbb{Q}$. Then we can use Feynman-Kac to price option under the risk-neutral measure.

Measure Change and Girsanov Theorem

So the problem is: How can we find such a measure $\mathbb{Q}$ under which the stock price dynamics are risk-neutral? First we need to introduce a concept of measure change.

Definition 4.4. *Assume a stochastic process L_T is $\mathcal{F}_T$-measurable and under measure $\mathbb{P}$ has expectation 1, then the measure $\mathbb{Q}$ defined by*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = L_T, \quad \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = L_t$$

*which is called the **Radon-Nikodym derivative**, is called the **measure change** from $\mathbb{P}$ to $\mathbb{Q}$.*

Measure change has the following properties:

Proposition 4.1. *For any $\mathbb{P}$-integrable random variable X, we have*

$$\mathbb{E}^{\mathbb{Q}}[X] = \mathbb{E}^{\mathbb{P}}[L_T X]$$

and

$$\mathbb{E}^{\mathbb{P}}[X] = \mathbb{E}^{\mathbb{Q}}\left[\frac{X}{L_T}\right]$$

If $s < t$, then

$$\mathbb{E}^{\mathbb{Q}}\left[X | \mathcal{F}_s \right] = \frac{1}{L_s} \mathbb{E}^{\mathbb{P}}\left[L_t X | \mathcal{F}_s \right]$$

Proof. The first equation comes from the definition of the expectation under a new measure:

$$\mathbb{E}^{\mathbb{Q}}[X] = \int X d\mathbb{Q} = \int X L_T d\mathbb{P} = \mathbb{E}^{\mathbb{P}}[L_T X]$$

The second equation is just dual to the first one.

For the third equation, recall that conditional expectation can be written as

$$\mathbb{E}^{\mathbb{Q}}(X|\mathcal{F}_s) = \int_{\omega \in \mathcal{F}_s} X \frac{d\mathbb{Q}}{d\mathbb{Q}|_{\mathcal{F}_s}}$$

Then we have

$$\mathbb{E}^{\mathbb{Q}}(X|\mathcal{F}_s) = \int_{\omega \in \mathcal{F}_s} X \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} \frac{d\mathbb{P}|_{\mathcal{F}_s}}{d\mathbb{Q}|_{\mathcal{F}_s}} \frac{1}{d\mathbb{P}|_{\mathcal{F}_s}} = \frac{1}{L_s} \mathbb{E}^{\mathbb{P}}\left[L_t X | \mathcal{F}_s \right]$$

□

Example 13. *Consider a standard Gaussian random variable $z \sim \mathcal{N}(0, 1)$, recall that the density function of z is*

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Consider a new random variable $\tilde{z}$ where $\lambda \in \mathbb{R}$:

$$\tilde{z} = z \exp\left(\lambda z - \frac{\lambda^2}{2} \right)$$

Deduce it's expectation:

$$\begin{aligned}\mathbb{E}[\tilde{z}] &= \mathbb{E}\left[ze^{\left(-\frac{1}{2}\lambda^2+\lambda z\right)}\right] \\ &= \int ze^{\left(-\frac{1}{2}\lambda^2+\lambda z\right)}\phi(z)dz \\ &= \frac{1}{\sqrt{2\pi}}\int ze^{\left(-\frac{1}{2}\lambda^2+\lambda z\right)}e^{-\frac{1}{2}z^2}dz \\ &= \frac{1}{\sqrt{2\pi}}\int ze^{-\frac{1}{2}(z-\lambda)^2}dz\end{aligned}$$

Where z was a standard $N(0,1)$ random variable; $\tilde{z}$ is $\mathcal{N}(\lambda,1)$. Naturally $\mathbb{E}[\tilde{z}]=\lambda$, and the magic of this result is that the multiplier $e^{\left(-\frac{1}{2}\lambda^2+\lambda z\right)}$ inside the expectation acts as if it changes the mean of the random variable z. In other words, if we apply this multiplier to a function $f(z)$, and compute the expectation, then:

$$\mathbb{E}\left[f(z)e^{\left(-\frac{1}{2}\lambda^2+\lambda z\right)}\right] = \frac{1}{\sqrt{2\pi}}\int f(z)e^{-\frac{1}{2}(z-\lambda)^2}dz = \mathbb{E}\left[f(z+\lambda)\right]$$

Now define $L := e^{\left(-\frac{1}{2}\lambda^2+\lambda z\right)}$, one can verify that $\mathbb{E}[L]=1$. Then we can define a new measure $\mathbb{Q}$ by the Radon-Nikodym derivative:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = L$$

Hence we have

$$\mathbb{E}^{\mathbb{Q}}[f(z)] = \mathbb{E}^{\mathbb{P}}[f(z+\lambda)]$$

So if z is a standard normal random variable, then $z-\lambda$ is a standard normal random variable under the measure $\mathbb{Q}$.

$$\mathbb{Q}(Z-\lambda<c) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{Z\le c+\lambda\}}\right] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{1}_{\{Z+\lambda\le c+\lambda\}}\right] = \mathbb{P}(X\le c) = \Phi(c),$$

where Φ is the cumulative distribution function of a $\mathcal{N}(0,1)$ random variable.

Similarly, a standard brownian motion W_t under true world measure $\mathbb{P}$ will not be a standard brownian motion under the risk-neutral measure $\mathbb{Q}$. We need to find a Radon-Nikodym derivative L_T to change the measure.

Theorem 4.4. ***Girsanov Theorem**: Let $W_t^{\mathbb{P}}$ be a standard Brownian motion under the measure $\mathbb{P}$, and L_t be an $\mathcal{F}_t$-measurable **martingale** with $\mathbb{E}[L_T]=1$, and L_t has dynamics:*

$$dL_t = k_tL_tdW_t^{\mathbb{P}}, \quad L_0 = 1$$

Define probability measure $\mathbb{Q}$ by the Radon-Nikodym derivative:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = L_T$$

Then the process $W_t^{\mathbb{Q}}$ defined by

$$dW_t^{\mathbb{Q}} = dW_t^{\mathbb{P}} - k_tdt$$

is a standard Brownian motion under the measure $\mathbb{Q}$.

Note 7. *L_t can be interpreted as:*

$$L_t = \exp\left\{\int_0^t k_sdW_s^{\mathbb{P}} - \frac{1}{2}\int_0^t |k_s|^2\, ds\right\}.$$

*If $\mathbb{E}^{\mathbb{P}}\left[\exp\left\{\frac{1}{2}\int_0^T |k_t|^2\,dt\right\}\right]<\infty$, then $\mathbb{E}^{\mathbb{P}}\left[L_T\right]=1$, which is called **Novikov condition**. It guarantees that the assumption of Girsanov Theorem (that is $\mathbb{E}^{\mathbb{P}}\left[L_T\right]=1$) is fulfilled.*

Change of Measure in Black-Scholes Model

Using Girsanov theorem, given k_t,

$$dW_t^{\mathbb{Q}} = dW_t^{\mathbb{P}} - k_tdt$$

is a standard Brownian motion under the measure $\mathbb{Q}$. In the Black-Scholes model, we have the dynamics of the stock price under the real world measure $\mathbb{P}$:

$$dS_t = \mu S_tdt + \sigma S_t(dW_t^{\mathbb{Q}} + k_tdt)$$

compare it to the dynamics under the risk-neutral measure $\mathbb{Q}$:

$$dS_t = rS_t dt + \sigma S_t dW_t^{\mathbb{Q}}$$

we can find that:

$$k_t = -\frac{\mu - r}{\sigma}$$

$\lambda := \dfrac{\mu - r}{\sigma}$ is called the **market price of risk** or **Sharpe ratio**. And hence we gain the Radon-Nikodym derivative:

$$\begin{aligned} L_T &= \exp\left\{\int_0^T k_t dW_t^{\mathbb{P}} - \frac{1}{2}\int_0^T k_t^2 dt\right\} \\ &= \exp\left\{-\int_0^T \lambda dW_t^{\mathbb{Q}} - \frac{1}{2}\int_0^T \lambda^2 dt\right\} \\ &= \exp\left\{-\lambda W_T^{\mathbb{Q}} - \frac{1}{2}\lambda^2 T\right\} \\ &= \exp\left\{-\frac{\mu - r}{\sigma} W_T^{\mathbb{Q}} - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 T\right\} \end{aligned}$$

Finally we state that the BS model is complete and show the risk-neutral variation formula.

Definition 4.5. *A market is said to be **complete** if there exists a unique risk-neutral measure $\mathbb{Q}$ under which all derivatives can be perfectly replicated by self-financing portfolios.*

Lemma 4.1. *In the Black-Scholes model, under the risk-neutral measure $\dfrac{S_t}{B_t}$ is a martingale.*

Proof. Use Multi-Dimensional Itô's Lemma[1], let $f(S_t, B_t) = \dfrac{S_t}{B_t}$, then

$$\frac{\partial f}{\partial S} = \frac{1}{B_t}, \quad \frac{\partial f}{\partial B} = -\frac{S_t}{B_t^2}, \quad \frac{\partial^2 f}{\partial S^2} = 0$$

and

$$dS_t dB_t = 0, \quad (dB_t)^2 = 0$$

Hence

$$\begin{aligned} d\left(\frac{S_t}{B_t}\right) &= \frac{1}{B_t} dS_t - \frac{S_t}{B_t^2} dB_t + \frac{1}{2}\cdot 0 \\ &= \frac{\sigma S_t}{B_t} dW_t \end{aligned}$$

There's no drift term in the dynamics of $\dfrac{S_t}{B_t}$, so $\dfrac{S_t}{B_t}$ is a martingale. □

Note 8. *To aviod use of Multi-Dimensional Itô's Lemma, just calculate*

$$f(t, W_t) = \frac{S_t}{B_t} = \frac{S_0}{B_0}\exp\left(-\frac{1}{2}\sigma^2 t + \sigma W_t\right)$$

and use simple Itô's Lemma.

Corollary 4.1. *A stochastic process X_t is a martingale under $\mathbb{Q}$, if and only if $X_t \cdot L_t$ is a martingale under $\mathbb{P}$.*

[1]It will be introduced in the 7th section

Proof. Use the third equation in proposition 4.1,

$$\mathbb{E}^{\mathbb{P}}(X_t L_t|\mathcal{F}_s) = L_s \mathbb{E}^{\mathbb{P}}\left(X_t \frac{L_t}{L_s}|\mathcal{F}_s\right) = L_s \mathbb{E}^{\mathbb{Q}}(X_t|\mathcal{F}_s) = L_s X_s$$

On the other hand,

$$\mathbb{E}^{\mathbb{Q}}(X_s|\mathcal{F}_s) = \mathbb{E}^{\mathbb{P}}\left(\frac{L_t}{L_s} X_t|\mathcal{F}_s\right) = \frac{L_s X_s}{L_s} = X_s$$

□

Note 9. *We can see*

$$\begin{aligned}\frac{S_t}{B_t} \cdot L_t &= \frac{S_t}{B_t} \cdot \exp\left\{-\frac{\mu - r}{\sigma} W_t^{\mathbb{P}} - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 t\right\} \\ &= \frac{S_0}{B_0} \exp\left(C_1 t + C_2 W_t\right)\end{aligned}$$

where

$$C_1 = \mu - r - \frac{1}{2}\sigma^2 - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2, \quad C_2 = \sigma - \frac{\mu - r}{\sigma}$$

One can verify $C_1 - \frac{1}{2}C_2^2 = 0$, cancelling out the drift term in Itô's Lemma so $\frac{S_t}{B_t} \cdot L_t$ is a martingale under $\mathbb{P}$.

Lemma 4.2. ***Martingale Representation Lemma:*** *Let W_t be a Wiener process and $\mathcal{F} := (\mathcal{F}_t)_{t \geq 0}$ the filtration generated by W_t. Let $M := (M_t)_{t \geq 0}$ be a martingale with respect to $\mathcal{F}$. Then there exists a process h adapted to $\mathcal{F}$ such that*

$$M_t = M_0 + \int_0^t h_s dW_s$$

Theorem 4.5. *The Black-Scholes model is **complete**.*

Proof. Define

$$M_t = \mathbb{E}^{\mathbb{Q}}\left[\frac{V_T}{B_T}\middle|\mathcal{F}_t\right]$$

Then M_t is a $\mathbb{Q}$-martingale, and $M_T = \frac{V_T}{B_T}$.

Our goal is to prove that $V_t = M_t B_t$ for all t.

By martingale presentation theory, there exists g_t such that

$$dM_t = g_t dW_t^{\mathbb{Q}}$$

Recall the lemma $\frac{S_t}{B_t} = S_t/B_t \sigma dW_t^{\mathbb{Q}}$, then

$$dM_t = g_t dW_t^{\mathbb{Q}} = \frac{g_t B_t}{\sigma S_t} d\left(\frac{S_t}{B_t}\right).$$

Now let's build a self-financing portfolio Π_t:

$$\Pi_t = \alpha_t B_t + \beta_t S_t = \left(M_t - \frac{S_t}{B_t}\beta_t\right) B_t + \beta_t S_t = M_t B_t$$

Use Itô's Lemma,

$$\begin{aligned}d\Pi_t &= B_t dM_t + M_t dB_t \\ &= B_t \frac{g_t B_t}{\sigma S_t} d\left(\frac{S_t}{B_t}\right) + \left(\alpha_t + \beta_t \frac{S_t}{B_t}\right) dB_t \\ &= B_t \frac{g_t B_t}{\sigma S_t}\left(\frac{dS_t}{B_t} - \frac{S_t}{B_t^2} dB_t\right) + \left(\alpha_t + \beta_t \frac{S_t}{B_t}\right) dB_t \\ &= \left(\alpha_t + \beta_t \frac{S_t}{B_t} - \frac{g_t}{\sigma}\right) dB_t + \frac{g_t B_t}{\sigma S_t} dS_t\end{aligned}$$

Hence

$$\beta_t = \frac{g_t B_t}{\sigma S_t}$$

Now we've already formed a portfolio Π_t which replicates the derivative V_t at time t. So the Black-Scholes model is complete. □

Put M_t into $V_t = M_t B_t$, we have the risk-neutral variation formula:

Theorem 4.6. ***Risk-Neutral Valuation Formula***: *In the Black-Scholes model, the price of a derivative V_t at time t is given by*

$$V_t = B_t \mathbb{E}^{\mathbb{Q}} \left[\left. \frac{V_T}{B_T} \right| \mathcal{F}_t \right]$$

where $\mathbb{Q}$ is the risk-neutral measure.

Note 10. *Actually, the risk-neutral variation formula is same with Feynman-Kac formula.*

4.4 Black-Scholes Formula and Greeks

Finally, we can derive the Black-Scholes formula by solving the Black-Scholes PDE. And thanks to Feynman-Kac, Girsonov Theorem and Risk-Neutral Variation Formula, we can just calculate the expectation of the terminal payoff under the risk-neutral measure.

Theorem 4.7. ***Black-Scholes Formula***: *The price of a European call option in the Black-Scholes model is*

$$C(S_t, t) = S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2)$$

where

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

and Φ is the cumulative distribution function of the standard normal distribution.

Proof. We compute $V(0, S_0)$, other time is similar.The goal is to compute

$$V(0, S_0) = \mathbb{E}\left[e^{-rT}(S_T - K)^+ \mid S_0\right]$$

under a lognormal distribution for the random variable:

$$S_T = S_0 e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z} \qquad Z \sim \mathcal{N}(0,1)$$

This is just calculus which will be laid out here diligently.

$$\begin{aligned}
V(0, S_0) &= e^{-rT}\mathbb{E}\left[(S_T - K)^+ \mid S_0\right] \\
&= e^{-rT}\mathbb{E}\left[\left.\left(S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z} - K\right)^+ \right| S_0\right] \\
&= e^{-rT}\int_{-\infty}^{\infty}\left(S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z} - K\right)^+ \phi(Z)dZ \\
&= e^{-rT}\int_{K^*}^{\infty}\left(S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z} - K\right)\phi(Z)dZ \\
&= S_0\int_{K^*}^{\infty}\left(e^{-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}Z}\right)\phi(Z)dZ - Ke^{-rT}\int_{K^*}^{\infty}\phi(Z)dZ
\end{aligned}$$

The integration variable that we use is Z and the condition $S_T > K$ translates to $Z > K^*$:

$$\begin{aligned}
S_T > K &\Longrightarrow S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z} > K \\
&\Longrightarrow Z > \frac{\log\frac{K}{S_0} - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} := K^*
\end{aligned}$$

For the first part,

$$\begin{aligned}
S_0 \int_{K^*}^{\infty} \left(e^{-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}Z}\right) \phi(Z) dZ &= \frac{S_0}{\sqrt{2\pi}} \int_{K^*}^{\infty} \left(e^{-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}Z - \frac{1}{2}Z^2}\right) dZ \\
&= \frac{S_0}{\sqrt{2\pi}} \int_{K^*}^{\infty} e^{-\frac{1}{2}(Z - \sigma\sqrt{T})^2} dZ \\
&= S_0 P\left(Z > K^* - \sigma\sqrt{T}\right) \\
&= S_0 \Phi\left(\sigma\sqrt{T} - K^*\right)
\end{aligned}$$

For the second part,

$$\begin{aligned}
Ke^{-rT} \int_{K^*}^{\infty} \phi(Z) dZ &= Ke^{-rT} P\left(Z > K^*\right) \\
&= Ke^{-rT} \Phi\left(-K^*\right)
\end{aligned}$$

This gives the famous Black Scholes option pricing formula for a call option on a lognormal non dividend paying asset:

$$\begin{aligned}
C &= S_0 \Phi\left(d_1\right) - Ke^{-rT} \Phi\left(d_2\right) \\
d_1 &= \frac{\log \frac{S_0}{K} + \left(r + \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}} \\
d_2 &= \frac{\log \frac{S_0}{K} + \left(r - \frac{1}{2}\sigma^2\right) T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}
\end{aligned}$$

□

Use the following Python code to calculate the price of a European call option.

```python
from scipy.stats import norm
def bs_call_price(S, K, T, r, sigma):
    d1 = (np.log(S / K) + (r + 0.5 * sigma ** 2) * T) / (sigma * np.sqrt(T))
    d2 = d1 - sigma * np.sqrt(T)
    return S * norm.cdf(d1) - K * np.exp(-r * T) * norm.cdf(d2)
```

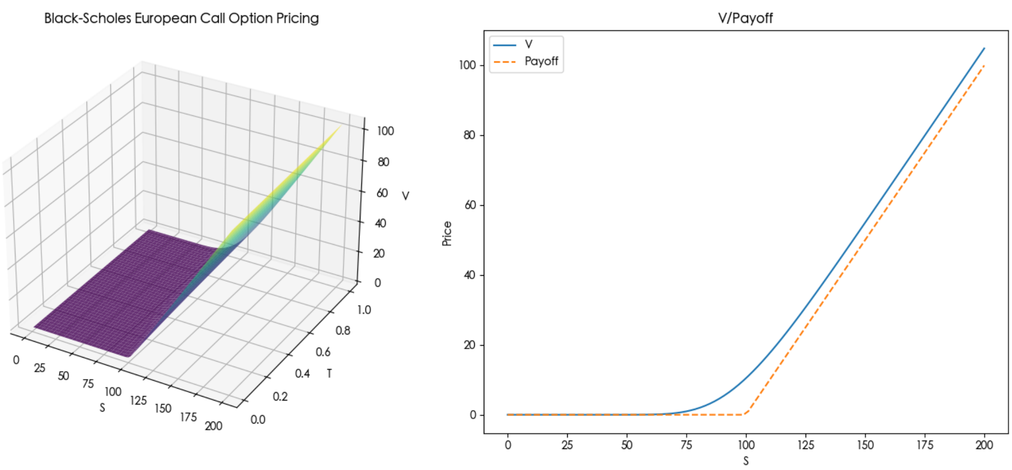

Figure 4.4: Price of a European call option

For European put option, we have the put-call parity:

$$C(t, S_t) - P(t, S_t) = S_t - Ke^{-r(T-t)}, \quad P(t, S_t) = Ke^{-r(T-t)}\Phi(-d_2) - S_t\Phi(-d_1)$$

Greeks

Definition 4.6. ***Greeks*** *are the sensitivities of the option price to changes in the parameters of the model.*

- ***Delta***: $\Delta = \dfrac{\partial C}{\partial S}$, *the sensitivity of the option price to changes in the stock price.*

- ***Gamma***: $\Gamma = \frac{\partial^2 C}{\partial S^2}$, *the sensitivity of the delta to changes in the stock price.*
- ***Theta***: $\Theta = \frac{\partial C}{\partial t}$, *the sensitivity of the option price to changes in time.*
- ***Vega***: $\nu = \frac{\partial C}{\partial \sigma}$, *the sensitivity of the option price to changes in the volatility.*
- ***Rho***: $\rho = \frac{\partial C}{\partial r}$, *the sensitivity of the option price to changes in the risk-free rate.*

For BS model and a European call option, we have:

$$\begin{aligned}
&\Delta = \Phi\left(d_1\right), \quad \Gamma = \frac{\phi\left(d_1\right)}{S_0 \sigma \sqrt{T}} \\
&\Theta = -\frac{S_0 \phi\left(d_1\right) \sigma}{2\sqrt{T}} - rKe^{-rT}\Phi\left(d_2\right), \quad \nu = S_0\sqrt{T}\phi\left(d_1\right) \\
&\rho = KTe^{-rT}\Phi\left(d_2\right)
\end{aligned}$$

where ϕ is the density function of the standard normal distribution.

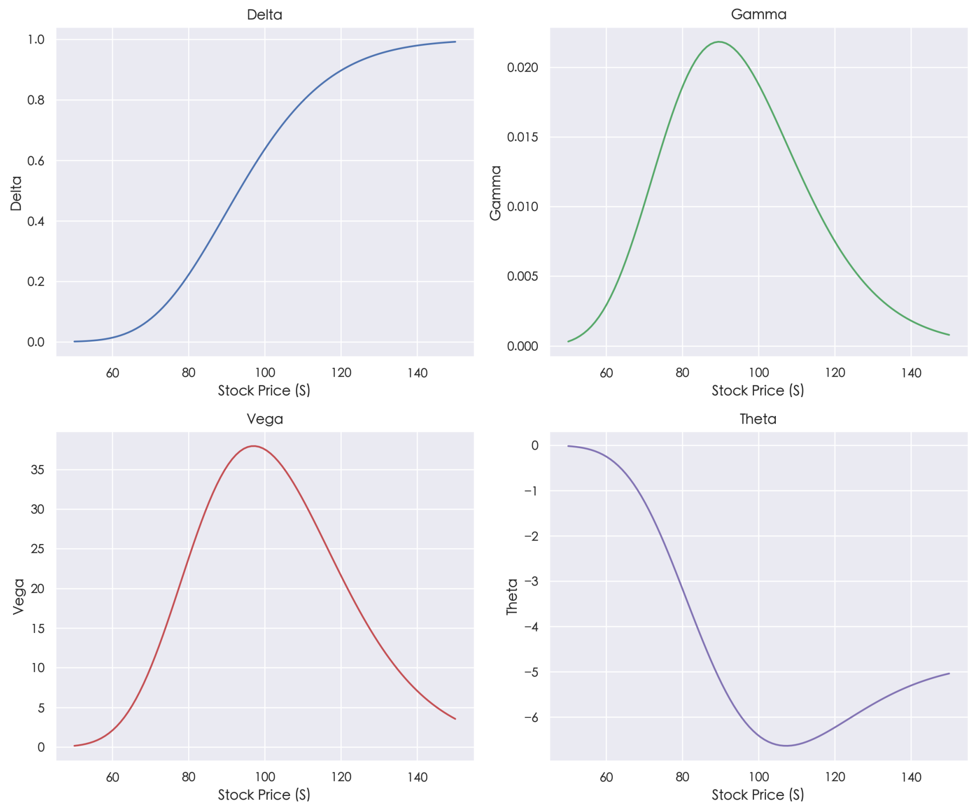

Figure 4.5: Greeks of a European call option

Definition 4.7. ***Delta Neutral Portfolio***: *A portfolio is said to be* ***delta neutral*** *if its delta is zero.*
Gamma Neutral Portfolio: *A portfolio is said to be* ***gamma neutral*** *if its gamma is zero.*

If I have sold a financial derivative with price $V\left(t, S_t\right)$, how many stocks do I need to be Δ-neutral (Δ-hedged)?

$$\frac{\partial}{\partial S}\left[-V(t, S) + \beta S\right] = 0 \text{ implies } \beta = \frac{\partial V}{\partial S}.$$

In a Δ-hedging strategy, the portfolio is constantly rebalanced to be Δ-neutral.
If Γ is high, the portfolio needs to be rebalanced frequently.

The stock has $\Gamma = 0$. So, if I have sold a financial derivative, I need to trade also a different financial derivative to hold a Γ-neutral portfolio.

Let $G(t, S_t)$ be the price of the second financial derivative.

$$\frac{\partial^2}{\partial S^2}[-V(t,S) + \gamma G(t,S)] = 0 \text{ implies } \gamma = \frac{\Gamma_V}{\Gamma_G}.$$

Adjusting the position in stocks leads then to a Δ - and Γ-neutral portfolio:

1. Choose $\gamma = \Gamma_V / \Gamma_G$;
2. Choose $\beta = \frac{\partial V}{\partial S} - \gamma \frac{\partial G}{\partial S}$;
3. Then the portfolio $-V(t,S) + \gamma G(t,S) + \beta S$ is Δ-and Γ-neutral.

4.5 Dividend-Paying Stock

How does the Black-Scholes equation change if the stock pays dividends? We now assume the stock pays a continuous stream of dividends $\{qS_t\}_{t\geq 0}$.

Let D_t be the cumulative dividends paid to the holder of one stock. The dividend process of the stock has dynamics

$$dD_t = S_t q dt$$

The gain process is $G_t = S_t + D_t$.

$$dG_t = (\mu + q)dS_t + \sigma S_t dW_t$$

What is the price at time t of a derivative that pays $\varphi(S_T)$ at maturity T ?

Similarly, we form a portfolio Π_t:

$$\Pi_t = V_t + \alpha_t S_t + \beta_t B_t$$

Then we should have

$$\beta_t B_t = -\alpha_t S_t - V_t$$

However, due to the dividends, so $d\Pi_t$ under self-financing will be:

$$\begin{aligned}
d\Pi_t &= dV_t + \alpha_t dG_t + \beta_t r B_t dt \\
&= \frac{\partial V}{\partial t}dt + \frac{\partial^2 V}{2\partial S^2}(\sigma S)^2 dt + \frac{\partial V}{\partial S}dS + \alpha_t(dS_t + qS_t dt) + \beta_t r B_t dt \\
&\overset{\alpha_t = -\frac{\partial V}{\partial S}}{=\!=\!=} \frac{\partial V}{\partial t}dt + \frac{\partial^2 V}{2\partial S^2}(\sigma S)^2 dt - q\frac{\partial V}{\partial S}S_t dt + \beta_t r B_t dt \\
&\overset{\beta_t B_t = \frac{\partial V}{\partial S}S_t - V_t}{=\!=\!=\!=\!=} \left(\frac{\partial V_t}{\partial t} + (r-q)S_t\frac{\partial V_t}{\partial S_t} + \frac{1}{2}(\sigma S_t)^2\frac{\partial^2 V_t}{\partial S_t^2} - rV_t\right)dt
\end{aligned}$$

This one equals to 0, hence we have the Black-Scholes PDE for a dividend-paying stock:

$$\frac{\partial V}{\partial t} + (r-q)S_t\frac{\partial V}{\partial S} + \frac{1}{2}(\sigma S_t)^2\frac{\partial^2 V}{\partial S^2} - rV = 0$$

Applying the Feynman-Kac Formula, we get that

$$V(t, S_t) = e^{-r(T-t)}\mathbb{E}_{t,s}^{\mathbb{Q}}[\varphi(S_T)],$$

where the dynamics of S_t are given by

$$dS_t = S_t(r-q)dt + S_t\sigma dW_t^{\mathbb{Q}}.$$

The stock price is

$$\begin{aligned}
S_T &= S_t \exp\left\{\left(r - q - \frac{\sigma^2}{2}\right)(T-t) + \sigma\left(W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}}\right)\right\} \\
&= S_t e^{-q(T-t)} \exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\left(W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}}\right)\right\}
\end{aligned}$$

Hence to obtain the pricing formulas for European-style derivatives on dividend-paying stocks, replace S_t with $S_t e^{-q(T-t)}$, and other part of calculation remains the same.

Theorem 4.8. ***Black-Scholes Formula for Dividend-Paying Stock***: *The price of a European call option on a dividend-paying stock is*

$$C(S_t, t) = S_t e^{-q(T-t)}\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r - q + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

4.6 Change of Numeraire

We've already seen the price of the stock **renormalized** by the value of the bond, $\frac{S_t}{B_t}$, is a martingale under $\mathbb{Q}$. Here we define B_t as a **numeraire**, a unit of account. So $\frac{S_t}{B_t}$ is the price of the stock in units of bonds.

With B_t as numeraire, risk-neutral valuation is

$$\frac{V_t}{B_t} = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{V_T}{B_T}\right].$$

Under risk-neutral measure $\mathbb{Q}$, the price of the derivative in units of the bond B is a martingale.

If S_t is the price of a financial asset (say, a stock), then

$$\frac{S_t}{B_t} = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{S_T}{B_T}\right].$$

If we choose another financial asset as the numeriare, say N_t, does it exist another probability measure $\mathbb{Q}^N$ under which the price of the derivative in units of the new numeraire is still a martingale?

Define $\mathbb{Q}^N$ by

$$L_T = \frac{d\mathbb{Q}^N}{d\mathbb{Q}} := \frac{B_0}{B_T}\cdot\frac{N_T}{N_0}$$

where Radon-Nikodym derivative L_T is the ratio of the final value of the numeraire under the two measures.

One can verify that L_T is positive and L_T has expectation 1:

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[L_T] &= \mathbb{E}^{\mathbb{Q}}\left[\frac{B_0}{B_T}\cdot\frac{N_T}{N_0}\right] = \frac{B_0}{N_0}\cdot\mathbb{E}^{\mathbb{Q}}\left[\frac{N_T}{B_T}\right]\\ &= \frac{B_0}{N_0}\cdot\frac{N_0}{B_0} = 1.\end{aligned}$$

For $t = 0$, we have

$$V_0 = B_0\mathbb{E}^{\mathbb{Q}}\left[\frac{V_T}{B_T}\right] = B_0\mathbb{E}^{\mathbb{Q}^N}\left[\frac{1}{L_T}\frac{V_T}{B_T}\right] = N_0\mathbb{E}^{\mathbb{Q}^N}\left[\frac{V_T}{N_T}\right]$$

To sum up, we have the following theorem:

Theorem 4.9. ***Change of Numeraire I***: *Let $B(t)$ be the bank account, $S(t), N(t)$ be 2 assets respectively. Under risk-neutral measure $\mathbb{Q}$, we have*

$$d\left(\frac{S_t}{B_t}\right) = \sigma_1\frac{S_t}{B_t}dW_t^{\mathbb{Q}}, \quad d\left(\frac{N_t}{B_t}\right) = \sigma_2\frac{N_t}{B_t}dW_t^{\mathbb{Q}}$$

Set $N(t)$ as the new numeraire, and define new measure $\mathbb{Q}^N$ by

$$L_T = \frac{d\mathbb{Q}^N}{d\mathbb{Q}} = \frac{B_0}{B_T}\frac{N_T}{N_0}$$

Then

$$dW_t^N = dW_t^{\mathbb{Q}} - \sigma_2 dt$$

is a Brownian motion under $\mathbb{Q}^N$. Let $X_t = \frac{S_t}{N_t}$ be the price of the asset S_t in units of the numeraire N_t, then

$$dX_t = (\sigma_1 - \sigma_2)X_t dW_t^{\mathbb{Q}^N}$$

For pricing a derivative, we have

Theorem 4.10. ***Change of Numeraire II**: Let V_t be the price of a derivative at time t with payoff V_T at time T. Let B_t and N_t be two numeraire processes. Then there **exists** a measure $\mathbb{Q}^N$ such that the price of the European style derivative at time t in units of the numeraire N_t is*

$$\frac{V_t}{N_t} = \mathbb{E}_t^{\mathbb{Q}^N}\left[\frac{V_T}{N_T}\right]$$

where V_t is the price of the derivative at time t.

Example 14. *Let the price dynamics of a non-dividend stock under $\mathbb{Q}$ be*

$$dS_t = rS_t dt + \sigma S_t dW_t^{\mathbb{Q}}$$

Derive the dynamics of S_t under the measure $\mathbb{Q}^S$ where S_t is the numeraire.

By definition, let $L_T = \frac{d\mathbb{Q}^S}{d\mathbb{Q}} = \frac{B_0}{B_T}\frac{S_T}{S_0}$, then

$$L^T = \exp\left\{\int_0^T \sigma dW_t - \frac{1}{2}\int_0^T \sigma^2 dt\right\}$$

Apply Girsanov Theorem, we have

$$dW_t^{\mathbb{Q}^S} = dW_t^{\mathbb{Q}} - \sigma dt$$

is a Brownian motion under $\mathbb{Q}^S$. Hence the dynamics of S_t under $\mathbb{Q}^S$ is

$$dS_t = rS_t + \sigma S_t\left(dW_t^{\mathbb{Q}^S} + \sigma dt\right) = (r+\sigma^2)S_t dt + \sigma S_t dW_t^{\mathbb{Q}^S}$$

which is a Geometric Brownian Motion with drift $(r+\sigma^2)$ and volatility σ.

$$S_t = S_0 \exp\left\{\left(r + \frac{1}{2}\sigma^2\right)t + \sigma W_t^{\mathbb{Q}^S}\right\}$$

1. *What's today's price of the derivative with payoff S_T if $S_T > K$ and 0 otherwise?*

   The payoff is $\phi(S_T) = S_T \mathbb{1}_{\{S_T > K\}}$, then we choose S_t as the numeraire, so the price of the derivative is

$$\begin{aligned}
V_t &= S_t \mathbb{E}^{\mathbb{Q}^S}\left[\frac{S_T}{S_T}\mathbb{1}_{\{S_T>K\}}\right] \\
&= S_t \mathbb{Q}^S\left[\left(r+\frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}Z > \log\left(\frac{K}{S_0}\right)\right] \\
&= S_t \mathbb{Q}^S\left[-Z < \frac{\log\left(\frac{K}{S_0}\right) + \left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&= S_t \Phi(d_1)
\end{aligned}$$

   where $d_1 = \dfrac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$.

2. *Find the price at time 0 of a European option that pays $S_T \mathbb{1}_{\{\log(S_T) \geq K\}}$ at maturity T.*

   The calculation is similar and answer is $S_t\Phi(d_1)$, but here the definition of d_1 is

$$d_1 = \frac{\log(S_0) - K + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$$

3. *Find the price at time 0 of a European option that pays $S_T\left(\log\left(S_T\right)-K\right)^+$ at maturity T. Hint: It could be useful to use that for the density function f of a standard normal distribution it holds $f'(x)=-xf(x)$ for all $x\in\mathbb{R}$.*

$$\begin{aligned}
V_0 &= S_0\mathbb{E}^{\mathbb{Q}^S}\left[(\log(S_T)-K)^+\right] \\
&= S_0\int_{-\infty}^{\infty}\left(\log(S_0e^{(r+\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}z})-K\right)^+\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
&= S_0\int_{-\infty}^{\infty}\left(\log(S_0)+\left(r+\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}z-K\right)^+\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
&= S_0\int_{-d_1}^{\infty}\left(\log(S_0)+\left(r+\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}z-K\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
&= S_0\left(\log(S_0)+\left(r+\frac{1}{2}\sigma^2\right)T-K\right)\Phi(d_1)+S_0\sigma\sqrt{T}\int_{-d_1}^{\infty}zf(z)dz \\
&= S_0\left(\log(S_0)+\left(r+\frac{1}{2}\sigma^2\right)T-K\right)\Phi(d_1)-S_0\sigma\sqrt{T}f(d_1) \\
&= S_0\sigma\sqrt{T}\left(d_1\Phi(d_1)-f(d_1)\right)
\end{aligned}$$

4.7 Multi-Dimensional Black-Scholes Model

4.7.1 PDE Framework

The Black Scholes model extended naturally to derivatives with payoff contingent on multiple assets (basket options are a typical example. Each underlying asset follows its own lognormal SDE, although the brownian increments may be correlated:

$$dS_i=\mu_iS_idt+\sigma_iS_idW_t^i \quad \left\langle dW_t^i,dW_t^j\right\rangle=\rho_{ij}dt \quad \forall i,j\in[1,d]$$

Note 11. *If W_t^1 and W_t^2 have correlation ρ, then there exists a Wiener process Z_t such that*

1. *Z_t is independent of W_t^1*
2. *$W_t^2=\rho W_t^1+\sqrt{1-\rho^2}Z_t$.*

Form the portfolio Π_t:

$$\Pi_t=V_t+\sum_{i=1}^{d}\alpha_iS_i+\beta B_t$$

Then we should have

$$\beta B_t=-\sum_{i=1}^{d}\alpha_iS_i-V_t$$

and self-financing condition gives

$$\sum_{i=1}^{d}S_id\alpha_i+B_td\beta=0$$

$$\text{i.e.}\quad d\Pi=dV_t+\sum_{i=1}^{d}\alpha_idS_i+\beta dB_t$$

dV follows Multi-Dimensional Itô's Lemma shown below:

Theorem 4.11. ***Multi-Dimensional Itô's Lemma**: Let $f(t,x_1,x_2,\dots,x_d)$ be a function of t and d independent variables $x_1,x_2,\dots,x_d$ that follow the SDEs:*

$$dx_i=\mu_ix_idt+\sigma_ix_idW_t^i \quad \left\langle dW_t^i,dW_t^j\right\rangle=\rho_{ij}dt \quad \forall i,j\in[1,d]$$

Then $f(t, x_1, x_2, \ldots, x_d)$ follows the SDE:

$$
\begin{aligned}
df &= \frac{\partial f}{\partial t}dt + \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}dx_i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}dx_i dx_j \\
&= \left(\frac{\partial f}{\partial t} + \sum_{i=1}^{d} \mu_i x_i \frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \sigma_i \sigma_j x_i x_j \rho_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \right) dt + \sum_{i=1}^{d} \sigma_i x_i \frac{\partial f}{\partial x_i} dW_t^i
\end{aligned}
$$

Hence

$$
dV_t = \frac{\partial V}{\partial t}dt + \sum_{i=1}^{d} \frac{\partial V}{\partial S_i}dS_i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 V}{\partial S_i \partial S_j}dS_i dS_j
$$

So

$$
\begin{aligned}
d\Pi_t &= dV_t + \sum_{i=1}^{d} \alpha_i dS_i + \beta dB_t \\
&= \frac{\partial V}{\partial t}dt + \sum_{i=1}^{d} \frac{\partial V}{\partial S_i}dS_i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 V}{\partial S_i \partial S_j}dS_i dS_j + \sum_{i=1}^{d} \alpha_i dS_i + \beta dB_t \\
&\xlongequal[\text{Delta-Hedging}]{\alpha_i = -\frac{\partial V}{\partial S_i}} \left(\frac{\partial V}{\partial t} + \beta r B_t \right) dt + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 V}{\partial S_i \partial S_j}dS_i dS_j \\
&\xlongequal[\text{Replication Condition}]{\beta B_t = \sum_{i=1}^{d} \frac{\partial V}{\partial S_i} S_i - V_t} \left[\frac{\partial V}{\partial t} + \left(\sum_{i=1}^{d} \frac{\partial V}{\partial S_i} S_i - V_t \right) r \right] dt + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 V}{\partial S_i \partial S_j}dS_i dS_j \\
&\xlongequal[\text{Unfold}]{dS_i dS_j} \left[\frac{\partial V}{\partial t} + \left(\sum_{i=1}^{d} \frac{\partial V}{\partial S_i} S_i - V_t \right) r \right] dt + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 V}{\partial S_i \partial S_j} \sigma_i \sigma_j S_i S_j \rho_{ij} dt
\end{aligned}
$$

Let $d\Pi_t = 0$, we have the Multi-Dimensional Black-Scholes PDE:

$$
\frac{\partial V}{\partial t} + \sum_{i=1}^{d} \frac{\partial V}{\partial S_i} \mu_i S_i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 V}{\partial S_i \partial S_j} \sigma_i \sigma_j \rho_{ij} S_i S_j - rV = 0
$$

Feynman-Kac says, the risk neutral dynamics that will fit in with the PDE is

$$
dS_i = rS_i dt + \sigma_i S_i dW_t^i
$$

Theorem 4.12. ***Multi-Dimensional Girsonov Theorem**: Let $(\theta_1, \theta_2, \cdots, \theta_n)$ be a constant vector, and let $W_{1,t}^{\mathbb{P}}, W_{2,t}^{\mathbb{P}}, \cdots, W_{n,t}^{\mathbb{P}}$ be independent Brownian motions under $\mathbb{P}$. Then, there exists a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that*

$$
\begin{aligned}
dW_{1,t}^{\mathbb{Q}} &= dW_{1,t}^{\mathbb{P}} - \theta_1 dt \\
dW_{2,t}^{\mathbb{Q}} &= dW_{2,t}^{\mathbb{P}} - \theta_2 dt \\
&\vdots \\
dW_{n,t}^{\mathbb{Q}} &= dW_{n,t}^{\mathbb{P}} - \theta_n dt
\end{aligned}
$$

are independent Brownian motions under $\mathbb{Q}$.

Change measure to $\mathbb{Q}$, we have

$$
V(t, \mathbf{S}_t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)} V(T, \mathbf{S}_T) \mid \mathbf{S}_t \right]
$$

Simulation using Monte Carlo

We can simulate the expectation using Monte Carlo method, that is, generate a large number of paths of the stock prices and calculate the average payoff of the derivative at maturity. For each underlying asset, we generate 10000 normal random variables and simulate the stock price path. Then we calculate the payoff of the derivative at maturity and take the average. The following Python code shows how to simulate the price of a European call option with a payoff of $(S_1 + S_2 - K)^+$.

```
import numpy as np
# Parameters
r, sigma, T, t0, K = 0.05, 0.2, 1.0, 0.0, 150
num_assets = 2  # number of assets
num_simulations = 10000  # number of Monte Carlo simulations

# Correlation matrix and Cholesky decomposition
corr_matrix = np.full((num_assets, num_assets), rho)
np.fill_diagonal(corr_matrix, 1.0)
L = np.linalg.cholesky(corr_matrix)

# Generate correlated random variables
np.random.seed(42)  # for reproducibility
Z_perp = np.random.normal(size=(num_assets, num_simulations))
Z = L @ Z_perp

# Define a grid of initial prices for asset 1 and asset 2
S1_range = np.linspace(0, 300, 100)
S2_range = np.linspace(0, 300, 100)
S1_grid, S2_grid = np.meshgrid(S1_range, S2_range)

# Calculate option prices for each combination of S1 and S2
def calculate_option_price(S_0):
    S_T = S_0[:, np.newaxis] * np.exp((r - 0.5 * sigma**2) * (T - t0) + sigma *
     ↪ np.sqrt(T - t0) * Z)
    payoff = np.maximum(np.sum(S_T, axis=0) - K, 0)
    option_price = np.exp(-r * (T - t0)) * np.mean(payoff) # Monte Carlo estimator
    return option_price

option_prices = np.array([[calculate_option_price(np.array([S1, S2])) for S1 in
 ↪ S1_range] for S2 in S2_range])
payoff = np.maximum(S1_grid + S2_grid - K, 0)
```

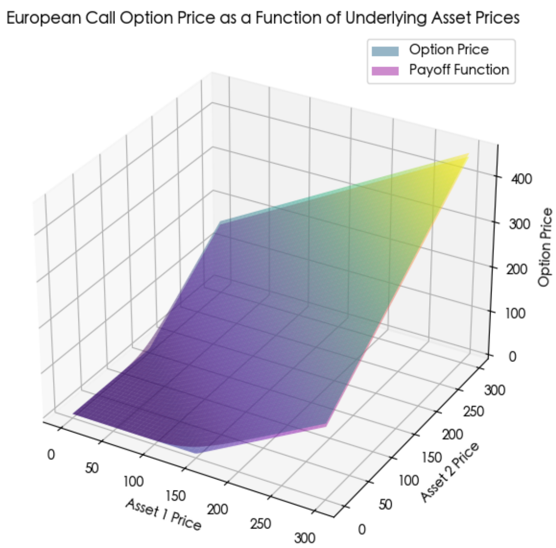

Figure 4.6: Price of a European call option with two underlying assets

Basket Options

Consider $d = 2$ and a derivative $V(t, S_1(t), S_2(t))$ which pays at expiry T a call on the marimurn of the two assets. Mathematically we have the terminal payoff:

$$V(T) = (M_T - K)^+ \qquad \text{where } M_T = \max(S_1(T), S_2(T))$$

So the risk-neutral framwork gives us

$$V\left(t_0, S_1\left(t_0\right), S_2\left(t_0\right)\right) = \mathbb{E}^0\left[e^{-r(T-t_0)}\left(M_T - K\right)^+ \mid S_1\left(t_0\right), S_2\left(t_0\right)\right]$$

For basket options, we define a Python Class:

```
class basket:
    def __init__(self, num_assets, S0, K, r, sigma, T, t0, rho, num_simulations):
        self.num_assets, self.S0, self.K, self.r, self.sigma, self.T, self.t0, self.rho, self.num_simulations = num_assets, S0, K, r, sigma, T, t0, rho, num_simulations

    def generate_Z(self):
        corr_matrix = np.full((self.num_assets, self.num_assets), self.rho)
        np.fill_diagonal(corr_matrix, 1.0)
        L = np.linalg.cholesky(corr_matrix)
        Z_perp = np.random.normal(size=(self.num_assets, self.num_simulations))
        Z = L @ Z_perp
        return Z

    def payoff(self):
        S_T = self.S0[:, np.newaxis] * np.exp((self.r - 0.5 * self.sigma**2) * (self.T - self.t0) + self.sigma * np.sqrt(self.T - self.t0) * self.generate_Z())
        M_0 = np.max(self.S0)
        M_T = np.max(S_T, axis=0)
        payoff_0 = np.maximum(M_0 - self.K, 0)
        payoff_T = np.maximum(M_T - self.K, 0)
        return [payoff_0, payoff_T]

    def price(self):
        payoff = self.payoff()[1]
        option_price = np.exp(-self.r * (self.T - self.t0)) * np.mean(payoff)
        return option_price
```

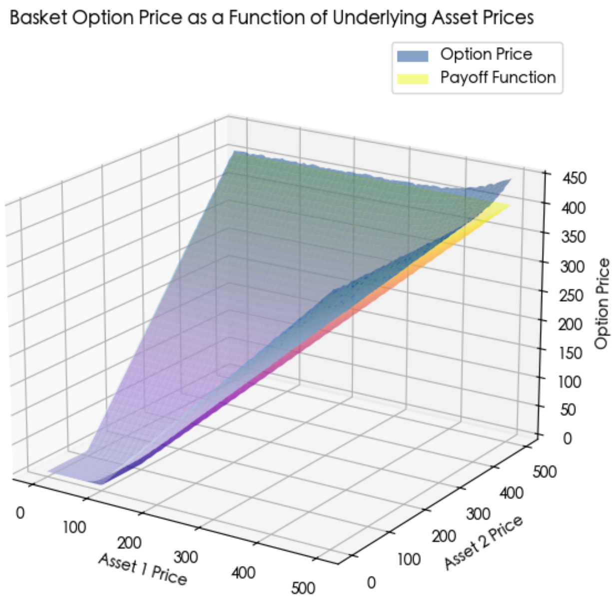

Figure 4.7: Price of a Basket option with two underlying assets

- If correlation is close to 1 then both assets will move in the same direction; either both assets will end up above the strike, triggering a payoff, or both of them will end up below the strike, resulting in no payoff.
- If correlation is close to -1 then both assets will move in opposite directions; one asset will end up above the strike triggering a payoff. It doesn't matter which one ends up above the strike, to trigger a payoff either one of them

needs to end up above the strike. In this case, the payoff is more likely to be triggered. So we expect that the option price to decrease as the correlation increases.

4.7.2 Exchange Options

Definition 4.8. *An **exchange option** is a derivative that gives the holder the right, but not the obligation, to exchange one asset for another at a predetermined exchange rate K at maturity T. Always let $K = 1$.*

The holder can choose to exercise the option and exchange asset 2 for asset 1. If the price of 1 is higher than the price of 2, the return is $S_1 - S_2$; If S_2 is higher than S_1, the option holder will not exercise the option and the return is zero. Therefore, the holder always obtains the more advantageous choice between the two.

Assume the price of these 2 assets follows the following SDEs:

$$\begin{aligned} dS_1 &= \mu_1 S_1 dt + \sigma_1 S_1 dW_t^1 \\ dS_2 &= \mu_2 S_2 dt + \sigma_2 S_2 dW_t^2 \\ d\langle W_t^1, W_t^2 \rangle &= \rho dt \end{aligned}$$

Consider $d = 2$ and a derivative $V(t, S_1(t), S_2(t))$ which pays at expiry T a call on the maximum between $S_1(T) - S_2(T)$ and 0. Mathematically we have the terminal payoff:

$$V(T) = (S_1(T) - S_2(T))^+ = \max(S_1(T) - S_2(T), 0)$$

We shall compute

$$V(t, S_1(t), S_2(t)) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-r(T-t)} \max(S_1(T) - S_2(T), 0)\right]$$

Since there're 2 assets, an intuitive method is to let S_2 be the numeraire. Let $X_t = \frac{S_1(t)}{S_2(t)}$, then

$$\begin{aligned} dX_t &= -\frac{S_1(t)}{S_2(t)^2} dS_2 + \frac{S_1(t)}{S_2(t)^3}(dS_2(t))^2 + \frac{1}{S_2(t)} dS_1 - \frac{1}{S_2(t)^2} dS_1 dS_2 \\ &= X_t\left((\mu_1 + \sigma_2^2 - \mu_2 - \sigma_1\sigma_2\rho)dt + (\sigma_1\rho - \sigma_2)dW_t^2 + \sigma_1\sqrt{1-\rho^2}dZ_t\right) \end{aligned}$$

Let $Y_t = \frac{B_t}{S_2(t)}$, then

$$dY_t = Y_t\left((r + \sigma_2^2 - \mu_2)dt - \sigma_2 dW_t^2\right)$$

Recap that the purpose of measure change is to make the price of one asset in units of the numeraire a martingale. So we need to find a measure $\mathbb{Q}^*$ such that X_t, Y_t are martingales. By Girsanov Theorem, there exists a measure $\mathbb{Q}^*$, under which

$$dZ_t^{1*} = dW_2^t - \theta_1 dt, \quad dZ_t^{2*} = dZ_t - \theta_2 dt$$

are independent Brownian motions. So under $\mathbb{Q}^*$, we have

$$\begin{aligned} dX_t &= X_t\left((\mu_1 + \sigma_2^2 - \mu_2 - \sigma_1\sigma_2\rho)dt + (\sigma_1\rho - \sigma_2)dW_t^2 + \sigma_1\sqrt{1-\rho^2}dZ_t\right) \\ &= X_t\left((\mu_1 + \sigma_2^2 - \mu_2 - \sigma_1\sigma_2\rho)dt + (\sigma_1\rho - \sigma_2)(dZ_t^{1*} + \theta_1 dt) + \sigma_1\sqrt{1-\rho^2}(dZ_t^{2*} + \theta_2 dt)\right) \end{aligned}$$

and

$$dY_t = Y_t\left((r + \sigma_2^2 - \mu_2)dt - \sigma_2(dZ_t^{1*} + \theta_1 dt)\right)$$

Choose

$$\theta_1 = \frac{r + \sigma_2^2 - \mu_2}{\sigma_2}, \quad \theta_2 = -\frac{\mu_1 + \sigma_2^2 - \mu_2 - \sigma_1\sigma_2\rho + (\sigma_1\rho - \sigma_2)\theta_1}{\sigma_1\sqrt{1-\rho^2}}$$

then X_t, Y_t are martingales under $\mathbb{Q}^*$.

$$dX_t = X_t(\sigma_1\rho - \sigma_2)dZ_t^{1*} + X_t\sigma_1\sqrt{1-\rho^2}dZ_t^{2*}$$

Since the feasible region of the option is $S_1 \geq S_2$, we have $\log(X_t) \geq 0$.

$$d\log(X_t) = \frac{1}{X_t}dX_t - \frac{1}{2X_t^2}(dX_t)^2 = (\sigma_1\rho - \sigma_2)dZ_t^{1*} + \sigma_1\sqrt{1-\rho^2}dZ_t^{2*} - \frac{1}{2}(\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho)dt$$

Let $\sigma^* = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho}$, then

$$\log(X_T) = \log(X_t) + \int_t^T (\sigma_1\rho - \sigma_2)dZ_t^{1*} + \sigma_1\sqrt{1-\rho^2}dZ_t^{2*} - \frac{1}{2}\sigma^{*2}dt$$

Hence

$$X_T = X_t \cdot e^U, U \sim \mathcal{N}\left(-\frac{1}{2}\sigma^{*2}(T-t), \sigma^{*2}(T-t)\right)$$

Finally we have

Theorem 4.13. ***Margrabes Formula**: The price of an exchange option is*

$$V(t, S_1(t), S_2(t)) = S_1(t)\Phi(d_1) - S_2(t)\Phi(d_2)$$

where

$$d_1 = \frac{\log\left(\frac{S_1(t)}{S_2(t)}\right) + \frac{\sigma^{*2}}{2}(T-t)}{\sigma^*\sqrt{T-t}}, \quad d_2 = d_1 - \sigma^*\sqrt{T-t}$$
$$\sigma^* = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho}$$

Proof. Let $L_T = \frac{B(t)}{B(T)}\frac{S_2(T)}{S_2(t)}$, recap that this defines a measure change from risk neutral measure $\mathbb{Q}$ to measure with S_2 as numeraire $\mathbb{Q}^*$. Then

$$\begin{aligned}
V_t &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[(S_1(T) - S_2(T))^+\right] \\
&= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}^*}\left[\frac{1}{L_T}(S_1(T) - S_2(T))^+\right] = S_2(t)\mathbb{E}_t^{\mathbb{Q}^*}[(X_t e^U - 1)^+] \\
&= \frac{1}{\sqrt{2\pi(T-t)}\sigma^*}S_2(t)\int_{-\log(X_t)}^{\infty}(X_t e^u - 1)\exp\left\{-\frac{\left(u + \frac{1}{2}(T-t)\sigma^{*2}\right)^2}{2(T-t)\sigma^{*2}}\right\}du \\
&= \frac{1}{\sqrt{2\pi}}S_2(t)X_t\int_{-d_2}^{\infty}\exp\left\{-\frac{1}{2}\left(y - \sqrt{T-t}\sigma^*\right)^2\right\}dy - \frac{1}{\sqrt{2\pi}}S_2(t)\int_{-d_2}^{\infty}\exp\left\{-\frac{1}{2}y^2\right\}dy \\
&= S_1(t)\Phi(d_1) - S_2(t)\Phi(d_2)
\end{aligned}$$

□

Note 12. *One can also derive this formula directly from applying ordinary Black-Scholes Formula, with $r = 0, S = X, K = 1, \sigma = \sigma^*$.*

$$V(t, S_1(t), S_2(t)) = S_2(t)\left(X_t\Phi(d_1) - \Phi(d_2)\right) = S_1(t)\Phi(d_1) - S_2(t)\Phi(d_2)$$

where

$$d_1 = \frac{\log\left(\frac{X_t}{1}\right) + \frac{\sigma^{*2}}{2}(T-t)}{\sigma^*\sqrt{T-t}}, \quad d_2 = d_1 - \sigma^*\sqrt{T-t}$$

Note 13. *In the process of changing measure to $\mathbb{Q}^*$ under which S_2 is the numeraire, we can also first change measure to risk neutral measure $\mathbb{Q}$, then change measure to $\mathbb{Q}^*$.*

Under $\mathbb{Q}$, we have

$$dX_t = X_t(\sigma_2^2 - \sigma_1\sigma_2\rho)dt + X_t(\sigma_1\rho - \sigma_2)dW_t^{\mathbb{Q}2} + X_t\sigma_1\sqrt{1-\rho^2}dZ_t^{\mathbb{Q}}$$

and

$$dY_t = \sigma_2^2 Y_t dt - \sigma_2 Y_t dW_t^{\mathbb{Q}2}$$

Define

$$\theta_1 = \sigma_2, \quad \theta_2 = 0$$

then

$$dZ_t^{1*} = dW_t^{\mathbb{Q}2} - \sigma_2 dt, \quad dZ_t^{2*} = dZ_t^{\mathbb{Q}}$$

are independent Brownian motions under $\mathbb{Q}^*$.

Hence we still derive

$$dX_t = X_t(\sigma_1\rho - \sigma_2)dZ_t^{1*} + X_t\sigma_1\sqrt{1-\rho^2}dZ_t^{2*}$$

5 Implied Volatility Skew(Smile)

This part includes:

1. Implied Volatility and Volatility Smile.
2. Market Implied density.
3. Local volatility model and Dupire's Equation.
4. Stochastic Volatility Model.

The most subjective parameter in the BS model is the volatility σ. All the others are either defined by the terms of the contract or by some economic indicator.

- In a lognormal model the volatility describes the variance of the Log-price over some time horizon.
- In the general case when **spot volaitility** is a deterministic function of time then the variance of the log-price is given by

  $$\int_0^t \sigma^2(s)ds$$

  This can be conveniently represented as some average volatility $\overline{\sigma}$ over the time interval $[0, T]$.

  $$\overline{\sigma} = \sqrt{\frac{1}{T}\int_0^T \sigma^2(s)ds}$$

5.1 Implied Volatility

Definition 5.1. ***Implied Volatility*** *is the volatility parameter that makes the theoretical price of an option equal to the market price of the option. Define market price is $M(S_0, r, T, K)$, then the implied volatility is the value σ_I such that BS price $\mathrm{BS}(\sigma_I; S_0, r, T, K)$ equals the market price:*

$$M(S_0, r, T, K) = \mathrm{BS}(\sigma_I; S_0, r, T, K)$$

Note 14. *The implied volatility for a call should be the same as for a put.*

Newton Raphson Method

The Newton-Raphson method is a root-finding algorithm that uses linear approximation to find the roots of a real-valued function. It is a very powerful method that converges quickly to the root. The formula for the Newton-Raphson method is:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

where $f(x)$ is the function whose root we are trying to find, and $f'(x)$ is the derivative of the function.

Let's define

$$f(\sigma_I) = \mathrm{BS}(\sigma_I; S_0, r, T, K) - M(S_0, r, T, K)$$

Then the Newton-Raphson method for finding the implied volatility is:

$$\sigma_I^{k+1} = \sigma_I^k - f(\sigma_I^k)\left(\left.\frac{\partial\,\mathrm{BS}}{\partial\sigma_I}\right|_{\sigma_I^k}\right)^{-1}$$

Once it has converged then we have the volatility for the time horizon $[0, T]$.

Volatility Term Structure

Definition 5.2. ***Volatility Term Structure*** *is the relationship between the implied volatility and the time to maturity.*

If we have a term structure of market prices, i.e. $M(S_0, r, s, K)\ \forall 0 < s < T_N$ then we can extract a function for implied volatility with expiry $s : \sigma_I(s)$. This may involve sampling a fine set of knot points for s, backing out implied volatility for those points and then fitting a function for $\sigma_I(s)$. But once it is available then a continuous bootsrapping algorithm can be used to extract the spot volatility $\sigma(s)$. By definition:

$$\sigma_I^2(t)t = \int_0^t \sigma^2(s)ds$$

Differentiate with respect to the variable t and we get:

$$\sigma(t) = \sqrt{\sigma_I^2(t) + 2t\sigma_I(t)\frac{d\sigma_I(t)}{dt}}$$

In practice, the term structure is not continuous but rather a discrete set of market prices for benchmark expiry dates $T_1, T_2, \cdots, T_N$. From these prices we extract the corresponding set of implied volatility $\sigma_I(T_1), \sigma_I(T_2), \cdots, \sigma_I(T_N)$. After that apply a discrete **bootstrapping algorithm** then produces a **piece-wise constant function** for $\sigma(t)$. The formulas below demonstrate the bootstrap algorithm for the interval (T_{k-1}, T_k). First generate the set of implied volatility for all $T_1, T_2, \cdots, T_N$:

$$\begin{aligned}\sigma_I^2(T_k)T_k &= \int_0^{T_{k-1}} \sigma^2(s)ds + \int_{T_{k-1}}^{T_k} \sigma^2(s)ds \\ &= \sigma_I^2(T_{k-1})T_{k-1} + \int_{T_{k-1}}^{T_k} \sigma^2(s)ds\end{aligned}$$

Then recursively extract piece-wise constant functions for the spot volatility for each subsequent interval (T_{k-1}, T_k):

$$\sigma(t) = \sqrt{\left(\frac{\sigma_I^2(T_k)T_k - \sigma_I^2(T_{k-1})T_{k-1}}{T_k - T_{k-1}}\right)} \quad t \in (T_{k-1}, T_k)$$

5.2 Volatility Smile

The pure Black Scholes model does not consider the interplay of volatility with strike; in the case of pricing vanilla call or put options the model did not discriminate against different strikes. The strike was pretty much a silent parameter of the framework, relevant only in so far that it defined the level of money-ness of derivative. It had no role in modelling.

So it is no surprise that, for a given expiry, the implied volatility of vanilla option prices for different strike computed from a Black Scholes model, plot a flat profile. Saying this another way: **in a pure Black Scholes world where options prices are determined by the Black Scholes formula, the implied volatility is the same for all strikes.**

The real world however, is not pure Black Scholes. If you extract call and put option prices at different strikes, but all for the same expiry, then the associated implied volatility is far from flat; **wing options** (that is options with strike away from the spot underlying price) have different implied volatility compared to options with strike close to the spot price. Furthermore, the effect is magnified the further out into the wings you go.

Table 5.1: Volatility Smile

Strike	Implied Volatility
$K << S_0$	$\sigma_{\text{OTM}} > \sigma_{\text{ATM}}$
$K \approx S_0$	σ_{ATM}
$K >> S_0$	$\sigma_{\text{ITM}} > \sigma_{\text{ATM}}$ [a]

[a] σ_{OTM} is the implied volatility for out-of-the-money options, σ_{ATM} is the implied volatility for at-the-money options, and σ_{ITM} is the implied volatility for in-the-money options.

In a **market crash scenario**, **fear and chaos** will ensue. Panicked investors will be more sensitive to stock price movements, selling and buying more aggressively than they otherwise would, which means the volatility of the stock prices will increase.

To hedge themselves against this scenario, many investors would buy low strike put options to help offset their loses on the stock positions. This pushes up the price of low strike put options, and by association the implied volatility of these options.

On the contrary, say, in the case of a **market rally**, investors are optimistic and happy to hold their stocks.

Use simple Python function to simulate the volatility smile:

```python
def volatility_smile(strike_prices, at_the_money_vol, skew, kurtosis):
    return at_the_money_vol * (1 + skew * (strike_prices - 100) / 100 + kurtosis *
    ↪ ((strike_prices - 100) / 100) ** 2)
```

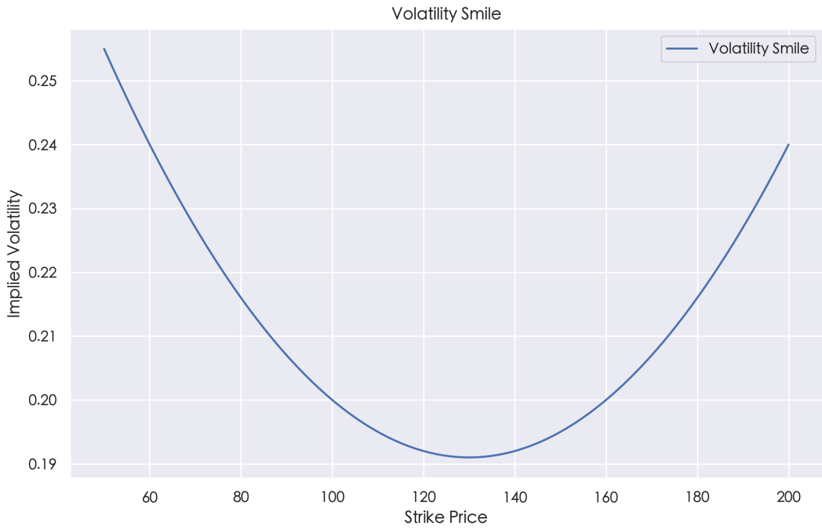

Figure 5.1: Volatility Smile

5.3 Market Implied Density

Lemma 5.1. ***Leibniz Integration Rule:*** *Let $f(x,y)$ be a continuous function, then*

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,y)dy = \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}dy + f(x,b(x))\frac{db}{dx} - f(x,a(x))\frac{da}{dx}$$

Note 15. *Simpler,*

$$\frac{\partial}{\partial x}\left(\int_{-\infty}^{x} f(x,y,t)dt\right) = f(x,y,x) + \int_{-\infty}^{x} \frac{\partial f(x,y,x)}{\partial x}dt$$

Recall that

$$\begin{aligned} C(t,S_t,K,T) &= \mathbb{E}\left[e^{-r(T-t)}\left(S_T-K\right)^+ \middle| S_t\right] \\ &= e^{-r(T-t)}\int_K^{\infty}\left(S_T-K\right)\phi\left(S_T\right)dS_T \\ &= e^{-r(T-t)}\int_K^{\infty}(s-K)\phi(s)ds \end{aligned}$$

Given a continuum of market prices $C(K,T)$ in the market at time t(i.e., fixing t), we back out the **market implied conditional density** $\phi(s)$ by differentiating the price with respect to the strike:

$$\begin{aligned} \frac{\partial C}{\partial K} &= \frac{\partial}{\partial K}\left(e^{-r(T-t)}\int_K^{\infty}(s-K)\phi(s)ds\right) \\ &= e^{-r(T-t)}\left(0\cdot\phi(K) - \int_K^{\infty}\phi(s)ds\right) \\ &= -e^{-r(T-t)}\left(\int_K^{\infty}\phi(s)ds\right) \end{aligned}$$

Applying the Leibniz rule again we get

$$\begin{aligned} \frac{\partial^2 C}{\partial K^2} &= e^{-r(T-t)}\phi(K) \\ \phi\left(S_T\right) &= e^{r(T-t)}\left(\frac{\partial^2 C}{\partial K^2}\right)\Bigg|_{K=S_T} \end{aligned}$$

A useful identity that follow from the derivation above is:

$$\int_K^\infty s\phi(s)ds = e^{r(T-t)}\left(C\left(S_0, K, T\right) - K\frac{\partial C}{\partial K}\right)$$

This comes from expanding the risk neutral expression for the price of a call option, and is explicitly used by Dupire.

We've derived the probability density function of the underlying asset from market information; it derived the market implied pdf of the asset.

Note 16. *This is the pdf that the market as a whole measures for the underlying asset. But the method didn't say anything about the dynamics that the underlying asset might follow. There're 2 approaches for deriving it, which can be made consistent with the market implied pdf:*

1. *Local Volatility Model.*
2. *Stochastic Volatility Model.*

5.4 Local Volatility Model: Dupire's Equation

Let the underlying asset have the following dynamics:

$$dS_t = rS_t dt + \sigma(S_t, t) S_t dW_t$$

where $\sigma(S_t, t)$ is the local volatility function. The local volatility model is a model where the volatility of the underlying asset is a deterministic function of the spot price and time.

Lemma 5.2. ***Fokker-Planck Equation**: Let $f(x, t \mid s, y)$ be a probability density function given $X_s = y$ of a stochastic process X_t that satisfies the SDE:*

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t$$

Then the probability density function $f(x, t)$ satisfies the Fokker-Planck equation:

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}(\mu f) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2 f\right)$$

Proof. Consider a smooth function $h(t, x)$ with compact support on $(s, T) \times \mathbb{R}$, particularly, $h(T, x) = h(s, x) = 0$. Then by Itô's Lemma:

$$h(T, X_T) = h(s, X_s) + \int_s^T \frac{\partial h}{\partial t} + \mu\frac{\partial h}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 h}{\partial x^2}dt + \int_s^T \sigma(t, X_t)\frac{\partial h}{\partial x}dW_t$$

Taking expectation $\mathbb{E}[\cdot \mid X_s = y]$:

$$\iint_{(s,T)\times\mathbb{R}} \frac{\partial h}{\partial t}f(t, x \mid s, y) + \mu\frac{\partial h}{\partial x}f(t, x \mid s, y) + \frac{1}{2}\sigma^2\frac{\partial^2 h}{\partial x^2}f(t, x \mid s, y)dxdt = 0$$

Taking integrals by parts yields:

$$\iint_{(s,T)\times\mathbb{R}} h(t, x)\left(-\frac{\partial f}{\partial t} - \frac{\partial(\mu f)}{\partial x} + \frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2 x)\right)dxdt = 0$$

Due to the compact support of $h(t, x)$, we have the Fokker-Planck equation:

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}(\mu f) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2 f\right)$$

□

By Fokker-Planck, the density $\phi(S_t, t)$ of stock price process $dS_t = rS_t dt + \sigma(S_t, t)S_t dW_t$ satisfies:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial}{\partial S}(rS\phi) + \frac{1}{2}\frac{\partial^2}{\partial S^2}\left(\sigma^2 S^2 \phi\right)$$

Now we can compute $\dfrac{\partial C}{\partial T}$:

$$\begin{aligned}
\frac{\partial C}{\partial T} &= \frac{\partial}{\partial T}\left(e^{-r(T-t)}\int_K^\infty (s-K)\phi(s)ds\right)\\
&= -rC + e^{-r(T-t)}\int_K^\infty (s-K)\frac{\partial \phi}{\partial T}ds\\
&\overset{\text{Fokker-Planck}}{=\!=\!=\!=} -rC - e^{-r(T-t)}r\int_K^\infty (s-K)\frac{\partial}{\partial s}\left(s\phi(s)\right)ds + \frac{1}{2}e^{-r(T-t)}\int_K^\infty (s-K)\frac{\partial^2}{\partial s^2}\left(s^2\sigma^2(s,T)\phi(s)\right)ds\\
&:= -rC + e^{-r(T-t)}\left(-rI_1 + \frac{1}{2}I_2\right)
\end{aligned}$$

The first integral is solved using integration by parts:

$$\begin{aligned}
I_1 &= \int_K^\infty (s-K)\frac{\partial}{\partial s}(s\phi)ds\\
&= s(s-K)\phi(s)\big|_{s=K}^\infty - \int_K^\infty s\phi(s)ds\\
&= -e^{r(T-t)}\left(C - K\frac{\partial C}{\partial K}\right)
\end{aligned}$$

The second integral is solved using integration by parts twice:

$$\begin{aligned}
I_2 &= \int_K^\infty (s-K)\frac{\partial^2}{\partial s^2}\left(s^2\sigma(t,s)^2\phi\right)ds\\
&= (s-K)\frac{\partial}{\partial s}\left(s^2\sigma(t,s)^2\phi\right)\Big|_{s=K}^\infty - \int_K^\infty \frac{\partial}{\partial s}\left(s^2\sigma(t,s)^2\phi\right)ds\\
&= 0 - s^2\sigma(t,s)^2\phi(s)\big|_{s=K}^\infty\\
&= K^2\sigma(t,K)^2\phi(K)\\
&= e^{r(T-t)}K^2\sigma(t,K)^2\left(\frac{\partial^2 C}{\partial K^2}\right)
\end{aligned}$$

Putting it all together:

$$\begin{aligned}
\frac{\partial C}{\partial T} + rC &= e^{-r(T-t)}\left[-rI_1 + \frac{1}{2}I_2\right]\\
&= r\left(C - K\frac{\partial C}{\partial K}\right) + \frac{1}{2}K^2\sigma(t,K)^2\left(\frac{\partial^2 C}{\partial K^2}\right)
\end{aligned}$$

Finally we arrive at Dupire's equation:

Theorem 5.1. ***Dupire's Equation***: *The local volatility function $\sigma(S,T)$ satisfies the following formula:*

$$\sigma^2(t,K) = \frac{\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}$$

Explicitly this means:

$$\sigma(t,s)^2 = \frac{\frac{\partial C}{\partial T}\big|_{T=t,K=s} + rK\frac{\partial C}{\partial K}\big|_{T=t,K=s}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}\big|_{T=t,K=s}}$$

5.5 Stochastic Volatility Model

Black Scholes assumed that the volatility is deterministic; Dupire went a step further and assigned it a function of time and the underlying asset (thereby making is random, but it doesn't study the dynamics). Stochastic volatility models explicitly **assign a stochastic differential equation to the volatility that drives the underlying asset**.

Set $B(\sigma) = \mathrm{BS}(T-t, S_t, K, \sigma)$ is the pure-BS price with usual constant lognormal volatility. Now introduce a random variable ν which satisfies $\mathbb{E}[\nu] = \sigma$. Since function B is convex in σ, we have by **Jensen's inequality**:

$$V := \mathbb{E}[B(\nu)] \geq B(\mathbb{E}[\nu]) = B(\sigma)$$

V is the price under a stochastic volatility model and we see above that it is larger than the BS price. Coresspondingly, the implied volatility of V is larger than the implied volatility of pure BS price. This gives us a model to create a **volatility smile**.

The simplest model is to propose a random variable α with some distribution $\mathcal{D}\left(\mu, \nu^2\right)$. μ will serve as a calibration parameter and ν will control the shape of the smile. The model works by calibrating μ to the **at-the-money vanilla call/put option price** $A\left(T-t, S_t\right)$. Explicitly, we compute it according to:

$$\alpha \sim \mathcal{D}(\mu, \nu^2), \quad \text{s.t.} \int_{\Omega_\alpha} \mathrm{BS}\left(T-t, S_t, K=S_t, \alpha\right) \phi(\alpha) d\alpha = A\left(T-t, S_t\right)$$

$A\left(T-t, S_t\right)$ can correspond to either a call or a put since they are related by Put-Call parity and share the same implied volatility. Now given μ we proceed to price non at-the-money options according to:

$$V\left(T-t, S_t, K\right) = \int_{\Omega_\alpha} \mathrm{BS}\left(T-t, S_t, K, \alpha\right) \phi(\alpha) d\alpha$$

Stochastic Volatility Pricing Framework

The pricing framework follows along similar lines as the standard Black Scholes; by deriving arbitrage free prices that drop out of a model which uses hedging strategies to replicate the derivative payoff.

There is one problem with stochastic volatility though; volatility itself is not a trade-able asset. We can not buy or sell volatility itself in the same way that we can buy and sell the underlying asset. We can however buy and sell some other derivative that already depends on volatility.

Generally, the model is:

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma_t S_t dW_t \\ d\sigma_t &= \mu_\sigma\left(t, \sigma_t\right) dt + b\left(t, \sigma_t\right) dZ_t \\ \langle dW_t, dZ_t \rangle &= \rho dt \end{aligned}$$

The value of a derivative at time t is now a function $V_t = V(t, S_t, \sigma_t)$, and the replication strategy uses the portfolio:

$$\Pi_t = V_t + \alpha_t S_t + \kappa_t V_t^1 + \beta_t B_t$$

The strategy to derive a general price for such a universal derivative is to **use an existing derivative which is already priced using stochastic volatility**.

The replication condition:

$$\beta_t B_t = -(V_t + \alpha_t S_t + \kappa_t V_t^1)$$

The self-financing condition:

$$d\Pi_t = dV_t + \alpha_t dS_t + \kappa_t dV_t^1 + \beta_t dB_t$$

Note that the SDE for V_t is now 2-dimensional with three contributing Itô terms:

$$dV_t = \frac{\partial V_t}{\partial t} + \frac{\partial V_t}{\partial S_t} dS_t + \frac{\partial V_t}{\partial \sigma_t} d\sigma_t + \frac{1}{2}\frac{\partial^2 V_t}{\partial S_t^2}\left(dS_t\right)^2 + \frac{1}{2}\frac{\partial^2 V_t}{\partial \sigma_t^2}\left(d\sigma_t\right)^2 + \frac{\partial^2 V_t}{\partial S_t \partial \sigma_t}\left(dS_t\right)\left(d\sigma_t\right)$$

But V_t^1 also follows its own SDE:

$$dV_t^1 = \frac{\partial V_t^1}{\partial t} + \frac{\partial V_t^1}{\partial S_t} dS_t + \frac{\partial V_t^1}{\partial \sigma_t} d\sigma_t + \frac{1}{2}\frac{\partial^2 V_t^1}{\partial S_t^2}\left(dS_t\right)^2 + \frac{1}{2}\frac{\partial^2 V_t^1}{\partial \sigma_t^2}\left(d\sigma_t\right)^2 + \frac{\partial^2 V_t^1}{\partial S_t \partial \sigma_t}\left(dS_t\right)\left(d\sigma_t\right)$$

To eliminate randomness in the portfolio Π, we need to ensure 2 factor $\Delta-$ hedging:

$$\begin{aligned} \alpha_t &= -\kappa_t \frac{\partial V_t^1}{\partial S_t} - \frac{\partial V_t}{\partial S_t} \\ \frac{\partial V_t}{\partial \sigma_t} &+ \kappa_t \frac{\partial V_t^1}{\partial \sigma_t} = 0 \end{aligned}$$

Putting altogether, we have:

$$\begin{aligned} d\Pi_t = &\left(\frac{\partial V_t}{\partial t} + \frac{1}{2}\frac{\partial^2 V_t}{\partial S_t^2}\sigma_t^2 S_t^2 + \frac{1}{2}\frac{\partial^2 V_t}{\partial \sigma_t^2} b^2(t, \sigma_t) + \frac{\partial^2 V_t}{\partial S_t \partial \sigma_t}\sigma_t S_t b(t, S_t)\rho\right) dt \\ &+\kappa_t\left(\frac{\partial V_t^1}{\partial t} + \frac{1}{2}\frac{\partial^2 V_t^1}{\partial S_t^2}\sigma_t^2 S_t^2 + \frac{1}{2}\frac{\partial^2 V_t^1}{\partial \sigma_t^2} b^2(t, \sigma_t) + \frac{\partial^2 V_t^1}{\partial S_t \partial \sigma_t}\sigma_t S_t b(t, S_t)\rho\right) dt \\ &- r(V_t + \alpha_t S_t + \kappa_t V_t^1 + \beta_t B_t) dt \end{aligned}$$

Define

$$\mathcal{L} := \frac{\partial}{\partial t} + \frac{1}{2}\sigma_t^2 S_t^2 \frac{\partial^2}{\partial S_t^2} + \frac{1}{2} b^2(t,\sigma_t)\frac{\partial^2}{\partial \sigma_t^2} + \rho\sigma_t S_t b(t,\sigma_t)\frac{\partial^2}{\partial S_t \partial \sigma_t}$$

Then let $d\Pi_t = 0$ we get

$$\mathcal{L}(V_t) + \kappa_t \mathcal{L}(V_t^1) - r\left(V_t - \kappa_t \frac{\partial V_t^1}{\partial S_t} S_t - \frac{\partial V_t}{\partial S_t} S_t - \left(\frac{\partial V_t}{\partial \sigma_t} / \frac{\partial V_t^1}{\partial \sigma_t}\right) V_t^1\right) = 0$$

Thus,

$$\frac{\mathcal{L}(V_t) + rS_t \frac{\partial V_t}{\partial \delta_t} - rV_t}{\frac{\partial V_t}{\partial \sigma_t}} = \frac{\mathcal{L}(V_t^1) + rS_t \frac{\partial V_t^1}{\partial S_t} - rV_t^1}{\frac{\partial V_t^1}{\partial \sigma_t}}$$

The Left-hand side is an expression for V_t, simultaneously the Right-hand side is an expression for V_t^1. The only way that this can hold if both sides equal some function, say $-a(t, S_t, \sigma_t)$ of the independent variables. This gives us the following system of PDE:

$$\frac{\partial V_t}{\partial t} + rS_t\frac{\partial V_t}{\partial S_t} + a(t,S_t,\sigma_t)\frac{\partial V_t}{\partial \sigma_t} + \frac{1}{2}\sigma_t^2 S_t^2\frac{\partial^2 V_t}{\partial S_t^2} + \frac{1}{2}b^2(t,\sigma_t)\frac{\partial^2 V_t}{\partial \sigma_t^2} + \rho\sigma_t S_t b(t,\sigma_t)\frac{\partial^2 V_t}{\partial S_t\partial\sigma_t} - rV_t = 0$$

Exhilaratingly, we can now use the Feynman-Kac theorem under risk-neutral dynamics

$$\begin{aligned} dS_t &= rS_t dt + \sigma_t S_t dW_t \\ d\sigma_t &= a(t,S_t,\sigma_t)\,dt + b(t,\sigma_t)\,dZ_t \\ \langle dW_t, dZ_t\rangle &= \rho dt \end{aligned}$$

Let $f(t, S_t, \sigma_t) = e^{-r(t-t_0)} V(t, S_t, \sigma_t)$, by simple calculus we have

$$\begin{gathered} \frac{\partial V}{\partial t} - rV = e^{r(t-t_0)}\frac{\partial f}{\partial t}, \quad \frac{\partial V}{\partial S_t} = e^{r(t-t_0)}\frac{\partial f}{\partial S_t}, \quad \frac{\partial V}{\partial \sigma_t} = e^{r(t-t_0)}\frac{\partial f}{\partial \sigma_t} \\ \frac{\partial^2 V}{\partial S_t^2} = e^{r(t-t_0)}\frac{\partial^2 f}{\partial S_t^2}, \quad \frac{\partial^2 V}{\partial \sigma_t^2} = e^{r(t-t_0)}\frac{\partial^2 f}{\partial \sigma_t^2}, \quad \frac{\partial^2 V}{\partial S_t\partial\sigma_t} = e^{r(t-t_0)}\frac{\partial^2 f}{\partial S_t\partial\sigma_t} \end{gathered}$$

Plug into the PDE above, we get

$$\frac{\partial f}{\partial t} + rS_t\frac{\partial f}{\partial S_t} + a(t,S_t,\sigma_t)\frac{\partial f}{\partial \sigma_t} + \frac{1}{2}\sigma_t^2 S_t^2\frac{\partial^2 f}{\partial S_t^2} + \frac{1}{2}b^2(t,\sigma_t)\frac{\partial^2 f}{\partial \sigma_t^2} + \rho\sigma_t S_t b(t,\sigma_t)\frac{\partial^2 f}{\partial S_t\partial\sigma_t} = 0$$

On the other hand, by multi-dimensional Itô's Lemma, we have

$$\begin{aligned} df &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S_t}dS_t + \frac{\partial f}{\partial \sigma_t}d\sigma_t + \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}(dS_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial \sigma_t^2}(d\sigma_t)^2 + \frac{\partial^2 f}{\partial S_t\partial\sigma_t}(dS_t)(d\sigma_t) \\ &\overset{\text{Plug the dynamics}}{=\!=\!=} \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S_t}(rS_t dt + \sigma_t S_t dW_t) + \frac{\partial f}{\partial \sigma_t}(a(t,S_t,\sigma_t)dt + b(t,\sigma_t)dZ_t) \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}\sigma_t^2 S_t^2 dt + \frac{1}{2}\frac{\partial^2 f}{\partial \sigma_t^2}b^2(t,\sigma_t)dt + \rho\sigma_t S_t b(t,\sigma_t)\frac{\partial^2 f}{\partial S_t\partial\sigma_t}dt \\ &= \left(\frac{\partial f}{\partial t} + rS_t\frac{\partial f}{\partial S_t} + a\frac{\partial f}{\partial \sigma_t} + \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}\sigma_t^2 S_t^2 + \frac{1}{2}\frac{\partial^2 f}{\partial \sigma_t^2}b^2(t,\sigma_t) + \rho\sigma_t S_t b(t,\sigma_t)\frac{\partial^2 f}{\partial S_t\partial\sigma_t}\right)dt + \frac{\partial f}{\partial S_t}\sigma_t S_t dW_t + \frac{\partial f}{\partial \sigma_t}b(t,\sigma_t)dZ_t \end{aligned}$$

The drift term in the above equation is exactly the same as the PDE, thus it'll be cancelled out. Hence

$$f(T, S_T, \sigma_T) = f(t_0, S_{t_0}, \sigma_{t_0}) + \int_{t_0}^{T}\frac{\partial f}{\partial S_t}\sigma_t S_t dW_t + \int_{t_0}^{T}\frac{\partial f}{\partial \sigma_t}b(t,\sigma_t)dZ_t$$

Take conditional expectation on both sides, we have

$$\mathbb{E}^{\mathbb{Q}}[f(T, S_T, \sigma_T) \mid \mathcal{F}_{t_0}] = f(t_0, S_{t_0}, \sigma_{t_0}) = V(t_0, S_{t_0}, \sigma_{t_0})$$

This leaves us with:

$$V(t, S_t, \sigma_t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}V(T, S_T) \mid S_t, \sigma_t\right]$$

Now we have made use of the fact that the price of the derivative **doesn't depend on volatility** and so

$$V(T, S_T) = V(T, S_T, \sigma_T)$$

A Monte Carlo algorithm will need to be used because the terminal value S_T depends on the **path** that σ_t takes over the interval (t_0, T). And so the problem is one of pricing a path-dependent derivative, like an Asian option.

Note 17. *There is however one problem that has not yet been addressed. What about the function a* (t, S_t, σ_t) *? We know that this function must exist mathematically but we do not have an expression for it, and we don't know what it is. The reason why end up in this situation is because we started off with a* ***chicken and egg problem****. We were only able to derive the pricing formula for* V_t *by assuming the existence of a derivative* V_t^1*; and we assumed that we could already trade this derivative to set up the delta hedging strategy. But what is the value of* V_t^1 *? Where does it come from? Who priced it?*

The answer is that the market priced it. We assume that the market prices V_t^1 *and we assume that these prices are sensible and arbitrage free (this is one of the prime assumptions of financial maths - that the markets are efficient and arbitrage free). From this, our pricing formula is a framework to price a new derivative* V_t *which eliminates arbitrage but only relative to the price of* V_t^1*. In other words, the framework prevents the introduction of any new arbitrage.*

Recall that the same pricing formula can be applied to V_t^1 *and* V_t*. So we calibrate the function a* (t, S_t, σ_t) *to the market prices of* V_t^1*; once the function is calibrated then we proceed to use it to price a new derivative* V_t*. This calibration is also called model fitting because we are fitting the stochastic volatility model to the existing prices of derivatives that are observed in the market, before using the same model to then extrapolate to new derivatives. It is a scenario that we encounter once again in a later chapter on Interest Rate derivatives.*

SABR and Heston Models

The **SABR model** is defined with the following stochastic differential equations for the underlying asset:

$$
\begin{aligned}
dF_t &= \sigma_t F_t^{\beta} dW_t \quad F_0 \text{ given} \\
d\sigma_t &= \nu \sigma_t dZ_t \quad \sigma_0 \text{ given} \\
\langle dW_t, dZ_t \rangle &= \rho dt
\end{aligned}
$$

The dynamics are written on the forward price $F_t = F(t, T)$; which converges to the spot asset $S(T)$ at $t = T$. So an option on $S(T)$ is the same as an option on $F(t, T)$ which expires at time T (i.e. the maturity of the forward must align with the expiry of the option).

Notice that the dynamics of F_t are some blend of Gaussian (when $\beta = 0$) and lognormal (when $\beta = 1$). The spot volatility of F_t is itself a random quantity which follows its own log-normal distribution. ν is the volatility of volatility and ρ controls the correlation between the two processes of asset and volatility.

The **Heston model** is defined with the following stochastic differential equations for the underlying asset:

$$
\begin{aligned}
dS_t &= \mu S_t dt + \sqrt{\sigma_t} S_t dW_t \quad S_0 \text{ given} \\
d\sigma_t &= \kappa (\theta - \sigma_t) dt + \eta \sqrt{\sigma_t} dZ_t \quad \sigma_0 \text{ given} \\
\langle dW_t, dZ_t \rangle &= \rho dt
\end{aligned}
$$

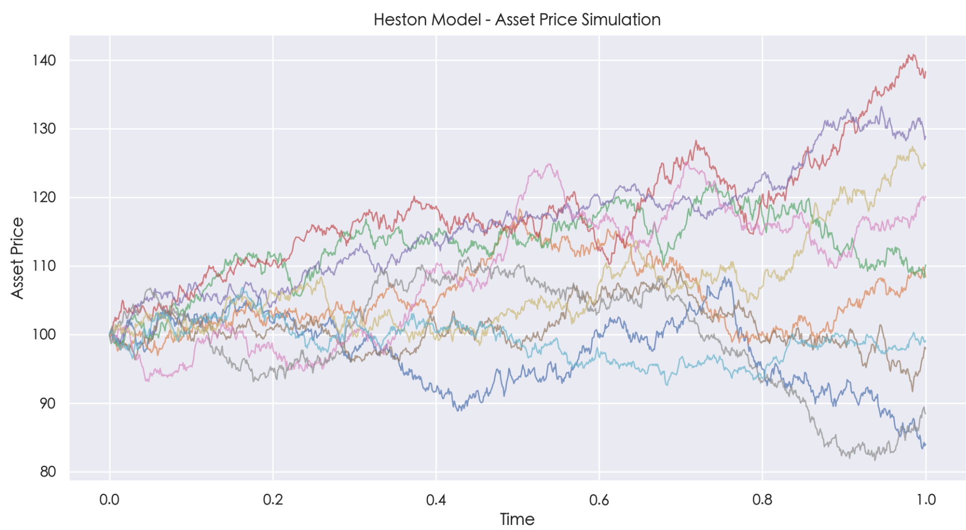

Figure 5.2: Heston Model

6 Exotic Options

Exotic options are specialized derivatives that differ from standard (”vanilla”) options by incorporating complex features such as non-traditional payoff structures, path-dependent pricing, or customized terms tailored to specific investor needs. Unlike vanilla options, which have standardized terms (e.g., fixed strike prices and expiration dates), exotic options often depend on multiple triggers, such as average prices over time (Asian options), predefined price barriers (barrier options), or even the performance of multiple assets (basket options). These instruments are primarily traded over-the-counter (OTC), allowing for greater flexibility in design but also introducing challenges like liquidity risk and valuation complexity.

The term ”exotic option” emerged in the late 20th century, popularized by financial economist Mark Rubinstein in his 1990 working paper *Exotic Options*. The naming is thought to draw inspiration from either ”exotic wagers” in horse racing (high-risk, high-reward bets) or geographic references like the ”Asian option” which was developed in Tokyo in 1987 by Bankers Trust executives Mark Standish and David Spaughton to price crude oil derivatives based on average prices. Rubinstein's work also highlighted the term's potential connection to Western-centric financial terminology, reflecting historical biases in academic discourse.

The 1980s and 1990s saw rapid innovation in exotic derivatives, driven by financial engineering and demand for tailored risk-management tools. For example, barrier options gained traction for their ability to limit costs by activating or deactivating based on market thresholds. Despite early skepticism—such as Federal Reserve Chairman Paul Volcker's 1980 critique of ”exotic” financial instruments—these options became integral to hedging and speculative strategies in global markets. Today, they remain a cornerstone of structured finance, offering solutions for nuanced risks like currency fluctuations or commodity price volatility.

This part includes:

1. Reflection Principle and Barrier Options.
2. Asian Options.
3. Foreign Exchange Options.

6.1 Barrier Options

Definition 6.1. *A **Barrier Option** is a type of option whose payoff depends on whether the underlying asset's price crosses (or not crosses) a certain level during a certain period of time. It consists of:*

- *A payoff $\varphi(S_T)$ at expiry T.*
- *A barrier level B.*

The barrier can be either **knock-in** or **knock-out**:

- A **knock-in** option is activated only if the underlying asset's price hits the pre-specified barrier level during the option's life.
- A **knock-out** option is deactivated if the underlying asset's price hits the pre-specified barrier level during the option's life.

And there're 2 types of barrier, which are lower barrier $B < S_0$ and upper barrier$B > S_0$.

There're 4 types of barrier options:

Example 15. *Today Sam buys an at-the-money call option with strike* 100\$ *and expiry at $T = 1$ year. It is up-and-in with barrier at* 120\$. *If the price never hits* 120\$, *the option is worthless, if the price touches the barrier between today and maturity, the option behaves as a regular call.*

Table 6.1: Types of Barrier Options

Type	Start	End	Payoff(Call)
Up-and-Out	Below B	Knock-out if $S_t > B$	$(S_T - K)\mathbb{1}_F$, $F = \{B > S_T \geq K, \max\limits_{0\leq t\leq T} S_t < B\}$
Up-and-In	Below B	Knock-in if $S_t \geq B$	$(S_T - K)\mathbb{1}_G$, $G = \{S_T \geq K, \max\limits_{0\leq t\leq T} S_t \geq B\}$
Down-and-Out	Above B	Knock-out if $S_t < B$	$(S_T - K)\mathbb{1}_F$, $F = \{S_T \geq K, \min\limits_{0\leq t\leq T} S_t > B\}$
Down-and-In	Above B	Knock-in if $S_t \leq B$	$(S_T - K)\mathbb{1}_G$, $G = \{S_T \geq K, \min\limits_{0\leq t\leq T} S_t \leq B\}$

Denote

$$c_{do}(S_0, B, K, T), \quad c_{di}(S_0, B, K, T), \quad c_{uo}(S_0, B, K, T), \quad c_{ui}(S_0, B, K, T)$$

as the price of down-and-out, down-and-in, up-and-out, up-and-in barrier options respectively. The price of a vanilla call option is $c(S_0, K, T)$. Consider a portfolio longing one down-and-out call as well as a down-and-in call, the payoff at maturity is:

$$\begin{cases} 0 + (S_T - K)^+, & \text{if Barrier gets hit} \\ (S_T - K)^+ + 0, & \text{if Barrier never gets hit} \end{cases}$$

Hence anyway we can get the payoff of a vanilla call option. Thus we have the following theorem in general:

Theorem 6.1. ***In-Out Parity***: *For any barrier option, we have knock-out option + knock-in option = vanilla option. Explicitly:*

$$c_{do}(S_0, B, K, T) + c_{di}(S_0, B, K, T) = c(S_0, K, T), \quad c_{uo}(S_0, B, K, T) + c_{ui}(S_0, B, K, T) = c(S_0, K, T)$$

Proof. It is obvious:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}\varphi(S_T)\mathbb{1}_F\right] + \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}\varphi(S_T)\mathbb{1}_G\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}\varphi(S_T)\right]$$

□

So we can just study the price of knock-in options and then use the parity to get the price of knock-out options.

6.1.1 Down-and-In Call Options, Reflection Principle

Pre-Pricing Down-and-In Call Options

By risk-neutral valuation,

$$c_{di}(S_0, B, K, T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}(S_T - K)\mathbb{1}_G\right] = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[S_T \cdot \mathbb{1}_G\right] - e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[K \cdot \mathbb{1}_G\right] := I_1 - I_2$$

where $G = \{S_T \geq K, \min\limits_{0\leq t\leq T} S_t \leq B\}$.

Recap the change of numeraire formula, we have

$$I_1 = B_0\mathbb{E}^{\mathbb{Q}}\left[\frac{S_T}{B_T} \cdot \mathbb{1}_G\right] = S_0\mathbb{E}^{\mathbb{Q}^S}\left(\mathbb{1}_G\right) = S_0\mathbb{Q}^S(G)$$

where $\mathbb{Q}$ is the risk-neutral measure under the numeraire B_T and $\mathbb{Q}^S$ is the risk-neutral measure under the numeraire S_T. Also,

$$I_2 = Ke^{-rT}\mathbb{E}^{\mathbb{Q}}\left(\mathbb{1}_G\right) = Ke^{-rT}\mathbb{Q}(G)$$

Recall, under $\mathbb{Q}^S$, the stock price follows:

$$d(\log S_t) = \left(r + \frac{1}{2}\sigma^2\right)dt + \sigma dW_t^S$$
$$d(\log S_t) = \left(r - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t$$

where W_t^S, W_t is a Brownian motion under $\mathbb{Q}^S$ and $\mathbb{Q}$ respectively.

In terms of log-prices, the area G becomes[1]:

$$\begin{aligned} G &= \left\{\log S_T \geq \log K, \min_{0\leq t\leq T} \log S_t \leq \log B\right\} \\ &= \left\{\log\left(\frac{S_T}{S_0}\right) \geq \log\left(\frac{K}{S_0}\right), \min_{0\leq t\leq T} \log\left(\frac{S_t}{S_0}\right) \leq \log\left(\frac{B}{S_0}\right)\right\} \\ &:= \{X_T \geq x, m_T \leq m\} \end{aligned}$$

where we define $X_t = \log\left(\frac{S_t}{S_0}\right), x = \log\left(\frac{K}{S_0}\right), m_T = \min_{0\leq t\leq T} X_t, m = \log\left(\frac{B}{S_0}\right)$.Note that we're considering a lower barrier so $B < S_0$ and hence $m < 0$, and

$$X_t = \begin{cases} \left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t, & \text{under } \mathbb{Q} \\ \left(r + \frac{1}{2}\sigma^2\right) t + \sigma W_t^S, & \text{under } \mathbb{Q}^S \end{cases} := \eta t + \sigma W_t^*$$

Our goal is compute $\mathbb{Q}^*(X_T \geq x, m_T \leq m)$, where $\mathbb{Q}^*$ is the measure under which W_t^* is a Brownian motion. To achieve this we first study the case $\eta = 0, \sigma = 1$: then we call

$$m_T = \min_{0\leq t\leq T} W_t$$

the minimum of the Wiener process. Later, we extend it to general η, σ by using Girsanov's theorem.

Case I: $\eta = 0, \sigma = 1$: Reflection Principle

Definition 6.2. *Define a random variable*

$$\tau = \inf\{t \geq 0 : W_t = m\}$$

*as the **first hitting time** of the Wiener process to the level m.*

Define

$$m_t = \min_{0\leq s\leq t} W_s, \quad M_t = \max_{0\leq s\leq t} W_s$$

*as the **running minimum** and **running maximum** of the Wiener process up to time t.*

Proposition 6.1. *According to the definition, we have*

$$\mathbb{P}(\tau \leq t) = \begin{cases} \mathbb{P}(m_t \leq m) & \text{if} \quad m < 0 \\ \mathbb{P}(M_t \geq m) & \text{if} \quad m > 0 \end{cases}$$

We have the following theorem:

Theorem 6.2. ***Reflection Principle**: The path after time τ has the same distribution as the path before time τ reflected at the level $m = W_\tau$. Explicitly, for any $t \geq 0$:*

$$\widetilde{W}_t = \begin{cases} W_t, & \text{if } t \leq \tau \\ 2m - W_t, & \text{if } t > \tau \end{cases}$$

is still a Brownian motion. And we have

$$\mathbb{P}(\widetilde{W}_T \geq x \mid \tau \leq T) = \mathbb{P}(W_T \leq 2m - x \mid \tau \leq T)$$

Proof. Let

$$Y_t = W_t \mathbb{1}_{t\leq\tau}, \quad Z_t = W_{t+\tau} - m$$

Due to Markov Property, Z_t is a Brownian motion independent of Y_t. Thus $-Z_t$ is also a Brownian motion. Hence (Y, Z) and $(Y, -Z)$ are two Brownian motions with same distribution. Define

$$\delta : (Y, Z) \to Y_t \mathbb{1}_{t\leq\tau} + (b + Z_{t-\tau})\mathbb{1}_{t>\tau}$$

Then $\delta(Y, Z) = W_t, \delta(Y, -Z) = \widetilde{W}_t$. Because $\delta(Y, Z)$ has same distribution with $\delta(Y, -Z)$, $\widetilde{W}_t$ is a Brownian motion. □

[1] $\min_{0\leq t\leq T} S_t \leq B \leftrightharpoons \exists 0 \leq t^* \leq T \quad \text{s.t.} S_t^* \leq B \leftrightharpoons \exists 0 \leq t^* \leq T \quad \text{s.t.} \log(S_t^*) \leq \log(B) \leftrightharpoons \min_{0\leq t\leq T} \log(S_t) \leq \log(B)$

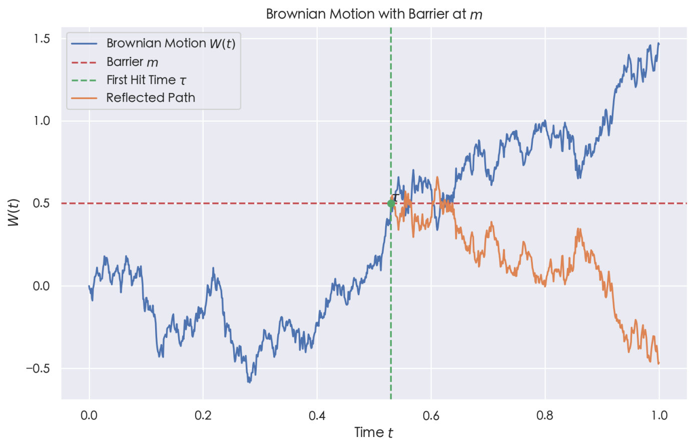

Figure 6.1: Reflection Principle

Theorem 6.3. *For every $m \geq 0$, we have*

$$\mathbb{P}(M_t \geq m) = 2\mathbb{P}(W_t \geq m)$$

For every $m \leq 0$, we have

$$\mathbb{P}(m_t \leq m) = 2\mathbb{P}(W_t \leq m)$$

Proof. Only prove the second one.

$$\mathbb{P}(m_t \leq m) = \mathbb{P}(m_t \leq m, W_t \leq m) + \mathbb{P}(m_t \leq m, W_t \geq m)$$

Consider

$$\begin{aligned}
\mathbb{P}(m_t \leq m, W_t > m) &= \mathbb{P}(\tau \leq t, W_t > m) \\
&= \mathbb{P}(W_t \geq m | \tau \leq t)\mathbb{P}(\tau \leq t) \\
&\overset{\text{Reflection Principle}}{=\!=\!=\!=} \mathbb{P}(\widetilde{W}_t \geq m | \tau \leq t)\mathbb{P}(\tau \leq t) \\
&= \mathbb{P}(2m - W_t \geq m | \tau \leq t)\mathbb{P}(\tau \leq t) \\
&= \mathbb{P}(W_t \leq m | \tau \leq t)\mathbb{P}(\tau \leq t) \\
&= \mathbb{P}(W_t \leq m, \tau \leq t)
\end{aligned}$$

Note that $\{W_t \leq m\} \subset \{\tau \leq t\}$ Hence

$$\mathbb{P}(m_t \leq m) = 2\mathbb{P}(W_t \leq m)$$

The first one can be proved similarly. □

Define

$$A := \{W_T \geq x, m_T \leq m\}, \quad B := \{W_T \leq 2m - x\}$$

assume $m \leq x$, and note that if $W_T \leq 2m - x \leq m$ then $\tau \leq T$, hence

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(W_T \geq x, \tau \leq T) \\
&\overset{\text{Reflection Principle}}{=\!=\!=\!=} \mathbb{P}(W_T \leq 2m - x, \tau \leq T) \\
&= \mathbb{P}(B) = \Phi\left(\frac{2m - x}{\sqrt{T}}\right)
\end{aligned}$$

It shows that if W_t hits the barrier m before time T, then W_T and $2m - W_T$ have the same law, more generally we have

Corollary 6.1. *Assume A, B are defined above, and $m \leq x$, then*

$$\mathbb{E}[\mathbb{1}_A g(W_T)] = \mathbb{E}[\mathbb{1}_B g(2m - W_T)]$$

If we define

$$A' := \{W_T \leq x, M_T \geq M\}, \quad B' := \{W_T \geq 2M - x\}$$

then similarly

$$\mathbb{P}(A') = \mathbb{P}(B') = \Phi\left(\frac{2M-x}{\sqrt{T}}\right)$$

and

$$\mathbb{E}[\mathbb{1}_A g(W_T)] = \mathbb{E}[\mathbb{1}_B g(2M - W_T)]$$

Case II: General η, σ

We've already study the case $\eta = 0, \sigma = 1$, now we consider the general case:

$$dX_t = \eta dt + \sigma dW_t, \quad m_t = \min_{0 \le s \le t} X_s$$

Write

$$X_T = \eta T + \sigma W_T = \sigma\left(\frac{\eta}{\sigma}T + W_T\right) := \sigma \widehat{W}_T$$

We need to find $\mathbb{P}(X_T \ge x, m_T \le m)$, the procedure is

1. Find a measure $\widehat{\mathbb{P}}$ such that $\widehat{W}_t$ is a Brownian motion.
2. Compute $\widehat{\mathbb{P}}(X_T \ge x, m_T \le m)$, applying what we gained from the case $\eta = 0, \sigma = 1$.
3. Change back to measure $\mathbb{P}$.

Define $dL_t = -\frac{\eta}{\sigma}L_t dW_t$, and the change of measure

$$\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}} = L_T = \exp\left(-\frac{\eta}{\sigma}W_T - \frac{\eta^2}{2\sigma^2}T\right)$$

by Girsonov's theorem, $\widehat{W}_t = W_t + \frac{\eta}{\sigma}$ is a Brownian motion under $\widehat{\mathbb{P}}$. Thus

$$\begin{aligned}\widehat{\mathbb{P}}(X_T \ge x, m_T \le m) &= \widehat{\mathbb{P}}\left(\widehat{W}_T \ge \frac{x}{\sigma}, \frac{m_T}{\sigma} \le \frac{m}{\sigma}\right) \\ &\overset{\text{Using Results from Case I}}{=\!=\!=} \widehat{\mathbb{P}}\left(\widehat{W}_T \le \frac{2m}{\sigma} - \frac{x}{\sigma}\right) = \Phi\left(\frac{2m-x}{\sigma\sqrt{T}}\right)\end{aligned}$$

Define set

$$\widehat{A} := \left\{\widehat{W}_T \ge \frac{x}{\sigma}, \frac{m_T}{\sigma} \le \frac{m}{\sigma}\right\}, \quad \widehat{B} := \left\{\widehat{W}_T \le \frac{2m}{\sigma} - \frac{x}{\sigma}\right\}$$

Hence

$$\begin{aligned}\mathbb{P}(X_T \ge x, m_T \le m) &= \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{\{X_T \ge x, m_T \le m\}}] = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{\widehat{A}}] = \mathbb{E}^{\widehat{\mathbb{P}}}\left(\mathbb{1}_{\widehat{A}}\frac{1}{L_T}\right) \\ &= \mathbb{E}^{\widehat{\mathbb{P}}}\left[\mathbb{1}_{\widehat{A}}\exp\left(-\frac{\eta^2}{2\sigma^2}T + \frac{\eta}{\sigma}\widehat{W}_T\right)\right] \\ &\overset{\text{Corollary}}{=\!=\!=} \mathbb{E}^{\widehat{\mathbb{P}}}\left[\mathbb{1}_{\widehat{B}}\exp\left(-\frac{\eta^2}{2\sigma^2}T + \frac{\eta}{\sigma}\left(\frac{2m}{\sigma} - \widehat{W}_T\right)\right)\right] \\ &= \frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{\frac{2m-x}{\sigma}} \exp\left[-\frac{\eta^2}{2\sigma^2}T + \frac{\eta}{\sigma}\left(\frac{2m}{\sigma} - z\right) - \frac{z^2}{2T}\right]dz \\ &= \frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{\frac{2m-x}{\sigma}} \exp\left[\frac{2\eta m}{\sigma^2} - \frac{1}{2T}\left(\frac{\eta}{\sigma}T + z\right)^2\right]dz \\ &= \exp\left(\frac{2\eta m}{\sigma^2}\right)\Phi\left(\frac{2m-x+\eta T}{\sigma\sqrt{T}}\right)\end{aligned}$$

Proposition 6.2. *For a stochastic process*

$$X_t = \eta t + \sigma W_t$$

define $m_T = \min_{0\le t\le T} X_t$, *then if* $m \le x, m < 0$,

$$\mathbb{P}(X_T \ge x, m_T \le m) = \exp\left(\frac{2\eta m}{\sigma^2}\right)\Phi\left(\frac{2m - x + \eta T}{\sigma\sqrt{T}}\right)$$

Pricing Down-and-In Call Options

First recall what we've gained in Pre-pricing section:

$$\begin{aligned} c_{di} &= S_0\mathbb{Q}^S(G) - Ke^{-rT}\mathbb{Q}(G) \\ G &= \{X_T \ge x, m_T \le m\} \\ X_T &= \log\left(\frac{S_T}{S_0}\right), x = \log\left(\frac{K}{S_0}\right), \\ m_T &= \min_{0\le t\le T} X_t, m = \log\left(\frac{B}{S_0}\right); \\ \eta &= \begin{cases} r - \frac{1}{2}\sigma^2, & \text{under } \mathbb{Q} \\ r + \frac{1}{2}\sigma^2, & \text{under } \mathbb{Q}^S \end{cases} \end{aligned}$$

Therefore,

$$\mathbb{Q}^S(G) = \exp\left(\frac{2\left(r + \frac{\sigma^2}{2}\right)\log\frac{B}{S_0}}{\sigma^2}\right)\Phi\left(\frac{\overbrace{\log\frac{B^2}{S_0K}}^{2m-x} + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)$$

and

$$\mathbb{Q}(G) = \exp\left(\frac{2\left(r - \frac{\sigma^2}{2}\right)\log\frac{B}{S_0}}{\sigma^2}\right)\Phi\left(\frac{\log\frac{B^2}{S_0K} + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)$$

Define

$$d_{\text{barrier}} = \frac{\log\frac{B^2}{S_0K} + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}, \quad \alpha = \frac{2r}{\sigma^2} + 1$$

Then we get

Theorem 6.4. *Price of Down-and-In Call Option:*

$$c_{di}(S_0, B, K, T) = S_0\left(\frac{B}{S_0}\right)^{\alpha}\Phi(d_{barrier}) - Ke^{-rT}\left(\frac{B}{S_0}\right)^{\alpha-2}\Phi\left(d_{barrier} - \sigma\sqrt{T}\right)$$

Note 18. *If we denote*

$$\text{BS}(S_0, K, T, \sigma, r) = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$

as the standard Black-Scholes formula, then the price of down-and-in call option can be written as

$$c_{di}(S_0, B, K, T) = \left(\frac{B}{S_0}\right)^{\alpha-2}\text{BS}\left(\frac{B^2}{S_0}, K, T, \sigma, r\right)$$

and hence the price of a down-and-out call option is

$$c_{do}(S_0, B, K, T) = \text{BS}(S_0, K, T, \sigma, r) - \left(\frac{B}{S_0}\right)^{\alpha-2}\text{BS}\left(\frac{B^2}{S_0}, K, T, \sigma, r\right)$$

Use the following Python code to calculate the price of down-and-in call option:

```python
from scipy.stats import norm
def standard_BSM_call(S, K, T, r, sigma):
    d1 = (np.log(S / K) + (r + 0.5 * sigma ** 2) * T) / (sigma * np.sqrt(T))
    d2 = d1 - sigma * np.sqrt(T)
    price = S * norm.cdf(d1) - K * np.exp(-r * T) * norm.cdf(d2)
```

```
    return price

def down_and_in_call(S, K, B, T, r, sigma):
    if S < B:
        return 0
    alpha = (2*r) / sigma ** 2+1
    return (B/S)**(alpha-2)* standard_BSM_call(B**2/S,K,T,r,sigma)

def down_and_out_call(S, K, B, T, r, sigma):
    if S<=B:
        return 0
    return standard_BSM_call(S, K, T, r, sigma) - down_and_in_call(S, K, B, T, r, sigma)

K,B,T,r,sigma = 100,80,1.0,0.05,0.2
S_range = np.linspace(50, 150, 100)
call_prices_in = [down_and_in_call(S, K, B, T, r, sigma) for S in S_range]
call_prices_out = [down_and_out_call(S, K, B, T, r, sigma) for S in S_range]
vanilla_call_prices = [standard_BSM_call(S, K, T, r, sigma) for S in S_range]
```

Figure 6.2: Down-and-In v.s. Down-and-Out Call Option

6.1.2 Up-and-Out Call Options

$$c_{uo} = S_0 \mathbb{Q}^S(F) - Ke^{-rT}\mathbb{Q}(F)$$

where $F = \{M > M_T > X_T > x\}$.

$$X_t = \log\left(\frac{S_t}{S_0}\right), x = \log\left(\frac{K}{S_0}\right), M_T = \max_{0\le t\le T} X_t, M = \log\left(\frac{B}{S_0}\right)$$

and assume $M > 0$, recall $X_t = \eta t + \sigma W_t$, where $\eta = r - \frac{1}{2}\sigma^2$ under $\mathbb{Q}$ and $\eta = r + \frac{1}{2}\sigma^2$ under $\mathbb{Q}^S$.

Define measure change as the same before, i.e.

$$\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}} = L_T = \exp\left(-\frac{\eta}{\sigma}W_T - \frac{\eta^2}{2\sigma^2}T\right)$$

then $\widehat{W}_t = W_t + \frac{\eta}{\sigma} = \frac{X_t}{\sigma}$ is a Brownian motion under $\widehat{\mathbb{P}}$. Hence under $\widehat{\mathbb{P}}$,

$$\widehat{\mathbb{P}}(F) = \widehat{\mathbb{P}}\left(\widehat{M} > \widehat{M}_T > \widehat{W}_T > \frac{x}{\sigma}\right), \quad \text{where} \frac{M_T}{\sigma} = \max_{0\le t\le T} \widehat{W}_t := \widehat{M}_T$$

to study this, we have a lemma

Lemma 6.1. *For a standard Brownian motion W_t, define $M_t = \max_{0\le s\le t} W_s \ge 0$, then (W_t, M_t) has joint density*

$$f_{W_t,M_t}(w,m) = \frac{2(2m-w)}{t\sqrt{2\pi t}} \exp\left(-\frac{(2m-w)^2}{2t}\right) \mathbb{1}_{\{w\le m, m\ge 0\}}$$

Proof. Recall

$$\mathbb{P}(W_t \le x, M_t \ge M) = \mathbb{P}(W_t \ge 2M - x) = \Phi\left(\frac{2M-x}{\sqrt{t}}\right)$$

hence,

$$\int_m^\infty \int_{-\infty}^w f_{W_T,M_T}(x,y)dxdy = \frac{1}{\sqrt{2\pi t}} \int_{2m-w}^\infty \exp\left(-\frac{z^2}{2t}\right)dz$$

calculate the derivative of the left hand side with respect to m first, gaining

$$-\int_{-\infty}^w f_{W_T,M_T}(x,m)dx = -\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(2m-w)^2}{2t}\right)$$

then calculate the derivative with respect to w, gaining

$$f_{W_T,M_T}(w,m) = \frac{2(2m-w)}{t\sqrt{2\pi t}} \exp\left(-\frac{(2m-w)^2}{2t}\right)$$

□

So under $\widehat{\mathbb{P}}$, $\left(\widehat{W}_T, \widehat{M}_T\right)$ has the joint density above. Now we change back to measure $\mathbb{P}$, and we have, under $\mathbb{P}$,

$$\begin{aligned}
\mathbb{P}\left(\widehat{W}_T \le w, \widehat{M}_T \le m\right) &= \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{\{\widehat{W}_T\le w, \widehat{M}_T\le m\}}] \\
&= \mathbb{E}^{\widehat{\mathbb{P}}}\left[\frac{1}{L_T}\mathbb{1}_{\{\widehat{W}_T\le w, \widehat{M}_T\le m\}}\right] \\
&= \int_{-\infty}^w \int_{-\infty}^m f^{\widehat{\mathbb{P}}}_{\widehat{W}_T,\widehat{M}_T}(x,y) \exp\left(\frac{\eta}{\sigma}w - \frac{\eta^2}{2\sigma^2}T\right)dxdy
\end{aligned}$$

Hence

$$f^{\mathbb{P}}_{\widehat{W}_T,\widehat{M}_T}(w,m) = \frac{\partial^2}{\partial m \partial w}\mathbb{P}\left(\widehat{W}_T \le w, \widehat{M}_T \le m\right) = \frac{2(2m-w)}{T\sqrt{2\pi T}} \exp\left(\frac{\eta}{\sigma}w - \frac{\eta^2}{2\sigma^2}T - \frac{(2m-w)^2}{2T}\right)$$

Now we have

$$\begin{aligned}
\mathbb{Q}^S(F) = \mathbb{Q}^S\left(\widehat{M} > \widehat{M}_T > \widehat{W}_T > \frac{x}{\sigma}\right) &= \int_{\frac{x}{\sigma}}^{\frac{M}{\sigma}} \int_w^{\frac{M}{\sigma}} \frac{2(2m-w)}{T\sqrt{2\pi T}} \exp\left(\frac{\eta}{\sigma}w - \frac{\eta^2}{2\sigma^2}T - \frac{(2m-w)^2}{2T}\right)dmdw \\
= -\int_{\frac{x}{\sigma}}^{\frac{M}{\sigma}} \frac{1}{\sqrt{2\pi T}} \exp\left(\frac{\eta}{\sigma}w - \frac{\eta^2}{2\sigma^2}T - \frac{(2m-w)^2}{2T}\right)\Bigg|_{m=w}^{m=\frac{M}{\sigma}} dw \\
= \frac{1}{\sqrt{2\pi T}}\int_{\frac{x}{\sigma}}^{\frac{M}{\sigma}} \exp\left(\frac{\eta}{\sigma}w - \frac{\eta^2}{2\sigma^2}T - \frac{w^2}{2T}\right)dw - \frac{1}{\sqrt{2\pi T}}\int_{\frac{x}{\sigma}}^{\frac{M}{\sigma}} \exp\left(\frac{\eta}{\sigma}w - \frac{\eta^2}{2\sigma^2}T - \frac{\left(2\frac{M}{\sigma}-w\right)^2}{2T}\right)dw \\
:= I_1 - I_2
\end{aligned}$$

Compute I_1 first, we have

$$\begin{aligned}
I_1 &= \frac{1}{\sqrt{2\pi T}}\int_{\frac{x}{\sigma}}^{\frac{M}{\sigma}} \exp\left(-\frac{1}{2T}\left(W - \frac{\eta}{\sigma}T\right)^2\right)dw \\
&= \frac{1}{\sqrt{2\pi}}\int_{\frac{x-\eta T}{\sigma\sqrt{T}}}^{\frac{M-\eta T}{\sigma\sqrt{T}}} \exp\left(-\frac{1}{2}z^2\right)dz \\
&= \Phi\left(\frac{M-\eta T}{\sigma\sqrt{T}}\right) - \Phi\left(\frac{x-\eta T}{\sigma\sqrt{T}}\right) = \Phi\left(\frac{\log\left(\frac{B}{S_0}\right)-\eta T}{\sigma\sqrt{T}}\right) - \Phi\left(\frac{\log\left(\frac{K}{S_0}\right)-\eta T}{\sigma\sqrt{T}}\right) \\
&= \Phi\left(\frac{\log\left(\frac{S_0}{K}\right)+\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right) - \Phi\left(\frac{\log\left(\frac{S_0}{B}\right)+\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)
\end{aligned}$$

and

$$
\begin{aligned}
I_2 &= \frac{1}{\sqrt{2\pi T}} \int_{\frac{x}{\sigma}}^{\frac{M}{\sigma}} \exp\left(\frac{\eta}{\sigma}w - \frac{\eta^2}{2\sigma^2}T - \frac{\left(2\frac{M}{\sigma} - w\right)^2}{2T}\right) dw \\
&= \exp\left(\frac{2M\eta}{\sigma^2}\right) \frac{1}{\sqrt{2\pi T}} \int_{\frac{x}{\sigma}}^{\frac{M}{\sigma}} \exp\left(-\frac{1}{2T}\left(w - \frac{\eta}{\sigma}T - 2\frac{M}{\sigma}\right)^2\right) dw \\
&= \exp\left(\frac{2M\eta}{\sigma^2}\right) \frac{1}{\sqrt{2\pi}} \int_{\frac{x-\eta T-2M}{\sigma\sqrt{T}}}^{\frac{-M-\eta T}{\sigma\sqrt{T}}} \exp\left(-\frac{1}{2}z^2\right) dz \\
&= \exp\left(\frac{2M\eta}{\sigma^2}\right)\left(\Phi\left(\frac{-M-\eta T}{\sigma\sqrt{T}}\right) - \Phi\left(\frac{x-\eta T-2M}{\sigma\sqrt{T}}\right)\right) \\
&= \exp\left(\frac{2M\eta}{\sigma^2}\right)\left(\Phi\left(\frac{\log\left(\frac{S_0}{B}\right) - \eta T}{\sigma\sqrt{T}}\right) - \Phi\left(\frac{\log\left(\frac{K}{S_0}\right) - \eta T - 2\log\left(\frac{B}{S_0}\right)}{\sigma\sqrt{T}}\right)\right) \\
&= \exp\left(\frac{2M\eta}{\sigma^2}\right)\left(\Phi\left(\frac{\log\left(\frac{B^2}{KS_0}\right) + \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right) - \Phi\left(\frac{\log\left(\frac{B}{S_0}\right) + \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)\right)
\end{aligned}
$$

The only difference between $\mathbb{Q}$ and $\mathbb{Q}^S$ is η, the rest is the same. Let

$$
\begin{aligned}
d_1^K &= \frac{\log\left(\frac{S_0}{K}\right) + \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}, \quad d_1^B = \frac{\log\left(\frac{S_0}{B}\right) + \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \\
d_{\text{barrier}}^K &= \frac{\log\left(\frac{B^2}{KS_0}\right) + \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}, \quad d_{\text{barrier}}^B = \frac{\log\left(\frac{B}{S_0}\right) + \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}
\end{aligned}
$$

Hence

$$
\begin{aligned}
c_{uo} &= S_0\mathbb{Q}^S(F) - Ke^{-rT}\mathbb{Q}(F) \\
&= S_0\Phi(d_1^K) - S_0\Phi(d_1^B) - S_0\left(\frac{B}{S_0}\right)^{\alpha}\Phi(d_{\text{barrier}}^K) + S_0\left(\frac{B}{S_0}\right)^{\alpha}\Phi(d_{\text{barrier}}^B) \\
&\quad -Ke^{-rT}\Phi\left(d_1^K - \sigma\sqrt{T}\right) + Ke^{-rT}\Phi\left(d_1^B - \sigma\sqrt{T}\right) + Ke^{-rT}\left(\frac{B}{S_0}\right)^{\alpha-2}\Phi\left(d_{\text{barrier}}^K - \sigma\sqrt{T}\right) - Ke^{-rT}\left(\frac{B}{S_0}\right)^{\alpha-2}\Phi\left(d_{\text{barrier}}^B - \sigma\sqrt{T}\right) \\
&= \mathrm{BS}(S_0,T,K) - \left(\frac{B}{S_0}\right)^{\alpha-2}\mathrm{BS}\left(\frac{B^2}{S_0},T,K\right) \\
&\quad - \left[\mathrm{BS}(S_0,T,B) - \left(\frac{B}{S_0}\right)^{\alpha-2}\mathrm{BS}\left(\frac{B^2}{S_0},T,B\right)\right]
\end{aligned}
$$

Theorem 6.5. *Price of Up-and-Out Call Option:*

$$
c_{uo}(S_0,B,K,T) = \mathrm{BS}(S_0,T,K) - \left(\frac{B}{S_0}\right)^{\alpha-2}\mathrm{BS}\left(\frac{B^2}{S_0},T,K\right) - \left[\mathrm{BS}(S_0,T,B) - \left(\frac{B}{S_0}\right)^{\alpha-2}\mathrm{BS}\left(\frac{B^2}{S_0},T,B\right)\right]
$$

6.1.3 PDE Framework

Consider a knock-out option. The option is alive and its price $V(t,S_t)$ at time t must satisfy the BS PDE:

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

Note that under barrier case, the boundary condition is $V(T,S_T) = (S_T-K)^+$ and the barrier condition is $V(t,B) = 0$. If it's down-out, then the solution domain is $[0,T]\times(B,\infty)$.

If it's a down-and-in option, the solution domain is $[0,T]\times[H,\infty)$, where H is the barrier. The boundary condition is $V(T,S_T) = (S_T-K)^+$.

Table 6.2: Conditions for Barrier Options

	Down-and-Out	**Down-and-In**	**Up-and-Out**	**Up-and-In**
Domain	$[0,T]\times(B,\infty)$	$[0,T]\times[H,\infty)$	$[0,T]\times(0,B)$	$[0,T]\times(0,H]$
Expiry Condition	$V(T,S_T)=(S_T-K)^+$			
Barrier Condition	$V(t,B)=0$	$V(t,H)=V_{\text{vanilla}}$	$V(t,B)=0$	$V(t,H)=V_{\text{vanilla}}$
Horizontal Condition	$\lim_{S\to\infty}V(t,S)=0$		$\lim_{S\to 0}V(t,S)=0$	

Lemma 6.2. *Let $V_{do}(t,S_t;B,\varphi)$ be the price at time t of a down-and-out option with final payoff $\varphi(S_T)$ and lower knock-out barrier B. Then*

$$V_{do}(t,S_t;B,\alpha\varphi+\beta\psi)=\alpha V_{do}(t,S_t;B,\varphi)+\beta V_{do}(t,S_t;B,\psi).$$

Proof.

$$\mathbb{E}^{\mathbb{Q}}[e^{-rT}(\alpha\varphi+\beta\psi)\mathbb{1}_F]=\alpha\mathbb{E}^{\mathbb{Q}}[e^{-rT}\varphi\mathbb{1}_F]+\beta\mathbb{E}^{\mathbb{Q}}[e^{-rT}\psi\mathbb{1}_F]$$

□

Theorem 6.6. ***Put-Call Parity for Barrier Options**:*

$$p_{do}(S,B,K,T)=K\cdot b_{do}(S,B,K,T)-s_{do}(S,B,K,T)+c_{do}(S,B,K,T),$$

where $p_{do}(S,B,K,T)$ is the price of a down-and-out put option with barrier B and strike K, b_{do} is the price of the down-and-out contract with payoff 1, and s_{do} is the price of the down-and-out contract with payoff S_T.

Proof. Note that

$$(K-S_T)^+=K-S_T+(S_T-K)^+$$

Hence use the lemma and the proof is completed. □

6.2 Asian Options

The name "Asian options" is somewhat misleading since these options were first used in 1987 by the Bankers Trust Tokyo office, hence the association with Asia. Their name stuck, and they became known as "Asian options" due to their initial association with the Asian financial markets.

The first Asian options were primarily used for pricing commodities, particularly oil. The aim was to provide better protection against price fluctuations, and they became popular with companies involved in commodity trading because of their ability to reduce the impact of price manipulation. For example, companies that wanted to hedge the cost of buying crude oil over several months preferred Asian options, as they offered a less volatile payoff due to averaging.

Today, people trade Asian options mainly for similar reasons. It provides with robust against sudden large shocks, useful if the stock is very volatile. It is also less vulnerable to price manipulation and cheaper than vanilla European options, which is why it is often used in the commodities market.

Definition 6.3. ***Asian Option** is an option whose payoff depends on the average price of the underlying asset over a certain period of time.*

There're different kinds of Asian Options:

1. Geometric Average v.s. Arithmetic Average

2. Continuous Average v.s. Discrete Average

3. Fixed Strike v.s. Floating Strike

$$A_T = \begin{cases} \frac{1}{n}\sum_{i=1}^{n} S_{t_i}, & \text{Arithmetic Average Discretely Sampled} \\ \left(\prod_{i=1}^{n} S_{t_i}\right)^{1/n}, & \text{Geometric Average Discretely Sampled} \\ \frac{1}{T}\int_0^T S_t dt, & \text{Arithmetic Average Continuously Sampled} \\ \exp\left(\frac{1}{T}\int_0^T \log S_t dt\right), & \text{Geometric Average Continuously Sampled} \end{cases}$$

The payoff funcItôn of Asian Call Option can be classified to 2 types:

1. Fixed Strike: $V_T = (A_T - K)^+$
2. Floating Strike: $V_T = (S_T - k \cdot A_T)^+$, where k is a constant approximately equal to 1.

Only continously sampled Asian options are considered in this note, and note that there's no closed-form solution for the price of Arithmetic Asian options.

6.2.1 Geometric Asian Options

Define the geometric average as

$$G_t = \exp\left(\frac{1}{t}\int_0^t \log S_u du\right) := e^{X_t}, \quad X_t = \frac{1}{t}\int_0^t \log S_u du$$

the payoff of an Asian Call option with fixed strike K is

$$V_T = (e^{X_t} - K)^+$$

hence our goal is to calculate

$$c_{ga}(S_t, K, T) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}\left(e^{X_T} - K\right)^+ |\mathcal{F}_t]$$

First calculate X_T:

$$\begin{aligned} X_T &= \frac{1}{T}\int_0^T \log S_u du \\ &= \underbrace{\frac{1}{T}\int_0^t \log S_u du}_{\text{Known at time } t} + \underbrace{\frac{1}{T}\int_t^T \log S_u du}_{\text{In the future: Random}} \\ &= \frac{t}{T}X_t + \frac{1}{T}\int_t^T \left(\log S_t + \left(r - \frac{1}{2}\sigma^2\right)(u-t) + \sigma(W_u - W_t)\right) du \\ &= \frac{t}{T}X_t + \frac{T-t}{T}\log S_t + \frac{1}{T}\left(r - \frac{1}{2}\sigma^2\right)\int_t^T (u-t)du + \frac{1}{T}\sigma\int_t^T (W_u - W_t)du \\ &= \frac{t}{T}X_t + \frac{T-t}{T}\log S_t + \frac{1}{T}\left(r - \frac{1}{2}\sigma^2\right)\frac{(T-t)^2}{2} + \frac{1}{T}\sigma\int_t^T (W_u - W_t)du \end{aligned}$$

So we need to study the distribution of $\int_t^T (W_u - W_t)du$. We have the following lemma:

Lemma 6.3. *For a standard Brownian motion W_t, the integral*

$$\int_t^T (W_u - W_t)du$$

*is **normally distributed** with mean 0 and variance $\frac{(T-t)^3}{3}$.*

Proof. Consider the partition $t = t_0 < t_1 < \cdots < t_n = T$, then

$$\begin{aligned}
\int_t^T (W_u - W_t)du &\approx \sum_{i=0}^{n-1} (W_{t_i} - W_t)(t_{i+1} - t_i) \\
&= \sum_{i=0}^{n-1} W_{t_i}(t_{i+1} - t_i) - W_t(T - t) \\
&= \sum_{i=1}^{n-1} W_{t_{i-1}} t_i + W_{t_{n-1}} t_n - \sum_{i=1}^{n-1} W_{t_i} t_i - W_{t_0} t_0 - W_t T + W_t t \\
&= \sum_{i=1}^{n-1} \left(W_{t_{i-1}} - W_{t_i}\right) t_i + T(W_{t_{n-1}} - W_t) \\
&= -\sum_{i=1}^{n-1} \Delta W_i t_i + T \sum_{i=1}^{n-1} \Delta W_i \\
&= \sum_{i=1}^{n-1} \Delta W_i (T - t_i), \quad \Delta W_i = W_{t_i} - W_{t_{i-1}}
\end{aligned}$$

Hence

$$\int_t^T (W_u - W_t)du = \lim_{n\to\infty} \sum_{i=1}^{n-1} \Delta W_i (T - t_i)$$

the left hand side is a limit of a sum of independent normal random variables, hence it's normally distributed. The mean is 0 and the variance is

$$\begin{aligned}
\mathbb{E}\left[\left(\int_t^T (W_u - W_t)du\right)^2\right] &\xlongequal{\text{Monotone Convergence}} \lim_{n\to\infty} \mathbb{E}\left[\left(\sum_{i=1}^{n-1} \Delta W_i (T - t_i)\right)^2\right] \\
&= \lim_{n\to\infty} \sum_{i=1}^{n-1} \mathbb{E}\left[(\Delta W_i (T - t_i))^2\right] = \lim_{n\to\infty} \sum_{i=1}^{n-1} (T - t_i)^2 \mathbb{E}\left[(\Delta W_i)^2\right] \\
&= \lim_{n\to\infty} \sum_{i=1}^{n-1} (T - t_i)^2 (t_i - t_{i-1}) \\
&= \int_t^T (T - u)^2 du = \frac{(T - t)^3}{3}
\end{aligned}$$

□

Now define

$$\overline{\mu} = \frac{1}{T}\left(r - \frac{1}{2}\sigma^2\right)\frac{(T - t)^2}{2}, \quad \overline{\sigma} = \frac{\sigma}{T}\sqrt{\frac{(T - t)^3}{3}}$$

then

$$X_T = \frac{t}{T} X_t + \frac{T - t}{T} \log S_t + \overline{\mu} + \overline{\sigma} Z$$

where $Z \sim \mathcal{N}(0, 1)$, hence

$$G_T = e^{X_T} = e^{\frac{t}{T} X_t} \cdot e^{\frac{T-t}{T} \log S_t} \cdot e^{\overline{\mu} + \overline{\sigma} Z} = G_t^{\frac{t}{T}} \cdot S_t^{\frac{T-t}{T}} \cdot e^{\overline{\mu} + \overline{\sigma} Z}$$

and the price of the Asian Call option is

$$c_{ga}(S_t, K, T) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}(G_T - K)^+ | \mathcal{F}_t] = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}\left[\left(G_t^{t/T} S_t^{(T-t)/T} \exp(\overline{\mu} + \overline{\sigma} Z) - K\right)^+ \middle| \mathcal{F}_t\right]$$

Theorem 6.7. ***Price of Geometric Asian Call Option:***

$$c_{ga}(S_t, K, T) = e^{-r(T-t)}\left[G_t^{t/T} S_t^{(T-t)/T} \exp\left(\overline{\mu} + \frac{1}{2}\overline{\sigma}^2\right) \Phi(d_1) - K\Phi(d_2)\right]$$

where

$$d_2 = \frac{1}{\overline{\sigma}}\left(\frac{t}{T}\log G_t + \frac{T-t}{T}\log S_t - \log K + \overline{\mu}\right), \quad d_1 = d_2 + \overline{\sigma}$$

$$\overline{\mu} = \frac{1}{T}\left(r - \frac{1}{2}\sigma^2\right)\frac{(T-t)^2}{2}, \quad \overline{\sigma} = \frac{\sigma}{T}\sqrt{\frac{(T-t)^3}{3}}$$

Proof. The proof is straightforward by simple calculus. Firstly, we have

$$G_t^{t/T} S_t^{(T-t)/T} e^{\overline{\mu}+\overline{\sigma}Z} - K > 0 \Leftrightarrow Z > \frac{1}{\overline{\sigma}}\left(\log K - \frac{t}{T}\log G_t - \frac{T-t}{T}\log S_t - \mu\right) = -d_2$$

Hence

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\left(G_t^{t/T} S_t^{(T-t)/T}\exp(\overline{\mu}+\overline{\sigma}Z) - K\right)^+ \middle| \mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{Q}}\left[\left(G_t^{t/T} S_t^{(T-t)/T}\exp(\overline{\mu}+\overline{\sigma}Z) - K\right)\mathbb{1}_{\{Z>-d_2\}} \middle| \mathcal{F}_t\right] \\
&= G_t^{t/T} S_t^{(T-t)/T}\mathbb{E}^{\mathbb{Q}}\left[\exp(\overline{\mu}+\overline{\sigma}Z)\mathbb{1}_{\{Z>-d_2\}} \middle| \mathcal{F}_t\right] - K\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{Z>-d_2\}} \middle| \mathcal{F}_t\right] \\
&= G_t^{t/T} S_t^{(T-t)/T}\frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty}\exp\left(\overline{\mu}+\overline{\sigma}z - \frac{z^2}{2}\right)dz - K\Phi(d_2) \\
&= G_t^{t/T} S_t^{(T-t)/T} e^{\overline{\mu}+\frac{1}{2}\overline{\sigma}^2}\int_{-d_2}^{\infty}\exp\left(-\frac{1}{2}(z-\overline{\sigma})^2\right)dz - K\Phi(d_2) \\
&= G_t^{t/T} S_t^{(T-t)/T} e^{\overline{\mu}+\frac{1}{2}\overline{\sigma}^2}\Phi(d_1) - K\Phi(d_2)
\end{aligned}$$

□

If at time 0, we can use the following Python code to calculate the price of Geometric Asian Call Option:

```python
def asian_call(S0,K,r,sigma,T):
    mu = T*(r-0.5*sigma**2)/2
    sigma_hat = sigma*np.sqrt(T)/np.sqrt(3)
    d2 = (np.log(S0/K)+mu)/sigma_hat
    d1 = d2 + sigma_hat
    return np.exp(-r*T)*S0*np.exp(mu+sigma**2/2)*norm.cdf(d1) -
    ↪ K*np.exp(-r*T)*norm.cdf(d2)
```

However, at other time t, we need to calculate G_t first, which contains an integral on geometric Brownian motion. Hence we need to use Monte Carlo simulation to calculate the price of Geometric Asian Call Option at time t.

```python
def simulate_geometric_brownian_motion(S0, r, sigma, T, N, M):
    """
    N: number of time steps
    M: number of paths for Monte Carlo simulation
    """
    dt = T / N
    paths = np.zeros((M, N + 1))
    paths[:, 0] = S0
    for t in range(1, N + 1):
        Z = np.random.standard_normal(M)
        paths[:, t] = paths[:, t - 1] * np.exp((r - 0.5 * sigma ** 2) * dt + sigma *
        ↪ np.sqrt(dt) * Z)
    return paths

def asian_call_simulate(St, r, sigma, tau, N, M, K):
    """
    tau: time to maturity T-t
    """
    paths = simulate_geometric_brownian_motion(St, r, sigma, T, N, M)
    geometric_average = np.exp(np.sum(np.log(paths[:, 1:]), axis=1) / N)
    asian_call_payoff = np.maximum(geometric_average - K, 0)
    asian_call_price = np.exp(-r * tau) * np.mean(asian_call_payoff)
    return asian_call_price
```

Obviously, asian options are cheaper than vanilla options.

Figure 6.3: Price of Geometric Asian Call Option

6.2.2 PDE Framework: For Path-Dependent Options

Arithmetic Asian options are path-dependent options whose payoff depends on the average price of the underlying asset over a certain period of time. But it doesn't have an explicit pricing formula. Our goal is to find the equivalent PDE for path-dependent options.

To generalize, define $V(t, S_t, I_t)$ as the price of a path-dependent option at time t with payoff $V_T = \varphi(S_T, I_T)$, where I_t is the path-dependent variable

$$I_t = \int_0^t f(S_u, u)du, \quad dI_t = f(S_t, t)dt$$

One can notice that $dI_t dI_t = 0, dS_t dI_t = 0$, apply multi-dimensional Itô's Lemma to $V(t, S_t, I_t)$, we have

$$dV_t = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \frac{\partial V}{\partial S} dS_t + \frac{\partial V}{\partial I} dI_t$$

Similar as before, form a portfolio

$$\Pi_t = V(t, S_t, I_t) + \kappa S_t + \beta B_t$$

then the replication condition is

$$\beta B_t = -V(t, S_t, I_t) - \kappa S_t$$

The portfolio has dynamics:

$$\begin{aligned} d\Pi_t &\xlongequal{\text{Self-Financing}} \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \frac{\partial V}{\partial S} dS_t + \frac{\partial V}{\partial I} dI_t + \kappa dS_t + \beta dB_t \\ &\xlongequal[\text{Delta-Hedging}]{\kappa = -\frac{\partial V}{\partial S}} \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \frac{\partial V}{\partial I} dI_t + \beta B_t r dt \\ &= \left(\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + f(S_t, t) \frac{\partial V}{\partial I} \right) dt - rV dt \end{aligned}$$

Let $d\Pi_t = 0$ yeilding the PDE:

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + f(S_t, t) \frac{\partial V}{\partial I} - rV = 0$$

6.2.3 Rogers-Shi Method

For the Arithmetic Asian, let $I_t = \int_0^t S_u du$, then $f(S_t, t) = S_t$,

$$A_t = \frac{1}{t} \int_0^t S_u du := \frac{1}{t} I_t$$

Note that

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}[(A_T-K)^+] &= \mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{1}{T}\int_0^T S_u du - K\right)^+\right]\\ &= \mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{1}{T}\int_t^T S_u du - \left(K - \frac{1}{T}\int_0^t S_u du\right)\right)^+\right]\\ &= S_t\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{1}{T}\int_t^T \frac{S_u}{S_t}du - X_t\right)^+\right]\end{aligned}$$

where $X_t = \frac{K-\frac{I_t}{T}}{S_t}$.

Note 19. *For $u > t$,*

$$\frac{S_u}{S_t} = \exp\left(\sigma(W_u - W_t) + \left(r - \frac{1}{2}\sigma^2\right)(u-t)\right)$$

is independent of the history of the asset price up to time t, hence

$$\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{1}{T}\int_t^T \frac{S_u}{S_t}du - X_t\right)^+\right]$$

is a function of X_t and t only.

Define $H(X_t,t) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{1}{T}\int_t^T \frac{S_u}{S_t}du - X_t\right)^+\right]$, then $V(t,S_t,I_t) = S_tH(X_t,t)$. Note that

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

and

$$X_t = \frac{K-\frac{I_t}{T}}{S_t}, \quad \frac{\partial X}{\partial S} = -\frac{X}{S}, \quad \frac{\partial X}{\partial I} = -\frac{1}{TS}$$

Hence

$$\frac{\partial V}{\partial S} = H + S\frac{\partial H}{\partial X}\frac{\partial X}{\partial S} = H - X\frac{\partial H}{\partial X}, \quad \frac{\partial V}{\partial I} = S\frac{\partial H}{\partial X}\frac{\partial X}{\partial I} = -\frac{1}{T}\frac{\partial H}{\partial X}$$

Substituting in the PDE for V gives

$$\frac{\partial H}{\partial t} - \frac{\partial H}{\partial X}\left(\frac{1}{T} + rX\right) + \frac{1}{2}\sigma^2X^2\frac{\partial^2 H}{\partial X^2} = 0, \quad X \in (-\infty,+\infty), t \in [0,T)$$

the terminal condition is

$$V(S_t,I_t,T) = \left(\frac{I_t}{T} - K\right)^+ = S(-X)^+ \Leftrightarrow H(X_T,T) = (-X_T)^+$$

For the case $X \le 0$, i.e., $A_t \ge K$, we can find an explicit solution. In this case,

$$\frac{1}{T}\int_t^T \frac{S_u}{S_t}du - X_t \ge 0$$

Hence,

$$\begin{aligned}H(X_t,t) &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{1}{T}\int_t^T \frac{S_u}{S_t}du - X_t\right)^+\right] = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\frac{1}{T}\int_t^T \frac{S_u}{S_t}du - X_t\right]\\ &= e^{-r(T-t)}\frac{1}{T}\int_t^T \mathbb{E}_t^{\mathbb{Q}}\left[\frac{S_u}{S_t}\right]du - e^{-r(T-t)}X_t\\ &= e^{-r(T-t)}\frac{1}{T}\int_t^T \frac{\mathbb{E}_t^{\mathbb{Q}}(S_u)}{S_t}du - e^{-r(T-t)}X_t\\ &= e^{-r(T-t)}\frac{1}{T}\int_t^T e^{-r(u-t)}du - e^{-r(T-t)}X_t\\ &= \frac{1-e^{-r(T-t)}}{rT} - e^{-r(T-t)}X_t\end{aligned}$$

Thus,

$$\begin{aligned} V(t,S_t,I_t) &= S_t H(X_t,t) = S_t\left(\frac{1-e^{-r(T-t)}}{rT} - e^{-r(T-t)}X_t\right) \\ &= S_t\frac{1-e^{-r(T-t)}}{rT} + e^{-r(T-t)}\left(\frac{I_t}{T}-K\right) \end{aligned}$$

If a portfoilo that is long one Arithmetic Asian call and short one Arithmetic Asian put, than the payoff

$$c(S_T,I_T,T) - p(S_T,I_T,T) = \left(\frac{I_T}{T}-K\right)^+ - \left(K-\frac{I_T}{T}\right)^+ = \frac{I_T}{T}-K$$

By risk-neutral valuation, the price of the portfolio at time t is

$$c(t,S_t,I_t) - p(t,S_t,I_t) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-r(T-t)}\left(\frac{I_T}{T}-K\right)\right]$$

the right-hand-side is actually what we have done before, hence we have the put-call parity for Arithmetic Asian options:

Theorem 6.8. *Put-Call Parity for Arithmetic Asian Options:*

$$c(t,S_t,I_t) - p(t,S_t,I_t) = S_t\frac{1-e^{-r(T-t)}}{rT} + e^{-r(T-t)}\left(\frac{I_t}{T}-K\right)$$

6.3 Lookback Options

Definition 6.4. ***Lookback options*** *are options whose payoffs depend on the maximum or the minimum of the underlying stock price over a certain period of time (the lookback period).*

Let the lookback period be $[0,T]$. Define the minimum value of the asset over the lookback period $[0,T]$ by

$$m_T := \min_{0\le t\le T} S_t$$

Define the maximum value of the asset over the lookback period $[0,T]$ by

$$M_T := \max_{0\le t\le T} S_t$$

A distinctive feature of lookback options is there's no regret on the time to exercise the option: as if the option was exercised at the optimal point. In other words, it reduces market timing risk, reduces the chance that the option will be worthless at expiry. However, they can be very expensive.

Obviously, lookback options are path-dependent options. Recall that The price $V(t,S,I)$ of a path-dependent option is

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial I}f(S,t) - rV = 0,$$

where

$$I_t = \int_0^t f\left(S_u,u\right)du$$

is the path-dependent variable. However, for lookback options, the path-dependent variable is

$$I_t = M_t = \max_{0\le u\le t} S_u$$

where no such function $f(S,t)$ satisfies $dI_t = f(S_t,t)dt$.

Theorem 6.9. *Let $f(x)$ be a continuous positive function on $[a,b]$. Then*

$$\max_{a\le x\le b} f(x) = \lim_{n\to\infty}\left(\int_a^b f^n(x)dx\right)^{1/n}.$$

Proof. First, note that $f(x)$ is a continuous positive function on $[a, b]$, and let:

$$m = \min_{a \le x \le b} f(x), \quad M = \max_{a \le x \le b} f(x).$$

Since $f(x)$ is continuous and positive on the closed interval $[a, b]$, $f(x)$ attains its maximum and minimum at some points within $[a, b]$. Therefore, we have $0 < m \le f(x) \le M$ for all $x \in [a, b]$.

Consider the expression:

$$I_n = \int_a^b f^n(x)dx$$

Since $f(x) \le M$ for all $x \in [a, b]$, we have:

$$f^n(x) \le M^n \quad \text{for all } x \in [a, b]$$

Thus:

$$(I_n)^{1/n} \le (M^n(b-a))^{1/n} = M((b-a))^{1/n}.$$

As $n \to \infty$, $((b-a))^{1/n} \to 1$, so:

$$\limsup_{n\to\infty} (I_n)^{1/n} \le M.$$

Now, for a lower bound, note that $f(x) \ge m > 0$ for all $x \in [a, b]$. Let x_0 be a point in $[a, b]$ where $f(x_0) = M$. Since $f(x)$ is continuous, for any $\epsilon > 0$, there exists a neighborhood U around x_0 such that $f(x) > M - \epsilon$ for all $x \in U$, where the measure of U is $|U| > 0$. Thus:

$$I_n = \int_a^b f^n(x)dx \ge \int_U f^n(x)dx > \int_U (M-\epsilon)^n dx = |U|(M-\epsilon)^n.$$

Taking the n-th root, we get:

$$(I_n)^{1/n} > (|U|(M-\epsilon)^n)^{1/n} = (M-\epsilon)|U|^{1/n}.$$

As $n \to \infty$, $|U|^{1/n} \to 1$, so:

$$\liminf_{n\to\infty} (I_n)^{1/n} \ge M - \epsilon$$

Since $\epsilon > 0$ is arbitrary, we have:

$$\liminf_{n\to\infty} (I_n)^{1/n} \ge M$$

Combining the upper and lower bounds, we have:

$$\lim_{n\to\infty} (I_n)^{1/n} = M$$

□

Corollary 6.2. *We have*

$$M_t = \max_{0 \le u \le t} S_u = \lim_{n\to\infty} \left(\int_0^t S_u^n du \right)^{1/n}.$$

$$m_t = \min_{0 \le u \le t} S_u = \lim_{n\to\infty} \left(\int_0^t S_u^{-n} du \right)^{-1/n}.$$

For any positive integer n, we define

$$I_{n,t} := \int_0^t S_u^n du,$$
$$M_{n,t} := I_{n,t}^{1/n}.$$

Then $M_t = \lim_{n\to\infty} M_{n,t}$.

Hence, the SDE for $I_{n,t}$ and $M_{n,t}$ are

$$\begin{aligned} dI_{n,t} &= S_t^n dt, \\ dM_{n,t} &= \frac{1}{n} I_{n,t}^{\frac{1-n}{n}} dI_{n,t} \\ &= \frac{1}{n} \frac{1}{M_{n,t}^{n-1}} dI_{n,t} \\ &= \frac{1}{n} \frac{S_t^n}{M_{n,t}^{n-1}} dt. \end{aligned}$$

Applying Itô's Lemma to $V(t, S_t, M_{n,t})$, we have

$$\begin{aligned} dV_t &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS_t + \frac{\partial V}{\partial M_{n,t}}dM_{n,t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS_t)^2 \\ &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{n}\frac{S_t^n}{M_{n,t}^{n-1}}\frac{\partial V}{\partial M_{n,t}}\right)dt + \frac{\partial V}{\partial S}dS_t. \end{aligned}$$

Form a portfolio

$$\Pi_t = V(t, S_t, M_{n,t}) + \kappa S_t + \beta B_t$$

After similar calculations, we have

$$\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{n}\frac{S_t^n}{M_{n,t}^{n-1}}\frac{\partial V}{\partial M_{n,t}} - rV = 0$$

If $M_t > S_t$, recall $M_t = \lim_{n\to\infty} M_{n,t}$, so for large n we have

$$M_{n,t}^{n-1} \geq S_t^{n-1}$$

Hence

$$\frac{1}{n}\frac{S_t^n}{M_{n,t}^{n-1}} \leq \frac{1}{n}S_t, \quad \lim_{n\to\infty}\frac{1}{n}\frac{S_t^n}{M_{n,t}^{n-1}} = 0$$

If $M_t = S_t$, then $\frac{\partial V}{\partial M_{n,t}} = 0$, hence finally, the European floating strike lookback put option satisfies

$$\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} - rV = 0$$

with the following boundary conditions:

$$\begin{aligned} V(T, S, M) &= M - S, \\ V(t, 0, M) &= e^{-r(T-t)}M, \\ \frac{\partial V}{\partial M}(t, M, M) &= 0. \end{aligned}$$

The solution domain is

$$\{(t, S, M) : 0 < S < M, 0 \leq t \leq T\}.$$

6.4 Foreign Exchange (FX) Options

Foreign Exchange (Forex or FX) options are financial derivatives that give the buyer the right, but not the obligation, to exchange a specific amount of one currency for another at a predetermined exchange rate (known as the strike price) on or before a specified expiration date. FX options are commonly used by investors, corporations, and financial institutions to hedge against currency risk or to speculate on currency movements.

Definition 6.5. *An* ***exchange rate*** *is the* ***price*** *at which one currency can be exchanged for another. Formally, it is defined as the value of one country's currency in terms of another currency. It indicates how much of one currency you need to pay to buy a unit of another currency.*

Mathematically, an exchange rate between two currencies, A and B, can be denoted as:

$$\text{Exchange Rate}_{A/B} = \frac{\text{Units of Currency B}}{\text{One Unit of Currency A}}$$

This rate is typically quoted in two ways:

1. **Direct Quotation**: The amount of domestic currency needed to buy one unit of foreign currency.
2. **Indirect Quotation**: The amount of foreign currency that can be obtained with one unit of the domestic currency.

Example 16. *If Y_t is the SGD-CNY exchange rate at time t, then Y_t can be thought as the CNY value of 1 unit of SGD. This is a direct quotation for Chinese investors and an indirect quotation for Singaporean investors. Recently, the SGD-CNY exchange rate is around 5.40, that is, 1 SGD can be exchanged for 5.40 CNY.*

Domestic Risk Neutral Measure

Domestic Risk Neutral Measure means the measure is defined on the domestic country. In this case, X_t is direct quotation for the domestic country, that means, X_t is **the domestic currency value of 1 unit of foreign currency**. Under domestic risk-neutral measure, X_t has dyanmics:

$$dX_t = rX_t dt + \sigma X_t dW_t$$

where r is the domestic interest rate and σ is the volatility of the exchange rate. However, like what dividend-paying stock behaves, we now have the interest rate of the foreign currency which is paid to the holder of the foreign currency. Hence the dynamics of the exchange rate under domestic risk-neutral measure is

$$dX_t = (r - r_f)X_t dt + \sigma X_t dW_t$$

where r_f is the foreign interest rate.

Hence the valuation of an FX option is then a natural reapplication of the risk-neutral pricing framework on the Black-Scholes model. But here we call it Garman-Kohlhagen model, specifically for FX options. The price of a European call option on an FX rate is

$$c(X_t, K, T) = X_t e^{-r_f(T-t)} N(d_1) - Ke^{-r(T-t)} N(d_2)$$

where

$$d_1 = \frac{\log(X_t/K) + (r - r_f + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

Foreign Risk Neutral Measure

$(X_T - K)^+$ is for domestic investors, say Singaporeans. For example, if the CNY-SGD rate grows up from 0.2 to 0.3 which means Singapore dollar fells and strike $K = 0.25$, then the payoff is 0.75 SGD.

To a CNY-based investor, this is equivalent to payoff

$$\frac{(X_T - K)^+}{X_T}$$

in Chinese Yuan. For a Chinese investor,

$$c_f(X_t, K, T) = e^{-r_f(T-t)} \mathbb{E}_t^{\mathbb{Q}^f}\left[\frac{(X_T - K)^+}{X_T}\right]$$

Note that, if we define $Y_t = \frac{1}{X_t}$, then Y_t is the SGD-CNY exchange rate, and the dynamics of Y_t must be

$$dY_t = (r_f - r)Y_t dt + \sigma_Y Y_t dW_t^f$$

in the CNY risk-neutral measure. Pricing the same option under the foreign risk-neutral measure is then:

$$c_f(X_t, K, T) = e^{-r_f T} \mathbb{E}_t^{\mathbb{Q}^f}\left[\frac{(X_T - K)^+}{X_T}\right] = Ke^{-r_f(T-t)} \mathbb{E}_t^{\mathbb{Q}^f}\left[\left(\frac{1}{K} - Y_T\right)^+\right]$$

Intuitively, it says that an option to buy 1 CNY with strike SGD-K is the same as K units of an option to sell 1 SGD with strike CNY-$\frac{1}{K}$.

Now the model is:

$$c_f(X_T) = K\left(\frac{1}{K} - Y_T\right)^+$$
$$\frac{1}{K}c_f(X_t) = \mathbb{E}^{\mathbb{Q}^f}\left[e^{-r_f(T-t)}\left(\frac{1}{K} - Y_T\right)^+\right]$$
$$\frac{1}{K}c_f(X_t) = \frac{1}{K}e^{-r_f(T-t)}\Phi\left(-d_2^f\right) - Y_0 e^{-r(T-t)}\Phi\left(-d_1^f\right)$$
$$d_1^f = \frac{\ln\frac{Y_0}{K} + \left(r_f - r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2^f = d_1^f - \sigma\sqrt{T-t}$$

Theorem 6.10. *If $c_f(X_t)$ and $c(X_t)$ are the prices of the same FX option under the foreign and domestic risk-neutral measures, then*

$$c_f(X_t) \cdot X_t = c(X_t)$$

Proof. Recall that $Y_t = \frac{1}{X_t}$, then $-d_1^f = d_2, -d_2^f = d_1$, then the result follows. □

Quanto Derivatives

Quanto derivatives are a type of financial derivative in which the underlying is denominated in a currency other than the currency in which the derivative is settled. The name "quanto" is short for "quantum of one currency into another." Quanto derivatives are used to hedge against foreign exchange risk, offering investors currency protection. They are also used to speculate on the exchange rate of one currency into another. Let's summarize all the components that we have discussed so far:

- B_t^d: Domestic currency bank account: $dB_t^d = r_d B_t^d dt$
- S_t^f: Foreign currency asset: $dS_t^f = \mu_S S_t^f dt + \sigma_S S_t^f dW_t^S$
- X_t: FX rate: $dX_t = \mu_X X_t dt + \sigma_X X_t dW_t^X$, $\langle dW_t^S, dW_t^X \rangle = \rho dt$
- $\overline{S}_t = S_t^f X_t$: Domestic currency value of foreign asset

$$d\overline{S}_t = S_t^f dX_t + X_t dS_t^f + dX_t dS_t^f$$
$$d\overline{S}_t = (\mu_S + \mu_X + \rho\sigma_S\sigma_X)\overline{S}_t dt + \sigma_S \overline{S}_t dW_t^S + \sigma_X \overline{S}_t dW_t^X$$

- B_t^f: Foreign currency bank account: $dB_t^f = r_f B_t^f dt$
- $\overline{B}_t = B_t^f X_t$: Domestic currency value of foreign bank account

$$d\overline{B}_t = B_t^f dX_t + X_t dB_t^f$$
$$d\overline{B}_t = (r_f + \mu_X)\overline{B}_t dt + \sigma_X \overline{B}_t dW_t^X$$

- V_t: Quanto derivative in currency d. This doesn't depend on the FX rate.

For example, a Singaporean(domestic) investor wants to buy Chinese stock, then he needs a Chinese bank account and holds stock in CNY. However, the investment is based on Singapore and the derivative is in SGD currency, then for him translating all financial assets into SGD is needed.

Set up a replication strategy in a portfolio Π_t. Then in the accounting currency d:

$$\Pi_t = V_t + \Delta_t^S \overline{S}_t + \Delta_t^X \overline{B}_t + \beta_t B_t^d$$

the replication condition is

$$\beta_t B_t^d = -V_t - \Delta_t^S \overline{S}_t - \Delta_t^X \overline{B}_t$$

moreover, by self-financing, we have

$$d\Pi_t = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_t^f} dS_t^f + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^{f2}} dS_t^{f2} + \Delta_t^S d\overline{S}_t + \Delta_t^X d\overline{B}_t + \beta_t dB_t^d$$

Put dynamics into it, we have

$$\begin{aligned} d\Pi_t = &\frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_t^f} dS_t^f + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^{f2}} \sigma_S^2 S_t^{f2} dt \\ &+ \Delta_t^S \left[S_t^f dX_t + X_t dS_t^f + dX_t dS_t^f\right] + \Delta_t^X \left[B_t^f dX_t + X_t dB_t^f\right] + \beta_t r_d B_t^d dt \end{aligned}$$

Note that we need to hedge out both the FX rate risk and the stock risk. Hence we need to have the hedging condition:

$$\Delta_t^S = -\frac{1}{X_t}\frac{\partial V}{\partial S_t^f}, \quad \Delta_t^X = -\frac{\Delta_t^S S_t^f}{B_t^f}$$

then the replication condition becomes

$$\beta_t B_t^d = -\left(V_t - S_t \frac{\partial V}{\partial S_t} + S_t \frac{\partial V}{\partial S_t}\right) = -V_t$$

and note that $dX_t dS_t^f = \rho\sigma_S\sigma_X S_t^f X_t dt = \rho\sigma_S\sigma_X \overline{S}_t dt$, then

$$\begin{aligned}
d\Pi_t &= \frac{\partial V}{\partial t}dt + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^{f2}}\sigma_S^2 S_t^{f2} dt + \Delta_t^S \rho\sigma_S\sigma_X \overline{S}_t dt + \Delta_t^X r_f \overline{B}_t dt + \beta_t r_d B_t^d dt \\
&= \left[\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^{f2}}\sigma_S^2 S_t^{f2} + \frac{\partial V}{\partial S_t}\left(r_f - \rho\sigma_S\sigma_X\right) S_t + \beta_t r_d B_t^d\right] dt \\
&\overset{\text{Replication}}{=\!=} \left[\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^{f2}}\sigma_S^2 S_t^{f2} + \frac{\partial V}{\partial S_t}\left(r_f - \rho\sigma_S\sigma_X\right) S_t - r_d V_t\right] dt
\end{aligned}$$

Hence the PDE for the Quanto derivative is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^{f2}}\sigma_S^2 S_t^{f2} + \frac{\partial V}{\partial S_t}\left(r_f - \rho\sigma_S\sigma_X\right) S_t - r_d V_t = 0$$

It turns out that the appropriate dynamics for S_t^f which solve the PDE above via an expectation is:

$$dS_t^f = \left(r_f - \rho\sigma_S\sigma_X\right) S_t^f dt + \sigma_S S_t^f dW_t$$

where $r_f - \rho\sigma_S\sigma_X$ is called **foreign interest rate in quanto adjustment**. By Feynman-Kac, the price of a Quanto derivative is

$$V(t, S_t^f) = e^{-r_d(T-t)}\mathbb{E}_t^{\mathbb{Q}^f}\left[V(T, S_T^f)\right]$$

By writing the dynamcis of the asset as

$$dS_t^f = (r_d - q)\, S_t^f + \sigma_S S_t^f dW_t, \quad q = r_d - r_f + \rho\sigma_S\sigma_X$$

this becomes nothing more than an equation that has already been covered: the risk neutral dynamics for a domestic asset which pays a dividend at rate q. Hence the formula is

$$V(t, S_t^f) = S_t^f e^{-q(T-t)} N(d_1) - K e^{-r_d(T-t)} N(d_2)$$

where

$$\begin{aligned}
d_1 &= \frac{\log(S_t^f/K) + (r_d - q + \sigma_S^2/2)(T-t)}{\sigma_S\sqrt{T-t}} \\
d_2 &= d_1 - \sigma_S\sqrt{T-t}
\end{aligned}$$

7 American Options

Definition 7.1. *An* ***American Option*** *is a type of financial derivative that gives the holder the right, but not the obligation, to buy or sell an underlying asset at a specified price (the strike price)* ***at any time before or on the expiration date****.*

If exercised at time $t < T$, an American call option has payoff $(S_t - K)^+$, an American put option has payoff $(K - S_t)^+$.

American Options were first traded on the Chicago Board Options Exchange (CBOE) in 1973. The flexibility of American-style options made them particularly attractive to investors seeking greater control over the timing of their decisions, especially in volatile markets.

A **Bermuda Option** is a type of option that falls between American and European options in terms of exercise flexibility. It can be exercised on a set of predetermined dates before the expiration, not just any time like an American Option. The exercise dates are typically spaced at regular intervals, such as monthly or quarterly, giving the holder more flexibility compared to a European Option but less than an American Option.

Bermuda Options are particularly useful for managing specific financial strategies where some, but not total, flexibility in timing is beneficial. They are often used in complex financial products and for hedging certain types of insurance or interest rate risks.

This part includes:

1. Early Exercise and Optimal Stopping.
2. American Put Options:
   - Linear Complemenarity Problem
   - Free Boundary
   - Smooth Pasting condition
3. Perpetual American Options.

Recall that we've already derived the bounds for American Option in Chapter 2:

$$\left(S_t - Ke^{-r(T-t)}\right)^+ \leq C_E\left(S_t, t\right) \leq C_A\left(S_t, t\right) \leq S_t$$
$$\left(Ke^{-r(T-t)} - S_t\right)^+ \leq P_E\left(S_t, t\right) \leq P_A\left(S_t, t\right) \leq K$$

where C_E and P_E are the European call and put option prices, C_A and P_A are the American call and put option prices.

7.1 Early Exercise and Optimal Stopping

Definition 7.2. *Let $\mathcal{F} := \left(\mathcal{F}_t\right)_{t\geq 0}$ be a filtration on probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A* ***stopping time*** *with respect to $\mathcal{F}$ is a non-negative random variable τ such that*

$$\{\omega \in \Omega \mid \tau(\omega) \leq t\} := \{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all } t \geq 0.$$

Example 17. *Let X_t be a stochastic process. The first time X_t hits a certain level K is a stopping time:*

$$\tau = \inf\{t \geq 0 \mid X_t = K\}$$

At any time t, we can determine whether the time τ has occurred or not.

For an American Option, the time of exercise is a stopping time. Take American Call as an example, we need to study

$$\arg\max_{0\leq\tau\leq T} \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}\left(S_\tau - K\right)^+\right]$$

if we define $Z_t := e^{-rt}\left(S_t - K\right)^+$, then the problem becomes: given Z_t, find τ that maximize $\mathbb{E}^{\mathbb{Q}}\left[Z_\tau\right]$.

Lemma 7.1. ***Optional Sampling Theorem***: *Let X_t be a martingale, and $\tau \in [t, T]$ be a stopping time. Then X_τ is integrable and*

- *If Z_t is a submartingale, then $Z_t \leq \mathbb{E}_t(Z_\tau)$.*
- *If Z_t is a supermartingale, then $Z_t \geq \mathbb{E}_t(Z_\tau)$.*
- *If Z_t is a martingale, then $Z_t = \mathbb{E}_t(Z_\tau)$.*

Hence if we want to solve $\sup_{0 \leq \tau \leq T} \mathbb{E}[Z_\tau]$. Hence,

1. If Z_t is a submartingale, then the maximizer is $\hat{\tau} = T$.
2. If Z_t is a supermartingale, then the maximizer is $\hat{\tau} = 0$.
3. If Z_t is a martingale, then any stopping time τ with $0 \leq \tau \leq T$ is optimal.

Definition 7.3. *For each stopping time, define the* ***value function*** *as*

$$J_t(\tau) = \mathbb{E}_t(Z_\tau)$$

Compare all stopping time, define

$$V_t = \sup_{t \leq \tau \leq T} J_t(\tau)$$

as the ***optimal value function****.*

Let a stopping time $\hat{\tau}_t \in [t, T]$ is optimal at time t with $J_t(\hat{\tau}_t) = V_t = \sup_{t \leq \tau \leq T} J_t(\tau)$, which means for all time $t \leq \tau \leq T$, strategy that stops at $\hat{\tau}_t$ gains the maximum expected payoff.

Now our goal is to characterize the optimal value function V_t.

Theorem 7.1. *Process $(V_t)_{0 \leq t \leq T}$ is a supermartingale.*

Proof. Take $u < t$. Compare 2 strategies from u to maturity T:

1. Use the optimal strategy $\tau_1 = \hat{\tau}_u$
2. Don't stop between u and t. After t, follow the strategy $\tau_2 = \hat{\tau}_t$

Strategy 1 has (by definition) value function V_u at time u. The value function of strategy 2 is

$$J_u(\tau_2) = \mathbb{E}\left[Z_{\tau_2} \mid \mathcal{F}_u\right] = \mathbb{E}\left[\mathbb{E}\left[Z_{\hat{\tau}_t} \mid \mathcal{F}_t\right] \mid \mathcal{F}_u\right] = \mathbb{E}\left[V_t \mid \mathcal{F}_u\right]$$

By definition we know strategy 1 is better than strategy 2 because $u < t$ and for all stopping time in $[u, T]$ and certainly including t, $J_u(\hat{\tau}_u) \geq J_u(\hat{\tau}_t)$. Hence

$$V_u = J_u(\hat{\tau}_u) \geq J_u(\hat{\tau}_t) = \mathbb{E}\left[V_t \mid \mathcal{F}_u\right]$$

Hence V_t is a supermartingale. □

We say that the process X **dominates** the process Y if $X_t \geq Y_t$ almost surely for all $t \geq 0$.

Theorem 7.2. *The optimal value function V_t is the smallest supermartingale dominating Z_t. Also called* ***Snell Envelope*** *of Z_t.*

Proof. Compare 2 strategies from t to maturity T:

1. Use the optimal strategy $\tau_1 = \hat{\tau}_t$
2. Stop at time t, i.e., $\tau_2 = t$

Strategy 1 has value function V_t at time t. The value function of strategy 2 is

$$J_t(\tau_2) = \mathbb{E}\left[Z_{\tau_2} \mid \mathcal{F}_t\right] = \mathbb{E}\left[Z_t \mid \mathcal{F}_t\right] = Z_t$$

By definition we know strategy 1 is better than strategy 2 because t is a stopping time and for all stopping time in $[t, T]$, $J_t(\hat{\tau}_t) \geq J_t(t)$. Hence V_t dominates Z_t. Let X be another supermartingale that dominates Z_t and let $\tau \geq t$ be a stopping time. Then

$$X_t \geq \mathbb{E}\left[X_\tau \mid \mathcal{F}_t\right] \geq \mathbb{E}\left[Z_\tau \mid \mathcal{F}_t\right] = J_t(\tau)$$

This holds for all stopping time $\tau \geq t$, hence X_t dominates V_t. Therefore, V_t is the smallest supermartingale that dominates Z_t. □

Optimal Stopping with Itô Processes

Let X_t be a process that solves the SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

where W_t is a Brownian motion. Consider a payoff function $\varphi(t, X_t)$ our goal is to study

$$\sup_{0 \leq \tau \leq T} \mathbb{E}\left[\varphi(\tau, X_\tau)\right]$$

Define $\hat{\tau}_t$ as a stopping time that realizes $V(t, X_t) = \sup_{t \leq \tau \leq T} \mathbb{E}_t\left[\varphi(\tau, X_\tau)\right]$. Assume V, μ, σ are sufficiently smooth. On the interval $[t, t + \Delta t]$, compare 3 strategies:

1. Use the optimal strategy $\tau_1 = \hat{\tau}_t$, value function $V(t, X_t)$;
2. Stop at once, i.e., $\tau_2 = t$, value function $\varphi(t, X_t)$;
3. Do nothing until $t + \Delta t$, then follow the optimal strategy $\tau_3 = \hat{\tau}_{t+\Delta t}$, value function

$$\mathbb{E}_t[\varphi(\tau_3, X_{\tau_3})] = \mathbb{E}_t[\mathbb{E}[\varphi(\tau_3, X_{\tau_3}) \mid \mathcal{F}_{t+\Delta t}]] = \mathbb{E}_t[V(t + \Delta t, X_{t+\Delta t})]$$

From the optimality of $\hat{\tau}_t$ we get

$$V(t, X_t) \geq \varphi(t, X_t)$$
$$V(t, X_t) \geq \mathbb{E}_t[V(t + \Delta t, X_{t+\Delta t})]$$

Apply Itô's formula to $V(t, X_t)$:

$$\begin{aligned} dV(t, X_t) &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial X_t}dX_t + \frac{1}{2}\frac{\partial^2 V}{\partial X_t^2}dX_t^2 \\ &= \left(\frac{\partial V}{\partial t} + \mu(t, X_t)\frac{\partial V}{\partial X_t} + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2 V}{\partial X_t^2}\right)dt + \sigma(t, X_t)\frac{\partial V}{\partial X_t}dW_t \end{aligned}$$

Hence

$$\begin{aligned} V(t + \Delta t, X_{t+\Delta t}) = V(t, X_t) &+ \int_t^{t+\Delta t}\left(\frac{\partial V}{\partial s} + \mu(s, X_s)\frac{\partial V}{\partial X_s} + \frac{1}{2}\sigma^2(s, X_s)\frac{\partial^2 V}{\partial X_s^2}\right)ds \\ &+ \int_t^{t+\Delta t}\sigma(s, X_s)\frac{\partial V}{\partial X_s}dW_s \end{aligned}$$

Take the conditional expectation $\mathbb{E}_t$ on both sides, we get

$$\begin{aligned} \mathbb{E}_t[V(t + \Delta t, X_{t+\Delta t})] &= V(t, X_t) + \int_t^{t+\Delta t}\mathbb{E}_t\left[\frac{\partial V}{\partial s} + \mu(s, X_s)\frac{\partial V}{\partial X_s} + \frac{1}{2}\sigma^2(s, X_s)\frac{\partial^2 V}{\partial X_s^2}\right]ds \\ &\leq V(t, X_t) \end{aligned}$$

It follow that

$$\frac{\partial V}{\partial s} + \mu(s, X_s)\frac{\partial V}{\partial X_s} + \frac{1}{2}\sigma^2(s, X_s)\frac{\partial^2 V}{\partial X_s^2} \leq 0$$

Summarize 2 conditions:

1. Strategy 1 is better than strategy 2, $V(t, X_t) \geq \varphi(t, X_t)$;

2. Strategy 1 is better than strategy 3,

$$\frac{\partial V}{\partial s} + \mu(s, X_s)\frac{\partial V}{\partial X_s} + \frac{1}{2}\sigma^2(s, X_s)\frac{\partial^2 V}{\partial X_s^2} \leq 0, \quad s \in [t, t+\Delta t]$$

At time t, it is optimal to stop if and only if $V(t, X_t) = \varphi(t, X_t)$. It is optimal not to stop if and only if $V(t, X_t) > \varphi(t, X_t)$.

Define $C := \{(t, X_t) : V(t, X_t) = \varphi(t, X_t)\}$, then C is the **continuation region**, which collects times and values where it is optimal to continue the option.

Proposition 7.1. *Under sufficent regularity conditions, the optimal value function solves*

$$\begin{aligned} &\frac{\partial V}{\partial t} + \mu(t, X_t)\frac{\partial V}{\partial X_t} + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2 V}{\partial X_t^2} = 0, \quad (t, X_t) \in C \\ &V(t, X_t) = \varphi(t, X_t), \quad (t, X_t) \in \partial C \end{aligned}$$

Finding C as well as its boundary ∂C is the key to solving the optimal stopping problem.

American Call Option

For an American Call Option, the payoff function is $\varphi(t, S_t) = (S_t - K)^+$. The optimal stopping probelm is

$$\sup_{0 \leq \tau \leq T} \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}(S_\tau - K)^+\right]$$

where S_t follows the dynamic $dS_t = S_t(rdt + \sigma dW_t)$.

Lemma 7.2. *If $\varphi(X)$ is increasing and convex, and X is a submartingale, then $\varphi(X)$ is a submartingale.*

Proof.

$$\begin{aligned} \mathbb{E}\left[\varphi(X_\tau) \mid \mathcal{F}_t\right] &\geq \varphi(\mathbb{E}[X_\tau \mid \mathcal{F}_t]) \\ &\geq \varphi(X_t) \end{aligned}$$

□

By this lemma, $Z_t = e^{-rt}(S_t - K)^+$ is a submartingale. Hence by previous result, the optimal stopping time is T.

Theorem 7.3. *If the stock pays no dividends, it is **NEVER** optimal to exercise an American Call Option early. The optimal stopping time is T.*

7.2 American Put Options

For an American Put Option, the payoff function is $\varphi(t, S_t) = (K - S_t)^+$. The optimal stopping probelm is

$$V(t, S_t) = \sup_{t \leq \tau \leq T} \mathbb{E}_t^{\mathbb{Q}}\left[e^{-r\tau}(K - S_\tau)^+\right] = \sup_{t \leq \tau \leq T} \mathbb{E}_t^{\mathbb{Q}}\left[\frac{F_\tau}{B_\tau}\right]$$

where F_τ is the price of the put option at exercise time τ and B_τ is the bank account at time τ.

In general, the price of the American put is strictly higher than the price of the European put. Let $t < T$ and assume the Black-Scholes model holds. If $S_t = 0$ then $S_T = 0$ and the price of the European put option is $P_E(t, S_t) = Ke^{-r(T-t)} < K = (K - S_t)^+$. By continuity, also for small $S_t > 0$, $P_E(t, S_t) < (K - S_t)^+$. Instead, the price of an American option is always above its exercise value.

The holder of a call option pays K when the call option is exercised. The holder of a put option receives K when the put option is exercised. Holding K earn interest. That's why it makes sense to exercise a put option early, and not a call option.

Table 7.1: American v.s. European

	Long		**Short**	**Situation**
European	dV_t	$=$	$-\kappa_t dS_t - \beta_t dB_t$	Must equal, or arbitrage exists
American	dV_t	$<$	$-\kappa_t dS_t - \beta_t dB_t$	Will early exercise immediatly

Linear Complementarity Problem and Free Boundary

Slimilary form a portfolio $\Pi_t = V_t + \kappa_t S_t + \beta_t B_t$, one can think it as long V_t and short $-\kappa_t S_t - \beta_t B_t$. The situation in American case is different from European case. Hence for American put option, we have

$$\frac{\partial F}{\partial t} + r\frac{\partial F}{\partial S}S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 - rF \leq 0$$

If it's not optimal to exercise, then at t $d\Pi_t = dV_t + \kappa_t dS_t + \beta_t dB_t = 0$, otherwise one will exercise immediately. Hence

$$\frac{\partial F}{\partial t} + r\frac{\partial F}{\partial S}S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 - rF = 0, \quad \text{Not optimal to exercise i.e. } (t, S_t) \in C$$

Summarize:

- Option not exercised:

$$F(t, S_t) > (K - S_t)^+$$
$$\frac{\partial F}{\partial t} + r\frac{\partial F}{\partial S}S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 - rF = 0$$

- Option exercised:

$$F(t, S_t) = (K - S_t)^+$$
$$\frac{\partial F}{\partial t} + r\frac{\partial F}{\partial S}S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 - rF \leq 0$$

Table 7.2: American v.s. European

	European	**American**
Payoff	$P_E(T, S_T) = (K - S_T)^+$	$P_A(\tau, S_\tau) = (K - S_\tau)^+$
Price	$P_E(t, S_t) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[P_E(T, S_T)]$	$P_A(t, S_t) = \sup_{t \leq \tau \leq T} e^{-r(\tau - t)}\mathbb{E}_t^{\mathbb{Q}}[P_A(\tau, S_\tau)]$
Optimal Value Function	$f(t, S_t) = e^{-rt}P_E(t, S_t)$	$\begin{aligned} V(t, S_t) &= \sup_{t \leq \tau \leq T} \mathbb{E}_t^{\mathbb{Q}}\left[e^{-r\tau}P_A(\tau, S_\tau)\right] \\ &= e^{-rt}P_A(t, S_t) \end{aligned}$
Dynkin	$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = 0$	$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \begin{cases} = 0, & (t, S_t) \in C \\ \leq 0, & \text{exercise when } V(t, S_t) = e^{-rt}(K - S_t)^+ \end{cases}$
Feynman-Kac	$\frac{\partial P_E}{\partial t} + rS\frac{\partial P_E}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P_E}{\partial S^2} - rP_E = 0$	$\frac{\partial P_A}{\partial t} + rS\frac{\partial P_A}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P_A}{\partial S^2} - rP_A \begin{cases} = 0, & (t, S_t) \in C, \quad P_A(t, S_t) > (K - S_t)^+ \\ \leq 0, & \text{exercise when} \quad P_A(t, S_t) = (K - S_t)^+ \end{cases}$

Proposition 7.2. ***Linear Complementarity Problem**: The price of an American put option satisfies the following linear complementarity problem: The price F (t, S_t) of the American option fulfils*

$$\min\left\{-\frac{\partial F}{\partial t} - r\frac{\partial F}{\partial S}S - \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 + rF, F - \varphi\right\} = 0$$
$$F(T, S) = \varphi(S) = (K - S)^+$$

where the domain is

$$D = [0, T) \times (0, +\infty)$$

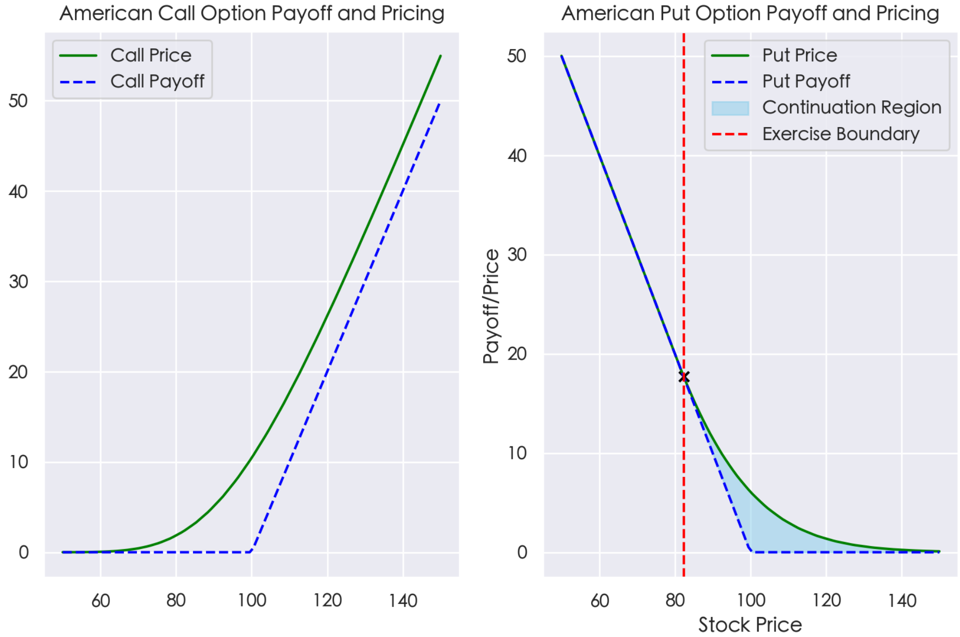

Figure 7.1: Continuation Region and Exercise Boundary for American Put Option

Proposition 7.3. *Let $S_*(t)$ be the optimal exercise boundary for an American put option.*

$$\frac{\partial F}{\partial t} + r\frac{\partial F}{\partial S}S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 - rF = 0, \quad S > S_*(t)$$
$$F(t, S_*(t)) = K - S_*(t).$$
$$F(T, S) = (K - S)^+.$$

Smooth Pasting Condition

Proposition 7.4. *At exercise boundary $S_*(t)$, the price of the American put option satisfies the smooth pasting condition:*

$$\frac{\partial F}{\partial S}(t, S_*(t)) = -1$$

Proof. If $\frac{\partial F}{\partial S}(t, S_*(t)) > -1$, then $S_*(t)$ is not the threshold for early exercise. If $\frac{\partial F}{\partial S}(t, S_*(t)) < -1$, then

$$P(t, S_*(t)) < K - S_*(t)$$

An arbitrage opportunity exists: buy the put option and exercise it immediately, then sell the stock short. The profit is $K - S_*(t) - P(t, S_*(t)) > 0$.

Hence $\frac{\partial F}{\partial S}(t, S_*(t)) = -1$. □

7.3 Perpetual American Options

A perpetual American option is an American option that can be exercised at any time, the option does not expire. Let P_∞ be the price of the perpetual American put option. In most cases there is no analytical solution to the free boundary value problem, but perpetual American option is an exception.

Assume the price P_∞ only depends on S, the free boundary problem reduces to

$$rS\frac{\partial}{\partial S}P_\infty + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2}P_\infty - rP_\infty = 0, \quad S > S^*$$

The general solution for this ordinary differential equation is

$$P_\infty = AS + BS^{-\frac{2r}{\sigma^2}} := AS + BS^{-\gamma}$$

A must be 0 because P_∞ is bounded for large S.

The continuity and smooth pasting conditions are

$$\begin{aligned} P_\infty(S^*) &= K - S^* \\ \frac{\partial P_\infty}{\partial S}(S^*) &= -1 \end{aligned}$$

the second condition gives $B = \frac{S^{*\gamma+1}}{\gamma}$, hence

$$S^* = \frac{2rK}{\sigma^2 + 2r} = \frac{\gamma K}{1+\gamma}; \quad P_\infty = \frac{K}{1+\gamma}\left(\frac{S^*}{S}\right)^\gamma$$

8 Interest Rate Modelling(An Introduction)

In previous chapters we introduced a simplified model for interest rates using the money market bank account where interest rate accrues deterministically. But interest rates are much more complex, and they are stochastic. At the end of 2021, as the world emerged from Covid, the world had experienced almost two years where interest rates were at rock bottom and hardly moved. They were for all intents and purposes anchored down, and deterministic.

But as the world opened up in 2022 interest rates increased significantly; and they increased much more than anyone predicted. They were the epitome of a stochastic process. In this chapter we will address the largest class of derivatives that make up the financial derivatives market: interest rate derivatives.

8.1 The Bank Account Revisited

In the previous chapters we introduced the bank account as a deterministic process. We will now introduce the bank account as a stochastic process. The bank account is a money market account that accrues interest at the short rate r_t. The value of such an account at time t is

$$B_t = e^{\int_0^t r_s ds}, \quad dB_t = r_t B_t dt$$

where r_t is the short rate at time t. And although there is no Brownian increment; the process for B_t is stochastic simply because $r(t)$ is stochastic. Independently to the process for B_t we can model $r(t)$ separately. It is often called the short rate, and for these notes we will consider three short rate models.

Ho-Lee Model

The Ho-Lee model is a simple model for the short rate. It is a one-factor model where the short rate is driven by a single Brownian motion. The dynamics of the short rate under the Ho-Lee model is

$$dr_t = \theta_t dt + \sigma dW_t$$

θ_t is interpreted as the speed of interest rate with time; and it can clearly be positive and negative to signal rising or falling interest rates. If it is zero then this signals a model where interest rates only fluctuate about the current level r_0.

The model is Gaussian,

$$\begin{aligned} \mathbb{E}\left[r_T\right] &= r_0 + \int_0^T \mu_t dt \\ \mathrm{Var}\left[r_T\right] &= \sigma^2 T \end{aligned}$$

Ornstein-Uhlenbeck: Vasicek Model

The Vasicek model is a one-factor model where the short rate is driven by a mean-reverting process. The dynamics of the short rate under the Vasicek model is

$$dr_t = \alpha(\theta - r_t)dt + \sigma dW_t$$

α is the speed of mean reversion, θ is the long-term mean of the short rate, and σ is the volatility of the short rate.

Solving the stochastic differential equation is not difficult and we have all the tools from stochastic calculus to do this, and we will use the integrating factor e^{at}.

$$
\begin{aligned}
dr_t &= \alpha\left(\theta - r_t\right)dt + \sigma dW_t \\
dr_t + \alpha r_t dt &= \alpha\theta dt + \sigma dW_t \\
d\left(r_t e^{\alpha t}\right) &= \alpha\theta e^{\alpha t}dt + \sigma e^{\alpha t}dW_t \\
r_T e^{\alpha T} - r_0 &= \alpha\theta\int_0^T e^{\alpha t}dt + \sigma\int_0^T e^{\alpha t}dW_t \\
r_T e^{\alpha T} - r_0 &= \theta\left(e^{\alpha T} - 1\right) + \sigma\int_0^T e^{\alpha t}dW_t \\
r_T &= r_0 e^{-\alpha T} + \theta\left(1 - e^{-\alpha T}\right) + \sigma\int_0^T e^{-\alpha(T-t)}dW_t
\end{aligned}
$$

We see that r_T is Gaussian,

$$
\begin{aligned}
\mathbb{E}\left[r_T\right] &= r_0 e^{-\alpha T} + \theta\left(1 - e^{-\alpha T}\right) \\
\operatorname{Var}\left[r_T\right] &= \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha T}\right) \\
\lim_{T\to\infty}\mathbb{E}\left[r_T\right] &= \theta
\end{aligned}
$$

Ornstein-Uhlenbeck: Hull-White Model

The Hull-White formulation is an upgrade of Vasicek to allow the parameters of the model to be time dependent. Although this extends to α, in practice it is held at a constant value.

$$dr_t = \alpha\left(\theta_t - r_t\right)dt + \sigma_t dW_t$$

The solution is a slightly more tedious repetition of the derivation from the previous section:

$$
\begin{aligned}
r_T &= r_0 e^{-\alpha T} + \alpha\int_0^T \theta_t e^{-\alpha(T-t)}dt + \int_0^T \sigma_t e^{-\alpha(T-t)}dW_t \\
\mathbb{E}\left[r_T\right] &= r_0 e^{-\alpha T} + \alpha\int_0^T \theta_t e^{-\alpha(T-t)}dt \\
\operatorname{Var}\left[r_T\right] &= \int_0^T \sigma_t^2 e^{-2\alpha(T-t)}dt
\end{aligned}
$$

And now $\lim_{T\to\infty}\mathbb{E}\left[r_T\right]$ depends on the form of θ_t.

All three models are Gaussian and this implies that they all result in a lognormal distribution for the money market account.

8.2 Tradeable Interest Rate (Bonds)

Interest rates are not assets; they are a measure of the cost of money. You can not buy or sell interest rates in the same way that you can buy or sell stocks. So interest rate derivatives will depart slightly from the no arbitrage technique that we have used so far in the course. No arbitrage is still the goal, but since we can not trade interest rate directly, we must first introduce the concept of an **interest rate asset**.

Definition 8.1. *A* ***zero coupon bond*** *is an instrument that pays one unit at a future maturity date. If interest rates were deterministic, the value of the zero coupon bond was simply* $e^{-r(T-t)}$.

Denote the value of such a derivative at time $t < T$ as $Z(t,T)$. It represents some 'average' value for $e^{-\int_t^T r(s)ds}$, which must come from replication.

More generally, bonds pay a strip of future cash flows called **coupons** to the bond holder. Denote the **notional** of the bond as 1, its maturity T_N, $i = 1, 2, \cdots, N$ to be the coupon payment dates. If the bond pays a coupon c at each coupon date, then the value of the bond at time t is

$$P(t,T) = \sum_{i=1}^{N} cZ(t,T_i)\delta_i + Z(t,T_N) \to \int_t^T cZ(t,s)ds + Z(t,T)$$

where δ_i is the time between coupon payments $[T_{i-1}, T_i)$ and c is the coupon rate.

If $P(t,T)$ comes out to precisely 1, then this bond is said to be **priced at par**.

Definition 8.2. *Given the price of a bond $P(t,T)$, the* ***yield to maturity*** *is the rate value y such that you retrieve the bond price via*

$$P(t,T) = \sum_{i=1}^{N} c\delta_i e^{-yT_i} + e^{-yT_N} \rightarrow \int_t^T ce^{-ys}ds + e^{-yT}$$

Theorem 8.1. *If a bond pays at par, i.e., the initial price of the bond equals to the par value 1 here, then the yield to maturity is the same as the coupon rate.*

Proof. We have

$$P(0,T) = \int_0^T ce^{-ys}ds + e^{-yT} = 1$$

yielding

$$\frac{c}{y}\left(1 - e^{-yT}\right) + e^{-yT} = 1$$

which implies $c = y$. □

8.3 Interest Rate Derivatives Modelling

Derivatives on interest rates will be denoted by $V(t, r_t)$ and the hedge instrument will be an already existing derivative $V_1(t, r_t)$ that can be bought and sold freely without friction as per the Black Scholes assumptions. A stochastic money market account is also assumed to exist where cash accrued at the prevailing (albeit random) interest rate. For now assume that interest rates follow the stochastic differential equation:

$$\begin{aligned} dr_t &= \mu(t, r_t)\, dt + \sigma(t, r_t)\, dW_t \\ dB_t &= r_t B_t dt \end{aligned}$$

We proceed as usual. We form a portfolio $\Pi_t = V(t, r_t) + \Delta_t V_1(t, r_t) + \psi_t B_t$ and require that the portfolio is self-financing. The value of the portfolio at time t is

$$\Pi_t = V(t, r_t) + \Delta_t V_1(t, r_t) + \psi_t B_t$$

As before, the replication condition ensures that $\Pi_t = d\Pi_t = 0, \forall t \in [0,T]$. The self financing strategy will also be invoked: $V_1(t, r_t)\, d\Delta_t + B_t d\psi_t = 0$. We have:

$$\begin{aligned} \psi_t B_t &= -\left[V(t, r_t) + \Delta_t V_1(t, r_t)\right] \\ dV(t, r_t) &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial r_t}dr_t + \frac{1}{2}\frac{\partial^2 V}{\partial r_t^2}(dr_t)^2 \\ dV_1(t, r_t) &= \frac{\partial V_1}{\partial t}dt + \frac{\partial V_1}{\partial r_t}dr_t + \frac{1}{2}\frac{\partial^2 V_1}{\partial r_t^2}(dr_t)^2 \end{aligned}$$

Setting $\Delta_t = -\left(\frac{\partial V}{\partial r_t}\right) / \left(\frac{\partial V_1}{\partial r_t}\right)$ eliminates all terms in dW_t. It is convenient to define the operator $\mathcal{L}$ as:

$$\mathcal{L} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma(t, r_t)^2 \frac{\partial^2}{\partial r_t^2}$$

We then have:

$$\begin{aligned} \mathcal{L}V - rV &= \Delta_t\left(\mathcal{L}V_1 - rV_1\right) \\ \frac{\mathcal{L}V - rV}{\frac{\partial V}{\partial r_t}} &= \frac{\mathcal{L}V_1 - rV_1}{\frac{\partial V_1}{\partial r_t}} \end{aligned}$$

The only way for this equation to be mathematically valid is for both left and right hand side to equal some function of (t, r_t). Of course this function is not known so we can call it $a(t, r_t)$.

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma(t, r_t)^2 \frac{\partial^2 V}{\partial r_t^2} + a(t, r_t)\frac{\partial V}{\partial r_t} - rV = 0$$

We can now also derive using Feynman-Kac theorem that the price at time t for some payoff V_T is:

Theorem 8.2. *Derivatives on stochastic interest rate have the following pricing formula:*

$$V_t = \mathbb{E}\left[e^{-\int_t^T r_t ds} V_T \mid r_t\right] \xlongequal{\text{Define } D(t)=e^{-\int_0^t r_t ds}} \frac{1}{D(t)} \mathbb{E}\left[D(T) V_T \mid r_t\right]$$

This being valid under the risk neutral dynamics for r_t :

$$dr_t = a\left(t, r_t\right) dt + \sigma\left(t, r_t\right) dW_t$$

Note 20. *Everything is great except for the problem that we do not have an expression for the function $a\left(t, r_t\right)$. The solution is to resort to calibration. Assume that the market is efficient and arbitrage free and that the value of $a\left(t, r_t\right)$ can be extracted from the prices of existing derivatives in the market. We already assumed from the replication strategy that some derivative $V_1\left(t, r_t\right)$ exists that we can trade. So the idea is to imply $a\left(t, r_t\right)$ such that it fits the prices of $V_1\left(t, r_t\right)$. Once $a\left(t, r_t\right)$ is known then it can be used to price any other derivative $V\left(t, r_t\right)$.*

The most natural derivatives that can be used for calibration turn out to be the very bonds that we set out to price: $Z(t, T)$ for all maturities T. Typically T goes out $30y$ into the future.

Term Structure Calibration(Ho-Lee)

Under risk-neutral dynamics, the model is now

$$dr_s = a(s)dt + \sigma dW_s, \quad r_0 \text{ is given}$$

And under this model, the price of a zero coupon bond with notional 1 is

$$Z\left(0, T\right) = \mathbb{E}\left[e^{-\int_0^T r_t dt} \middle| r_0\right]$$

What we are doing here is calibrating the model, so we extract $Z\left(0, T\right)$ from the market, and reverse solve the equation above to extract $a(s)$. For this we need to writ down the pricing formula for $Z\left(0, T\right)$ in terms of $a(s)$. We have:

$$\begin{aligned}
Z\left(0, T\right) &= \mathbb{E}\left[\exp\left(-\int_0^T r_t dt\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\int_0^T \left(r_0 + \int_0^t a(s)dt + \sigma \int_0^t dW_s\right) dt\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\int_0^T r_0 dt - \int_0^T \int_0^t a(s)dsdt - \sigma \int_0^T \int_0^t dW_s dt\right)\right] \\
&= \mathbb{E}\left[\exp\left(-r_0 T - \int_0^T \int_0^t a(s)dsdt - \sigma \int_0^T \int_s^T dtdW_s\right)\right] \\
&= \mathbb{E}\left[\exp\left(-r_0 T - \int_0^T \int_0^t a(s)dsdt - \sigma \int_0^T [T-s]dW_s\right)\right] \\
&= \exp\left(-r_0 T - \int_0^T \int_0^t a(s)dsdt + \frac{\sigma^2}{6} T^3\right)
\end{aligned}$$

A neater expression for the price of the zero coupon bond at any time $t < T$ is

$$\begin{aligned}
Z(t, T) &= \exp\left(-r_t[T-t] - \int_t^T \int_t^u a(s)dsdu + \frac{\sigma^2}{6}(T-t)^3\right) \\
&= \exp\left(-r_t[T-t] - \int_t^T a(s)[T-s]ds + \frac{\sigma^2}{6}[T-t]^3\right)
\end{aligned}$$

It is worth noting that as $t \to T$ each term in the exponential goes to zero and so we will have the desired feature of $\lim_{t \to T} Z\left(t, T\right) = 1$. This is what must happen economically as well because a risk free zero coupon bond will always redeem at par.

We can also now deduce the expression for the function $a(t)$ under this model which calibrates to the price of zero coupon bonds:

$$
\begin{aligned}
\ln Z(0,T) &= -r_0 T - \int_0^T \int_0^t a(s) ds dt + \frac{\sigma^2}{6} T^3 \\
\frac{\partial}{\partial T} (\ln Z(0,T)) &= -r_0 - \int_0^T a(s) ds + \frac{\sigma^2}{2} T^2 \\
\frac{\partial^2}{\partial T^2} (\ln Z(0,T)) &= -a(T) + \sigma^2 T \\
a(t) &= -\left. \frac{\partial^2}{\partial T^2} (\ln Z(0,T)) \right|_{T=t} + \sigma^2 t
\end{aligned}
$$

Ho-Lee Model Bond Price Dynamics

The thing that remains is to compute the SDE for the price of a zero coupon bond under the Ho-Lee model. **This allows us to model any other derivatives on zero coupon bonds**. Denote the price of a bond $Z(t,T) := e^{\Psi(t,T)}$ where

$$
\Psi(t,T) = -r_t[T-t] - \int_t^T a(s)[T-s]ds + \frac{\sigma^2}{6}[T-t]^3
$$

and the SDE for $\Psi(t,T)$ is derived using differential calculus; note that in general we must preserve the second order term that is required by Ito's Lemma, although there is no Ito term in this specific case since $\Psi(t,T):$ is first order linear in r_t.

$$
\begin{aligned}
d\Psi(t,T) &= \frac{\partial \Psi}{\partial t} dt + \frac{\partial \Psi}{\partial r_t} dr_t + \frac{1}{2} \frac{\partial^2 \Psi}{\partial r_t^2} (dr_t)^2 \\
&= r_t dt - [T-t] dr_t + a(t)[T-t]dt - \frac{\sigma^2}{2}[T-t]^2 dt \\
&= r_t dt - \frac{\sigma^2}{2}(T-t)^2 dt - (T-t)\sigma dW_t
\end{aligned}
$$

where we have also put in the dynamics of the Ho-Lee model: $dr_t = a_t dt + \sigma dW_t$. The risk neutral dynamics of the bond are now derived using Ito's Lemma

$$
\begin{aligned}
dZ(t,T) &= \frac{\partial Z(t,T)}{\partial \Psi} d\Psi + \frac{1}{2} \frac{\partial^2 Z(t,T)}{\partial \Psi^2} (d\Psi)^2 \\
\frac{dZ(t,T)}{Z(t,T)} &= r_t dt - \sigma(T-t) dW_t
\end{aligned}
$$

Proposition 8.1. ***Hyper-Generalized Geometric Brownian Motion***: *The price of a zero coupon bond under the Ho-Lee model is*

$$
Z(t,T) = Z(0,T) \exp\left\{ \left(\int_0^t r_s ds - \frac{1}{2}\sigma^2 \int_0^t (T-s)^2 ds \right) - \sigma \sqrt{\int_0^T (T-s)^2 ds} Z \right\}, \quad Z \sim \mathcal{N}(0,1)
$$

Example 18. *A call option on a T maturity bond, where the option expires on date $T^* < T$, can be evaluated as a risk-neutral expectation of the payoff at expiry date. At this date, the distribution of the zero coupon bond is modelled as*

$$
Z(T^*,T) = Z(0,T) \exp\left\{ \left(\int_0^{T^*} r_t dt - \frac{1}{2}\sigma^2 \int_0^{T^*} (T-t)^2 dt \right) - \sigma \sqrt{\int_0^T (T-t)^2 dt} Z \right\}
$$

Then the call-option on this zero coupon bond has price

$$
V(0) = \mathbb{E}\left[e^{-\int_0^{T^*} r_t dt} (Z(T^*,T) - K)^+ \middle| r_0, Z(0,T) \right]
$$

8.4 Forward

8.4.1 Forward Pricing

Now that the zero coupon bond has been introduced we can replace the money market account. Under stochastic interest rates the money market account grows randomly and this destroys the credibility of being to replicate a fixed payout in the future.

Formally define $D(t) = e^{-\int_0^t r_s ds}$ as the discount factor process. It's value is not known until all interest rates for the period $[0, T]$ have been realised. So it can not possibly be the fair value of anything at time 0. However, we do have the means to accrue money deterministically; this is the zero coupon bond $Z(0, T)$. \$1 invested into the bond at time 0 accrues to $\frac{1}{Z(0,T)}$ by time T. Selling this bond at time 0 and then buying it back at time T is the means to take out a loan for the period $[0, T]$. These are like loans whose terms are negotiated and finalized at time t event hough the loan term is T . $Z(t, T)$ ends up playing the role of the money market account when interest rates are stochastic.

Definition 8.3. *A **forward contract** is an agreement between two parties to buy or sell an asset at a future date T for a price K agreed upon today. The price agreed upon today is called the **forward price**. **Forward price must be such that the contract has zero value** at time* 0.

Theorem 8.3. *Let* $\mathrm{For}(t, T)$ *be the price of a forward contract at time t with maturity T. The price of a forward contract is*

$$K = \mathrm{For}(t, T) = \frac{S(t)}{Z(t, T)}$$

Proof. An inverstor sell a forward contract today and gain 0. Simultanously, he shorts $S(t)/Z(t, T)$ zero coupon bonds and hence gains $S(t)$, immediately buys the asset. At time T, he recieves K due to the contract and loses asset. Note that borrowing $S(t)$ from zero coupon bonds means he will return back $S(t)/Z(t, T)$ at time T. Hence at time T, his net gain is $K - S(t)/Z(t, T)$. If this is larger than 0, then an arbitrage opportunity exists, if this is smaller than 0, one can buy the forward contract and short the zero coupon bonds to make a profit. Hence the forward price must be $K = S(t)/Z(t, T)$. ☐

Note 21. *The value of a forward contract at time t is*

$$\frac{1}{D(t)}\mathbb{E}\left[D(T)(S(T) - K)|\,\mathcal{F}_t\right] = \frac{1}{D(t)}\mathbb{E}[D(T)S(T)|\mathcal{F}_t] - \frac{K}{D(t)}\mathbb{E}[D(T)|\mathcal{F}_t] = S(t) - KZ(t, T)$$

this must be zero by definition.

8.4.2 Forward Measure

Definition 8.4. *Let T be the maturity of a forward contract. The **forward measure** is the measure under which the price of the forward contract is a martingale. The forward measure is denoted as* $\mathbb{P}^T$.

$$L(t, T) = \frac{D(T)}{D(t)} \cdot \frac{Z(T, T)}{Z(t, T)} = \frac{D(T)}{D(t)Z(t, T)}$$

Note that the zero coupon bond is a kind of financial asset, so its discounting price under $\mathbb{Q}$ is a martingale:

$$d(D(t)Z(t, T)) = \sigma^* D(t)Z(t, T)dW_t^{\mathbb{Q}}$$

by the change of measure theorem I, we have

$$dW_t^{\mathbb{P}^T} = dW_t^{\mathbb{Q}} - \sigma^* dt$$

is a Brownian motion under the forward measure $\mathbb{P}^T$.

The price of the forward contract is actually the price of asset S in unit of zero coupon bond $Z(t, T)$, so the price of the forward contract is a martingale under the forward measure.

Theorem 8.4. ***Pricing under Forward Measure**: The price of a derivative V at time t is*

$$V(t) = Z(t, T)\mathbb{E}^{\mathbb{P}^T}[V(T) \mid \mathcal{F}_t]$$

Proof. We have seen

$$V_t = \frac{1}{D(t)}\mathbb{E}^{\mathbb{Q}}\left[D(T)V_T \mid \mathcal{F}_t\right]$$

By the proposition of measure change,

$$V_t = \frac{1}{D(t)}\mathbb{E}^{\mathbb{Q}}\left[D(T)V_T \mid \mathcal{F}_t\right] = \frac{1}{D(t)}\mathbb{E}^{\mathbb{P}^T}\left[\frac{1}{L(t, T)}D(T)V_T \mid \mathcal{F}_t\right] = Z(t, T)\mathbb{E}^{\mathbb{P}^T}[V(T) \mid \mathcal{F}_t]$$

☐

Black-Scholes Formula under Stochastic Interest Rates

Classic Black-Scholes formula is derived under the assumption of deterministic interest rates. We will now derive the Black-Scholes formula under stochastic interest rates.

In classic BS model, we are considering bank account B_t and under risk-neutral measure, the stock price S_t follows

$$d\left(\frac{S_t}{B_t}\right) = \sigma \frac{S_t}{B_t} dW_t^{\mathbb{Q}}$$

Under stochastic interest rates, **we have discussed that the money market account is replaced by the zero coupon bond**. Hence we're considering the stock price S_t in unit of zero coupon bond $Z(t,T)$, and under the **forward measure** $\mathbb{P}^T$, we have

$$\begin{aligned} d\left(\frac{S_t}{Z(t,T)}\right) &= \sigma \frac{S_t}{Z(t,T)} dW_t^{\mathbb{P}^T}, \quad \text{i.e.,} \\ d\,\mathrm{For}(t,T) &= \sigma\,\mathrm{For}(t,T) dW_t^{\mathbb{P}^T} \end{aligned}$$

Theorem 8.5. ***Black-Scholes Formula under Stochastic Interest Rates****: Let S_t be the underlying asset and $\mathrm{For}(t,T)$ be the forward price of the asset at time t with maturity T with dynamics above. The price of a European call option with strike K and maturity T is*

$$C(t) = S(t)\Phi(d_1) - KZ(t,T)\Phi(d_2)$$

where

$$\begin{aligned} d_1 &= \frac{\log\left(\frac{\mathrm{For}(t,T)}{K}\right) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned}$$

Note 22. *If the rate is deterministic, then the price of zero coupon bond is $e^{-r(T-t)}$, the forward price $\mathrm{For}(t,T) = S(t)e^{-r(T-t)}$, and the Black-Scholes formula will be the same as the classic one.*

Forward Rate Agreement

Assuming that we have a continuum of bond prices $Z(0,T)$, which are observed from the market at time 0.

Consider two dates in the future T_1 and T_2 where $T_1 < T_2$. If we set up a portfolio at time 0 which is composed of a long position in bond $Z(0,T_1)$ and short α units of bond $Z(0,T_2)$ then at time T_1 this portfolio pays the portfolio owner 1 unit of currency, and at time T_2 requires the portfolio holder to pay back α units of currency. The value $\alpha = \dfrac{Z(0,T_1)}{Z(0,T_2)}$ makes this portfolio zero at time 0. The portfolio then constructs a risk free loan to borrow 1 unit at T_1 and pay back $\dfrac{Z(0,T_1)}{Z(0,T_2)}$ at T_2. This is now a loan whose terms are finalized at time 0, even though the term of the loan is the future period $[T_1, T_2]$. These are called forward interest rates:

$$f(0,T_1,T_2) = \frac{1}{\delta}\left[\frac{Z(0,T_1)}{Z(0,T_2)} - 1\right]$$

δ is the day count fraction for time interval (T_1, T_2), using some specified day count convention. This ensures that the rates are always computed as annualized amount regardless of how big the time interval (T_1, T_2) is.

Definition 8.5. ***Forward Rate Agreements*** *(FRAs) are derivative contracts which allow an investor to lock in the interest on a loan or deposit at some future date.*

The logic above is the mathematics behind why the arbitrage free value for such an interest rate is the forward interest rate that is implied from the Zero Coupon bonds. Now consider the forward rate $f(t,s,s+\Delta)$ and take the limit $\Delta \to 0$.

$$\begin{aligned} f(t,s) &= \lim_{\Delta\to 0} f(0,s,s+\Delta) \\ &= \lim_{\Delta\to 0} \frac{1}{\Delta}\left[\frac{Z(0,s) - Z(t,s+\Delta)}{Z(0,s+\Delta)}\right] \\ &= -\frac{1}{Z(t,s)}\frac{\partial Z(t,s)}{\partial s} \\ &= -\frac{\partial}{\partial s}\ln Z(t,s) \end{aligned}$$

In the same way that $r(t)$ is an instantaneous interest rate that is applied to the time interval t to $t+dt$, we introduce the **forward interest rate** $f(t,s)$, which is the interest rate that you can lock in for time interval s to $s+ds$, using bonds that are trading in the market on date t. We can now write:

$$Z(t,T) = e^{-\int_t^T f(t,s)ds}$$

8.5 Swaps

Floating Rate Note

Where fixed coupon bonds pay a fixed deterministic coupon on each future coupon date, a **floating rate note** pays a coupon whose value is not known until the start of the coupon period.

Consider a note which pays coupon on dates T_i for coupon periods $[T_{i-1}, T_i]$ where $i = 1, 2 \cdots n$. T_n would correspond to the maturity of the bond when the principal is repaid. The rates that form the coupon are the deposit rates that are observed at each date T_{i-1}. We've seen that this is equal to the interest rate $f(T_{i-1}, T_{i-1}, T_i)$, which are stochastic right until date T_{i-1}.

However, the way to replicate this coupon at time 0 follows from the arguments above and we can compute the arbitrage free time 0 value of this coupon as $f(T_{i-1}, T_{i-1}, T_i)$.

$$f(0, T_{i-1}, T_i) = \frac{1}{\delta_i}\left[\frac{Z(0, T_{i-1})}{Z(0, T_i)} - 1\right]$$

δ_i is the day count fraction associated with the coupon period $[T_{i-1}, T_i)$ and by convention these are a regular interval like quarterly, semi-annually or annually.

Now, each coupon of the floating rate note is a mini-derivative whose fair value at time 0 is known. This is just the discounted value of the future implied cash flows, and so the price at time t_0 of all the floating coupons is just:

$$\sum_{i=1}^{n} \frac{1}{\delta_i}\left[\frac{Z(0, T_{i-1})}{Z(0, T_i)} - 1\right] Z(0, T_i)\,\delta_i$$

The floating coupon note also pays a redemption of the notional at maturity T_n and so the price of the floating rate note is:

$$\begin{aligned}&\sum_{i=1}^{n} \frac{1}{\delta_i}\left[\frac{Z(0, T_{i-1})}{Z(0, T_i)} - 1\right] Z(0, T_i)\,\delta_i + Z(0, T_n)\\ &= Z(0, T_0)\end{aligned}$$

So the price of the note is almost at par; and this is a general result that is valid at all times for a floating rate note if you only count the future coupons which have not yet started their life cycle.

Overnight Floating Index

Consider coupon periods $(T_{i-1}, T_i]$ where the interest rate that was applied was a deposit rate which is measured on date T_{i-1}.

Before 2020 this was the standard way that **floating rate legs** operated; (T_{i-1}, T_i) was typically either 1-month, 3 -month, 6 -moth or 12 -month. The associated interest rates for these periods, which were measured at the start of the period (T_{i-1}) were called **libor rates**. Since 2020 the standard market treatment for a floating leg has changed, and the floating interest rate for the coupon period is a now a one-day interest rate that is daily compounded between $(T_{i-1}, T_i]$. They are called **overnight rates**, e.g. SOFR1D and ESTR1D are examples for the US and EU economies.

Consider again the coupon period $[T_{i-1}, T_i]$, but now for the case of a floating coupon note which pays overnight interest. We will label each day $d_k k = 0, 1 \cdots m$ within the coupon period, so that d_0 corresponds to T_{i-1}, while d_m corresponds to T_i. The way it works is that on each day d_k in the period $(T_{i-1}, T_i]$ the interest rate compounds by an additional day, at the prevailing overnight interest rate that is observed on day d_k and applicable for day d_k to day d_{k+1}. This is essentially a 1-day deposit rate for $[d_k, d_{k+1}]$:

$$f(0, d_k, d_{k+1}) = \frac{1}{\delta_k}\left[\frac{Z(0, d_k)}{Z(0, d_{k+1})} - 1\right]$$

At time T_{i-1} the initial deposit is \$1, and this then compounds over the coupon period into the following value by T_i :

$$\begin{aligned}\prod_{k=0}^{m-1}\left(1 + \delta_k f(0, d_k, d_{k+1})\right) &= \prod_{k=0}^{m-1}\left(\frac{Z(0, d_k)}{Z(0, d_{k+1})}\right)\\ &= \frac{Z(0, d_0)}{Z(0, d_m)} = \frac{Z(0, T_{i-1})}{Z(0, T_i)}\end{aligned}$$

Overnight interest rates are rates of interest offered to borrowers on loans which extend for one day only; and even then there is frequently some collateral posting that is required to protect the interests of the lender. The risk to the lender is minimal and this is as close to risk free as we we can get. Libor by contrast is a longer term deposit rate (1-month, 3-month, 6-moth or 12-month) which is also unsecured. There is a much higher risk to the lender in this case. And by simple economic reasoning, lender are not prepared to extend loans on these term using the same level of interest as they would on overnight secured lending. **Libor interest rates were higher than overnight interest rates.**

Interest Rate Swaps

Definition 8.6. *An* ***interest rate swap*** *(IRS) is a financial derivative tool that allows two different market participants to exchange a series of cash flows based on different interest rates.*

This transaction's main purpose is to manage interest rate risk, reduce financing costs, or speculate on market interest rate changes.

In an interest rate swap, two parties agree to exchange a series of interest payments over a certain period of time, which are usually based on the same notional principal (**Notional Principal**), but with different interest structures. The most common form of interest rate swap is the **fixed-for-floating interest rate swap (vanilla)**, where one party pays a fixed rate and the other party pays a floating rate.

Example 19. *For example, Party A agrees to pay Party B a fixed rate of 5% per year, and Party B agrees to pay Party A a floating rate based on the three-month LIBOR. The notional principal amounts are the same (but the principal amounts are usually not actually exchanged), and cash flows are exchanged at regular intervals (such as every six months).*

Imagine a company that has taken out a loan for $100mio on which their lender is asking them to pay floating rate of interest. This is fairly typical even on mortgage debt. The company accepts the loan, but because of the uncertainty of what future floating rates of interest will be they might prefer to fix their interest rates so that they have cash flow certainty to help them with better financial planning.

The converse is also true; if an investor currently pays fixed rate of interest on a loan, but they think that interest rates will go down, then they may wish to switch there repayments from fixed interest to floating interest so that they can save on interest costs in the future.

A swap starts on date T_S and matures on date T_E. Between these dates are a set of regular dates to denote the fixed coupon periods: $\left[T_0^{fl}, T_1^{fl} \dots T_N^{fl}\right]$; and a second set of regular dates to denote the floating coupon periods: $\left[T_0^{fx}, T_1^{fx} \dots T_M^{fx}\right]$. The two sets of dates are not necessarily the same, although they would typically intersect on some common subset, and since both legs start and end on the same dates we also have $T_0^{fx} = T_0^{fl}$ and $T_M^{fx} = T_N^{fl}$.

The floating rate leg for a swap which starts on date T_S and ends on date T_E is expressed as a floating rate note:

$$\sum_{i=1}^{N} \frac{1}{\delta_i} \left[\frac{Z\left(0, T_{i-1}^{fl}\right)}{Z\left(0, T_i^{fl}\right)} - 1 \right] Z\left(0, T_i^{fl}\right) \delta_i + Z\left(0, T_N^{fl}\right) = Z\left(0, T_S\right)$$

The fixed leg, which pays a fixed coupon $Y\%$ is:

$$\sum_{i=1}^{M} YZ\left(0, T_i^{fx}\right) \bar{\delta}_i + Z\left(0, T_M^{fx}\right)$$

Putting it all together, a payer swap (where the investor pays the fixed coupon and receives the floating coupon) has value:

$$\begin{aligned} PV_0 &= \sum_{i=1}^{N} \frac{1}{\delta_i} \left[\frac{Z\left(0, T_{i-1}^{fl}\right)}{Z\left(0, T_i^{fl}\right)} - 1 \right] Z\left(0, T_i^{fl}\right) \delta_i - \sum_{i=1}^{M} YZ\left(0, T_i^{fx}\right) \bar{\delta}_i \\ &= \left[Z\left(0, T_S\right) - Z\left(0, T_E\right)\right] - \sum_{i=1}^{M} YZ\left(0, T_i^{fx}\right) \bar{\delta}_i \end{aligned}$$

A **par swap rate** is the fixed rate Y that makes the value of the swap zero at time 0. This is the rate that makes the swap fair to both parties.

$$Y(0, T_S, T_E) = \frac{Z(0, T_S) - Z(0, T_E)}{\sum_{i=1}^{M} Z(0, T_i^{fx}) \bar{\delta}_i}$$

8.6 Zero Curve Bootstrapping

Let $z_i = Z(0, T_i)$ be the price of a zero coupon bond at time 0 with maturity T_i. Let forward rates are $(\hat{y}_1, \cdots, \hat{y}_N)$, where $\hat{y}_i$ is the forward rate for the period $[T_{i-1}, T_i]$. The zero coupon bond price is (annual compounding):

$$
\begin{aligned}
z_k &= (1+\hat{y}_k)^{-(T_k - T_{k-1})} \prod_{i=1}^{k-1} (1+\hat{y}_i)^{-(T_i - T_{i-1})} \\
z_k &= (1+\hat{y}_k)^{-(T_k - T_{k-1})} z_{k-1} \\
\hat{y}_k &= \left(\frac{z_{k-1}}{z_k}\right)^{(T_k - T_{k-1})} - 1
\end{aligned}
$$

and this would extract a *piece-wise constant function* of the forward interest rate yield with time.

The algorithm is mathematically functional but in practice zero coupon bonds are not traded frequently enough for us to have a sensible idea of what their prices are.

Regular fixed coupon bonds have far higher liquidity, and usually these are the instruments that we use to extract the yield curve. From a pure theoretical perspective it should not matter which one we use since both must generate the same yield curve(otherwise there is an arbitrage opportunity); but the practicalities mean that only the fixed coupon bonds are really observable.

Extracting forward zero rates (forward yields) from fixed coupon bond prices $P(0, T_i)$ for $i \in 1, 2 \cdots N$ uses the following logic where it is assumed that $(\hat{y}_1, \hat{y}_2 \cdots \hat{y}_k)$ have already been extracted from $(P(0, T_1), P(0, T_2) \cdots P(0, T_k))$. To simplify the explanation assume that the coupon payment dates for all the bonds overlap on some subset of a set of regular coupon payment dates: $d_1, d_2 \cdots d_m$. The set $(T_1, T_2 \cdots T_N)$ are a subset of this set as well. Relaxing this assumption is easy to implement but makes the equations below very messy.

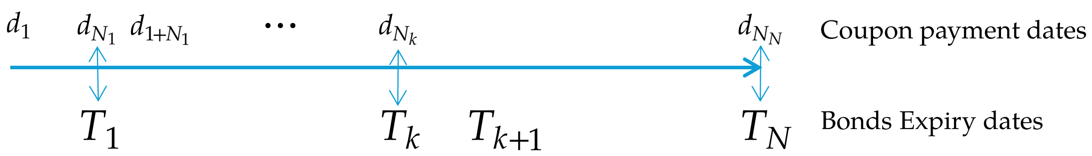

Figure 8.1: Payments dates presence

$$
\begin{aligned}
P(0, T_k) &= c_k \sum_{i=1}^{N_k} Z(0, d_i)\, \delta_i + Z(0, T_k) \\
c_k &= \frac{P(0, T_k) - Z(0, T_k)}{\sum_{i=1}^{N_k} Z(0, d_i)\, \delta_i}
\end{aligned}
$$

This can be used in the expression for the price of a T_{k+1} maturity fixed coupon bond:

$$
\begin{aligned}
P(0, T_{k+1}) &= c_{k+1} \sum_{i=1}^{N_{k+1}} Z(0, d_i)\, \delta_i + Z(0, T_{k+1}) \\
&= c_{k+1} \sum_{i=1}^{N_k} Z(0, d_i)\, \delta_i + c_{k+1} \sum_{i=1+N_k}^{N_{k+1}} Z(0, d_i)\, \delta_i + Z(0, T_{k+1}) \\
&= \frac{c_{k+1}}{c_k} \left[P(0, T_k) - Z(0, T_k)\right] + c_{k+1} \sum_{i=1+N_k}^{N_{k+1}} Z(0, d_i)\, \delta_i + Z(0, T_{k+1})
\end{aligned}
$$

The last two terms are the elements that contain the value $\hat{y}_{k+1}$ that is to be bootstrapped, via:

$$
\sum_{i=1+N_k}^{N_{k+1}} Z(0, d_i)\, \delta_i + Z(0, T_{k+1}) =
$$

$$
\sum_{i=1+N_k}^{N_{k+1}} Z(0, T_k)\, (1+\hat{y}_{k+1})^{d_i - T_k}\, \delta_i + Z(0, T_k)\, (1+\hat{y}_{k+1})^{T_{k+1} - T_k}
$$

This is a polynomial in $\hat{y}_{k+1}$ and can be solved for $\hat{y}_{k+1}$.

8.7 Introduction to Credit Risk

Credit Spreads

Definition 8.7. ***Survival Probability***: *Given an exponential distribution for the time to default, the probability that a corporate will not default by time T is called the survival probability and is denoted as $S(t,T)$.*

$$S(t,T) = \exp\left(-\int_t^T \lambda(u)du\right)$$

Let $B(0,T)$ be price of zero coupon bonds but now with credit risk. In general we will find that for any particular maturity T the yield on the risky bonds is higher than that on the risk free bond: $y_T < y'_T$. Clearly the risk of default gives rise to an enhanced yield.

$$\begin{aligned}
Z(0,T) &= \exp\left(-\int_0^T \hat{y}(u)du\right)\\
B(0,T) &= \exp\left(-\int_0^T \hat{y}'(u)du\right)\\
&= \exp\left(-\int_0^T [\hat{y}'(u) - \hat{y}(u)] + \hat{y}(u)du\right)\\
&= \exp\left(-\int_0^T [\hat{y}'(u) - \hat{y}(u)]\,du\right) Z(0,T)\\
&= \exp\left(-\int_0^T s(u)du\right) Z(0,T)
\end{aligned}$$

where we call $s(u) = \hat{y}'(u) - \hat{y}(u)$ the **credit spread**.

The spread over the risk free yield that investors will ask for the taking on risky debt. Notice however that the exponential term which invokes the credit spread has the same form as the **survival probability**. This means that the interpretation of **the risky zero coupon bond** $B(0,T)$ **is that of a risk free zero coupon bond multiplied by the survival probability of the corporate that issued bond** B**.** The credit spread $s(t)$ plays the role of the default intensity and quite critically it is implicit in this logic that upon default the investor loses everything. There is no recovery.

This is not a dynamic model. We have not proposed or derived any dynamics for the credit spread, or default events. All that we have done is extract information on the additional risk that is inherent in credit risky bonds and cast that to an exponential distribution for the probability of default. Like when we first postulated the money market account with the value of risk free interest rate as a parameter of the valuation framework, we now propose a credit spread which defined the survival probability of an issuing entity.

At this time we have not considered the possibility of hedging a default event; we have not yet even considered the possibility of hedging the risk of a change in derivative value because the credit spread might be a random process. Because of this the risk neutral framework is still valid. We have:

$$B(0,T) = \mathbb{E}\left[e^{-\int_0^T r(s)ds}\mathbb{1}_{\tau>T}\right]$$

where τ is the time of default, $\tau > T$ means that the default has not occurred by time T.

Credit Recovery

In general, when a default event occurs the investors do have some recourse in the form of a recovery fraction of the amount that they are owed. It is rare that an investor will lose everything. Here we assign the value $0 \leq R \leq 1$ as the fraction that an investor can recover in the event of default. This changes the mapping from credit spread to hazard rate.

A risk neutral valuation for the value of a risky zero coupon bond is now the expression:

$$\begin{aligned}
B(0,T) &= \mathbb{E}\left[e^{-\int_0^T r(s)ds}\left(\mathbb{1}_{\tau>T} + R\cdot\mathbb{1}_{r<T}\right)\right] \\
&= \mathbb{E}\left[e^{-\int_0^T r(s)ds}\left(\mathbb{1}_{\tau>T}[1-R] + R\right)\right] \\
&= \mathbb{E}\left[e^{-\int_0^T r(s)ds}\left(e^{-\int_0^T \lambda(s)ds}[1-R] + R\right)\right] \\
&\sim \mathbb{E}\left[e^{-\int_0^T r(s)ds}\left(\left(1-\int_0^T \lambda(s)ds\right)(1-R) + R\right)\right] \\
&\sim \mathbb{E}\left[e^{-\int_0^T r(s)ds}\left(1-\int_0^T (1-R)\lambda(s)ds\right)\right] \\
&\sim \mathbb{E}\left[\left(e^{-\int_0^T r(s)ds}\right)\left(e^{-\int_0^T (1-R)\lambda(s)ds}\right)\right]
\end{aligned}$$

So to first order the credit spread is related to the intensity process via:

$$s(t) = \lambda(t)(1-R)$$

Note that as $R \to 1$ then $s(t) \to 0$; because even in a default event the investor is able to recover all of their investment. Such a thing is unheard of in practice.

Credit Default Swaps

Definition 8.8. *A **credit default swap** (CDS) is a financial derivative that allows an investor to **hedge or speculate** on the risk of a corporate bond defaulting. The insurance buyer and the insurance seller to insure against losses in the event that a third party defaults.*

It is a swap with two legs: a premium leg and a protection leg. The premium leg is a stream of regular premium cash flows that the protection buyer makes to purchase the insurance policy (like monthly premium made on a car insurance policy); the premiums continue until the maturity of the swap or the default event, whichever comes first. The protection leg is a payout that is made to the insurance buyer in the event that a credit event occurs with the third party. If no credit event occurs then the protection buyer does not get anything for the premiums that they have paid; this is again like a standard car insurance scheme.

The CDS has a stated notional size to denote the size of the protection policy; premiums and default payouts are both determined by the notional. The protection leg is further defined in terms of the expected loss which is defined as $(1-R)$ for some value of R that is written into the contract. Typically $R = 40\%$.

Since the CDS is so tied to the default of the third party, it would come as no surprise that the survival curve $Q(t)$ plays an enormous role in the valuation of a CDS. We also assume the availability of an appropriate risk free discount curve $Z(0,s)\ \forall s \in [T_s, T_e]$ where T_s and T_m are the CDS start and maturity dates respectively.

The protection leg is simply the probability that a default event occurs at some point during the life of the swap; weighted by $(1-R)$. Since default can occur at some future date, the present value of this default event is represented as a present value using the risk free discount factor. This inherently assumes that while the third party is credit risky, the protection seller is not a credit risky institution.

We assume a set of regular dates which correspond to the premium payment dates: T_i where $i = 1, 2 \dots n$, and $T_0 = T_s$ while $T_n = T_e$. The protection leg is then:

$$\begin{aligned}
PV_{\text{prot}} &= -(1-R)\int_{T_s}^{T_e} Z(0,s)\,dQ_s \\
&= -(1-R)\sum_{i=1}^{n} Z(0,\bar{T}_i)\left[Q_{T_i} - Q_{T_{i-1}}\right]
\end{aligned}$$

where $\bar{T}_i = \frac{T_i + T_{i-1}}{2}$ approximates the default event that may occur in the interval to the middle of the interval.

The premium leg is a little more tricky. First consider any premium coupon P_i that is paid on date T_i and then use the linearity of expectation. This premium is paid in full if the default event has not yet occurred at time T_i (i.e. the event $\tau > T_i$), while it is paid in part if default occurs in the period $[T_{i-1}, T_i]$ (paid on a pro-rata basis). If default occurs prior to T_{i-1} then this premium payment never activates.

$$P_i = cQ_{T_i}Z(0,T_i)\cdot\delta_i - c\int_{T_{i-1}}^{T_i}\left(\frac{s-T_{i-1}}{T_i - T_{i-1}}\right)Z(0,s)\,dQ_s$$

where δ_i is the usual day count fraction for $[T_{i-1}, T_i]$. Now we know that the premium leg is the sum over all the coupons; $\sum_{i-1}^{n} P_i$.

$$PV_{\text{prem}} = \sum_{i=1}^{n} \left(cQ_{T_i} Z\left(0, T_i\right) \cdot \delta_i - c \int_{T_{i-1}}^{T_i} \left(\frac{s - T_{i-1}}{T_i - T_{i-1}} \right) Z\left(0, s\right) dQ_s \right)$$

A payer CDS is a swap where the investor pays for protection. The PV is $\left(PV_{\text{prot}} - PV_{\text{prem}}\right)$ and the par CDS spread is the value of c which gives this a valuation of zero.

9 Problem Set

9.1 Set 1

Problem 1. Consider the particular 2-period model with $S_0 = 4, u = 2$ and $d = \frac{1}{2}$. Let the interest rate $r = \log \frac{5}{4}$.

1. Write down the stock price state matrix for this model.
2. Consider a European Call option with strike price $K = 4$ that expires at time 2. What is the price of this option at time zero?
3. Consider a Look-back Call option that expires at time 2. The payoff of this option at time 2 is $\left(\max_{0\le k\le 2} S_k - S_2\right)^+$. What is the price of this option at time zero? Please derive the price via the risk-neutral pricing formula and the replicating portfolio method.

Problem 2. (Asian option) Consider a 3-period model, with parameters given in the previous question. For $n = 0, 1, 2, 3$, define $Y_n = \sum_{k=0}^{n} S_k$ to be the sum of the stock prices between times zero and n. Consider an Asian call option that expires at time three and has strike $K = 4$ i.e., whose payoff at time three is $\left(\frac{1}{4}Y_3 - 4\right)^+$. This is like a European call, except the payoff of the option is based on the average stock price rather than the final stock price. Let $v_n(s, y)$ denote the price of this option at time n if $S_n = s$ and $Y_n = y$. In particular, $v_3(s, y) = \left(\frac{1}{4}y - 4\right)^+$.

1. Develop an algorithm for computing v_n recursively. In particular, write a formula for v_n in terms of v_{n+1}.
2. Apply the algorithm developed in the previous question to compute $v_0(4, 4)$, the price of the Asian option at time zero.
3. Provide a formula for $\delta_n(s, y)$, the number of shares of stock that should be held by the replicating portfolio at time n if $S_n = s$ and $Y_n = y$.
4. Find the time-zero price and optimal exercise policy (optimal stopping time) for the path-dependent Asian-Like American Put whose intrinsic value at each time $n, n = 0, 1, 2, 3,$ is $\left(4 - \frac{1}{n+1}\sum_{j=0}^{n} S_j\right)^+$. This intrinsic value is a put on the average stock price between time zero and time n.

Problem 3. Consider an N-period binomial,discrete discounting model.

1. Let $M_0, M_1, \cdots, M_N$ and $M'_0 \cdot M'_1, \cdots, M'_N$ be martingales under the risk neutral measure $\widetilde{\mathbb{P}}$. That is,
$$p_u = \frac{1 + r - d}{u - d}, \quad p_d = \frac{u - 1 - r}{u - d}$$
Show that if $M_N = M'_N$ then, for each n between 0 and N, we have $M_n = M'_n$.
2. Let V_N be the payoff at time N of some derivative security. This is a random variable that can depend on all N coin tosses. Define recursively $V_{N-1}, V_{N-2}, \ldots, V_0$.
$$V_{i,j} = \frac{1}{1+r}\left(p_u V_{i,j+1} + (1 - p_u)V_{i-1,j+1}\right), \quad j = N-1, N-2, \cdots, 0, i \le j$$

Show that the discounting process of V_n:

$$V_0, \frac{V_1}{1+r} \cdots \frac{V_{N-1}}{(1+r)^{N-1}}, \frac{V_N}{(1+r)^N}$$

is a martingale under $\widetilde{\mathbb{P}}$.

Problem 4. (Dividend Paying) We define a random variable Y_n with probability p_u of being u and probability p_d of being d.

$$Y_{n+1} = \begin{cases} u, \text{if the stock goes up at time } n+1 \\ d, \text{if the stock goes down at time } n+1 \end{cases}$$

In the binomial model, $Y_{n+1}S_n$ was the stock price at time $n+1$. In a dividend-paying model considered here, we have another random variable A_{n+1}, taking values in $(0,1)$, and the dividend paid at time $n+1$ is $A_{n+1}Y_{n+1}S_n$. After the dividend is paid, the stock price at time $n+1$ is

$$S_{n+1} = (1 - A_{n+1})\, Y_{n+1} S_n$$

An agent who begins with initial portfolio X_0 and at each time n takes a position of Δ_n shares of stock, where Δ_n depends only on the first n coin tosses, has a portfolio value governed by the wealth equation

$$\begin{aligned} X_{n+1} &= \Delta_n S_{n+1} + (1+r)\,(X_n - \Delta_n S_n) + \Delta_n A_{n+1} Y_{n+1} S_n \\ &= \Delta_n Y_{n+1} S_n + (1+r)\,(X_n - \Delta_n S_n) \end{aligned}$$

1. Show that the discounted wealth process is a martingale under the risk neutral measure $\widetilde{\mathbb{P}}$.
2. Show that the risk-neutral pricing formula still applies.
3. Show that the discounted stock price is not a martingale under the risk neutral measure.However, if A_{n+1} is a constant $a \in (0,1)$, then $\dfrac{S_n}{(1-a)^n(1+r)^n}$ is a martingale under the risk-neutral measure.

Problem 5. (Chooser option)

1. Write down put-call parity for European options at time $m < N$, using discrete discounting.
2. Let $1 \le m \le N-1$ and $K > 0$ be given. A chooser option is a contract sold at time zero that confers on its owner the right to receive either a call or a put at time m. The owner of the chooser may wait until time m before choosing. The call or put chosen expires at time N with strike price K. Show that the time-zero price of a chooser option is the sum of the time-zero price of a put, expiring at time N and having strike price K, and a call, expiring at time m and having strike price $\dfrac{K}{(1+r)^{N-m}}$.

9.2 Set 2

Problem 6. Write down the answers without specifying the details of the calculations.

1. Derive the expression for $\mathbb{P}(X \le \nu)$ when $Y \sim \mathcal{LN}(\mu, \sigma^2)$, i.e., Y is log-normally distributed. From this derive the probability density function of Y.
2. Derive the SDE for:
   (a) $Y_t = t - W_t^2$
   (b) $Y_t = e^{W_t}$
   (c) $Y_t = e^{-rt}W_t$
   (d) $X_t = W_t - \beta \int_0^t e^{-\beta(t-s)} W_s ds$.
3. Let $x, \mu \in \mathbb{R}, \sigma > 0$. If $X_t = x + \mu t + \sigma W_t, t \ge 0$ and

$$Y_t = h \int_0^t X_s ds, \quad t \ge 0.$$

   Derive $\mathbb{E}Y_t$ as well as $\mathrm{Var}(Y_t)$.

Problem 7. Suppose $M(t), 0 \leq t \leq T$, is a martingale with respect to some filtration $\mathcal{F}(t), 0 \leq t \leq T$. Let $\Delta(t), 0 \leq t \leq T$, be a simple process adapted to $\mathcal{F}(t)$ i.e., there is a partition $\Pi = \{t_0, t_1, \ldots, t_n\}$ of $[0, T]$ such that, for every $j, \Delta\left(t_j\right)$ is $\mathcal{F}\left(t_j\right)$-measurable and $\Delta(t)$ is constant in t on each sub-interval $\left[t_j, t_{j+1}\right)$. For $t \in [t_k, t_{k+1})$, define the stochastic integral

$$I(t) = \sum_{j=0}^{k-1} \Delta\left(t_j\right)\left[M\left(t_{j+1}\right) - M\left(t_j\right)\right] + \Delta\left(t_k\right)\left[M(t) - M\left(t_k\right)\right]$$

We think of $M(t)$ as the price of an asset at time t and $\Delta\left(t_j\right)$ as the number of shares of the asset held by an investor between times t_j and t_{j+1}. Then $I(t)$ is the capital gains that accrue to the investor between times 0 and t. Show that $I(t), 0 \leq t \leq T$, is a martingale.

Problem 8. Verify that the European call price $C_t = S_t \Phi\left(d_1\right) - K e^{-r(T-t)} \Phi\left(d_2\right)$ satisfies the Black-Scholes PDE.

Problem 9. Now consider a date t_0 between 0 and T, and consider a chooser option, which gives the right at time t_0 to choose to own either the call or the put.

1. Show that at time t_0 the value of the chooser option is

$$C\left(t_0\right) + \max\left\{0, -F\left(t_0\right)\right\} = C\left(t_0\right) + \left(e^{-r(T-t_0)} K - S\left(t_0\right)\right)^+ .$$

2. Show that the value of the chooser option at time 0 is the sum of the value of a call expiring at time T with strike price K and the value of a put expiring at time t_0 with strike price $e^{-r(T-t_0)} K$.

Problem 10. (Cost of carry) Consider a commodity whose unit price at time t is $S(t)$. Ownership of a unit of this commodity requires payment at a rate a per unit time (cost of carry) for storage. Note that this payment is per unit of commodity, not a fraction of the price of the commodity. Thus, the value of a portfolio that holds $\Delta(t)$ units of the commodity at time t and also invests in a money market account with constant rate of interest r has dynamics

$$dX(t) = \Delta(t) dS(t) - a\Delta(t) dt + r(X(t) - \Delta(t) S(t)) dt.$$

We must choose the risk neutral measure so that the discounted portfolio value $e^{-rt} X(t)$ is a martingale. The dynamics of S_t is given by

$$dS(t) = rS(t) dt + \sigma S(t) d\widetilde{W}(t) + a dt$$

where $\widetilde{W}(t)$ is a Brownian motion under the risk-neutral measure $\widetilde{\mathbb{P}}$.

1. Show that the discounted portfolio value process $e^{-rt} X(t)$ is a martingale under $\widetilde{\mathbb{P}}$.

2. Define

$$Y(t) = \exp\left\{\sigma \widetilde{W}(t) + \left(r - \frac{1}{2}\sigma^2\right) t\right\}$$

Verify that, for $0 \leq t \leq T$,

$$dY(t) = rY(t) dt + \sigma Y(t) d\widetilde{W}(t)$$

that $e^{-rt} Y(t)$ is a martingale under $\widetilde{\mathbb{P}}$, and that

$$S(t) = S(0) Y(t) + Y(t) \int_0^t \frac{a}{Y(s)} ds$$

satisfies the dynamics of $S(t)$.

3. For $0 \leq t \leq T$, derive a formula for $\widetilde{\mathbb{E}}[S(T) \mid \mathcal{F}(t)]$ in terms of $S(t)$ by writing

$$\begin{aligned} \widetilde{\mathbb{E}}[S(T) \mid \mathcal{F}(t)] = & S(0) \widetilde{\mathbb{E}}[Y(T) \mid \mathcal{F}(t)] + \widetilde{\mathbb{E}}[Y(T) \mid \mathcal{F}(t)] \int_0^t \frac{a}{Y(s)} ds \\ & + a \int_t^T \widetilde{\mathbb{E}}\left[\left.\frac{Y(T)}{Y(s)}\right| \mathcal{F}(t)\right] ds \end{aligned}$$

and then simplifying the right-hand side of this equation.

4. The process $\mathbb{E}[S(T) \mid \mathcal{F}(t)]$ is the futures price process for the commodity i.e., $\left(\mathrm{Fut}_S(t, T) = \widetilde{\mathbb{E}}[S(T) \mid \mathcal{F}(t)]\right)$. This must be a martingale under $\widetilde{\mathbb{P}}$. To check the formula you obtained in 3, differentiate it and verify that $\widetilde{\mathbb{E}}[S(T) \mid \mathcal{F}(t)]$ is a martingale under $\widetilde{\mathbb{P}}$.

5. Let $0 \le t \le T$ be given. Consider a forward contract entered at time t to purchase one unit of the commodity at time T for price K paid at time T. The value of this contract at time t when it is entered is

$$\widetilde{\mathbb{E}}\left[e^{-r(T-t)}(S(T)-K) \mid \mathcal{F}(t)\right]$$

The forward price $\mathrm{For}_S(t,T)$ is the value of K that makes the contract value equal to zero. Show that $\mathrm{For}_S(t,T) = \mathrm{Fut}_S(t,T)$.

Problem 11. Consider the stochastic differential equation

$$dX(u) = (a(u) + b(u)X(u))du + (\gamma(u) + \sigma(u)X(u))dW(u),$$

where $W(u)$ is a Brownian motion relative to a filtration $\mathcal{F}(u), u \ge 0$, and we allow $a(u), b(u), \gamma(u)$, and $\sigma(u)$ to be processes adapted to this filtration. Fix an initial time $t \ge 0$ and an initial position $x \in \mathbb{R}$. Define

$$Z(u) = \exp\left\{\int_t^u \sigma(v)dW(v) + \int_t^u \left(b(v) - \frac{1}{2}\sigma^2(v)\right)dv\right\},$$
$$Y(u) = x + \int_t^u \frac{a(v) - \sigma(v)\gamma(v)}{Z(v)}dv + \int_t^u \frac{\gamma(v)}{Z(v)}dW(v)$$

1. Show that $Z(t) = 1$ and

$$dZ(u) = b(u)Z(u)du + \sigma(u)Z(u)dW(u), u \ge t.$$

2. By its very definition, $Y(u)$ satisfies $Y(t) = x$ and

$$dY(u) = \frac{a(u) - \sigma(u)\gamma(u)}{Z(u)}du + \frac{\gamma(u)}{Z(u)}dW(u), u \ge t.$$

Show that $X(u) = Y(u)Z(u)$ solves the stochastic differential equation and satisfies the initial condition $X(t) = x$.

Problem 12. (Zero-strike Asian call). Consider a zero-strike Asian call whose payoff at time T is

$$V(T) = \frac{1}{T}\int_0^T S(u)du$$

1. Suppose at time t we have $S(t) = x \ge 0$ and $\int_0^t S(u)du = y \ge 0$. Use the fact that $e^{-ru}S(u)$ is a martingale under $\widetilde{\mathbb{P}}$ to compute

$$e^{-r(T-t)}\widetilde{\mathbb{E}}\left[\frac{1}{T}\int_0^T S(u)du \middle| \mathcal{F}(t)\right].$$

Call your answer $v(t,x,y)$.

2. Verify that the function $v(t,x,y)$ you obtained satisfies the Black-Scholes PDE equation

$$v_t(t,x,y) + rxv_x(t,x,y) + xv_y(t,x,y) + \frac{1}{2}\sigma^2x^2v_{xx}(t,x,y) = rv(t,x,y)$$
$$0 \le t < T, x \ge 0, y \in \mathbb{R}$$

and the boundary conditions

$$v(t,0,y) = e^{-r(T-t)}\left(\frac{y}{T} - K\right)^+, 0 \le t < T, y \in \mathbb{R}$$
$$v(T,x,y) = \left(\frac{y}{T} - K\right)^+, x \ge 0, y \in \mathbb{R}$$

3. Determine explicitly the process $\Delta(t) = v_x(t,S(t),Y(t))$, and observe that it is not random.

4. Use the Itô formula to show that if you begin with initial capital $X(0) = v(0,S(0),0)$ and at each time you hold $\Delta(t)$ shares of the underlying asset, investing or borrowing at the interest rate r in order to do this, then at time T the value of your portfolio will be

$$X(T) = \frac{1}{T}\int_0^T S(u)du$$

Problem 13. (Portfolios under change of numéraire). Consider two assets with prices $S(t)$ and $N(t)$ given by

$$S(t) = S(0)\exp\left\{\sigma\widetilde{W}(t) + \left(r - \frac{1}{2}\sigma^2\right)t\right\}$$
$$N(t) = N(0)\exp\left\{\nu\widetilde{W}(t) + \left(r - \frac{1}{2}\nu^2\right)t\right\}$$

where $\widetilde{W}(t)$ is a one-dimensional Brownian motion under the risk-neutral measure $\widetilde{\mathbb{P}}$ and the volatilities $\sigma > 0$ and $\nu > 0$ are constant, as is the interest rate r. We define a third asset, the money market account, whose price per share at time t is $M(t) = e^{rt}$.

Let us now denominate prices in terms of the numéraire N, so that the re-denominated first asset price is

$$\hat{S}(t) = \frac{S(t)}{N(t)}$$

and the re-denominated money market account price is

$$\widehat{M}(t) = \frac{M(t)}{N(t)}$$

1. Show $d\hat{S}(t) = (\sigma - \nu)\hat{S}(t)d\widehat{W}(t)$, where $\widehat{W}(t) = \widetilde{W}(t) - \nu t$.
2. Compute the differential of $\dfrac{1}{N(t)}$.
3. Compute the differential of $\widehat{M}(t)$, expressing it in terms of $d\widehat{W}(t)$.

Consider a portfolio that at each time t holds $\Delta(t)$ shares of the first asset and finances this by investing in or borrowing from the money market. According to the usual formula, the differential of the value $X(t)$ of this portfolio is

$$dX(t) = \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt.$$

We define

$$\Gamma(t) = \frac{X(t) - \Delta(t)S(t)}{M(t)}$$

to be the number of shares of money market account held by this portfolio at time t and can then rewrite the differential of $X(t)$ as

$$dX(t) = \Delta(t)dS(t) + \Gamma(t)dM(t)$$

Note also that by the definition of $\Gamma(t)$, we have

$$X(t) = \Delta(t)S(t) + \Gamma(t)M(t)$$

We re-denominate the portfolio value, defining

$$\widehat{X}(t) = \frac{X(t)}{N(t)}$$

so that we have

$$\widehat{X}(t) = \Delta(t)\hat{S}(t) + \Gamma(t)\widehat{M}(t)$$

Use stochastic calculus to show that

$$d\widehat{X}(t) = \Delta(t)d\hat{S}(t) + \Gamma(t)d\widehat{M}(t)$$

which implies that the portfolio is self-financing under the numéraire N.

Problem 14. Consider the SDE

$$dS_t = S_t\left(\mu dt + \sigma dW_t^{\mathbb{P}}\right)$$

where $W_t^{\mathbb{P}}$ is a Brownian motion under $\mathbb{P}$. Find a probability measure $\mathbb{Q}$ such that the inverse process S^{-1} is a martingale.

Problem 15. Let $S(t)$ and $N(t)$ be the prices of two assets, denominated in a common currency, and let σ and ν denote their volatilities, which we assume are constant. We assume also that the interest rate r is constant. Then

$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}_1(t)$$
$$dN(t) = rN(t)dt + \nu N(t)d\widetilde{W}_3(t)$$

where $\widetilde{W}_1(t)$ and $\widetilde{W}_3(t)$ are Brownian motions under the risk-neutral measure $\widetilde{\mathbb{P}}$. We assume these Brownian motions are correlated, with $d\widetilde{W}_1(t)d\widetilde{W}_3(t) = \rho dt$ for some constant ρ.

1. Show that $S^{(N)}(t) = \dfrac{S(t)}{N(t)}$ has volatility $\gamma = \sqrt{\sigma^2 - 2\rho\sigma\nu + \nu^2}$. In other words, show that there exists a Brownian motion $\widetilde{W}_4$ under $\widetilde{\mathbb{P}}$ such that
$$\frac{dS^{(N)}(t)}{S^{(N)}(t)} = (\text{ Something })dt + \gamma d\widetilde{W}_4(t)$$

2. Show how to construct a Brownian motion $\widetilde{W}_2(t)$ under $\widetilde{\mathbb{P}}$ that is independent of $\widetilde{W}_1(t)$ such that $dN(t)$ may be written as
$$dN(t) = rN(t)dt + \nu N(t)\left[\rho d\widetilde{W}_1(t) + \sqrt{1-\rho^2}d\widetilde{W}_2(t)\right]$$

3. Determine the volatility vector of $S^{(N)}(t)$. In other words, find a vector (v_1, v_2) such that
$$dS^{(N)}(t) = S^{(N)}(t)\left[v_1 d\widetilde{W}_1^{(N)}(t) + v_2 d\widetilde{W}_2^{(N)}(t)\right]$$
and show that
$$v_1^2 + v_2^2 = \sigma^2 + \nu^2 - 2\rho\sigma\nu$$

Problem 16. Assume no dividends. Let $C(S,t)$ and $P(S,t)$ denote the time- t price of an American call and American put options respectively with the same maturity date T and strike price K. Show that

$$C(S,t) - P(S,t) \geq S - K$$

Hint: Suppose the equality does not hold and construct a strategy that leads to arbitrage.

9.3 Set 3

Problem 17. In this problem, you will code up a PDE solver for pricing simple equity derivatives using BS model. Assume that $S_0 = 100, \sigma = 15\%, r = 5\%, T = 1y$. The Black-Scholes PDE for a European call option is given by

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

1. Rewrite the PDE using reverse time $\tau = T - t$. Then write down the initial conditions that would be applicable in the case of a European Call and European Put.

2. Encapsulate the initial condition inside a function `payoff(type,S,K,...)` where the function argument `type` is a string that can take values `'c'` or `'p'`. The function should return the payoff of the option at maturity. By encapsulating the payoff your code will become more versatile and can be easily extended to other types of options. You can check your function by plotting the payoffs using `matplotlib`.

3. Use the Dirichlet boundary conditions for this problem. Recall that this sets the value of V at the 2 boundary points. In this exercise we will configure the boundaries to be some lower and upper points $S_{\min} = 20$ and $S_{\max} = 180$; though in practice these should be dynamically set depending on the parameters of the problem. Write down the mathematical expression for the Dirichlet boundary conditions in the case of a European Call and Put to reflect the approximation that when the option is either deep in or out of the money then its value is just the intrinsic value.
$$\lim_{S\to 0} V(S,\tau) = (S-K)^+ = 0, \quad \lim_{S\to\infty} V(S,\tau) = (S-K)^+ = \infty$$
Here, because we have set the boundaries to be $S_{\min}$ and $S_{\max}$, the conditions reduce to:
$$V(S_{\min},\tau) = (S_{\min} - K)^+, \quad V(S_{\max},\tau) = (S_{\max} - K)^+$$

4. In chapter 4 we derived the explicit finite difference approximations for the Black Scholes PDE. Use this to discretize the PDE on a grid $\mathbf{V}[i,j]$ where i represents a discretization node in the τ direction and j the discretization in the S direction. The grid spacings $\Delta S = (S_{i+1} - S_i)$ and $\Delta\tau = (\tau_{i+1} - \tau_i)$ are $\frac{S_u - S_l}{N_S}$ and $\frac{T}{N_T}$ respectively. Like $S_{\min}$ and $S_{\max}$, N_S and N_T are configuration parameters; they correspond to the number of intervals that the grid is discretized into in each dimension. Show that the explicit finite difference numerical scheme that can be used to sequentially solve for V$[i,j]$ as i is incremented from 1 to N_T can be written algebraically as:
$$\mathbf{V}_{i+1}^{\top} = \mathbf{M}\cdot\mathbf{V}_i^{\top} + \mathbf{P} \quad \forall i = 0, 1\ldots(N_T - 1)$$

where $\mathbf{V}_i$ is the row vector of length (N_S+1) that picks row i from $\mathbf{V}[i,j]$, $\mathbf{P}$ is a length (N_S+1) vector which holds the boundary conditions at the first and last index. The other entries to this vector are zero. And $\mathbf{M}$ is an $(N_S+1)\times(N_S+1)$ (almost) tridiagonal matrix of the form:

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ a_1 & b_1 & c_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & b_2 & c_2 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & b_3 & c_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_N & b_N & c_N \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned} N &= (N_S - 1) \\ a_i &= k\sigma^2 S_i^2 - rkS_i\Delta S \\ b_i &= 1 - 2k\sigma^2 S_i^2 - r\Delta T \\ c_i &= k\sigma^2 S_i^2 + rkS_i\Delta S \\ k &= \frac{\Delta T}{2\Delta S^2} \end{aligned}$$

5. Implement the PDE solver to price options under the configuration $N_S = N_T = 20$, and compare the prices you see for at the money call/put options for a few different values of S_0 against an online Black-Scholes pricer.

6. Increase the time to expiry of the call option from 1 Y to 5 Y , keeping everything else unchanged and price options using the PDE solver. Explain the observation and suggest how the numerical scheme could be updated. The Growth factor G for this problem is of the form $G = C\left(\frac{\Delta t}{\Delta S^2}\right)$ where C is some constant.

7. Now switch the payoff to American style by adding a condition on each node $[i,j]$ to compare the price of the from the algorithm against the immediate exercise value (called intrinsic value). Modify the algorithm to incorporate the optimal exercise decision. Do you see the same result for call options as in the case of the Binomial Model?

Problem 18. Write a Monte Carlo simulation to calculate the area of a circle in the top right hand quadrant. Check the convergence of the estimate against the true value for an increasing number of sample paths:

$$[100, 500, 1000, 5000, 10000, 50000, 1000000]$$

draw a plot of the error against the number of sample paths.

Problem 19. In this question you will write your own Python function to return an array of N independent samples from a bimodal distribution $Z \sim 0.3\mathcal{N}(1, 0.8) + 0.7\mathcal{N}(4, 0.5)$ using the acceptor-rejector algorithm. Restrict the range of the distribution of Z to $[-5, 5]$ so that your algorithm will never return a value outside of this range; this is a reasonable approximation even though $Z \in \mathbb{R}$ because $P(|Z| > 5) \sim 0$. Your function should apply an iterative algorithm based on the outline below.

1. Choose $\mathcal{N}(0,1)$ as the proposal distribution, choose proper M.

2. For iteration i create a tuple (a_i, b_i) where $a_i \sim N(0,1)$ and $b_i \sim U(0,1)$.

3. For iteration i compute $y_i = \phi(a_i)$ where $\phi(x)$ is the probability density function of $\mathcal{N}(0,1)$. Compute $M\phi(a_i)$. Also, compute $z_i = f(a_i)$ where $f(x)$ is the probability density function of the bimodal distribution.

4. Then apply the acceptor-rejector criteria on each (b_i, z_i); If (b_i, z_i) passes the criteria

$$b_i \leq \frac{z_i}{M\phi(a_i)}$$

then add the relevant value to the results array; otherwise reject this iteration and move on.

5. Once the results array has N entries then terminate the algorithm and return the array.

6. Check the results of your array visually by plotting a histogram and superimposing the true probability density function of Z.

Problem 20. In this coding exercise, you will price Black-Scholes European Call and Put options in a Monte Carlo framework. Recall first,

$$S_T = S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z}, \quad Z \sim \mathcal{N}(0,1)$$

under risk neutral valuation, the price of a European Call option is given by

$$C_0 = e^{-rT}\mathbb{E}[(S_T - K)^+]$$

hence you can simulate the price by generating N samples of S_T and taking the average of the discounted payoffs.

1. Write a function to evaluate the expectation above, using a Monte-Carlo simulation using N sample paths. The function should also take as input the parameters: `option_type,S_0,r,sigma,T`.

2. Investigate the convergence of the price to the closed form price for increasing $N : [2^{10}, 2^{11}, \cdots, 2^{17}]$. Use $S_0 = 100, K = 100, r = 0.05, \sigma = 0.2, T = 1$ and tabulate the absolute relative error between the Monte Carlo price and the closed form price.

3. Look at the convergence of a put option with strike 50. One should notice a significant deterioration in the convergence rate (A tremendously bad case). Explain why this is the case.

4. Evaluate the expectation using **Quasi-Monte Carlo** methods. Compare the convergence of the Quasi-Monte Carlo method against the standard Monte Carlo method. Recall that Quasi random variables are random numbers that evenly cover the sample space according to the probability distribution on the sample space. As such, you need to generate Zs which evenly cover the sample range of the normal distribution $(-\infty, \infty)$. Do this by generating QRNs from the $\mathcal{U}(0,1)$ and invert them into QRNs from the $\mathcal{N}(0,1)$ using the inverse CDF of the normal distribution.

```python
from scipy.stats import norm
def inverse_normal_cdf(u):
    return norm.ppf(u)
```

5. Asian options award a payout that is contingent on some average value of the underlying asset over the life of the option. Here consider the geometric average Asian option. This sub question will work with it averaging of the stock at N equally spaced discrete time points over the entire life $[0, T]$.

$$S_{av} = \left(\prod_{i=1}^{N} S_{i\cdot\delta t}\right)^{\frac{1}{N+1}}, \quad \delta t = \frac{T}{N}$$

payout is given by $(S_{av} - K)^+$. The task in this sub question is to implement code to price Asian call, under a lognormal model for S_t, using a Monte Carlo simulation. To update stock price at each time step, use the **Euler-Maruyama** scheme.

$$S_{t+\delta t} = S_t\left(1 + r\delta t + \sigma\sqrt{\delta t}Z\right), \quad Z \sim \mathcal{N}(0,1)$$

```python
def geom_asian_option(optType, S_0, strike, expiry, r, sigma, npaths, nsteps):
    delta_t = expiry / nsteps
    N_rands = npaths * nsteps

    z_s = iter(np.random.normal(0, 1, N_rands)) # iter function to create an iterator
    V_T = 0
    for _ in range(npaths):
        S_t = S_0
        S_av = S_0
        for _ in range(nsteps):
            z = next(z_s)
            """
            Your codes here, update S_t and S_av
            """
        S_T = S_t # Final stock price
        S_av = S_av ** (1/(nsteps+1)) # Average stock price
        V_T += max( S_av - strike, 0.0) if optType==CALL else max( strike - S_av, 0.0)
        # print(f"{S_T}, {S_av}")
```

```python
        """
        Your codes here, calculate the price of the option
        """
        return (V_0)
```

Problem 21. Below are the forward option premiums for forward at the money EUR-USD fx options at different expiry. EUR-USD spot rate is currently at 1.08(€1 = \$1.08), you can assume that USD and EUR interest rates are deterministic and constant at 5% and 3% respectively (both continuous compounded). Note that the premium below are forward premium, in each case they are forwarded to the expiry date. The strike for each option is the forward fx rate (this is what is meant by forward at the money).

T (years)	Premium (cents)
1	5.2716
2	7.0007
3	8.0169
5	8.9944
7	9.7660
10	10.7318
12	10.9863

1. Compute the strikes associated with each option listed in the table above. Remember the strike is just the forward FX rate.

$$K = x_0 e^{(r_d - r_f)T}, \quad x_0 = 1.08, \quad \text{Here, domesitic rate is 0.05 (American), T is the time to expiry}$$

```python
r_usd = 0.05
r_eur = 0.03
x_0 = 1.08
expiries = [1.0, 2.0, 3.0, 5.0, 7.0, 10.0, 12.0]
fwd_prem_cents = [5.2716, 7.0007, 8.0169, 8.9944, 9.7660, 10.7318, 10.9863]

# Cents need to be converted to dollars!!
fwd_premium = [ fpc/100 for fpc in fwd_prem_cents]
"""
Your codes here, calculate the strikes: strikes=[...]
"""
```

2. Derive mathematical expression for the forward price of call options on forward EUR-USD under a lognormal model for forward FX, and constant interest rates. This is otherwise known as the forward Black formula. Implement this formula into code as a function `black formula()`.

$$d_1 = \frac{\log\left(\frac{F}{K}\right) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}$$

and

$$C = F\Phi(d_1) - K\Phi(d_2)$$

```python
def black_formula(f, k, expiry, vol):
    """
    Your codes here, gain the d1 and d2 values
    """
    return f*norm.cdf(d1) - k*norm.cdf(d2)
```

3. Write a central finite difference scheme to compute the vega of an at the money option. Implement this into code as function `black_vega()`.

```python
def black_vega(f, k, expiry, vol):
    epsilon = 1e-4
    """
    Your codes here, gain the vega value by central finite difference
    """
    return vega
```

4. Use this to write a numerical algorithm to extract the implied volatility from the forward option premiums. You can use any root finding algorithm you wish, but the solutions will use **Newton-Raphson**.

```
def implied_vol(prem, f, k, expiry, init_guess):
    max_iter = 10
    ivol = init_guess # initial guess of implied volatility
    error = black_formula(f, k, expiry, ivol) - prem  # ivol is the root of
     ↪ this function
    vega = black_vega(f, k, expiry, ivol)

    count = 0
    while ( abs(error) > 1e-6 and count < max_iter ):
        """
        Your codes here, Newton-Raphson update: ivol = ivol - error/vega,
         ↪ error = ..., vega = ...
        """

    if count == max_iter:
        raise ValueError('Exceeded max iterations')
    else:
        return ivol
```

5. (**Volatility term structure**) Write an algorithm to **bootstrap** the forward volatility from the spot starting implied volatility, i.e. extract the $0-1y$ volatility, the 1y $-$ 2y volatility, the 2y $-$ 3y volatility... 10y $-$ 12y volatility. You should notice something interesting with one of the forward volatility outputs. Use this to propose a sensible trade that could make money. Recall that

$$\sigma(t) = \sqrt{\frac{\sigma_I^2(T_k)T_k - \sigma_I^2(T_{k-1})(T_{k-1})}{T_k - T_{k-1}}}$$

where $\sigma_I^2 = \frac{1}{T}\int_0^T \sigma^2(t)dt$ measures the average volatility over the life of the option, and thus the implied volatility at the start of the option. $\sigma(t)$ is a piecewise constant function that represents the forward volatility, measuring the term structure of volatility.

```
trades = list(zip(expiries, strikes, fwd_premium))
print(trades)


ivols = [ implied_vol(fp, k, k, t, 0.1 ) for t, k, fp in list(trades) ]
print(f'Implied vols are : {ivols}')

outright_vols =  list( zip(expiries, ivols) )

prev_t, prev_ivol = outright_vols[0]
fwd_vols = []
for i in range(1, len(outright_vols)):
    prev_t, prev_ivol = outright_vols[i-1] # previous implied vol
    curr_t, curr_ivol = outright_vols[i] # current implied vol
    try:
        """
        Your codes here, calculate the forward volatilities: fwd_vol = ...
        """
    except:
        fwd_vol = None
    fwd_vols.append(fwd_vol)
print(f'Implied fwd vols are : {fwd_vols}')
```

One can find that the last implied `fwd_vol` (for period 10y-12y) is returned as "None" because the implied variance for this period is negative. It suggests that the 12y option is too cheap and should increase in value relative to the 10y option. The trade to do is sell 10y option, buy 12y option in delta neutral amounts.

A Expectation, Information and Conditioning

A.1 Expectation

Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We would like to compute an "average value" of X, where we take the probabilities into account when doing the averaging. If Ω is finite, we simply define this average value by

$$\mathbb{E}X = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega)$$

If Ω is **countably infinite**, its elements can be listed in a sequence $\omega_1, \omega_2, \omega_3, \cdots$, and we can define $\mathbb{E}X$ as an infinite sum:

$$\mathbb{E}X = \sum_{k=1}^{\infty} X(\omega_k)\,\mathbb{P}(\omega_k)$$

However, if Ω is **uncountably infinite**, we cannot list its elements in a sequence. Uncountable sums cannot be defined.

Riemann Integral Revisited

If $f(x)$ is a continuous function defined for all x in the closed interval $[a, b]$, we define the Riemann integral $\int_a^b f(x)dx$ as follows. First partition $[a, b]$ into subintervals $[x_0, x_1], [x_1, x_2], \cdots, [x_{n-1}, x_n]$, where $a = x_0 < x_1 < \cdots < x_n = b$. We denote by $\Pi = \{x_0, x_1, \cdots, x_n\}$ the set of partition points and by

$$\|\Pi\| = \max_{1 \le k \le n} (x_k - x_{k-1})$$

the **length** of the longest subinterval in the partition. For each subinterval $[x_{k-1}, x_k]$, we set $M_k = \max_{x_{k-1} \le x \le x_k} f(x)$ and $m_k = \min_{x_{k-1} \le x \le x_k} f(x)$. The upper Riemann sum is

$$\text{RiemannSum}_{\Pi}^{+}(f) = \sum_{k=1}^{n} M_k (x_k - x_{k-1}),$$

and the lower Riemann sum is

$$\text{RiemannSum}_{\Pi}^{-}(f) = \sum_{k=1}^{n} m_k (x_k - x_{k-1})$$

Definition A.1. *As $\|\Pi\|$ converges to zero, the upper Riemann sum and the lower Riemann sum converge to the same limit, which we call $\int_a^b f(x)dx$. This is the Riemann integral.*

Lebesgue Integral

Recap that a random variable is a function $X(\omega) : \Omega \to \mathbb{R}$, where Ω is the sample space. The problem we have with imitating this procedure to define expectation is that the random variable X, unlike the function f in the previous paragraph which is a function of real numbers, is a function of $\omega \in \Omega$, and Ω is often not a subset of $\mathbb{R}$. In Figure, the " x-axis" is not the real numbers but some abstract space Ω. There is no natural way to partition the set Ω as we partitioned $[a, b]$ above. Therefore, we partition instead the y-axis.

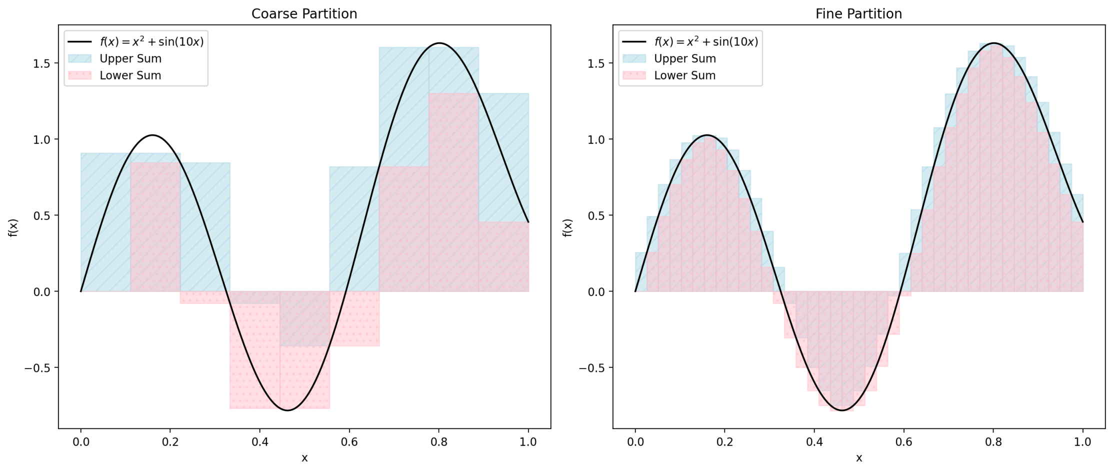

Figure A.1: Riemann Integral

Assume for the moment that $0 \le X(\omega) < \infty$ for every $\omega \in \Omega$, and let $\Pi = \{y_0, y_1, y_2, \cdots\}$, where $0 = y_0 < y_1 < y_2 < \cdots$. For each subinterval $[y_k, y_{k+1}]$, we set

$$A_k = \{\omega \in \Omega; y_k \le X(\omega) < y_{k+1}\}.$$

In Riemann integral, the measure of $x-$axis is the length of the subinterval, and the measure of $y-$axis is the value of the function. In this case, the measure of $x-$axis is the probability of the event. Therefore, we define the lower Lebesgue sum to be

$$\text{LebesgueSum}_{\Pi}^{-}(X) = \sum_{k=1}^{\infty} y_k \mathbb{P}\left(A_k\right)$$

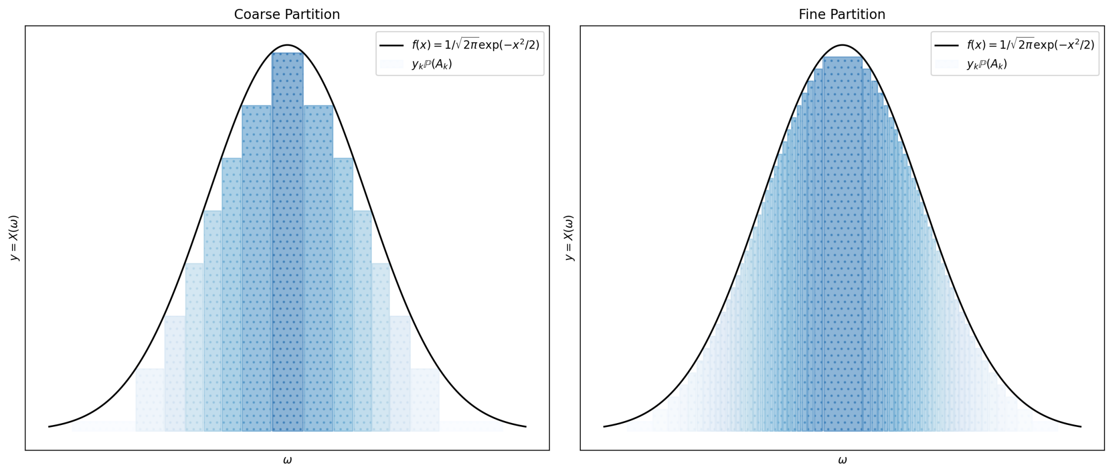

Figure A.2: Lebesgue Integral

This lower sum converges as $\|\Pi\|$, the maximal distance between the y_k partition points, approaches zero, and we define this limit to be the Lebesgue integral $\int_{\Omega} X(\omega) d\mathbb{P}(\omega)$, or simply $\int_{\Omega} X d\mathbb{P}$. The Lebesgue integral might be ∞, because we have not made any assumptions about how large the values of X can be.

We assumed a moment ago that $0 \le X(\omega) < \infty$ for every $\omega \in \Omega$. If the set of ω that violates this condition has zero probability, there is no effect on the integral we just defined. If $\mathbb{P}\{\omega; X(\omega) \ge 0\} = 1$ but $\mathbb{P}\{\omega; X(\omega) = \infty\} > 0$, then we define $\int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \infty$.

Finally, we need to consider random variables X that can take both positive and negative values. For such a random variable, we define the positive and negative parts of X by

$$X^{+}(\omega) = \max\{X(\omega), 0\}, \quad X^{-}(\omega) = \max\{-X(\omega), 0\}.$$

Both $X^{+}$and $X^{-}$are nonnegative random variables, $X = X^{+} - X^{-}$, and $|X| = X^{+} + X^{-}$. Both $\int_{\Omega} X^{+}(\omega) d\mathbb{P}(\omega)$ and $\int_{\Omega} X^{-}(\omega) d\mathbb{P}(\omega)$ are defined by the procedure described above, and provided they are not both ∞, we can define

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

If $\int_\Omega X^+(\omega)d\mathbb{P}(\omega)$ and $\int_\Omega X^-(\omega)d\mathbb{P}(\omega)$ are both finite, we say that X is integrable, and $\int_\Omega X(\omega)d\mathbb{P}(\omega)$ is also finite. If $\int_\Omega X^+(\omega)d\mathbb{P}(\omega) = \infty$ and $\int_\Omega X^-(\omega)d\mathbb{P}(\omega) = \infty$, we say $\int_\Omega X(\omega)d\mathbb{P}(\omega) = \infty$.

Proposition A.1. *Let X, Y be random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the Lebesgue integral $\int_\Omega X(\omega)d\mathbb{P}(\omega)$ is defined if and only if X is integrable. In this case, the following properties hold:*

1. *If X takes only finitely many values $y_0, y_1, y_2, \cdots, y_n$, then*

$$\int_\Omega X(\omega)d\mathbb{P}(\omega) = \sum_{k=0}^{n} y_k \mathbb{P}\{X = y_k\}$$

2. *(**Integrability**) The random variable X is integrable if and only if*

$$\int_\Omega |X(\omega)|d\mathbb{P}(\omega) < \infty.$$

3. *(**Comparison**) If $X \leq Y$ almost surely (i.e., $\mathbb{P}\{X \leq Y\} = 1$), and if $\int_\Omega X(\omega)d\mathbb{P}(\omega)$ and $\int_\Omega Y(\omega)d\mathbb{P}(\omega)$ are defined, then*

$$\int_\Omega X(\omega)d\mathbb{P}(\omega) \leq \int_\Omega Y(\omega)d\mathbb{P}(\omega).$$

In particular, if $X = Y$ almost surely and one of the integrals is defined, then they are both defined and

$$\int_\Omega X(\omega)d\mathbb{P}(\omega) = \int_\Omega Y(\omega)d\mathbb{P}(\omega)$$

4. *(**Linearity**) If α and β are real constants and X and Y are integrable, or if α and β are nonnegative constants and X and Y are nonnegative, then*

$$\int_\Omega (\alpha X(\omega) + \beta Y(\omega))d\mathbb{P}(\omega) = \alpha \int_\Omega X(\omega)d\mathbb{P}(\omega) + \beta \int_\Omega Y(\omega)d\mathbb{P}(\omega)$$

Definition A.2. *Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The expectation (or expected value) of X is defined to be*

$$\mathbb{E}X = \int_\Omega X(\omega)d\mathbb{P}(\omega).$$

This definition makes sense if X is integrable, i.e.; if

$$\mathbb{E}|X| = \int_\Omega |X(\omega)|d\mathbb{P}(\omega) < \infty$$

or if $X \geq 0$ almost surely. In the latter case, $\mathbb{E}X$ might be ∞.

A.2 Convergence of Integrals

Theorem A.1. ***Monotone Convergence Theorem*** *Let $X_1, X_2, \cdots$ be a sequence of nonnegative random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_1(\omega) \leq X_2(\omega) \leq \cdots$ for every $\omega \in \Omega$. Let $X(\omega) = \lim_{n\to\infty} X_n(\omega)$ for every $\omega \in \Omega$. Then*

$$\lim_{n\to\infty} \int_\Omega X_n(\omega)d\mathbb{P}(\omega) = \int_\Omega X(\omega)d\mathbb{P}(\omega).$$

i.e.

$$\lim_{n\to\infty} \mathbb{E}X_n = \mathbb{E}X$$

Another important convergence theorem is the **Dominated Convergence Theorem**.

Theorem A.2. ***Dominated Convergence Theorem*** *Let $X_1, X_2, \cdots$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_n(\omega) \to X(\omega)$ for every $\omega \in \Omega$. Suppose there is a random variable Y such that $|X_n(\omega)| \leq Y(\omega)$ for* *every n and ω. If Y is integrable, then*

$$\lim_{n\to\infty} \mathbb{E}X_n = \mathbb{E}X$$

A.3 Information and Conditional Expectation

Definition A.3. *Let Ω be a nonempty set. Let T be a fixed positive number, and assume that for each $t \in [0, T]$ there is a σ-algebra $\mathcal{F}(t)$. Assume further that if $s \le t$, then every set in $\mathcal{F}(s)$ is also in $\mathcal{F}(t)$. Then we call the collection of σ-algebras $\mathcal{F}(t), 0 \le t \le T$, a* ***filtration****.*

A filtration tells us the **information** we will have at future times. More precisely, when we get to time t, we will know for each set in $\mathcal{F}(t)$ whether the true ω lies in that set. Many times, we will be interested in the information we have at time t about some event that will happen at a later time $s > t$. This is the subject of **conditional expectation**.

Definition A.4. *Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{G}$ be a sub-sigma algebra of $\mathcal{F}$. The conditional expectation of X given $\mathcal{G}$ is a random variable denoted by $\mathbb{E}(X|\mathcal{G})$, and it is defined by the following properties:*

1. *$\mathbb{E}(X|\mathcal{G})$ is $\mathcal{G}$-measurable;*
2. *For every $A \in \mathcal{G}$, we have*

$$\int_A \mathbb{E}(X|\mathcal{G}) d\mathbb{P} = \int_A X d\mathbb{P}$$

Note 23. *If $\mathcal{G}$ is the σ-algebra generated by some other random variable Y (i.e., $\mathcal{G} = \sigma(Y)$), we generally write $\mathbb{E}[X \mid Y]$ rather than $\mathbb{E}[X \mid \sigma(Y)]$.*

There're 3 important properties of conditional expectation:

Proposition A.2. (i) (***Taking out what is known***) *If X and Y are integrable random variables, Y and XY are integrable, and X is $\mathcal{G}$-measurable, then*

$$\mathbb{E}[XY \mid \mathcal{G}] = X\mathbb{E}[Y \mid \mathcal{G}]$$

(ii)(***Snake Formula***) *If $\mathcal{H}$ is a sub- σ algebra of $\mathcal{G}$ ($\mathcal{H}$ contains less information than $\mathcal{G}$) and X is an integrable random variable, then*

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

(iii) (***Independence***) *If X is integrable and independent of $\mathcal{G}$, then*

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

Definition A.5. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\mathcal{F}(t), 0 \le t \le T$, be a filtration of sub- σ-algebras of $\mathcal{F}$. Consider an adapted stochastic process $X(t), 0 \le t \le T$. Assume that for all $0 \le s \le t \le T$ and for every nonnegative, Borel-measurable function f, there is another Borel-measurable function g such that*

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

Then we say that the X is a ***Markov process****.*

Markov property is a strong form of conditional independence. It says that the future of the process, given the present, is independent of the past.

B Monte Carlo Simulation

Consider the problem of evaluating the integral of a smooth continuous function over a given domain: $\int_a^b f(x)dx$. There are three options available to us.

1. Use Calculus: If the function is simple enough, we can use calculus to evaluate the integral.
2. Use Numerical Integration on a grid.

$$\lim_{N\to\infty}\left(\sum_{i=0}^{N-1} f(x_i)\,\Delta x\right) \to \int_a^b f(x)dx$$

3. Use Stochastic Integration. Rewrite the integrand as $f(x) = g(x)\phi(x)$, where $\phi(x)$ is a probability density function. Then the integral can be written as

$$\int_a^b f(x)dx = \int_a^b g(x)\phi(x)dx = \mathbb{E}\left[g(X)\right]$$

where X is a random variable with density $\phi(x)$. We can then estimate the integral by generating random samples of X and evaluating $g(X)$ at each sample point.

For the 3rd technique to work $\phi(x)$ has to describe a random variable whose domain is $[a,b]$ and $\phi(x) \neq 0$ for all $x \in [a,b]$. And so the original deterministic integral can now be solved as a stochastic Monte Carlo integral, which amounts to sampling N outcomes from the distribution $\phi(x)$ and evaluating the function $g(x)$ at each of these points. The integral is then approximated by the average of these evaluations:

$$S_N = \frac{1}{N}\sum_{i=1}^{N} g(x_i), \quad x_i \sim \phi(x), \quad i = 1, 2, \ldots, N$$

The computational algorithm for Monte Carlo integration is as follows:

1. Generate a sequence of random samples $x_1, x_2, \ldots, x_N$ from the distribution $\phi(\mathbf{x})$. In general, x_i is a vector in $\mathbb{R}^d$.
2. For each random sample invoke the function $g(\mathbf{x})$ to evaluate the integrand at that point.
3. Form the average S_N of the function evaluations.
4. Estimate the variance of $g(\mathbf{x})$ using the formula

$$\mathrm{Var}[g(\mathbf{x})] = \frac{1}{N}\sum_{i=1}^{N}\left(g(\mathbf{x}_i) - S_N\right)^2$$

5. Report back the value of S_N as the Monte Carlo estimate of the integral and the variance as an estimate of the error.

C Another Way to Derive Black-Scholes PDE

Assume the stock price follows the geometric Brownian motion

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

and the bank account follows the deterministic process

$$dB_t = rB_t dt$$

Definition C.1. *The portfolio strategy including bank account and underlying asset $h_t = (a_t, \beta_t)$ is self-financing if*

$$d\Pi_t = \beta_t dB_t + \kappa_t dS_t,$$

where

$$\Pi_t = \beta_t B_t + \kappa_t S_t.$$

Lemma C.1. *If a value process of a self-financing portfolio has dynamic*

$$d\Pi_t = k_t \Pi_t dt$$

where k_t is adapted, then $k_t = r$.

Proof. If $k_t > r$: borrow to invest in the portfolio and earns a riskless profit

$$\Pi_t \left(k_t - r\right) dt > 0$$

If $k_t < r$: short the portfolio and invest the proceed with rate r ; this earns a riskless profit

$$\Pi_t \left(r - k_t\right) dt > 0$$

□

Consider a portfolio of γ_t derivatives with price V_t and κ_t stocks with price S_t. The value of the portfolio is $\Pi_t = \gamma_t V_t + \kappa_t S_t$. Because of the self-financing condition, V_t has dynamics:

$$\begin{aligned} d\Pi_t &= \gamma_t dV_t + \kappa_t dS_t \\ &= \gamma_t \left(\left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2 \right) dt + \frac{\partial V}{\partial S} dS \right) + \kappa_t dS_t \end{aligned}$$

Choose $\kappa_t = -\gamma_t \frac{\partial V}{\partial S}$. Then

$$\begin{aligned} d\Pi_t &= \gamma_t \left(\left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2 \right) dt + \frac{\partial V}{\partial S} dS \right) - \gamma_t \frac{\partial V}{\partial S} dS_t \\ &= \gamma_t \left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2 \right) dt \\ &\overset{\gamma_t = \frac{\Pi_t}{V_t - \frac{\partial V}{\partial S} S_t}}{=\!=\!=\!=} \frac{\Pi_t}{V_t - \frac{\partial V}{\partial S} S_t} \left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2 \right) dt \\ &:= k_t \Pi_t dt \end{aligned}$$

where

$$\left(V_t - \frac{\partial V}{\partial S} S_t\right) k_t = \frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2$$

by the lemma $k_t = r$ we gain the Black-Scholes PDE

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

D An Abbreviated Proof of Girsanov's Theorem

Lemma D.1. ***Levy's Theorem**: If $M_t, t \geq 0$ is a martingale with respect to the filtration $\mathcal{F}_t$ and M_t is continuous and $M_0 = 0$. Then M_t is a Brownian motion if the quadratic variation of M_t is t.*

$$\langle M, M \rangle [T] = T$$

Proof. The Ito Lemma also holds for M_t. For any function $f(t, x)$ that is continuous and has continuous partial derivatives, we have

$$df(t, M_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial M_t} dM_t + \frac{1}{2} \frac{\partial^2 f}{\partial M_t^2} dt$$

Hence

$$f(t, M_t) = f(0, M_0) + \int_0^t \frac{\partial f}{\partial t} ds + \int_0^t \frac{\partial f}{\partial M_s} dM_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial M_s^2} ds$$

Take expectation on both sides

$$\mathbb{E}[f(t, M_t)] = f(0, M_0) + \mathbb{E}\left[\int_0^t \frac{\partial f}{\partial t} ds\right] + \frac{1}{2} \mathbb{E}\left[\int_0^t \frac{\partial^2 f}{\partial M_s^2} ds\right]$$

For any fixed number u, define

$$f(t, x) = \exp\left(ux - \frac{1}{2} u^2 t\right)$$

then

$$\frac{\partial f}{\partial x} = uf, \quad \frac{\partial^2 f}{\partial x^2} = u^2 f, \quad \frac{\partial f}{\partial t} = -\frac{1}{2} u^2 f$$

Thus

$$\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0$$

Hence

$$\mathbb{E}\left[\exp\left(uM_t - \frac{1}{2} u^2 t\right)\right] = 1$$

Implying

$$\mathbb{E}\left[\exp\left(uM_t\right)\right] = \exp\left(\frac{1}{2} u^2 t\right)$$

This is the moment generating function of a normal distribution with mean 0 and variance t. Therefore, M_t is a Brownian motion. □

Theorem D.1. ***Girsanov's Theorem**: Let W_t be a Brownian motion under the probability measure $\mathbb{P}$ and k_t be a predictable process. Define a new probability measure $\mathbb{Q}$ by*

$$L_t = \frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^t k_s dW_s - \frac{1}{2} \int_0^t k_s^2 ds\right)$$

Then $\widetilde{W}_t = W_t - \int_0^t \theta_s ds$ is a Brownian motion under the probability measure $\mathbb{Q}$.

The Novikov condition is

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T k_s^2 ds\right)\right] < \infty$$

which ensures that the Radon-Nikodym derivative has expectation 1.

Proof. First note that

$$d\widetilde{W}_t d\widetilde{W}_t = (dW_t - \theta_t dt)(dW_t - \theta_t dt) = dt$$

It remains to show that $\widetilde{W}(t)$ is a martingale under $\mathbb{Q}$. We first observe that L_t is a martingale under $\mathbb{P}$. With

$$X_t = \int_0^t k_s dW_s - \frac{1}{2}\int_0^t k_s^2 ds$$

and $f(x) = e^x$, we have $L_t = e^{X_t}$. By Ito's formula

$$dL_t = df(X_t) = k_t Z_t dW_t$$

Obviously, L_t is a martingale under $\mathbb{P}$ and $\mathbb{E}[L_t] = 1$. We can now apply Levy's theorem to show that $\widetilde{W}_t$ is a Brownian motion under $\mathbb{Q}$.

Next show that $\widetilde{W}_t L_t$ is a martingale under $\mathbb{P}$.

$$\begin{aligned} d(\widetilde{W}_t L_t) =& \widetilde{W}_t dZ_t + Z_t d\widetilde{W}_t + d\widetilde{W}_t dZ_t \\ =& (\widetilde{W}_t k_t + 1) L_t dW_t \end{aligned}$$

Hence $\widetilde{W}_t L_t$ is a martingale under $\mathbb{P}$.

Finally, we show that $\widetilde{W}_t$ is a martingale under $\mathbb{Q}$.

Now let $0 \leq s \leq t \leq T$ be given. Imply proposition 4.1, we have

$$\widetilde{\mathbb{E}}[\widetilde{W}_t \mid \mathcal{F}_s] = \frac{1}{L(s)}\mathbb{E}[\widetilde{W}_t Z_t \mid \mathcal{F}_s] = \frac{1}{L_s}\widetilde{W}_s Z_s = \widetilde{W}_s$$

This shows that $\widetilde{W}(t)$ is a martingale under $\mathbb{Q}$.

By the Levy's theorem, $\widetilde{W}(t)$ is a Brownian motion under $\mathbb{Q}$. □

E Bilingual Table

Table E.1: Bilingual Table

English	Chinese
Financial Derivatives	金融衍生品
Stochastic Calculus	随机微积分
Brownian Motion	布朗运动
Quadratic Variation	二次变差
Ito's Formula	伊藤引理
Stochastic Differential Equation(SDE)	随机微分方程
Asset Price	资产价格
Arbitrage	套利
Replicating Portfolio	复制投资组合
Risk-Neutral Measure	风险中性测度
Speculation	投机
Forward Contract	远期合约
Exercising	行权
At the Money	平价期权
In the Money	实值期权
Out of the Money	虚值期权
Long Position	多头头寸
Short Position	空头头寸
Put-Call Parity	看涨看跌平价
Call/Put Option	看涨/跌期权
European Option	欧式期权
American Option	美式期权
Asian Option	亚式期权
Lookback Option	回望期权
Barrier Option	敲出期权
Knock-in Option	敲入期权
Knock-out Option	敲出期权
Leverage	杠杆
Binomial (Tree) Model	二叉树模型
Law of One Price	一价定律
Intrinsic Value	内在价值
Stochastic Process	随机过程
Conditional Expectation	条件期望
Filtration	域流,信息流
σ-Algebra	σ-代数
Martingale	鞅
Supermartingale	超鞅
Submartingale	次鞅
Doob Martingale	杜布鞅
Donsker's Theorem	东斯克定理
Markov Property	马尔科夫性
Simple Process	简单过程
Ito Isometry	伊藤等距
Cauchy Sequence	柯西序列

Geometry Brownian Motion	几何布朗运动
ARIMA Model	差分滑动平均自回归模型
Linear Gaussian Model	线性高斯模型
Log-normal Distribution	对数正态分布
OU Process	奥恩斯坦伯格过程
Constant Elasticity of Variance	恒弹性方差
Cox-Ingersoll-Ross Process	科克斯-英格索尔-罗斯过程
Black-Scholes Model	布莱克-斯科尔斯模型
Self-Financing Portfolio	自融资投资组合
Underlying Asset	标的资产
Delta-Hedging	Delta对冲
Boundary Condition	边界条件
Forward Euler Method	前向欧拉法
Backward Euler Method	后向欧拉法
Crank-Nicolson Method	克兰克-尼科尔森法
Discretization Error	离散化误差
Dynkin's Formula	丁金公式
Feynman-Kac Formula	费曼-卡茨公式
Girsanov's Theorem	吉尔萨诺夫定理
Randon-Nikodym Derivative	拉东-尼科蒂姆导数
Sharpe Ratio	夏普比率
Complete Market	完全市场
Risk-Neutral Valuation	风险中性定价
Greeks	希腊字母
Delta-Neutral Portfolio	Delta对冲投资组合
Gamma-Neutral Portfolio	Gamma对冲投资组合
Dividend Paying Stock	分红股
Numeraire	计价单位
Multi-dimensional Ito's Lemma	多维伊藤引理
Black-Scholes PDE	布莱克-斯科尔斯偏微分方程
Monte Carlo Simulation	蒙特卡洛模拟
Basket Option	篮子期权
Exchange Option	交换期权
Feasible Region	可行域
Implied Volatility	隐含波动率
Volatility Term Structure	波动率期限结构
Piece-wise Constant Function	分段常数函数
Volatility Smile	波动率微笑
Volatility Skew	波动率偏斜
Kurtosis	峰度
Spot Price	现货价格
Spot Rate	即期利率
Spot Volatility	即期波动率
Bootstrapping	自举法
Market Crash	市场崩盘
Market Rally	市场反弹
Newton-Raphson Method	牛顿-拉弗森法
Leibniz Rule	莱布尼茨法则
Local Volatility Model	局部波动率模型
Dupire Equation	杜拜尔方程
Stochastic Volatility Model	随机波动率模型
Exotic Option	奇异期权
First Hitting Time	首达时
Running Maximum	最大游程
Reflection Principle	反射原理
Geometric Asian Option	几何亚式期权
Arithmetic Asian Option	算术亚式期权
Path-dependent Option	路径依赖期权
Rogers-Shi Method	罗杰斯-史方法
Foreign Exchange Option	外汇期权
Domestic	国内

Direct Quotation	直接报价
Indirect Quotation	间接报价
Quanto Derivative	双币种衍生品
Stopping Time	停时
Optimal Stopping	最优停时
Value Function	值函数
Optimal Value Function	最优值函数
Snell Envelope	包络
Dominate	支配
Linear Complementarity Problem	线性互补问题
Free Boundary	自由边界
Smooth Pasting Condition	光滑粘贴条件
Perpetual American Option	永续美式期权
Ordinary Differential Equation	常微分方程
Bond	债券
Coupon	票息
Zero Coupon Bond	零息债券
Yield to Maturity	到期收益率
Par Value	面值
Notional	名义金额(=面值)
Accrue	累积
Deposit Rate	存款利率
Forward Measure	远期测度
Forward Rate Agreement	远期利率协议
Forward Yield	远期收益率
Instantaneous Rate	瞬时利率
Floating Rate Note	浮动利率票据
Overnight Floating Index	隔夜浮动利率
Floating Rate Leg	浮动利率腿
LIBOR Rate	伦敦银行同业拆借利率
SOFR Rate	美国隔夜拆借利率
ESTR Rate	欧洲隔夜拆借利率
SHIBOR Rate	上海银行间同业拆放利率
Interest Rate Swaps	利率互换
Par Swap Rate	平价互换利率
Zero Curve Bootstrapping	零息曲线自举法
Credit Risk	信用风险
Spot Starting Yield	即期起始收益率
Default	违约
Survival Probability	生存概率
Credit Spread	信用利差