Quantitative Methods in Finance

William Arnaud, Charaf Zguiouar, Hugo Michel

hugo.michel1@etu.univ-paris1.fr

william.arnaud1@etu.univ-paris1.fr

charaf.zguiouar@etu.univ-paris1.fr

Professor: Dr Eric Vansteenberghe

Derivatives

-

M2 Finance Technology Data - Sorbonne School Of Economics

Université Paris 1 Panthéon-Sorbonne

Paris, France, Fall 2023

Contents

1 Introduction

1.1 Overview of Derivatives Pricing

Derivatives are financial instruments whose value is dependent on the value of an underlying asset (stocks, bonds, commodities).

It exsits several types of derivatives:

- **Forwards:** The simplest derivatives are forwards, which are used to fix today the price of a transaction that will take place at a future date.
- **Swaps:** Swaps are used to exchange a series of financial pairs (for example, a series of euro payments against a series of dollar payments).
- **Option:** Options are non-linear contracts that give the buyer greater freedom. For example, a call option on a financial asset gives the buyer the right, but not the obligation, to buy that asset at a future date.

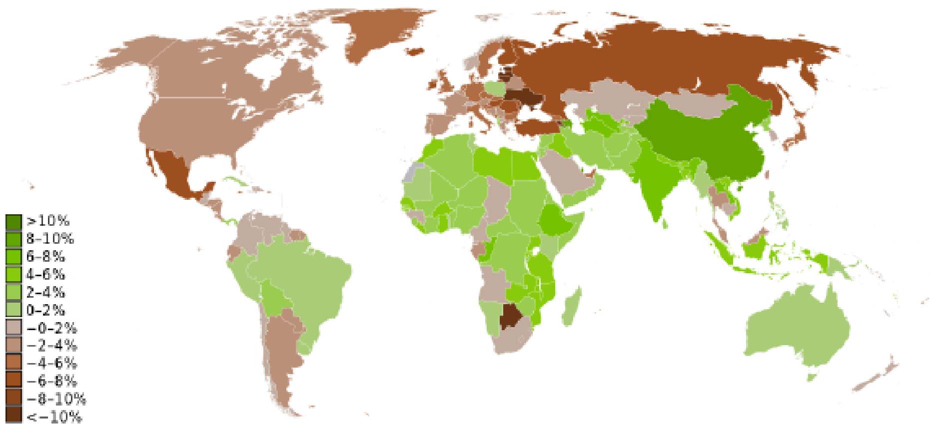
The main reason for the existence of derivative products is the transfer of financial risks between investors, or more generally between economic agents. Derivatives can be purchased to hedge future risks, to speculate (i.e. to take the risk by betting on the rise or fall of the price of a security) or to arbitrage (to take advantage of a market imbalance without taking risk). For example, a manufacturer can use a futures contract to hedge against fluctuations in commodity prices. An investment fund may use options either to avoid excessive losses, or as part of an aggressive management strategy. A company listed on the stock exchange can unlist its employees by paying out part of their remuneration in the form of stock options, thereby making them share in the risks associated with its operations.

Financial risks are those associated with fluctuations in the prices of financial or physical assets, interest rates or exchange rates (market risk), or more generally with financial transactions: non-repayment of a debt (default risk or counterparty risk), inability to sell an asset or trade on the market (liquidity risk). Derivative products generally cover market or credit risk. They may themselves be traded on an organized market (as is the case for stock/index options), or sold by major banks (e.g. currency or interest-rate options).

1.2 Financial crisis 2007-2008

Nothing illustrates the dangers of using financial models outside their scope of applicability as well as the financial crisis of 2007-2008. This crisis shook the world economy and caused massive losses in GDP for all developed countries, was largely caused by the disproportionate growth of a particular type of credit derivative: CDOs (Collateralised Debt Obligations), and by major failings in the understanding and management of the associated risks.

Figure 1.1: Stylised example of an investment bank's balance sheet. Source: McNeil, A. J., Frey, R., and Embrechts, P. (2015). Quantitative risk management: Concepts, techniques and tools. Princeton university press.



An unprecedented bubble: The CDO market

The Fed's rate cut with US government policies encourage to become homeowners, helped create a bubble in the real estate market to meet the growing demand for real estate loans, banks securitization, or the conversion of loans into securities. The principle of securitizing real estate loans is to sell them to a special purpose. This company issues bonds, which are sold to investors. Investors receive relatively liquid products with attractive returns, and the banks are relieved of the risk, as the receivables no longer appear on their balance sheets, and free up capital to grant more loans. These mortgage-backed bonds are called MBS (mortgage-backed securities) or CMO (collateralized mortgage obligations).

These bonds are classified into several categories or tranches, which determine the order in which they are affected by losses in the initial receivables portfolio. The senior tranche is the least remunerated. It generally represents between 80% and 90% of the total amount and is the last to suffer losses. It is therefore less risky than the receivables in the initial portfolio, because it will not be affected by portfolio losses until all other tranches have been destroyed. The next tranche (mezzanine) is riskier, as the bondholders in this tranche are paid only once the senior tranche has been repaid. It therefore more attractive remuneration. Finally, the last tranche (equity), which comprises residual assets, carries the most risk and is no longer really considered debt. The mezzanine tranches of MBS are rated by the rating agencies, and before the crisis the AAA rating, which corresponds to a historical default rate of less than 0.02% per year, as their diversification made them virtually risk-free.

While the senior tranches of MBS were easy to sell, the riskier tranches failed to attract investors despite their more attractive remuneration. To compensate for this lack of interest, an important innovation appeared: CDOs (collateralised debt obligations). CDOs are constructed according to the same principle, only instead of being backed directly by residential loans, they are backed by more varied debt, including (in most cases) the risky tranches of MBS. The senior tranches of CDOs also received a very high credit rating and equity tranches were sold to investors, while the mezzanine and equity tranches were sometimes repackaged into other structures of the same type, called CDO, whose senior tranches were also highly rated.

Because of their high ratings and attractive remuneration, CDOs were very much in demand by investors, particularly certain pension funds, which by law can only invest in AAA-rated assets. The banks that structured the CDOs also retained a large quantity of these assets as part of their proprietary investment activity. The demand for CDOs in turn fuelled the demand for property loans, which led to a reduction in the requirements for a potential borrower, and to the proliferation of variable-rate loans, with initial payment, and more generally so-called 'subprime' loans, i.e. loans with a high risk of non-repayment. Like other loans, subprime loans were restructured as CDOs and given AAA ratings, which led to the greatest aberration of the crisis: a portfolio of low-quality loans was used to manufacture bonds, 90% of which were assumed to be virtually risk-free.

How the crisis unfolded

With the increase in the Fed's rate in 2006, variable-rate borrowers were no longer able to repay their loans and their homes were sold at auction; With the market saturated due to a large number of houses being built, this caused first a slowdown and then a sharp fall in prices on the market, which further increased the non-repayment rate, as buyers could no longer repay their loans.

Figure 1.2: Structure of MBS, CDO and CDO2 pricing. Source: International Monetary Fund

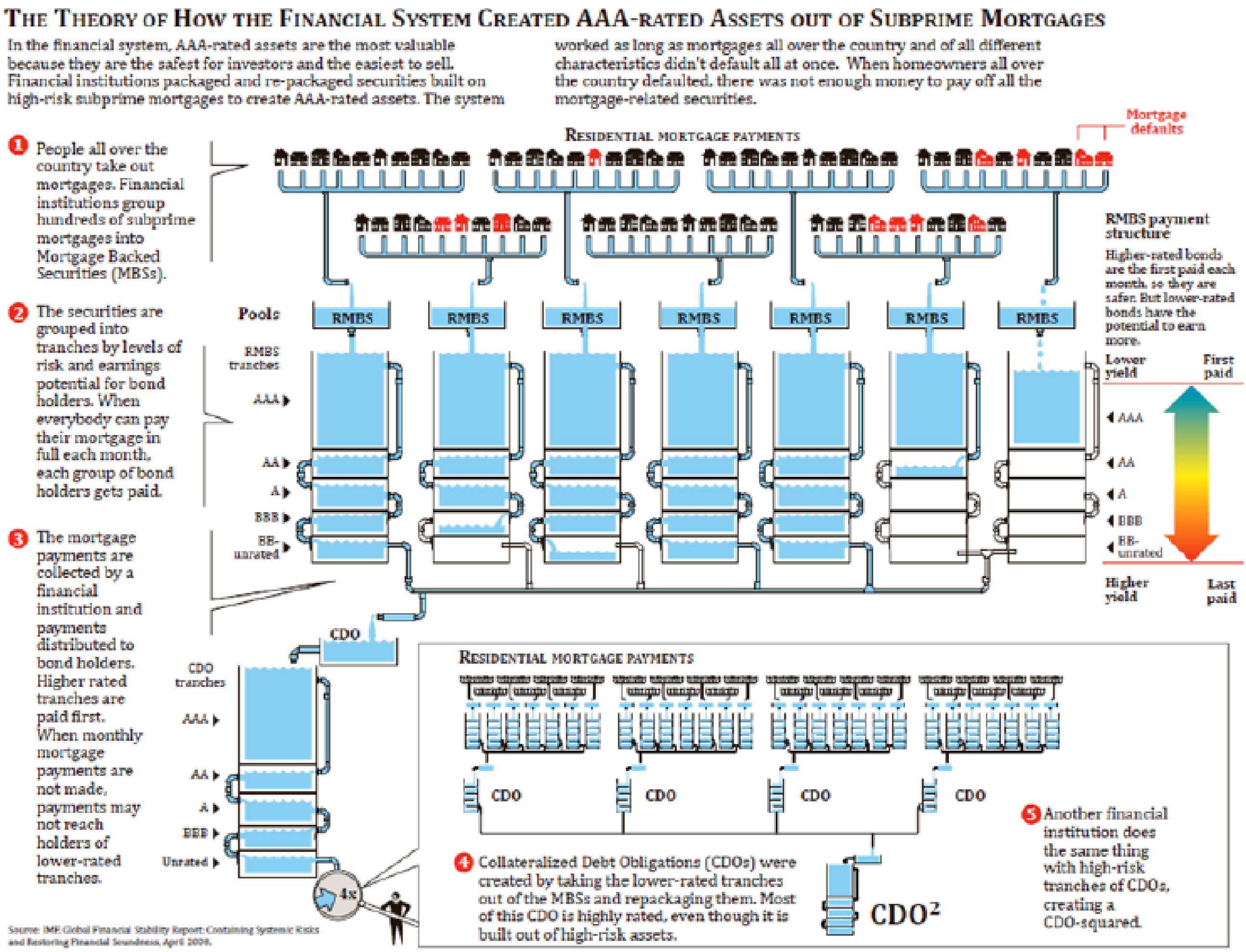
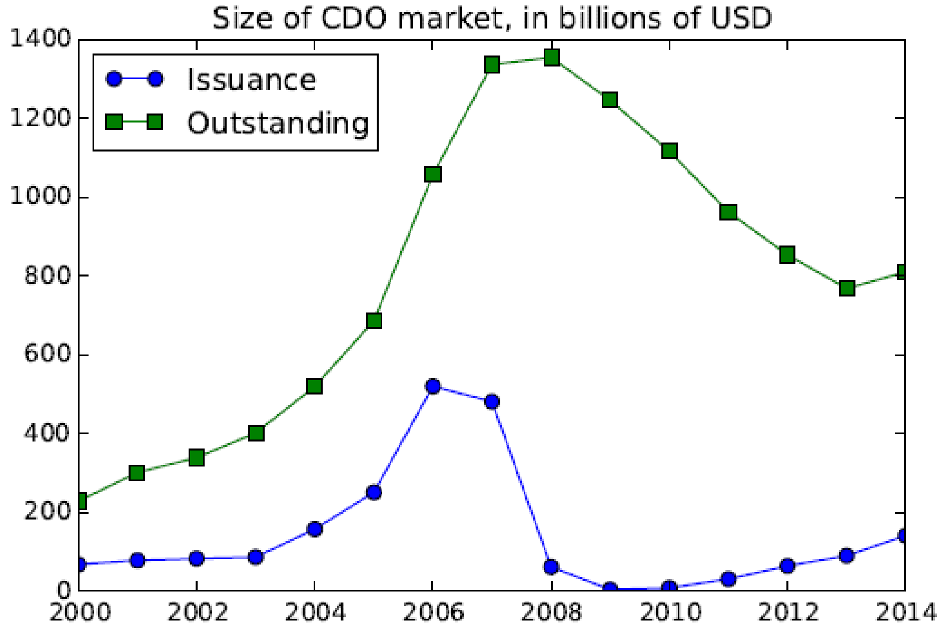


Figure 1.3: CDO market before and after the subprime crisis. Data source: Bank of International Settlements.



Credit rating and rating agencies

The credit quality of an issuer is measured by the credit rating: an evaluation of the issuer's solvency calculated internally by the bank or by an external rating agency. Sovereign, corporate and CDO ratings are calculated by international rating agencies, the best known of which are Standard Poor's, Moody's and Fitch ratings. For example, Standard and Poor's rating scale is AAA, AA, A, BBB, BB, B, CCC, CC, C, D, where AAA is the highest possible rating and D is the default. Ratings from AA to CCC can be specified by adding a plus (+) or a minus (-). Bonds rated at least BBB- are considered suitable for long-term investment ('investment grade'), while the others are considered speculative investments. Rating agencies present their ratings as mere 'opinions', whereas in reality they play a very important role in banking regulation: some money market funds can only invest in AAA-rated bonds, and when calculating the capital ratio, the rating determines the weighting of assets. For example, under Basel II, an AAA rating corresponds to a weighting of 0.6% and a BBB rating corresponds to a weighting of 4.6%. This means that a bank with AAA-rated assets in its portfolio will have to hold almost 8 times less regulatory capital than a bank with BBB-rated assets.

The rating agencies note an increase in the rate of non-repayment and are downgrading rating of MBS and CDOs. This had an immediate impact on CDO valuations in the balance sheets of banks and investment funds, as since 2006, financial institutions have been obliged to record their assets at market value, even if they do not intend to sell them in the short term. they do not intend to sell them immediately. The first victim was the American investment bank Bear Stearns, which in July 2007 announced unprecedented losses in two of its funds heavily invested in CDOs. In August, BNP Paribas temporarily suspended capital withdrawals from three of these CDO funds. In October 2007, it was the turn of Merrill Lynch announced $7.9 billion in losses linked to its CDO business. The panic triggered by these losses reduced the liquidity of the CDO market to almost zero. In fact, as the prices of these "toxic assets" were constantly being revised downwards, it became almost impossible to get rid of them.

The liquidity crisis on the CDO market then triggered a crisis of confidence on the interbank market: it became increasingly difficult for investment banks to finance themselves on a day-to-day basis. The first victim was Bear Stearns: to avoid the worst, in March 2008 the Fed urgently negotiated the purchase of this bank by JP Morgan. The rescue of Bear Stearns shareholders with taxpayers' money put the Fed under pressure. In the UK, Northern Rock was nationalized in the spring 2008 after a run on its deposits. The crisis peaked in September 2008 with the bankruptcy of Lehman Brothers and the emergency takeover of Merrill Lynch by Bank of America. Two other major investment banks, Goldman Sachs and Morgan Stanley, converted to retail to gain access to the Fed's emergency funding. Building societies Fanny Mae and Freddie Mac are taken over by the government, and insurer AIG receives 100 billion from the federal government. Fortis Bank is nationalized in Belgium on September 29. September. During the week of October 6 to 10, 2008, several major indices such as the S&P 500 and the CAC40 lost more than 20% of their value. Finally, in October 2008, a law an unprecedented $700 billion bank rescue program. billion.

CDO risk management failures: Too much reliance on simple models

The poor assessment and management of CDO risk by investors, banks and rating agencies played a key role in triggering the crisis. CDOs are extremely complicated financial products. Taking into account references to other documents documents, the number of pages in the prospectus for some of these products exceeds one million! In this context, most banks were using models that were too simple to account for the complex reality of simultaneous defaults affecting the pay-off of a CDO. For example, D. Li's Gaussian copula model [Li, 2000, Salmon, 2009] reduces the complexity of the problem to a single correlation value. Other buyers of CDOs relied on the ratings issued by the rating agencies at the request of the sellers. After all, to assign a rating to a CDO, the rating agencies received a commission of up to $500,000 per issue, so they were expected to analyse the risks of these products in depth. The problem was that even the rating agencies couldn't really grasp the complexity of CDOs and were forced to make major simplifications in their modelling, which later proved disastrous.

The risk of a CDO backed by MBS depends on the probability of non-repayment of the individual loans that make up the CDO via MBS, and on the correlations between the individual loans. However, in their models for analyzing the credit risk of CDOs, rating agencies never went down to the level of individual loans, these models were based exclusively on the ratings assigned to MBS tranches by the agencies themselves. The MBS ratings were calculated using data from a fairly short historical period, since the MBSs themselves had not existed for very long. Although they took into account the non-repayment of individual borrowers due to particular circumstances, the systemic risk, linked to a global downturn in the property market, was largely underestimated. To take this into account, we would have had to go back to the 1930s and the Great Depression, which nobody did. Similarly, the correlation between the default events of the different tranches of MBS making up a CDO was not modelled correctly in the absence of accurate data. For example, in June 2005, Moody's did not have a systematic procedure for estimating the correlation of MBS tranches and used 'expert opinions' to select correlation values. In June 2005, such a model was introduced, but it used only 20 years of historical data, which corresponds to a period of rising property prices, during which defaults were mainly due to individual circumstances and correlations were low.

There are several possible explanations for the failure of rating agencies. On the one hand, despite the staggering profits from the CDO rating business (Moody's alone took in 887 million dollars for its structured product rating services in 2006 alone), the quantitative teams remained relatively small, overworked and less well paid than investment bank employees [Commission, 2011, page 149]. Moreover, and this is probably more important, there was a major conflict of interest since rating agencies are paid for their services by the banks whose products they assess. Even if more prudent methodologies were developed within the agencies, they were slow to be applied for fear of assigning less attractive ratings and seeing an important client leave for a less scrupulous competitor.

In summary, the ratings assigned to CDOs by the rating agencies did not correctly reflect the risk of these products, due to a gross underestimation of the systemic risk associated with the general downturn in the real estate market, both in the calculation of the individual default probabilities of MBS tranches and in the calculation of their correlation. This is exactly the risk that materialized in 2006-2007, when rating agencies massively downgraded CDO ratings, reducing the liquidity of these instruments to almost zero and triggering the panic on the interbank market. So, it's not the mathematical models that are responsible for the crisis, but their use outside their scope of application and without proper calibration by the rating agencies, and the blind, common-sense belief in these models on the part of banks and investors.

Chronology of the crisis

- **1987:** Issuance of a first CDO by Drexel Burnham Lambert Inc.
- **2001:** The Fed's interest rate falls below 2
- **2006:** The Fed rate reached 5.75%. Rates on variable subprime mortgages followed the rise. The property market is shrinking.
- **1st half of 2007:** Wave of bankruptcies in organisations specialising in subprime loans; the subprime market was estimated at 1,300 billion in March 2007.
- **July 2007:** Bear Sterns, a US asset management giant, announces unprecedented losses in two of its funds invested heavily in CDOs
- **August 2007:** First liquidity crisis as illiquid CDOs are uncovered in the portfolios of numerous banks and funds around the world (including BNP Paribas). To calm the panic, the central banks made 400 billion euros available to the interbank market.
- **October 2007:** Merrill Lynch announces losses of 8.4 billion dollars linked to its CDO business
- **March 2007:** Bear Stearns sold to JP Morgan with federal government guarantees to avoid bankruptcy
- **April 2008:** OECD estimates subprime losses at 422 billion dollars
- **September 2008:** The two mortgage giants, Mae" and Mac", were effectively nationalised and received 200 billion dollars from the American government, while the insurer AIG, which sold protection against the default of CDOs, received 85 billion dollars.
- **14-15 september 2008:** The collapse of Lehman Brothers, the collapse of AIG and the takeover of Merrill Lynch by Bank of America.
- **Octobre 2008:** The American rescue plan is adopted: 700 billion dollars to recapitalise banks and guarantee assets under the Troubled Assets Relief Program.
- **6 October 2008:** The CAC 40 records its biggest daily fall since its creation: down 9%.
- **13 October 2008:** The French government opens a €10.5 billion credit facility to the country's 6 largest private banks; the CAC 40 beats its all-time daily record: +11%.
- **4 November 2008:** Election de Barack Obama aux Etats-Unis
- **16 december 2008:** Lowest US central bank rate since 1954
- **19 december 2008:** Adoption of a €26 billion recovery plan in France
- **March 2009:** The former world number one in insurance, AIG, announces a loss of 100 billion dollars for 2008
- **5 March 2009:** The ECB cut its main rate to 1.5%, the lowest level since the creation of the euro in 1999.

2 Historical Context and Evolution

2.1 Absence of Arbitrage (AoA)

The central notion of the course is the absence of arbitrage. An arbitrage is an investment strategy with zero initial cost, which has a positive or zero payoff at a future date whatever the market scenario, and a strictly positive payoff in certain scenarios. With a positive probability, it enables you to make a profit without initial investment and without risk. The absence of arbitrage means that the "fair price" of an asset is the price that does not lead to an arbitrage opportunity. For some assets, calculating this no-arbitrage price doesn't require the use of a model, while for others, assumptions must be made about asset dynamics. For example, the famous Black-Scholes formula for call and put option prices assumes that the price of the underlying asset can be described by geometric Brownian motion. Similarly, for some assets and in some models, the no-arbitrage assumption allows us to determine the unique price; in other situations, we obtain a non-empty interval of prices, all of which are compatible with the absence of arbitrage.

2.2 Hedging

Another key concept is hedging. Hedging is the use of a trading strategy in liquid assets trading strategy in liquid assets to minimize or cancel out the risk of a financial position (a priori, the sale of an option). If the selling price is not within the range of prices compatible with the absence of arbitrage (for example, if the bank has managed to sell the option at a price higher than the fair price), then hedging makes it possible to exploit the arbitrage opportunity and make a risk-free gain.

2.3 Players in the world of derivatives products

Futures are often traded on organized markets. Initially developed to enable farmers to hedge the risk associated with uncertain production, from the 1970s onwards these markets extended their scope to financial futures (interest rate/exchange rate/share). Swaps are in most cases traded OTC, and in the case of options there are markets for the simplest contracts (vanilla options), but exotic contracts are sold OTC. Options are often sold by investment banks as part of their market-making activities. They can also be bought by financial institutions such as hedge funds or banks themselves (for speculation or arbitrage purposes), or by companies and even local authorities, to optimize cash flow management. The conclusion of a forward contract or, more generally, of a derivative contract creates obligations which may be difficult for the signatories of the contract to honor in the event of a change in the market situation, leading to a significant counterparty risk. For example, the seller of a call option may not be able to deliver the underlying asset on the expiration date in the event of a sharp rise in market prices.

Financial risks are those associated with fluctuations in the prices of financial or physical assets, interest rates or exchange rates (market risk), or more generally with financial transactions: non-repayment of a debt (default risk or counterparty risk), inability to sell an asset or trade on the market (liquidity risk). Derivative products generally cover market or credit risk. They may themselves be traded on an organized market (as is the case for stock/index options), or sold outright by major banks (e.g. currency or interest-rate options). Table 1.1 shows the market sizes of the various derivative product categories in December 2012. For credit derivatives, the figure compares with 58.243 billion in December 2007, at the start of the financial crisis.

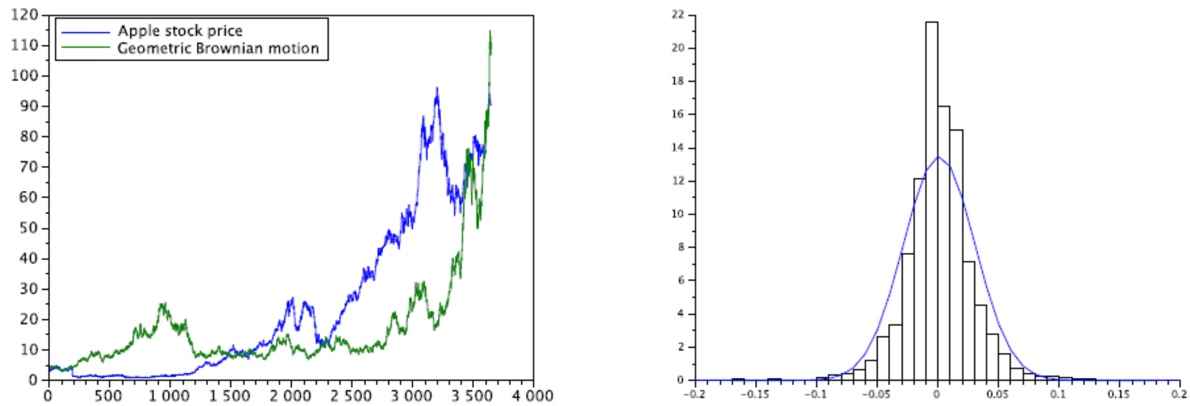
2.4 Risk quantification and mathematics

The values of derivatives depend on the future prices of the underlying assets, and these future prices are uncertain. It is therefore natural to try to quantify this uncertainty with a mathematical model, more precisely with a probabilistic model. To give an example, graph 1.1 compares Apple's price evolution over a 10-year period with a geometric Brownian motion trajectory, whose parameters have been chosen so that the two trajectories resemble each other as closely as possible. Clearly, the two graphs are almost indistinguishable to the naked eye. Similarly, the histogram of daily returns over the same period is very close to the Gaussian distribution. This means that the geometric Brownian motion model provides a fairly accurate description of Apple's daily price behavior, and that it can therefore provide information on the prices of products derived from this asset. In fact, the Black-Scholes model, based on the geometric Brownian motion hypothesis, has been extremely successful on the options market.

However, financial modelling is fundamentally different from modelling in the natural sciences. In physics, we're dealing with laws of nature, which don't change, whereas in finance we're trying to model asset prices, which result from the behavior and perceptions of market agents, which can change according to the economic situation or even according to the models used (because models affect market prices). We must therefore be extremely cautious with financial models, especially those that attempt to make long-term forecasts or analyses. Any model in finance is fundamentally flawed in the sense that, alongside 'quantifiable' risks, there are always uncertainties to which we cannot reasonably associate a probability.

In financial risk management, probabilistic tools such as Value at Risk, which provide a view of risk under 'normal' market conditions, are always complemented by stress scenario analysis, which seeks to identify plausible loss scenarios without giving a precise probability of their occurrence, and to take precautions to avoid the bank's failure should these scenarios materialize. Models made by quants are often used by people other than those who developed them. Consequently, the users of the models may not be aware of all the assumptions/limitations, or may deliberately decide not to take them into account due to competitive pressure, for example (if everyone else in the market is making benefices using a false model, it's hard not to do the same).

Figure 2.1: Price of an Apple share over a 10-year period, compared with a geometric Brownian motion trajectory.



2.5 Banking regulation

The way a bank operates naturally gives rise to conflicts of interest. While customers are primarily interested in the solvency and solidity of the bank, management and traders may be encouraged to take greater risks in order to maximise their remuneration in the event of a gain. Gains are rewarded with substantial bonuses, while losses are borne by shareholders, or even by the public authorities, who can intervene in extreme situations and make a profit.

Figure 2.2: Stylised example of an investment bank's balance sheet. Source: McNeil, A. J., Frey, R., and Embrechts, P. (2015). Quantitative risk management: Concepts, techniques and tools. Princeton university press.

Assets (actif)			Liabilities (passif)	
Cash and central bank balance		10M	Customer deposits	80M
Securities :			Bond issues :	
- bonds	*Trading book*	20M	- senior bond issues	25M
- stocks	*Trading book*	20M	- subordinated bond issues	15M
- derivatives	*Trading book*	10M	Short-term borrowing	30M
Loans and mortgages :			Reserves (for losses on loans)	20M
- corporates	*Banking book*	30M		
- retail	*Banking book*	40M		
- government	*Banking book*	30M	Debt (sum of the above)	170M
Other assets				
- property		10M		
- investment in companies		10M	Equity (fonds propres)	30M
Short-term lending		20M		
Total		200M	Total	200M

to strengthen banking in order to safeguard economic stability. The presence of these conflicts of interest and the importance of banks to the economy as a whole make banking regulation necessary, the main aim of which is to ensure bank solvency. In Europe, the tasks of banking regulation are shared between the ECB and national regulators (Autorite de Contrôle Prudentiel et de Resolution in France). The principles of banking regulation are developed in the Basel Agreements, which are progressively implemented by national regulators. Regulators monitor and impose limits on a number of regulatory ratios, the most important of which is the capital ratio. The recommendation of the Basel agreements is to have a capital ratio of at least 8%.

$$\text{Capital ratio} = \frac{\text{Capital}}{\text{Risk-qeighted asset}} \geq 8\%$$

To value the bank's capital, the market value of its assets must be estimated (marked to market), which can be a real problem for certain derivatives with low liquidity.

Basel agreements

The first Basel agreement (1988) is based on the banking book. It sets out the definition of capital:

$$\text{Capital} \approx \text{Equity}(Tier1) + Reserves(Tier2)$$

and introduces simple weightings for asset classes in the calculation of risk-weighted assets:

- The 1996 Market Risk Amendment introduces a special method for calculating capital for trading book risk. This document introduces an important distinction between the standardised approach and the internal model approach. For banks using an internal model, the capital charge is determined, broadly speaking, by the 99% Value at Risk (a measure of risk which is defined as the quantile of the loss distribution for a given horizon) calculated for a 10-day horizon.
- The second Basel agreement (2005) specifies the rules for calculating risk indicators and extends the possibility of using internal models to the banking book.
- The third Basel Accord (2011), adopted in the wake of the 2007-2008 crisis, strengthens the capital and liquidity constraints.
- The Fundamental review of trading book (2016) modifies the rules for calculating risk measures and, in particular, introduces the use of Expected Shortfall.

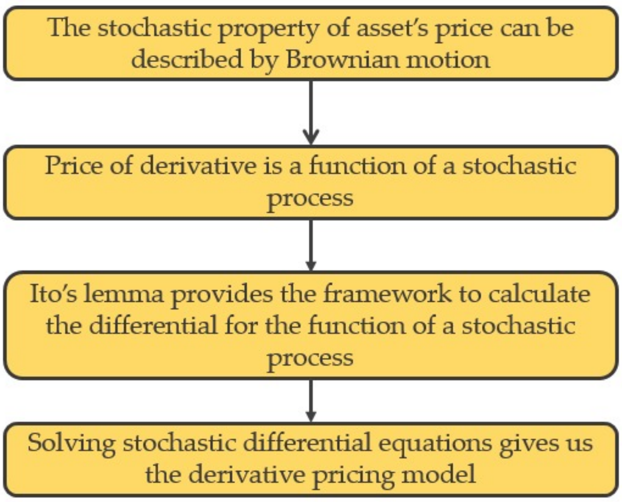
3 Basics of Stochastic Calculus in Finance

3.1

The concepts such as the Brownian motion, Ito's lemma, and stochastic differential equations are integral elements of this system, seamlessly linked with each together and perfectly fitted in the framework of stochastic calculus.

To put it simple, the (standard) Brown motion is the simplest continuous stochastic process. It is the basic model for describing the stochastic nature of asset prices. For options and other derivatives, their prices are functions of the underlying asset price. Since asset price is a stochastic process, the derivative price is a function of that stochastic process. The Ito's lemma provides a framework to differentiate the functions of stochastic process and this is of particular significance to derivative pricing (before Ito's work, people did not know how to do it). Ito's lemma allows us to derive the stochastic differential equation (SDE) for the price of derivatives. Solving such SDEs gives us the derivative pricing models. The derivation of the BS formula is one simple example of this procedure.

Figure 3.1: B-S formula derivation procedure



In the following section, we will explain the Brownian motion and its properties, as well as present the basic form of the Ito's lemma. We tried hard to reveal the properties of the Brownian motion and its implications to stock price movement. In the second part of this series, we will start with a more general form of the Ito's lemma, and apply it to solve for the geometric Brownian motion, and finally derive the BS equation and the BS formula.

3.2 Markov Chain

Markov Chains are a fundamental concept in probability theory and stochastic processes. They are used to model systems that undergo transitions from one state to another, with the key property that the future state depends only on the current state and not on the sequence of events that preceded it. This property is known as the **Markov property**. Indeed, history and how the present emerged from a past trajectory are irrelevant. Share prices are generally supposed to follow Markov processes: predictions about the value of a share have no relation to its price the previous week or the previous year. The only relevant information is the current price.

The foundational work of A.A. Markov in the early 20th century, particularly his paper "An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains" (Markov, 1913), laid the groundwork for the theory of Markov Chains.

1. **Discrete-Time Markov Chains (DTMCs):**

   - In a DTMC, changes of states occur at discrete time steps. Each state transition is characterized by a probability, often represented in a transition matrix.
   - The transition matrix P of a Markov chain, where P_{ij} represents the probability of moving from state i to state j, must satisfy the conditions:

   $$\sum_j P_{ij} = 1 \quad \forall i$$

   meaning the probabilities from any given state sum up to 1.

2. **Continuous-Time Markov Chains (CTMCs):**

   - In CTMCs, state changes can occur at any continuous time. They are defined by transition rates rather than probabilities.
   - The transition rate from state i to state j is denoted by q_{ij}, and the time spent in each state follows an exponential distribution.

3. **Classification of States:** States in a Markov chain can be transient or recurrent, and recurrent states can be further classified as null or positive recurrent. If every state can be reached from every other state, the chain is said to be irreducible.

4. **Stationary Distribution:** A Markov chain has a stationary distribution if the probability distribution over states does not change over time, i.e., $\pi P = \pi$, where π is the stationary distribution.

5. **Ergodicity:** A Markov chain is ergodic if it is possible to reach any state from any state in a finite number of steps, and there is a positive probability of return within a finite number of steps. Ergodic Markov chains converge to a unique stationary distribution regardless of the initial state.

Markov Chains have wide applications in various fields including economics, finance, engineering, biology, and computer science. They are particularly useful in queueing theory, inventory management, and in modeling random processes like stock prices or consumer behavior.

Markov Chains provide a versatile and powerful framework for modeling and analyzing systems characterized by stochastic transitions. Their ability to capture the essence of random processes with memoryless transitions makes them a cornerstone in the study of stochastic processes.

The Brownian motion is a particular case of a Markov process.

3.3 Stochastic Processes and Stochastic Calculus

Considérons une variable suivant un processus de Markov; supposons que sa valeur présente soit de 10 et que sa variation d'une année suive $\Phi(0,1)$ où $\Phi(\mu,\sigma)$ est une loi normale de moyenne μ et d'écart-type $sigma$. La variation sur deux ans est la somme de deux lois normales, chacunes d'elles ayant une moyenne nulle et un écart-type égale à 1. Or, l'addition de deux lois normales indépendantes suit encore une loi normale dont la moyenne est la somme des moyennes et la variance est la sommes des variances. La moyenne de la variation sur deux ans est donc nulle et la variance est égale à 2. La variation sur deux ans suit, par conséquent $\Phi(0,\sqrt{2})$.

The Wiener process is used in physics to describe the motion of a particle subject to a large number of molecular shocks.

The variation $\Delta_z = \epsilon\sqrt{\Delta_t}$ with ϵ follows a normal distribution $\phi(0,1)$.

Consider the increase z over a relatively long period T. It can be written $z(T) - z(0)$. It can be decomposed into the sum of the increments of z in N small intervals of length Δ_t where:

$$N = T/\Delta_t$$

then

$$z(t) - z(0) = \sum_{i=1}^{N} \epsilon_i \sqrt{\Delta_t}$$

The standard Brownian motion stands as the most straightforward continuous stochastic process, serving as the fundamental model for characterizing the stochastic behavior of asset prices. In the realm of options and other derivatives, their prices hinge on the underlying asset's price, which, being a stochastic process, renders the derivative price dependent on this randomness. Ito's lemma establishes a framework for differentiating functions of stochastic processes, a pivotal development in derivative pricing. Prior to Ito's work, the methodology for this differentiation was unknown. Ito's lemma enables the derivation of stochastic differential equations (SDE) for derivative prices, and solving these SDEs yields models for derivative pricing.

3.3.1 Brownian Motion - mathematical definition

The Brownian motion, also known as the Wiener process, represents a continuous-time stochastic process or a stochastic process in continuous space-time. Unlike discrete-time processes where the index variable assumes distinct values, the Brownian motion involves an index variable that takes on a continuous set of values.

The one-dimensional standard Brownian motion is defined as follows. There exists a probability distribution over the set of continuous functions $B : \mathbf{R} \longrightarrow \mathbf{R}$ satisfying the following conditions:

- $B(0) = 0$. Brownian motion starts from some original point at $t = 0$

- **Stationarity:** For all $0 \leq s < t$, the distribution of $B(t) - B(s)$ is the normal distribution of mean 0 and variance $t - s$. In any given finite time interval δt, $B(\delta t)$ satisfies a normal distribution with mean 0 and variance δt, where the variance increases linearly with time.

- **Independent increment:** If the intervals $[s_i, t_i]$ are nonoverlapping, the random variables $B(t_i) - B(s_i)$ exhibit mutual independence. This leads directly to the assertion that Brownian motion is a Markov process, signifying that the future movement after time t solely depends on the position at t and is unrelated to the historical path before t. Essentially, the current value $B(t)$ at time t encompasses all the information required to forecast future outcomes.

In summary, Brownian motion has continuous paths, meaning that the particle's position changes continuously over time. This contrasts with the discrete steps in a simple random walk. The changes in position over fixed time intervals follow a normal (Gaussian) distribution. The increments are independent of each other.

3.3.2 Properties of the Brownian motion

Information concerning Brownian motion that holds significant implications for modeling stock price movements using this stochastic process:

- The path crosses the x-axis (time axis) infinitely often.

- $B(t)$ has a very close relation with the curve $x = y^2$. At any time t, it does not deviate from this curve too much. As time passes by, the process will not deviate far from $B(0) \pm \sqrt{t}$

- Let $M(t)$ be $Max_{0 \leq s \leq t} B(t)$, it can show that $\mathbf{P}(M(t) \geq a) = 2 \cdot \mathbf{P}(B(t) \geq a)$. This shows how to derive the probability model for the extreme values of Brownian Motion. Since $B(t)$ satisfies the normal distribution $N(0, t)$, using the propriety, we can derive $\mathbf{P}(M(t) \geq a)$ easily,

  $\mathbf{P}(M(t) \geq a) = 2 \cdot \mathbf{P}(B(t) \geq a) = 2 \cdot (1 - \phi(\frac{a-0}{\sqrt{t}})) = 2 - 2\phi(\frac{a}{\sqrt{t}})$.

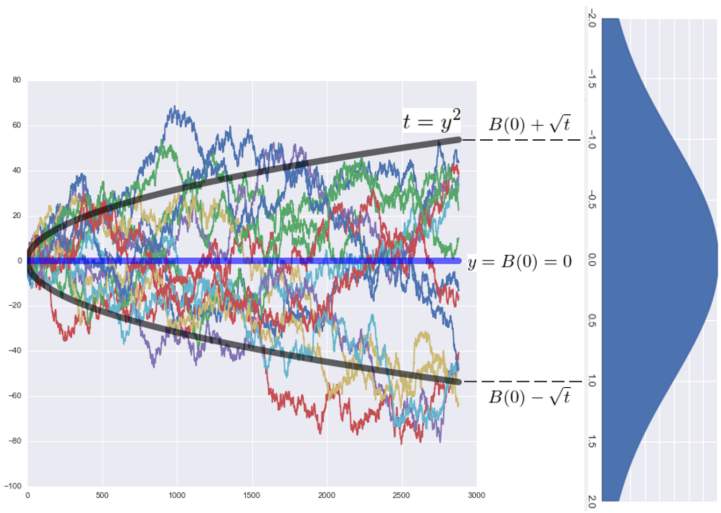
  This can be proved by the Markovian property and the reflection principle. Likewise, let $m(t)$ be the $Min_{0 \leq s \leq t} B(t)$, it can be shown that, $\mathbf{P}(m(t) \leq a) = 2 - 2\phi(\frac{a}{\sqrt{t}})$.

- **It is nowhere differentiable.** This aspect is pivotal in understanding Brownian motion as a stochastic process. It signifies that while Brownian motion is continuous, it lacks differentiability everywhere. The paths exhibit fluctuations, moving up and down in a random manner. The trajectory of Brownian motion differs entirely from the continuous and smooth trajectories familiar to us. Consequently, classical calculus is ineffective for analyzing Brownian Motion. Ito's calculus, developed to address this issue, serves as the cornerstone of modern financial mathematics.

Why GBM do not deviate too much from $t = y^2$?

To explain the first two properties, the following figure shows 15 sample paths of 15 standard Brownian motion in time interval 0 to t. Each path crosses y = 0 (the time axis) multiple times with the exception that only very few paths appear to be on the same side of y = 0 for the entire period. However, they will eventually cross the x-axis as t increases. The black parabola is the curve of $t = y^2$. We can see that although each sample path shows a distinct randomness, at any time $t' \leq t$, they do not deviate too far from the parabola curve which is $B(0)$ +/- the squart root of t. On the right of the figure is the probability density function of the normal distribution at t whose mean is 0 and variance is t. The range of the parabola corresponds to one standand deviation above and below the mean of the normal distribution.

Figure 3.2: GBM do not deviate too much from $x = y^2$



Suppose we choose to use the Brownian motion to describe the high frequency intraday price movement (later in this article we will point out that a more accurate model is the geometric Brownian motion with drift, but let's now use the Brownian motion for a short while), then the two properties above mean that the stock price will fluctuate around the open price and as time passes by, it will not deviate too much from the open price +/- the square root of t times the standard deviation of the price. These properties are important to high frequency traders.

The third property shows how to derive the probability model for the extreme values of the Brownian motion. Since B(t) satisfies the normal distribution $N(0, t)$, using $P(M(t) \geq a) = 2 \times Prob(B(t) \geq a)$, we can derive $Prob(M(t)a)$ easily :

$$\mathbf{P}(M(t) \geq a) = 2 \cdot \mathbf{P}(B(t) \geq a) = 2 \cdot (1 - \phi(\frac{a-0}{\sqrt{t}})) = 2 - 2\phi(\frac{a}{\sqrt{t}})$$

where Φ is the cumulative density function of the standard normal distribution. This can be proved by the Markovian property and the reflection principle. Likewise, let m(t) be the minimum value of $B(t)$ in $[0, t]$, i.e., $m(t) = min_{0 \leq s \leq t} B(t)$. It can be shown that :

$$P(m(t) \leq -a) = P(M(t) \geq a) = 2 - 2\phi(\frac{a}{\sqrt{t}})$$

These results can be used to quantify the probability distribution about the extreme values of the stock price, which can be of great importance in risk management.

The last property is a crucial nature of the Brownian motion as a stochastic process. It says thatalthough the Brownian motion is continuous, it is not differentiable everywhere (this can be proved by contradiction with the usage of the mean value theorem and the third property). This is very intuitive to understand. Let's take a look at those 15 sample paths of the Brownian motion. Each of them has been fluctuating up and down, demonstrating its randomness. It is clear that the trajectory of the Brownian motion is completely different from any continuous and smooth trajectory that we are familiar with.

The non-differentiability means that classical calculus is useless in analyzing the Brownian motion. This was undoubtedly frustrating because people finally come up with a simple random process (to model stock price), but lacked the tools to further study it. However, the pazzle was solved with the development of Ito calculus. It is no exaggeration to say that Ito calculus laid the foundation of modern financial mathematics.

3.3.3 Quadratic Variation

For a partition $\Pi = \{t_0 = 0 < t_1 < \ldots < t_N = T\}$ of an interval $[0, T]$ and a continuous function $f(t)$, its quadratic variation is defined as $\sum_{i=0}^{N-1}[f(t_{i+1}) - f(t_i)]^2$.

if f is continuously differentiable in $[0, T]$, then by using the mean value theorem of classical calculus, it can be shown that :

$\sum_i[f(t_{i+1}) - f(t_i)]^2 \leq \sum_i[t_{i+1} - t_i]^2 f'(s_i)^2 \leq Max_{s\in[0,T]} f'(s)^2 \sum_i[t_{i+1} - t_i]^2$

$\leq Max_{s\in[0,T]} f'(s)^2 Max_i\{t_{i+1} - t_i\} \cdot T$

This implies that as the partition becomes increasingly finer, i.e., as $max_i t_{i+1} - t_i$ approaches 0, the quadratic variation of $f(t)$ converges to zero.

Now, consider replacing $f(t)$ with the non-differentiable Brownian motion $B(t)$. In relation to the quadratic variation of $B(t)$, the following theorem applies:

For a partition $\Pi = \{t_0 = 0 < t_1 < \ldots < t_N = T\}$ of an interval $[0, T]$ let $\mid \Pi \mid = Max_i\{t_{i+1} - t_i\}$. A Brownian motion $B(t)$ satisfies the following equation with certainty:

$Lim_{|\Pi| \longrightarrow 0} \sum_i[B(t_{i+1}) - B(t_i)]^2 = T$

This can be demonstrated through the law of large numbers, which states that as a stochastic process, the quadratic variation of Brownian motion is T, unlike the case for continuous differentiable functions where it is 0.

This implies that the inherent randomness of $B(t)$ causes it to exhibit significant variability, regardless of how small the partitioning interval becomes. The cumulative sum of the fluctuations in $B(t)$ from these tiny intervals does not tend toward 0. Instead, the limit approaches T, which is essentially the length of the interval.

The quadratic variation of the Brownian motion can also be written in the infinitesimal difference form: $(dB)^2 = dt$. This nonzero quadratic variation of the Brownian motion has significant implications in the derivation of Ito's lemma.

3.3.4 Geometric Brownian Motion and Stock Prices

Now, we add a drift μt and a scaling parameter σ. This leads to the Brownian motion with drift : $X(t) = \mu t + \sigma B(t)$.

$X(t)$ follows a normal distribution with mean μt and variance $\sigma^2 t$. In its infinitesimal form, it becomes : $dX(t) = \mu dt + \sigma dB(t)$.

This is a stochastic differential equation. It differs from a regular differential equation in that it has at least one stochastic term (in this case $B(t)$).

Despite incorporating drift into the Brownian motion, it remains suboptimal for modeling stock price movements. This is attributed to the fact that X(t) or B(t) may assume negative values as time progresses, whereas stock prices cannot be negative. Conversely, stock returns can be either negative or positive. Hence, X(t) can be employed to model stock returns.

Consider $S(t)$ as the stock price, and let $dS(t)$ quantify the change in S within an infinitesimal time interval. Consequently, $\frac{dS(t)}{S(t)}$ represents the stock return in this interval, and we can express this as: $\frac{dS(t)}{S(t)} = \mu dt + \sigma dB(t)$.

This gives the SDE of $S(t)$: $dS(t) = \mu S(t)dt + \sigma S(t)dB(t)$.

A stochastic process S(t) that adheres to the stochastic differential equation (SDE) is referred to as a **geometric Brownian motion**. It is commonly employed to model stock prices due to the following reasons:

- **Normal distribution:** Empirical observations indicate that the continuous compound return of stocks approximately conforms to a normal distribution.

- **Markovian property:** The properties of Brownian motion readily demonstrate that S(t) is a Markov process. This implies that the current stock price at time t encompasses all the information necessary for predicting the future, aligning with the weak form of the efficient-market hypothesis.

- **Continuous and non differentiable Paths:** The continuity of $B(t)$ and consequently $S(t)$, coupled with their non-differentiability, aligns with the actual movement patterns observed in stock prices. GBM paths are continuous but not everywhere differentiable. It portrays smooth paths yet exhibit randomness and unpredictability: although the path is continuous in time, you can't find a tangent to the path at almost any point—it's too "rough.

  If a continuous function is differentiable at a specific time point, t, it indicates that *we could ascertain the derivative by observing how the function acts just prior to that point, at some moment slightly before t.*

- **Exponential Growth with Stochastic Fluctuations:** GBM is described mathematically as an exponential function multiplied by a stochastic process, leading to an exponential mean growth trend with stochastic fluctuations.

To analyze the stock price using $S(t)$, it is imperative to solve the aforementioned stochastic differential equation (SDE) and obtain a closed form for $S(t)$. This task can be accomplished with the assistance of Ito calculus.

Note : $E(X(t)) = \mu t$ because $E(B(t)) = 0$. $X(t)$ is denominated by μt, for all $\varepsilon > 0$, after long enough time, $X(t)$ will always be between $y = (\mu - \varepsilon)t$ and $y = (\mu + \varepsilon)t$.

3.3.5 Ito's Lemma basic form

Brownian motion provides a means to examine stock prices. However, concerning financial derivatives, their prices are contingent on underlying assets. Consider a continuous and smooth function of Brownian motion, denoted as $f(B_t)$. In financial mathematics, a significant area of study revolves around understanding the changes in $f(B_t)$ within infinitesimal time intervals, specifically focusing on the properties of df.

Classical calculus is useless in analyzing df, but Ito's calculus effectively resolves the issue and lays a solid foundation for stochastic analysis. To understand this, we need to understand why classical calculus does not work.

1. Why classical calculus does not work, first method :

To find df, where f is a continuous and smooth function of B_t, we apply the chain rule: $df = (\frac{dB_t}{dt} f'(B_t))dt$. Since B_t is not differentiable, the differentiation $\frac{dB_t}{dt}$ does not exit.

2. Why classical calculus does not work, second method :

One potential approach to address this issue is to attempt describing the difference df in terms of the difference dB_t, rather than relying on $\frac{dB_t}{dt}$. As previously mentioned, dB_t has a distinct meaning as it represents the change in B_t over an infinitesimal time interval. We have $df = f'(B_t)dB_t$. In this expression, both $f'(B_t)$ and dB_t can be computed. However, this does not work for the following reason :

Let's consider the Taylor expansion of $f(x)$ and it gives us :

$f(x+\Delta x) - f(x) = f'(x)(\Delta x) + \frac{f''(x)}{2!}(\Delta x)^2 + \ldots + \frac{f^{(n)}(x)}{n!}(\Delta x)^n$

As Δx approaches 0, the dominant term is the first term, while all other terms are of a smaller order of magnitude and can be disregarded. Then $df = f'(x)dx$. **However, this is not true for $x = B_t$:**

$\Delta f = f(B_t + \Delta B_t) - f(B_t) = f'(B_t)(\Delta B_t) + \frac{f''(B_t)}{2!}(\Delta B_t)^2 + \ldots + \frac{f^{(n)}(B_t)}{n!}(\Delta B_t)^n$

The second term can no longer be disregarded due to the quadratic variation, which states $(dB)^2 = dt$. Then, $df(B_t) = f'(B_t)dB_t + \frac{f''(B_t)}{2!}dt$.

This is the basic form of Ito's lemma.

In a broader context, contemplate a smooth function $f(t,x)$ that is dependent on two variables. In traditional calculus, we would obtain:

$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx$

With x replaced by B_t and according to Ito calculus, we obtain :

$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial B_t}dB_t + \frac{\partial^2 f}{2\partial B_t^2}(dB_t)^2$

$df = (\frac{\partial f}{\partial t} + \frac{\partial^2 f}{2\partial B_t^2})dt + \frac{\partial f}{\partial B_t}dB_t$, since $(dB_t)^2 = dt$

Upon comparing the outcomes above, it becomes evident that, due to the non-zero quadratic variation of Brownian motion, obtaining df requires incorporating an additional term to the outcomes derived from classical calculus. This supplementary term involves the second-order derivative of f with respect to B_t (if f depends solely on B_t) or the partial second-order derivative of f with respect to both B_t and t (if f depends on both B_t and t). This conclusion is transformative, allowing the application of calculus in the realm of stochastic processes.

3.3.6 Ito's Lemma general form

As stated in previous part, a Brownian motion with drift and diffusion satisfies the following stochastic differential equation (SDE), where μ and σ are some constants : $dX(t) = \mu dt + \sigma dB(t)$.

In a broader context, the drift and diffusion coefficients can vary as functions of both X(t) and t, rather than being constants. Denote these as $\mu = a(X(t), t)$ and $\sigma = b(X(t), t)$, representing the drift and diffusion coefficients, respectively. We define a stochastic process that satisfies the following SDE the Ito dirft-diffusion process: $dX = a(X(t), t)dt + b(X(t), t)dB$.

Let $f(X(t), t)$ be a continuous and smooth function of $X(t)$ and t. With Ito's lemma, we derive: $df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X}dX + \frac{\partial^2 f}{2\partial X^2}(dX)^2$.

We will substitute substitute $dX = a(X(t), t)dt + b(X(t), t)dB$ into the expansion:

$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X}(adt + bdB) + \frac{\partial^2 f}{2\partial X^2}(adt + bdB)^2$

$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X}(adt + bdB) + \frac{\partial^2 f}{2\partial X^2}(a^2dt^2 + b^2dB^2 + 2abdtdB)^2$

With :

- $\frac{\partial f}{\partial t}dt$: The change in f due to a small change in time.
- $\frac{\partial f}{\partial X}(adt + bdB)$: The change in f due to changes in X, where X follows a stochastic process with drift term adt and diffusion term bdB.
- $\frac{1}{2}\frac{\partial^2 f}{\partial X^2}(a^2dt^2 + b^2dB^2 + 2abdtdB)$: The second-order term accounting for the curvature of f in X, including the interaction between the drift and diffusion components.

We will apply properties of dt and dB

- dt: An infinitesimal increment in time, treated as a very small but non-random quantity.
- dB: An increment of Brownian motion, with $(dB)^2 \approx dt$ and higher-order terms like dB^3 being negligible.
- Higher-order terms of dt like dt^2, $dtdB$ are negligible because they are of a smaller order of magnitude compared to dt and dB.

So, in order to simplify the Expression

- Neglect a^2dt^2 and $2abdtdB$ because they are of higher order and thus negligible.
- Replace b^2dB^2 with b^2dt because $(dB)^2 \approx dt$.

The expression now simplifies to:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X}(adt + bdB) + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}b^2 dt$$

Which can be rearranged as :

$$df = \left(\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial X} + \frac{1}{2}b^2\frac{\partial^2 f}{\partial X^2}\right)dt + b\frac{\partial f}{\partial X}dB$$

Analyzing the stochastic differential equation (SDE) of f reveals that, as a function of X and t, f itself constitutes an Ito drift-diffusion process. Notably, upon comparing the expressions for dX and df, it becomes apparent that both X and f are influenced by the same underlying source of uncertainty, dB. In essence, the randomness inherent in the stochastic processes X and f originates from the same Brownian motion. This observation holds significant importance in the derivation of the Black-Scholes formula.

3.3.7 Geometric Brownian Motion and Ito's Lemma

We represent the price of a certain stock using geometric Brownian motion. If we denote the stock price as S, it adheres to the following stochastic differential equation (SDE): $dS = \mu S dt + \sigma S dB$

Here, μ represents the annual rate of return, and σ is the standard deviation of that return. Consequently, S qualifies as an Ito process with $a = \mu S$ and $b = \sigma S$. To solve for S, consider $f = Ln(S)$ and apply Ito's lemma to df in order to derive the stochastic differential equation for $Ln(S)$.

To derive the stochastic differential equation for $\ln(S)$ using Ito's Lemma, we start from the SDE of the stock price S and apply Ito's Lemma to the function $f(S,t) = \ln(S)$.

For $f(S,t) = \ln(S)$, we apply Ito's Lemma:

$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}dS + \frac{\partial^2 f}{2\partial S^2}(dS)^2$

With :

- $\frac{\partial \ln(S)}{\partial S} = \frac{1}{S}$

- $\frac{\partial^2 \ln(S)}{\partial S^2} = -\frac{1}{S^2}$
- Since $\ln(S)$ does not explicitly depend on t, $\frac{\partial f}{\partial t} = 0$.

Replacing partial derivatives, and dS with its SDE :

$df = 0 \cdot dt + \frac{1}{S} dS - \frac{1}{2}\frac{1}{S^2}(dS)^2$

$df = \frac{1}{S}(\mu S dt + \sigma S dB) - \frac{1}{2}\frac{1}{S^2}(\mu S dt + \sigma S dB)^2$

$df = \mu dt + \sigma dB - \frac{1}{2}\sigma^2 dt$ since $(dt)^2$ and $dt \cdot dB$ are negligible and $(dB)^2$ is approximated by dt.

$df = d(\ln(S)) = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dB$

To understand how $\ln(S_T)$ follows a certain normal distribution, we need to integrate the stochastic differential equation (SDE) for $\ln(S)$ and then apply the properties of Brownian motion.

We will integrate both sides of the equation over the interval $[0, T]$:

$\int_0^T d(\ln(S_t)) = \int_0^T (\mu - \frac{\sigma^2}{2})dt + \int_0^T \sigma dB_t$

Solving the integrals, we get:

$\ln(S_T) - \ln(S_0) = (\mu - \frac{\sigma^2}{2})T + \sigma \int_0^T dB_t$

Properties of Brownian Motion :

- The integral of dB_t over $[0, T]$ is normally distributed with mean 0 and variance T. This is because dB_t represents the increment of a Wiener process (Brownian motion), which has these properties.
- Therefore, $\sigma \int_0^T dB_t$ is normally distributed with mean 0 and variance $\sigma^2 T$.

Rearranging the equation, we find the expression for $\ln(S_T)$:

$\ln(S_T) = \ln(S_0) + (\mu - \frac{\sigma^2}{2})T + \sigma \int_0^T dB_t$

Given that $\ln(S_0)$ is a constant and $\sigma \int_0^T dB_t$ is normally distributed as mentioned, the sum of these terms implies that $\ln(S_T)$ is also normally distributed with mean $\ln(S_0) + (\mu - \frac{\sigma^2}{2})T$ and Variance $\sigma^2 T$

Thus, $\ln(S_T)$ follows a normal distribution $\mathcal{N}\left(\ln(S_0) + (\mu - \frac{\sigma^2}{2})T, \sigma^2 T\right)$, denoted as:

$$\ln(S_T) \sim \Phi \left[\ln(S_0) + (\mu - \frac{\sigma^2}{2})T, \sigma^2 T \right]$$

A variable follows a lognormal distribution if the natural logarithm of the variable conforms to a normal distribution. Consequently, when utilizing a geometric Brownian motion to model stock prices, the stock price adheres to a lognormal distribution.

By integrating the stochastic differential equation (SDE) of $\ln(S)$ and subsequently exponentiating the result, it becomes straightforward to ascertain how S fluctuates with T:

$e^{\ln(S_T)-\ln(S_0)} = e^{\left(\mu-\frac{\sigma^2}{2}\right)T+\sigma\int_0^T dB_t}$

$\frac{S_T}{S_0} = e^{\left(\mu-\frac{\sigma^2}{2}\right)T+\sigma\int_0^T dB_t}$

$S_T = S_0 e^{\left(\mu-\frac{\sigma^2}{2}\right)T+\sigma\int_0^T dB_t}$

Now, recall that $B(T)$ is the value of the Brownian motion at time T, so we can replace $\int_0^T dB_t$ with $B(T)$:

$S_T = S_0 e^{\left(\mu-\frac{\sigma^2}{2}\right)T+\sigma B(T)}$

This final expression represents the geometric Brownian motion model for the stock price at time T.

At first glance, this may appear counterintuitive. We initially assumed that the annual rate of return of the stock is μ. However, in the formula above, if we disregard the randomness from $B(T)$ and solely focus on the first part, it implies that the stock price grows at a rate of $\mu - 0.5\sigma^2$ rather than μ.

The correct interpretation is that $\mu - 0.5\sigma^2$ is the continuously compounded rate of return per annum. This can be clarified as follows. Let x be the continuously compounded rate per annum, and then $S(T)$ follows $S(T) = S(0)e^{xT}$ or $x = \frac{1}{T} \times (\ln S(T) - \ln S(0))$. Since S is lognormally distributed, we know that $\ln S(T) - \ln S(0)$ follows a normal distribution with mean $(\mu - 0.5\sigma^2)T$ and variance $(\sigma^2)T$. We can then solve for x, and it satisfies the following normal distribution: $x \sim \Phi(\mu - \frac{\sigma^2}{2}, \frac{\sigma^2}{T}$ mdr ou

3.4 Monte Carlo Simulations

Monte Carlo simulations are a class of computational algorithms that rely on repeated random sampling to estimate mathematical functions and mimic the operations of complex systems and obtain numerical results. They are widely used to model and analyze complex systems that are probabilistic in nature. Their application in finance is based on the statistical properties of asset returns and the probabilistic nature of market movements. In finance, Monte Carlo simulations are employed to evaluate and predict the behavior of financial instruments, portfolios, and investment strategies. The foundations of Monte Carlo methods in finance are well established in the seminal work of Boyle (1977), who applied the technique to option pricing.

Monte Carlo methods hinge on the Law of Large Numbers and the Central Limit Theorem. The Law of Large Numbers suggests that as the number of trials in a simulation increases, the average of the results obtained from the simulation will converge to the expected value. The Central Limit Theorem underpins the distribution of the sample means, which, for a large number of simulations, will tend to follow a normal distribution regardless of the underlying distribution of the asset returns.

Mathematically, if X is a random variable with a known probability distribution function $f(x)$ and an unknown expected value $E[X]$, then the Monte Carlo estimate $\hat{E}[X]$ based on N independent samples $x_1, x_2, ..., x_N$ is given by:

$$\hat{E}[X] = \frac{1}{N} \sum_{i=1}^{N} x_i$$

The theoretical basis for Monte Carlo simulations is grounded in probability theory and statistical inference. By simulating a large number of scenarios, Monte Carlo methods approximate the solutions to mathematical problems that may be deterministic in principle but are too complex for analytical solutions. Glasserman (2004) provides an extensive exploration of these principles in "Monte Carlo Methods in Financial Engineering."

In the context of derivative pricing, Monte Carlo simulations are particularly useful for valuing instruments with multiple sources of uncertainty or with **path-dependent features**, such as Asian options or American options. The simulation involves creating a large number of potential future paths for the underlying asset prices, discounting the payoffs along each path to present value, and then averaging these values to estimate the derivative's price. Clewlow and Strickland (1998) offer practical insights into implementing these techniques in "Implementing Derivatives Models." In derivative pricing, Monte Carlo simulations are used to estimate the expected payoff of complex derivatives by simulating the underlying asset paths. For a European call option, the expected payoff under risk-neutral valuation can be written as:

$$C_0 = e^{-rT}\hat{E}[\max(S_T - K, 0)]$$

where:

- C_0 is the current option price,
- e^{-rT} is the discount factor at the risk-free rate r over the option's life T,
- $\max(S_T - K, 0)$ represents the option payoff at maturity,
- S_T is the simulated asset price at maturity,
- K is the strike price.

Monte Carlo simulations are highly flexible and can be adapted to a wide variety of financial problems. They are especially advantageous when dealing with high-dimensional problems or when the payoff of an investment is dependent on the path of the underlying variables. However, they can be computationally intensive and may require sophisticated variance reduction techniques to improve efficiency and accuracy. Boyle, Broadie, and Glasserman (1997) discuss these challenges and present advanced methods in their paper "Monte Carlo Methods for Security Pricing." The accuracy of a Monte Carlo simulation in financial models depends on the number of paths simulated and the time steps within each path. Variance reduction techniques such as antithetic variates, control variates, and importance sampling are employed to improve the efficiency and precision of the simulations. For instance, antithetic variates involve running a pair of simulations with negatively correlated paths to reduce variance without increasing the number of simulations.

Key mathematical constructs used in Monte Carlo simulations include stochastic processes like the Geometric Brownian Motion (GBM), represented by:

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

where:

- S_t is the stock price at time t,
- μ is the drift coefficient,
- σ is the volatility coefficient,
- dW_t is the increment of a Wiener process or standard Brownian motion.

For a simulation, the discrete-time approximation of GBM can be used for the asset price S at a future time T based on the current price S_0:

$$S(T) = S_0 \exp\left((\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}Z\right)$$

where Z is a standard normal random variable. This equation reflects the random walk hypothesis and incorporates the randomness of the price movements through the term involving Z.

Monte Carlo simulations have become an indispensable tool in financial analysis, providing a powerful means for valuing complex derivatives and assessing risks under uncertainty. Their use has expanded beyond derivatives pricing to risk management, portfolio optimization, and strategic financial planning. Despite their computational intensity, advancements in computing power and algorithmic efficiency continue to enhance their practicality and precision in financial applications. Monte Carlo simulations provide a potent tool for financial analysts to price derivatives, assess risks, and evaluate investment strategies under uncertainty. By incorporating statistical and probability theories, these simulations offer a nuanced understanding of financial phenomena through numerical experimentation and stochastic modeling.

3.5 Stochastic process in Layman's terms

Stochastic processes describe the random evolution of a variable over time. A Markov process is one in which only the present value is useful for predicting the future distribution.

The history of the variable and how the present value was obtained are irrelevant.

A standard Wiener process z, is a process describing the evolution of a variable whose variations are normally distributed. The drift, or central tendency parameter, of this process is zero, and its variance parameter is equal to 1 per unit time. This means that if the variable is x_0 at date T, it follows a normal distribution with mean x_0 and standard deviation $qrtT$.

A general Wiener process describes the evolution of a normal variable with a drift equal to a and a variance parameter equal to b^2, with a and b constant. If, as before, the variable is equal to x_0 at date 0, then it follows a normal distribution with mean $x_0 + aT$ and standard deviation $b\sqrt{T}$ at date T.

An Itô process is a process x whose drift and variance parameter can be a function of both time and x itself. The variation of x in a very small time interval can be assumed to be normal, to a good approximation, but its variation over a longer period is likely not to be.

Simulating the evolution of a variable can be a good way of gaining an intuitive understanding of its behavior. This method involves dividing the study period into numerous intervals of very short duration and, for each of them, drawing a random number from among the possible trajectories of the variable. The probability distribution of the variable at a future date can then be calculated.

Itô's lemma is a way of determining the stochastic process followed by a function of the variable under study. One of the key points is that the Wiener process, z, governing the stochastic process of the variable under consideration is the same as that governing the evolution of the function dependent on this variable. In other words, both the variable and its function are subject to the same source of uncertainty.

A common assumption is that share prices evolve according to a geometric Brownian movement. Under these conditions, the profitability obtained by the holder of the share, in a short time interval, is distributed according to a normal distribution, and the profitabilities of two non-overlapping periods are independent. The value of the share at a future date follows a lognormal distribution.

3.6 Variance Reduction Techniques in Monte Carlo Simulations

Variance reduction techniques are essential in enhancing the efficiency and accuracy of Monte Carlo simulations, particularly in financial modeling. These techniques are designed to decrease the statistical variance of simulation results, enabling more accurate estimates with fewer simulation runs. Understanding these methods is crucial, as outlined in Glasserman's "Monte Carlo Methods in Financial Engineering" (Glasserman, 2004).

- **Antithetic Variates :** This technique involves generating pairs of negatively correlated random variables to reduce variance. If X is a random variable used in a simulation, the antithetic variable is $-X$. The average of X and $-X$ will have a lower variance than X alone. The variance reduction is given by:

$$\mathrm{Var}\left(\frac{X - X'}{2}\right) = \frac{1 - \mathrm{Cov}(X, X')}{2}\mathrm{Var}(X)$$

  where X' is the antithetic variate of X.

- **Control Variates :** This method uses the known expected value of a related variable to reduce variance. If Y is a control variate with a known expected value μ_Y, and X is the primary variable of interest, then the variance-reduced estimator is:

$$\hat{X}_{cv} = \hat{X} + \beta(\mu_Y - \hat{Y})$$

  where $\hat{X}$ and $\hat{Y}$ are the sample means, and β is a coefficient usually chosen to minimize variance.

- **Importance Sampling :** This approach changes the probability distribution to sample more frequently from the important parts of the domain. The original expectation $E[X]$ is estimated by:

$$E[X] = E[g(X)h(X)]$$

  where $g(X)$ is a new probability density function and $h(X)$ is a weighting function that corrects for this change in distribution.

- **Stratified Sampling :** This technique divides the domain of the random variables into strata and samples each stratum separately. This approach ensures a more uniform sampling across the entire domain. The variance reduction depends on the heterogeneity between the strata.

In financial modeling, these techniques are particularly beneficial for pricing complex derivatives and managing risk in portfolios where direct analytical solutions are infeasible or computationally intensive. They are effectively used in scenarios with high-dimensional integrals or in models sensitive to tail risks.

Variance reduction techniques are fundamental in optimizing Monte Carlo simulations in finance. They enable more accurate and efficient estimations, which are critical in high-stakes financial decision-making and risk assessment.

3.7 Diffusion model

3.7.1 Jump-Diffusion

The jump diffusion model is a financial model used to describe the movement of an asset's price, incorporating both continuous diffusion (Brownian motion) and discontinuous jumps. Traditional models, such as the Black-Scholes model, assume continuous and smooth price movements. However, financial markets often experience sudden and discontinuous movements, known as jumps, due to unexpected events or news. The jump diffusion model accounts for these abrupt changes due to various factors such as earnings announcement or unexpected events in asset prices. The jump diffusion model introduces a more realistic representation of market dynamics by acknowledging the presence of jumps. The jump diffusion model enhances the ability to capture the fat-tailed distribution of returns observed in financial markets. Traditional models often assume normal distribution, but jumps introduce the possibility of extreme price movements, leading to a more accurate representation of actual market behavior. In the context of option pricing, the jump diffusion model is valuable for pricing financial derivatives, such as options, in situations where discontinuous movements in the underlying asset's price can significantly affect the option's value. By accounting for jumps, the model provides more accurate pricing estimates for options.

For all this reason, (Merton, 1976) suggest to add jumps to the diffusion process to better account for the excess kurtosis found in most financial time series, adding a Poisson-distributed 'event' which is the arrival of an important piece of information about the stock. It is assumed that the arrivals are independently and identically distributed. Writing a compound Poisson process:

$$Q_t = \sum_{i=1}^{N_t} \left(\frac{V_{T_i}}{V_{T_{i-}}} - 1\right)$$

where $\frac{V_{Ti}}{V_{Ti-}}$ is the ratio of the price just after and before the jump and we assume that $log(\frac{V_{Ti}}{V_{Ti-}}) \sim \mathcal{N}(m, \sigma^2)$ and it can be demonstrated $E[Q] = \lambda kt$ with λ the average number of jumps per period and the average jump size measued as a percentage of the asset price $k = e^{m+\frac{\delta^2}{2}} - 1$

$$dV_t = V_t((+)dt + \sigma dW_t)V_{t-} - dQ_t$$

A solution of this previous equation is $V_t = V_0 \exp(\mu t + \sigma W_t - \frac{\sigma^2}{2}t - \lambda kt) \prod_{i=1}^{N} {}_t \frac{V_{Ti}}{V_{Ti-}}$

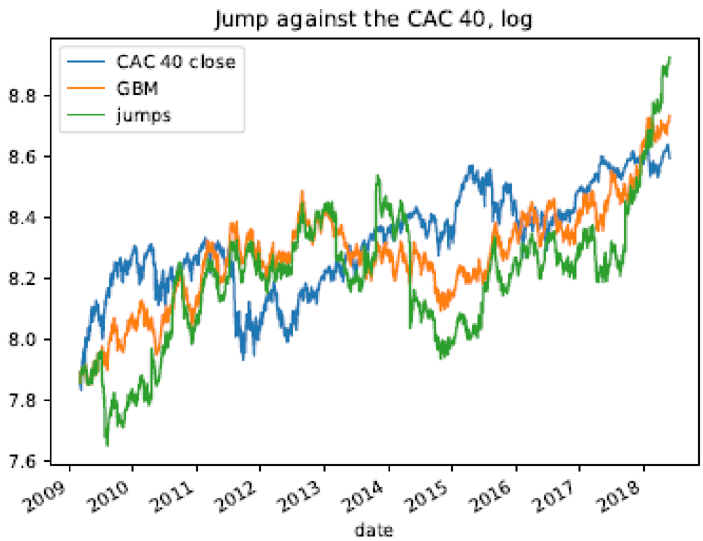
This model can be implemented using Monte Carlo simulations to determine the number and the size of the jumps. For illustration purpose, comparing with a Geometric Brownian Motion taken from the same Wiener process, we provide an arbitrary model with jumps of size 0.5% occurring on average every 100 days and set so that the jump distribution is centered:

Derivation of jump-diffusion model

Suppose X_t is a jump-diffusion process with evolution given by
$X_t = X_0 + \int_0^t a_s\, ds + \int_0^t b_s\, dW_s + \sum_{i=1}^{N_t} \Delta X_i$
where a_t is the drift term, b_t is the volatility term, and ΔX_i corresponds to jump i in the stock price.
Then, according to Cont and Tankov, for a function $f(X_t, t)$:
$df(X_t, t) = \frac{\partial f(X_t,t)}{\partial t}\, dt + a_t \frac{\partial f(X_t,t)}{\partial x}\, dt + b_t^2 \sigma^2 \frac{\partial^2 f(X_t,t)}{\partial x^2}\, dt + b_t \frac{\partial f(X_t,t)}{\partial x}\, dW_t + (f(X_{t-} + \Delta X_t) - f(X_{t-}))$.
Using this theorem for $f(\cdot) = \ln(\cdot)$ and S_t described by the stochastic differential equation (below):
$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t + dJ_t$
We get
$d\ln(S_t) = \frac{\partial \ln(S_t)}{\partial t}\, dt + \mu \frac{\partial \ln(S_t)}{\partial S_t}\, dt + \sigma \frac{\partial \ln(S_t)}{\partial t}\, dW_t + (\ln(J_t S_t) - \ln(S_t))\, \mu \frac{1}{S_t}\, dt +$

$\frac{\sigma^2 S_t^2}{2}\left(1-\frac{1}{S_t^2}\right)dt+\sigma S_t\,dW_t+(\ln(J_t)+\ln(S_t)-\ln(S_t))$
This is solved to give
$\ln(S_t)=\ln(S_0)+\left(\mu-\frac{1}{2}\right)t+\sigma W_t+\sum_{i=1}^{N_t}\ln(J_t)$.

Figure 3.3: Jump diffusion vs. GBM vs. CAC 40



3.7.2 Lévy Process

A Lévy process is a mathematical model used to describe the stochastic dynamics of a random variable over time. It is characterized by the properties of independence, stationarity, and increments with stationary and independent increments. In the context of financial modeling, Lévy processes are particularly relevant for simulating price movements due to their ability to capture both continuous and discontinuous variations in asset prices.

The general form of a Lévy process can be expressed as:

$$X(t)=X(0)+\mu t+\sigma W(t)+\int_{-\infty}^{t}\int_{|y|<1}y\,N(ds,dy)$$

$$X(t) : \text{The value of the process at time } t.$$

$$X(0) : \text{The initial value of the process.}$$

$$\mu : \text{The constant drift.}$$

$$t : \text{The time variable.}$$

$$\sigma : \text{The constant volatility.}$$

$$W(t) : \text{A standard Brownian motion.}$$

$$\int_{-\infty}^{t} \int_{|y|<1} y\, N(ds, dy) : \text{is a Lévy jump process with } N \text{ jumps occurring}[0, t] \text{ with sizes } y_i \text{ and arrival ti}$$

The integral representing a Lévy jump process with N jumps occurring in the time interval $[0, t]$ with sizes y_i and arrival times s_i.

The term involving the integral captures the impact of jumps on the process. The jumps occur at random times and have random sizes, which makes Lévy processes particularly suitable for modeling financial asset prices, where abrupt price changes (jumps) are observed.

The characteristics of Lévy processes make them flexible and powerful for capturing various types of price movements. Commonly used Lévy processes in finance include the following:

- **Brownian Motion (Wiener Process):** Continuous and smooth random motion.
- **Poisson Process:** Discrete jumps occurring at random times.
- **Compound Poisson Process:** Discrete jumps with random sizes occurring at random times.
- **Normal Inverse Gaussian (NIG) Process:** A Lévy process with finite moments and the ability to capture fat-tailed distributions.
- **Variance Gamma (VG) Process:** A process that generalizes Brownian motion and allows for jumps in both volatility and price.

1. **Poisson Process: Discrete jumps occurring at random times:**

- **Definition:** - A Poisson Process is a stochastic process that represents a sequence of discrete events occurring over continuous time. These events, or "jumps," happen at random, and the time between consecutive events follows an exponential distribution.

- **Key Characteristics:**

- **Independence:** The occurrence of events is independent of the past, meaning the process has no memory.
- **Stationarity:** The process has stationary increments, meaning the probability of a jump happening in a small time interval is constant.
- **Memoryless Property:** The time until the next event follows an exponential distribution, implying that the process has no memory of past events.

- **Mathematical Representation:** - If $N(t)$ represents the number of events up to time t, a Poisson Process has the following properties:

$$P(N(t+h) - N(t) = k) = e^{-\lambda h} \frac{(\lambda h)^k}{k!}$$

where λ is the intensity (average rate) of events.

2. **Compound Poisson Process: Discrete jumps with random sizes occurring at random times:**

- **Definition:** - A Compound Poisson Process is an extension of the Poisson Process where each jump has a random size. The jump sizes are typically independent and identically distributed (i.i.d.) random variables.

- **Key Characteristics:**

- **Poisson Process Base:** The occurrence of jumps follows a Poisson Process with a certain intensity.
- **Random Sizes:** Each jump has a random size drawn from a specified distribution.
- **Independence:** The jump occurrences and sizes are independent.

- **Mathematical Representation:** - If $N(t)$ represents the number of jumps up to time t, and Y_i represents the size of the i-th jump, a Compound Poisson Process is often represented as:

$$X(t) = \sum_{i=1}^{N(t)} Y_i$$

where $N(t)$ follows a Poisson Process and Y_i are i.i.d. random variables representing the sizes of the jumps.

- **Example:** - If $N(t)$ follows a Poisson Process with intensity λ and Y_i are i.i.d. exponential random variables with mean μ, then the Compound Poisson Process $X(t)$ is often denoted as:

$$X(t) = \sum_{i=1}^{N(t)} Y_i$$

4 Fundamentals of Derivatives

4.1 Definition and Types of Derivatives

Derivatives are financial instruments whose value is dependent on, or derived from, the value of an underlying asset, group of assets, or benchmark. The primary function of derivatives is to provide a mechanism for price risk management related to fluctuations in the underlying asset. The concept and use of derivatives are thoroughly examined in Hull's foundational text, "Options, Futures, and Other Derivatives" (Hull, 2017).

Fundamental Types of Derivatives :

- **Forwards and Futures :** Both are contracts to buy or sell an asset at a predetermined future date for a price agreed upon today. Futures are standardized contracts traded on exchanges, while forwards are customized contracts traded over-the-counter. The valuation frameworks for these instruments are explored in depth by Kolb and Overdahl in "Futures, Options, and Swaps" (Kolb & Overdahl, 2007).

- **Options :** Options provide the buyer the right, but not the obligation, to buy (call option) or sell (put option) the underlying asset at a specified strike price up to a specified date. The theoretical framework for options pricing, including the seminal Black-Scholes model, is covered extensively by Black and Scholes in their groundbreaking paper, "The Pricing of Options and Corporate Liabilities" (Black & Scholes, 1973).

- **Swaps :** Swaps are contracts in which two parties agree to exchange cash flows or other financial instruments. Common types include interest rate swaps and currency swaps. The valuation and use of swaps in corporate finance are detailed by Smith, Smithson, and Wilford in "Managing Financial Risk" (Smith, Smithson, & Wilford, 1995).

- **Credit Derivatives :** These include financial contracts such as credit default swaps (CDS) and collateralized debt obligations (CDOs), which are used to manage exposure to credit risk. The evolution and impact of credit derivatives are critically reviewed in Tett's "Fool's Gold" (Tett, 2009), which provides an account of their role in the financial crisis.

Apart from the standard derivatives, there are also exotic derivatives, which include complex products such as knock-in/knock-out options, digital options, rainbow options, asian options and compound options. These instruments are often customized for specific hedging or investment strategies. The pricing and risk management of exotic options are explored by Wilmott in "Paul Wilmott Introduces Quantitative Finance" (Wilmott, 2007).

Derivatives are integral to modern financial markets, offering tools for hedging, speculation, and arbitrage. They range from standardized exchange-traded contracts to bespoke over-the-counter agreements tailored to the needs of the parties involved. The academic literature provides a wealth of knowledge on the pricing, application, and regulation of derivatives, reflecting their complexity and significance in the global financial ecosystem.

4.2 Underlying Assets

Underlying assets are the financial assets, indexes, or other economic variables upon which derivatives' values are based. The performance of these assets directly influences the price and execution of derivative contracts. Fabozzi et al. provide a detailed explanation of underlying assets in "The Handbook of Fixed Income Securities" (Fabozzi et al., 2012).

Categories of Underlying Assets :

- **Equities :** Stocks or equity indexes serve as underlying assets for derivatives like equity options and index futures. The dynamics of equity as an underlying asset are analyzed in "Equity Derivatives: Corporate and Institutional Applications" (Cameron, 2017).

- **Fixed Income Securities :** Bonds or interest rates form the underlying assets for fixed income derivatives, including interest rate swaps and bond futures. Tuckman and Serrat elaborate on fixed income instruments and their applications in "Fixed Income Securities: Tools for Today's Markets" (Tuckman & Serrat, 2011).

- **Commodities :** Physical goods such as agricultural products, metals, and energy resources are the underlying assets for commodities derivatives. These instruments are critical for price risk management in sectors where commodity prices are volatile.

"Commodities and Commodity Derivatives" by Geman (2005) offers a comprehensive look at these markets.

- **Foreign Exchange:** Currencies are underlying assets for derivatives like currency swaps and foreign exchange options. The interplay between FX derivatives and international financial markets is covered in "The Forex Options Course" (Cofnas, 2008).
- **Credit :** For credit derivatives, the underlying asset is the creditworthiness of a corporate or sovereign entity. This is reflected in instruments such as credit default swaps. The complexities of credit derivatives are explored in "Credit Derivatives: Understanding and Working with the 2007 ISDA Definitions" (Hardy, 2009).
- **Other Economic Variables :** This can include anything from weather conditions, used in weather derivatives, to housing prices, used in real estate derivatives, the volatility of assets, etc. The broadening scope of derivative underlying assets is discussed in "Derivatives: The Tools That Changed Finance" (Dash, 2004).

The value of derivatives is inherently linked to the price movements of their underlying assets. Pricing models, such as the Black-Scholes model for equity options and the Black model for commodity options, incorporate these variables to determine fair value for derivatives contracts. These models are further scrutinized in "The Concepts and Practice of Mathematical Finance" (Joshi, 2003).

Underlying assets are the cornerstone upon which the derivatives market is built. They provide the foundation for a wide range of financial instruments designed for hedging, speculation, and arbitrage. The diversity of underlying assets reflects the expansive reach of derivatives in global finance.

4.3 Usage and Significance

Derivatives are pivotal financial instruments used by individuals, companies, and governments to manage and allocate risk. Their usage extends beyond simple risk management to include speculation, arbitrage, and market efficiency. The significance of derivatives in modern finance is underscored by the extensive body of research, such as in "Derivatives Markets" by McDonald (2006), which details their functions and impacts.

- **Risk Management :** One of the primary uses of derivatives is to hedge against risks associated with price movements of assets. By locking in prices or rates, derivatives like futures and options can protect against adverse movements in commodity prices, interest rates, exchange rates, and equity markets. The principle of risk hedging through derivatives is elaborated upon by Hull (2017) in "Risk Management and Financial Institutions," where the strategies and instruments used for hedging are thoroughly explained.

- **Speculation and Arbitrage :** Speculators use derivatives to bet on the future direction of market prices to gain profits. Unlike hedging, speculation involves taking on risk with the expectation of a reward. Meanwhile, arbitrageurs use derivatives to exploit price discrepancies in different markets, thus contributing to market efficiency. The dual role of speculation and arbitrage in derivatives markets is analyzed by Chance and Brooks (2015) in "An Introduction to Derivatives and Risk Management."

Derivatives contribute to market efficiency by facilitating the process of price discovery. As information is incorporated into asset prices, derivatives help reveal market sentiments and expectations about future price movements. The linkage between derivatives and market efficiency is explored by Shiller (2003) in "The New Financial Order: Risk in the 21st Century."

Derivatives enable market participation for assets or markets that may be otherwise inaccessible due to various constraints such as high costs, regulations, or geographical barriers. Through derivatives, investors can gain exposure to a wide array of assets, as highlighted by Tett (2009) in the analysis of synthetic assets in "Fool's Gold."

Derivatives are significant in financial innovation, providing new ways to transfer, manage, and price risk. They also contribute to economic growth by allowing for more efficient capital allocation. The broader economic implications of derivatives are considered by Stulz (1996) in "Rethinking Risk Management."

Derivatives play a multifaceted role in the financial system. They are integral tools for risk management, but also serve critical functions in speculation, arbitrage, price discovery, market access, and financial innovation. The breadth of their usage underlines their significance in fostering a robust and dynamic financial landscape.

4.4 Hedging fundamentals

Hedging is a risk management strategy employed to offset risks in financial markets. It involves taking a position in a financial instrument that is inversely correlated with an existing position in order to mitigate potential losses. This strategy is particularly significant in the realm of derivatives trading.

The concept of hedging is grounded in financial theories such as Modern Portfolio Theory (MPT) and the Capital Asset Pricing Model (CAPM). MPT, introduced by Harry Markowitz in his 1952 paper "Portfolio Selection" (Journal of Finance), suggests that diversification can reduce portfolio risk. CAPM, developed by Sharpe (1964), Lintner (1965), and Mossin (1966), further extends this idea by quantifying the trade-off between risk and return.

Hedging Techniques :

- **Delta Hedging :** This involves offsetting the price risk inherent in an options position by taking an opposite position in the underlying asset. The delta of an option measures the rate of change of the option's price relative to a change in the underlying asset's price. Delta hedging seeks to make the overall position delta-neutral.
- **Basis Risk Hedging :** This is commonly used in commodity markets. The basis is the difference between the spot price of an asset and the futures price. A basis risk hedge involves taking an opposite position in the futures market to offset the risk of price fluctuation in the spot market.
- **Interest Rate Hedging :** Employed to mitigate the risk of changing interest rates. Instruments like interest rate swaps, where two parties exchange interest rate payments on a specified principal amount, are used.
- **Currency Hedging :** Involves taking positions in foreign exchange derivatives to protect against currency risk, particularly useful for companies dealing in multiple currencies.

The effectiveness of hedging strategies in practice depends on factors like market volatility and the correlation between the hedging instrument and the underlying asset. Studies such as "Derivatives and Corporate Risk Management: Participation and Volume Decisions in the Insurance Industry" by Cummins, Phillips, and Smith in the "Journal of Risk and Insurance" (2001) provide empirical insights into hedging practices in specific sectors.

While hedging can reduce risk, it's not without its limitations and costs. As noted in the "Journal of Economic Perspectives" (2003) by Stulz, hedging can lead to a false sense of security and potentially increase risk if not properly managed.

Hence, Hedging is a nuanced and vital component of financial risk management, deeply rooted in economic theory and widely applied in various financial sectors. Its successful implementation requires a thorough understanding of both the instruments involved and the market dynamics.

5 No arbitrage theory for pricing

5.1 The time value of money

Everyone pretty much understands that money today is worth more than money tomorrow (with the notable exception of some recent negative interest rate environments in Europe).

5.2 Risk-neutral valuation

In finance, the risk-neutral world is a theoretical framework used for pricing financial derivatives, such as stock options. The concept assumes that investors are indifferent to risk when making investment decisions. In other words, they are willing to accept a certain return without requiring compensation for taking on risk.

Risk-neutral valuation says that when valuing derivatives like stock options, you can simplify by assuming that all assets grow—and can be discounted—at the risk-free rate.

The risk-neutral world simplifies the valuation of financial instruments by assuming that investors are risk-neutral, allowing for more straightforward mathematical models. This concept is particularly useful in the context of the Black-Scholes-Merton (BSM) model, a widely used model for pricing European-style options.

- **Risk-Neutral Probability:** In the risk-neutral world, the expected return on any asset is equal to the risk-free rate. This implies that the probability-weighted average of possible future payoffs is discounted at the risk-free rate.
- **Risk-Free Rate:** The risk-free rate serves as the discount rate in the risk-neutral world. This rate represents the return on an investment with zero risk, typically associated with government bonds.
- **Martingale Property:** The risk-neutral world assumes that the price of the underlying asset follows a stochastic process and possesses the martingale property. This means that, under the risk-neutral probability measure, the expected value of the future price is equal to the current price, discounted at the risk-free rate.
- **Derivative Pricing:** In the context of options, the risk-neutral pricing approach simplifies the valuation process. The expected payoff of the option is calculated using

the risk-neutral probability measure, and then this expected payoff is discounted back to the present value using the risk-free rate.

- **Black-Scholes-Merton Model:** The BSM model is a prominent example of a model based on the risk-neutral world assumption. It provides a formula for calculating the theoretical price of European-style options, assuming constant volatility, a risk-free rate, and efficient markets.

- **European vs. American Options:** The risk-neutral world is more applicable to European-style options, which can only be exercised at expiration. American-style options, which can be exercised at any time before expiration, introduce additional complexities.

It's important to note that the risk-neutral world is a theoretical construct, and in reality, investors do care about risk. However, the risk-neutral approach is a useful tool for simplifying option pricing calculations and gaining insights into the factors influencing option prices.

5.3 Arbitrage free

An arbitrage opportunity a zero-cost strategy (i.e risk free strategy) which has non negative pay-off in all state of nature and strictly positive pay-off in at least one state.

5.4 Complete market

A complete market in finance refers to a situation where there exists a set of financial securities that allows for the replication of any possible payoff or cash flow pattern at any point in the future. In a complete market, investors can create portfolios using the available securities to match precisely the payoffs of any other investment or combination of investments. In a complete market, it is not possible to find any arbitrage opportunities, meaning there is no way to create a portfolio that guarantees a profit without incurring any risk. Prices in a complete market reflect all available information, and there is no room for any undetected mispricing or anomalies.

5.5 European and American options

5.5.1 European call options

A *European call option* on the asset S^i is a contract where the seller promises to deliver the risky asset S^i at the maturity T for some given *exercise price or strike*, $K > 0$. A time T, the buyer has the possibility (and not the obligation) to exercise the option, i.e. to buy the risky asset from seller at strike K. Of course, the buyer would exercise the option only if the price which prevails at time T is larger than K. Therefore, the gain of the buyer out of this contract is:

$$B = (S^i_T - K)^+ = max(S^i_T - K, 0)$$

i.e. if the time T price of the asset S^i is larger than the strike K, then the buyer receives the payoff $S^i_T - K$ which corresponds to the benefit from buying the asset from the seller of the contract rather than on the financial market. If the time T price of the asset S^i is smaller than the strike K, the contract is worthless for the buyer.

5.5.2 European put options

A *European put option* on the asset S^i is contract where the seller promises to purchase the risky asset S^i at the maturity T for some given *exercise price* or *strike*, $K > 0$. At time T, the buyer has the possibility, and not the obligation, to exercise the option, i.e to sell the risky asset at strike K. Of course, the buyer would exercise the option only if the seller at strike K. Of course, the buyer would exercise the option only if the price which prevails at time T is smaller than K. Therefore, the gain of the buyer out of is this contract is:

$$B = (K - S^i_T)^+ = max(K - S^i_T, 0)$$

i.e. if the time T the price of the asset S^i is smaller than the strike K, then the buyer reveives the payoff $K - S^i_T$ which corresponds to the benefit from selling the asset to the seller of the contract rather than on the financial market. If the time T price o the asset

S^i is larger than the strike K, the contract is worthless for the buyer, as he can sell risky asset for a larger price on the financial market.

5.5.3 American call options

An *American call (resp put) option* with maturity T and strike $K > 0$ differs from the corresponding European contract in that it offers the possibility to be exercised at any time before maturity (and not only at the maturity).

The seller of a derivative requires a compensation for the risk that he is bearing. In other words, the buyer must pay the price or the premium for the benefit of the contract.

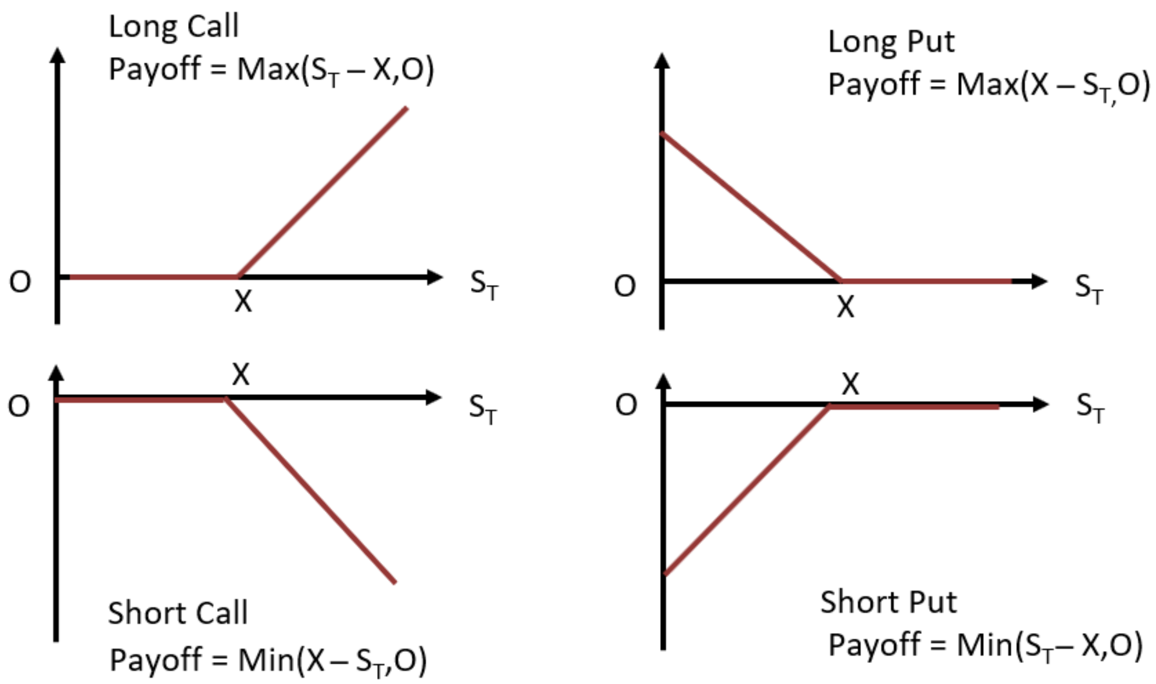
At every time $t < T$, The American call price (resp put) option is denoted by $C(t, S_t, T, K)$ and $P(t, S_t, T, K)$ and European call option (resp put) option is denoted by $c(t, S_t, T, K)$ and $p(t, S_t, T, K)$

The intrinsic value of the call and the put options are respectively: $C(t, S_t, T, K) = c(t, S_t, T, K) = (S_t - K)^+$ $P(t, S_t, T, K) = p(t, S_t, T, K) = (K - S_t)^+$

i.e the value received upon immediate exercising the option (payoff). An option is said to be **in-the-money** (resp. *out-the-money* if its intrincis value is positive. If $K = S_t$, the option is said to be *at-the-money*. Thus a call option is *in-the-money* if $S_t > K$, while a put option is in *in-the-money* if $S_t < K$;

5.6 Payoff Graphs

Figure 5.1: Payoff graph



5.7 Exotic option

More complex options, known as exotic options, some examples of which are given below, are often denominated directly in terms of their pay-off.

- Barrier options: payment takes place (or does not take place) if the underlying has exceeded a contractual level (the barrier) before that date. contractual level (the barrier) before that date. $H_t = (S_T - K)^+ 1_{M_t < B}$ où $M_T = max(S_u)$ avec $t \geq u \geq T$ The advantage of this option is that it is cheaper than the standard call but offers very similar guarantees in a normal situation (if the barrier is high enough).

Like futures and forwards, options can be used to transfer risk, to take bets on the evolution of underlying assets, and possibly to exploit potential arbitrage, while offering much greater freedom and flexibility in these uses. The following two examples illustrate the use of options to control and limit the risks associated with stock market fluctuations.

Example : Protective put :

A protective put is a combination of a long position in an asset and a put option on the same asset. The pay-off at option expiry is then given by $H_T = (K - S_T)^+ + S_T = max(S_T, K)$. This arrangement makes it possible to limit the pay-off to a desired level K.

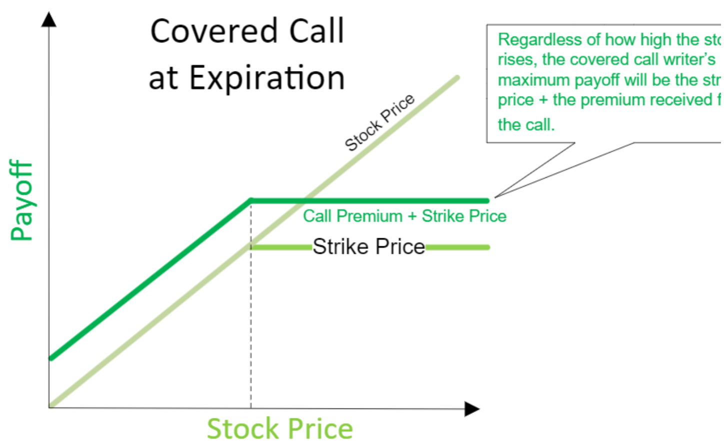
5.8 Optional strategy

Buying a **Call** allows you to bet on the underlying's rise, and buying a **Put** on its fall, with much more leverage, but also much more risk than buying the underlying itself.

5.8.1 Covered Call

The simplest option strategy is the covered call, which simply involves writing a call for stock already owned. If the call is unexercised, then the call writer keeps the premium, but retains the stock, for which he can still receive any dividends. If the call is exercised, then the call writer gets the exercise price for his stock in addition to the premium, but he foregoes the stock profit above the strike price. If the call is unexercised, then more calls can be written for later expiration months, earning more money while holding the stock. A more complete discussion can be found at Covered Calls.

Figure 5.2: Covered Call



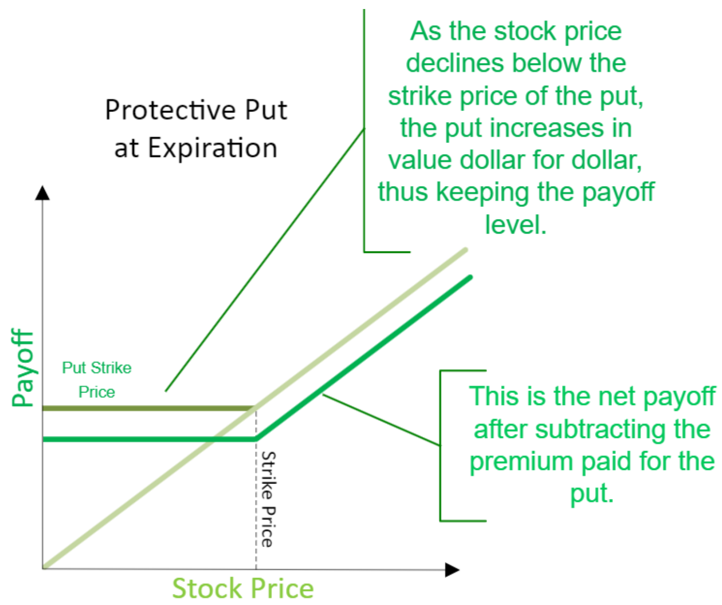
Example: Covered Call

On October 6, 2006, you own 1,000 shares of Microsoft stock, which is currently trading at \$27.87 per share. You write 10 call contracts for Microsoft with a strike price of \$30 per share that expire in January, 2007. You receive .35 per share for your calls, which equals \$35.00 per contract for a total of \$350.00. If Microsoft doesn't rise above \$30, you get to keep the premium as well as the stock. If Microsoft is above $30 per share at expiration, then you still get $30,000 for your stock, and you still get to keep the $350 premium.

5.8.2 Protective Put

A stockholder buys protective puts for stock already owned to protect his position by minimizing any loss. If the stock rises, then the put expires worthless, but the stockholder benefits from the rise in the stock price. If the stock price drops below the strike price of the put, then the put's value increases 1 dollar for each dollar drop in the stock price, thus, minimizing losses. The net payoff for the protective put position is the value of the stock plus the put, minus the premium paid for the put.

Figure 5.3: Protective Put



5.8.3 Example: Protective Put

Using the same example above for the covered call, you instead buy 10 put contracts at $0.25 per share, or $25.00 per contract for a total of $250 for the 10 puts with a strike of $25 that expires in January, 2007. If Microsoft drops to $20 a share, your puts are worth $5,000 and your stock is worth $20,000 for a total of $25,000. No matter how far Microsoft drops, the value of your puts will increase proportionately, so your position will not be worth less than $25,000 before the expiration of the puts — thus, the puts protect your position.

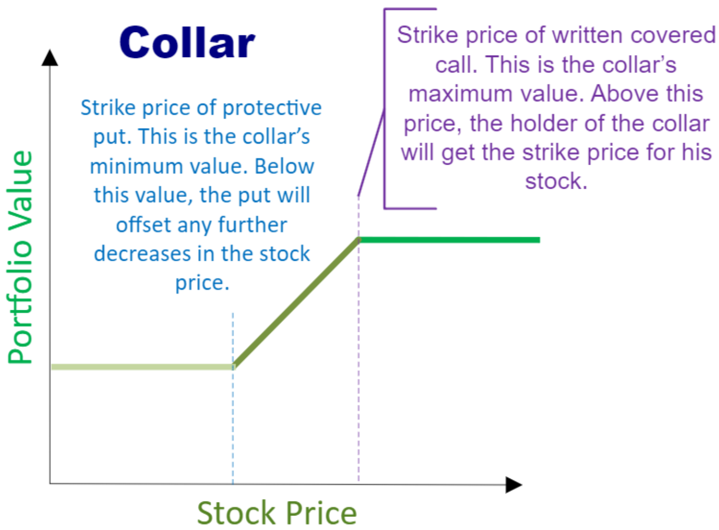
Protective Put Payoff = Stock Value + Put Value Put Premium

5.8.4 Collar

A collar is the use of a protective put and covered call to collar the value of a security position between 2 bounds. A protective put is bought to protect the lower bound, while a call is sold at a strike price for the upper bound, which helps pay for the protective put. This position limits an investor's potential loss, but allows a reasonable profit. However, as with the covered call, the upside potential is limited to the strike price of the written call.

Collars are one of the most effective ways of earning a reasonable profit while also protecting the downside. Indeed, portfolio managers often use collars to protect their position, since it is difficult to sell so many securities in a short time without moving the market, especially when the market is expected to decline. However, this tends to make puts more expensive to buy, especially for options on the major market indexes, such as the S&P 500, while decreasing the amount received for the sold calls. In this case, the implied volatility for the puts exceeds that for the calls.

Figure 5.4: Collar Put



5.8.5 Bull and Bear spread

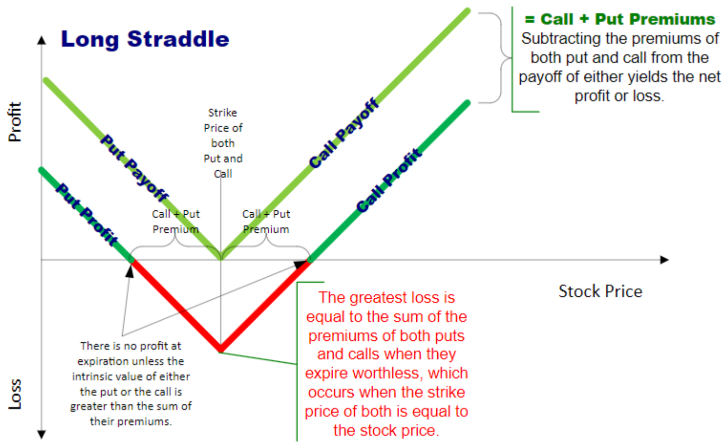
Bull spread (buy a Call strike $K1$, sell a Call strike $K2 > K1$, same expiry) and Bear spread (sell a Put strike $K1$, buy a Put strike $K2 > K1$, same expiry) are also directional bets; they have the advantage of being less expensive than the options themselves, but the gains are limited.

5.8.6 Straddles

A long straddle is established by buying both a put and call on the same security at the same strike price and with the same expiration. This investment strategy is profitable if the stock moves substantially up or down, and is often done in anticipation of a big movement in the stock price, but without knowing which way it will go. For instance, if an important court case will soon be decided that will substantially impact the stock price, but whether the price will rise or fall is not known beforehand, then the straddle would be a good investment strategy. The greatest loss for the straddle is the premiums paid for the put and call, which will expire worthless if the stock price doesn't move enough.

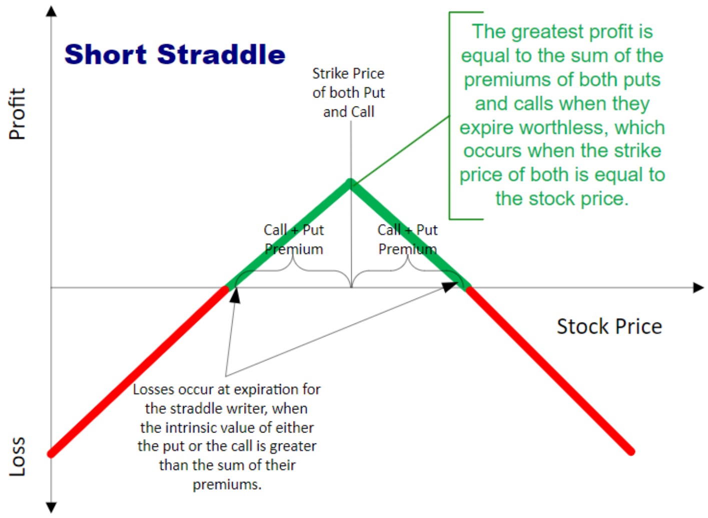
To be profitable, the price of the underlying must move substantially before the expiration date of the options; otherwise, they will expire either worthless or for a fraction of the premium paid. The straddle buyer can only profit if the value of either the call or the put exceeds the cost of the premiums of both options.

Figure 5.5: Long Straddle



A short straddle is created when one writes both a put and a call with the same strike price and expiration date, which one would do if she believes that the stock will not move much before the expiration of the options. If the stock price remains flat, then both options expire worthless, allowing the straddle writer to keep both premiums.

Figure 5.6: Short Straddle



Example: Long and Short Strangle

Merck is embroiled in potentially thousands of lawsuits concerning VIOXX, which was withdrawn from the market. On October, 31, 2006, Merck's stock was trading at $45.29, near its 52-week high. Merck has been winning and losing the lawsuits. If the trend goes one way or the other in a definite direction, it could have a major impact on the stock price, and you think it might happen before 2008, so you buy 10 puts with a strike price of $40 and 10 calls with a strike price of $50 that expire in January, 2008. You pay $2.30 per share for the calls, for a total of $2,300 for 10 contracts. You pay $1.75 per share for the puts, for a total of $1,750 for the 10 put contracts. Your total cost is $4,050 plus commissions. On the other hand, your sister, Sally, decides to write the strangle, receiving the total premium of $4,050 minus commissions.

Let's say, that, by expiration, Merck is clearly losing; it's stock price drops to $30 per share. Your calls expire worthless, but your puts are now in the money by $10 per share, for a net value of $10,000. Therefore, your total profit is almost $6,000 after subtracting the premiums for the options and the commissions to buy them, as well as the exercise commission to exercise your puts. Your sister, Sally, has lost that much. She buys the 1,000 shares of Merck for $40 per share as per the put contracts that she sold, but the stock is only worth $30 per share, for a net $30,000. Her loss of $10,000 is offset by the $4,050 premiums that she received for writing the strangle. Your gain is her loss. (Actually, she lost a little more than you gained, because commissions have to be subtracted from your gains and added to her losses.) A similar scenario would occur if Merck wins, and the stock rises to \$60 per share. However, if, by expiration, the stock is less than \$50 but more than \$40, then all of your options expire worthless, and you lose the entire \$4,050 plus the commissions to buy those options. For you to make any money, the stock would either have to fall a little below \$36 per share or rise a little above \$54 per share to compensate you for the premiums for both the calls and the puts and the commission to buy them and exercise them.

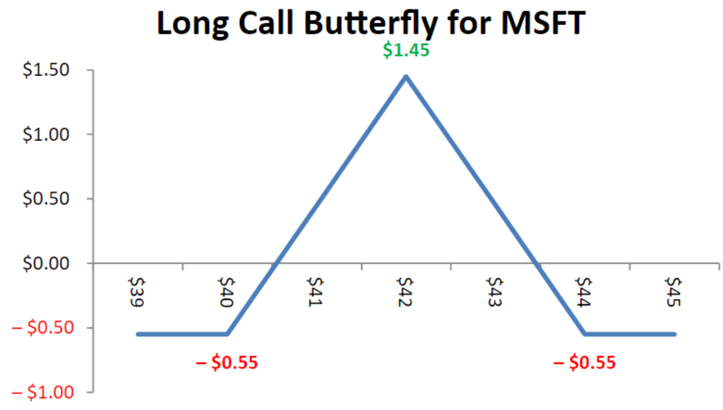
5.8.7 Butterfly

The Butterfly spread (purchase of a Call with strike $K - h$ and a Call with strike $K + h$; sale of 2 Calls with strike K, all with the same expiry date | where strike K is typically equal to the underlying's present value) is used to bet on a fall in volatility: the pay-off is maximum if the underlying remains close to its value.

A butterfly spread consists of either all calls or all puts at 3 consecutive strike prices. The 2 intermediate options share the same strike price, and have a position, either short or long, that is opposite of the outer strikes. The butterfly (aka fly) spread takes its name from the shape of the graph, where the 2 inner options are considered the body of the butterfly while the outer options are considered its wings. A butterfly can also be viewed as 2 adjacent vertical spreads where the intermediate options share the same strike price. Likewise, a butterfly can also be viewed as a short straddle bounded by a long strangle, or vice versa.

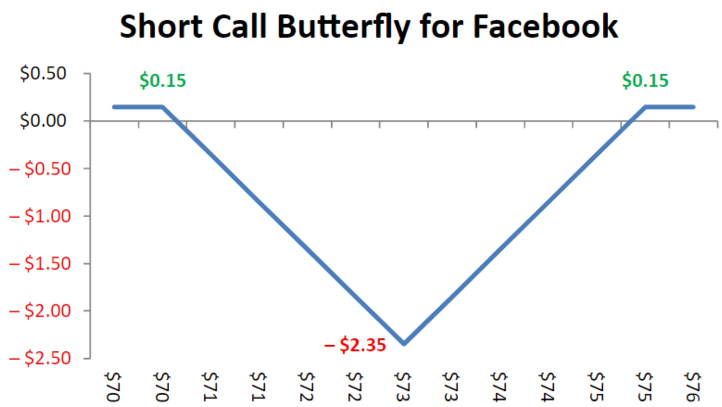
A long butterfly is established, as they say, by buying the wings and selling the body. With a long call butterfly, the long lower call is generally in the money, which is offset by the cost of the 2 middle calls, which are sold. To limit upside risk from the 2 short options, another long call is bought at a higher strike. With a long put butterfly, the highest strike put is generally in the money, while the lowest strike put is bought to offset the risk of the inner short puts.

Figure 5.7: Long Butterfly



The short butterfly profits when the underlying stock price is expected to be either lower than the bottom strike or higher than the top strike and is established by selling the 2 outer options and buying the 2 inner options. The maximum profit = the credit received for establishing the short butterfly.

Figure 5.8: Short Butterfly

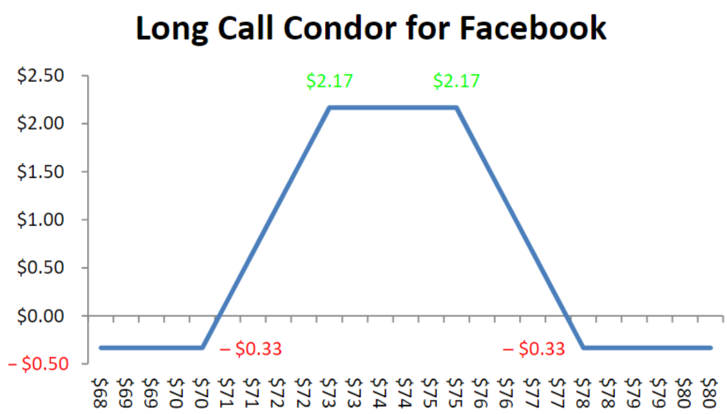


5.8.8 Condor

The condor option strategy is so-called because it is considered to have wider wings that results from using options with 4 consecutive strikes instead of the 3 used in a butterfly. The condor has wider breakeven points and can remain profitable over a longer range of the underlying stock price. However, the maximum profit will be less than for an equivalent butterfly.

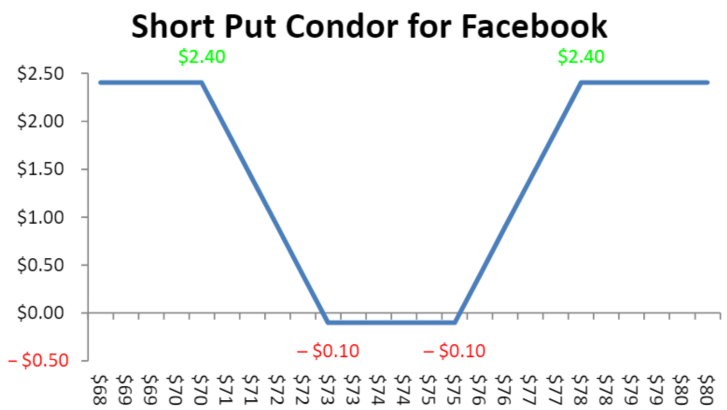
A long condor is a nondirectional market strategy consisting of all calls or all puts, where the 2 inner options are at consecutive strike prices and the lower outer long option is bought at the strike price below the 2 inner that are sold and another long option at the next strike price above those that are sold. If some strike prices are skipped between the inner short options and the outer long options, then this strategy is called a pterodactyl, for its wider wingspan. A long condor can also be thought of as being 2 verticals, a combination of a bull vertical and a bear vertical.

Figure 5.9: Long Call Condor



A short condor, like the short butterfly, is used when the underlying price is expected to move sharply upward or downward. The maximum risk occurs when the market meanders, without direction.

Figure 5.10: Short Put Condor



5.8.9 Calendar spread

The calendar spread (purchase of a call with strike K and maturity $T1$ and sale of a call with strike K and maturity $T2 < T1$) also allows you to bet on rising volatility.

5.9 No dominance principle

In the context of options, the "principle of non-dominance" is a concept from decision theory and economics. It is based on the idea that when a series of options is presented, none of them should dominate the others by being systematically better from all points of view. This principle is important to ensure a fair and balanced set of choices.

For example, Notice that, choosing to exercise the American option at the maturity T provides the same payoff as the European counterpart. Then the portfolio consisting of a long position in the American option and a short position in the European counterpart has at least a zero payoff at the maturity T. It then follows from the dominance principle that American calls and puts are at least as valuable as their European counterparts:

$C(t, S_t, T, K) \geq c(t, S_t, T, K)$ and $C(t, S_t, T, K) \geq c(t, S_t, T, K)$

5.10 Put Call parity

The equation $CP = SK \exp(-r(T-t))$ establishes a connection between European call options C and put options P having the same strike price K and expiration date T at time t, originating from arbitrage.

By simultaneously purchasing a call option, selling a put option, and engaging in a short sale of the underlying stock, we can evaluate the potential payoffs at the options' expiration.

1. If the stock price is above the strike price $S > K$ at expiration, the call option will yield a payoff of $S - K$, the put option will be worthless 0, and the short stock position will pay $-S$. The net payoff will be $(S - K + 0 - S) = -K$.

2. If the stock price is below the strike price (S < K) at expiration, the call option will be worthless 0, the short stock position will pay $-S$, and the short put option will have a payoff of $-(K - S)$. The net payoff will be $(0 - S - K + S) = -K$.

Reminder: a long put is the right but not the obligation to sell at the strike price the underlying. To get this right, someone needs to have an obligation (the put writer), to buy at the strike price if the long put holder exercises its right to ... sell the underlying at the same strike price. So if the put ends up in the money the writer of the put will have to pay the difference $K - S$ to its counterpart (the long put holder).

As a consequence, regardless of the stock price at expiration, the portfolio will consistently have a total value of $-K$.

This guaranteed amount known at expiration must be discounted back using the risk-free interest rate r.

Thus, we arrive at the put-call parity equation:

$$C - P - S = S - K\exp(-r(T-t))$$

Put-call parity is a fundamental concept in options trading, revealing the relationship between European call and put options with the same strike and expiration.

Notice that this argument is specific to European options. We shall see in fact that the corresponding result does not hold for American options.

5.11 Parameter dependence of option prices

- The price of a Call is decreasing in relation to the strike (and the price of a Put is increasing)

  $\mathrm{K}_1 \leq K_2 \implies Call(T, K_1) \geq Call(T, K_2)$

  This property derives from the existence of the Bull spread strategy, which consists of buying a call with strike K_1 and selling a call with strike K_2. As this strategy has a positive payoff, its price must be positive. Furthermore, the payoff of a Bull spread or Bear spread is bounded by $|K_2 - K_1|$.

- Calls/Puts prices are convex in relation to the strike. This property corresponds to the Butterfly spread strategy. We check that this strategy also has a positive payoff in all states of nature, which implies convexity.

- The price of a Call increases with maturity: $T_1 \geq T_2$ implies $Call(T_1, K) \leq Call(T_2, K)$. This property corresponds to the Calendar spread strategy: buy a Call with maturity T_2 and sell a Call with the same strike with maturity T_1. At date T_1, this strategy has a positive payoff.

6 Equity Derivatives Pricing Models

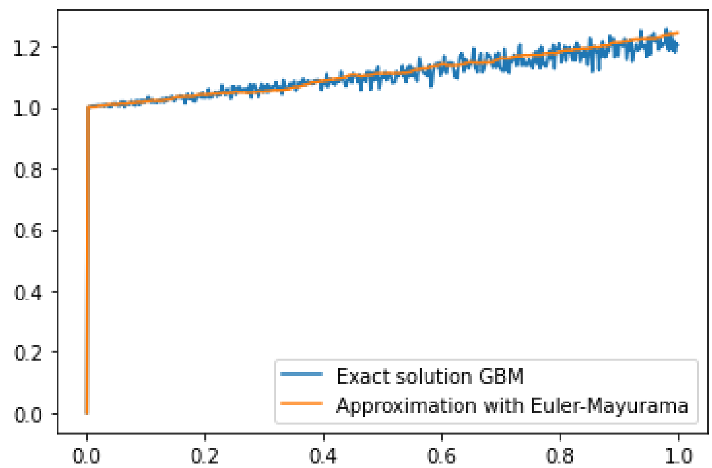
6.1 Euler-Mayurama Metho (EM)

The Euler-Maruyama method is a numerical scheme to approximate the solution of Geometric Brownian Motion SDE. The method is an extension of the classical Euler method used for ordinary differential equations.

In the context of asset pricing, we suggest to use the Euler-Maruyama method to generate a signal that satisfy the Black and Scholes model, using the Itô stochastic integral:

$$V_j = V_{j-1} + \mu V_{j-1}dt + \sigma V_{j-1}[W(t_{j-1+dt}) - W(t_{j-1})]$$

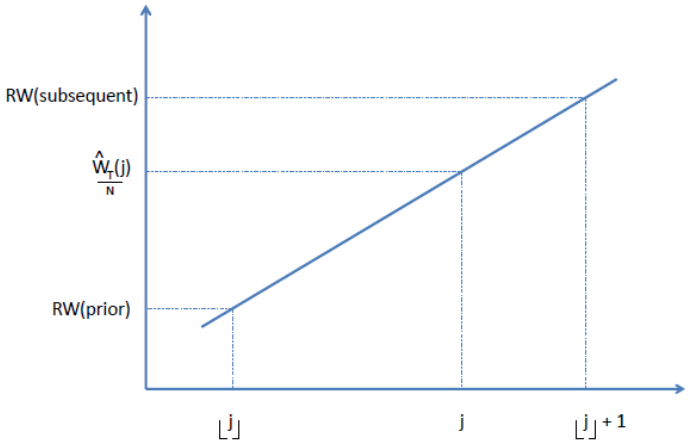
Figure 6.1: Euler-Mayurama approximation of the GBM



and $t_j - t_{j-1} = dt$. With the properties of a Wiener process, we can be demonstrated that $W(t_{j-1)} = W(dt) \approx \frac{\hat{W}(Ndt)}{\sqrt{N}}$ with a piecewise linear function of the interpolation of a random fortune.

Here to approximate a Wiener process for a period T decomposed in M time steps, we create a fortune RW of a given length. At each point x, we find the position just before and just after and approximate our Wiener process with a linear interpolation and we use the property that $\hat{W}(\frac{T}{N}j) = \sqrt{\frac{T}{N}}\hat{W}_{\frac{T}{N}}(j)$

Figure 6.2: EM Interpolation

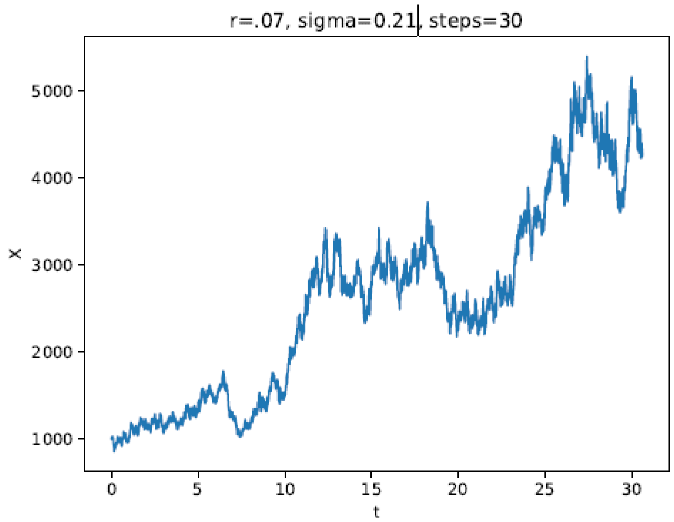


We have the for the linear interpolation:

$$\frac{\hat{W}_{\frac{T}{N}}(j) - RW(prior)}{j - |j|} = \frac{RW(subsequent) - RW(prior)}{|j| + 1 - |j|}$$

One realization we can get with the annual return and standard deviation of the CAC 40:

Figure 6.3: Annual return of the CAC 40 with Euler-Mayurama method



6.2 Finite Element Methods Analysis

Cox, Ingersoll, and Ross (1985) showed that under the assumption that the underlying asset follows a generalized Wiener process, the option pricing problem will be the solution to a resulting parabolic partial differential equation. Even though option pricing problems continue to increase in complexity, OTC markets continue to create innovative option structures that preclude analytic solutions. In a study of several popular numerical techniques, Geske and Shastri (1985) concluded that no one method is best, each has advantages and disadvantages in regard to certain applications. Practitioners generally favor lattice-based methods, as they are intuitive and fairly efficient. Finite difference methods and binomial models are generally easy to specify for most option problems with one or two state variables. In most cases, the results from these methods are reasonably accurate for most applications. Certain exotic options prove difficult for standard lattice-based methods to handle, however. For instance, path dependency or extreme nonlinearity of the pricing function may be difficult to handle with standard lattice methods.

Brennan and Schwartz (1978) provided a discussion of both the explicit and implicit finite difference methods as they relate to one-dimensional option pricing problems. Geske and Shastri (1985) provided a comparison of several lattice approaches, including the explicit and implicit finite difference methods, concluding that the explicit finite difference method with log transformation is the most efficient method for valuing large numbers of options. Hull and White (1990) modified the explicit finite difference method to guarantee convergence when valuing derivative securities.

Finite Element Analysis (FEA) is a numerical technique used in engineering and physics to analyze the behavior of structures, components, or materials under various conditions. It involves dividing a complex system into smaller, simpler elements and then solving equations for each element to simulate the overall behavior of the entire system. In fact, the approaches developed to solve PDEs can be clustered into several groups. Among them, *Finite Difference Methods (FDMs)*, *Finite Element Methods (FEMs)*, and *Finite Volume Methods (FVMs)* are numerical schemes generally adopted. These three methods simplify the computation of numerical solutions to PDEs by breaking down the problem into smaller elements. By solving these smaller components, one can approximate the overall solution to the PDE numerically.

Mathematical Formulation

- FEM: FEM formulates the problem by discretizing the domain into finite elements and approximating the solution using piecewise polynomial functions over these elements.

- FDM: FDM discretizes the domain in a regular grid, replacing derivatives with finite difference approximations.
- FVM: FVM divides the domain into control volumes and approximates the solution by integrating the governing equations over these volumes.

6.3 Finite Element Method (FEM)

For finite element analysis to perform its necessary simulations, a mesh – containing millions of small elements that together form the shape of a structure – must be created. Calculations must be performed on every single element; the combination of each of these individual answers provides the final result for the full structure.

In FEM, the domain of interest is divided into smaller, non-overlapping subdomains called elements. These elements are connected at discrete points called nodes, forming a mesh. The behavior of the system within each element is approximated using a set of interpolation functions, known as shape functions. The PDE is then transformed into a system of algebraic equations by applying variational principles, and these equations are solved numerically. FEM often involves transforming the original PDE into a weak or variational form, making it well-suited for problems with mixed boundary conditions and a wide range of material properties. In the FEM, the domain of the problem is divided into subdomains or finite elements. Each finite element relates values in the domain of the element to nodal values on the boundary and interior of the element. Commonly, the relationship between nodal values and values in the element domain are expressed as polynomial interpolation functions or shape functions. The FEM is done in such a way as to ensure (at least) the approximating function is continuous where the elements meet (nodal boundary points). Element shape functions may be specified to allow continuity of the derivatives for the function as well.

In the FEM, the solution function u to the set of differential equations is replaced with an approximate function of the form:

$$u \approx \sum j = 1 N u_j N_j$$

where u_j is the approximate value of u (or possibly derivative) at each element node and N_j is the interpolation function (usually a piecewise polynomial function). A difference between the FEM and other numerical methods, such as the method of finite differences or classical variational methods, is that the approximation for u is established on subdomains or finite elements.

Over each element, the function of interest, u, is related to nodal points, u_i (for a finite number i), through interpolation functions, also known as shape functions. The number and location of nodal points per element will depend on the geometry of the element (dimensions, shape, and properties of the problem to be retained) and the type of interpolation. Generally, the interpolation used is a polynomial expansion based on the coordinate system (known variables). The number of terms used in the polynomial expansion will determine the number of nodal unknowns needed. The shapes, dimensions, and properties of the problem that need to be retained will help determine the location of the nodes for the element

6.4 Finite Difference Method (FDM)

Finite-difference methods (FDM) is the generic term for a large number of techniques that can be used for solving differential equations, approximating derivatives with finite differences. It consists of a discretization of both spatial (underlying price in our case) and time intervals to obtain a finite grid. We find solutions for a differential equation by approximating every partial derivative with numerical methods and we apply the solution to the points of the grid.

The finite difference technique tends to converge faster than lattices and approximates complex exotic options very well.

The PDE is approximated at each grid point using finite difference approximations for the derivatives. The original PDE is replaced with a system of algebraic equations based on these approximations, and these equations are solved iteratively. FDM directly discretizes the original PDE, which can be advantageous for simpler problems with well-defined boundary conditions.

To solve a PDE by finite differences working backward in time, a discrete-time grid of size M by N is set up to reflect asset prices over a course of time. So S and t take on the following values at each point on the grid:

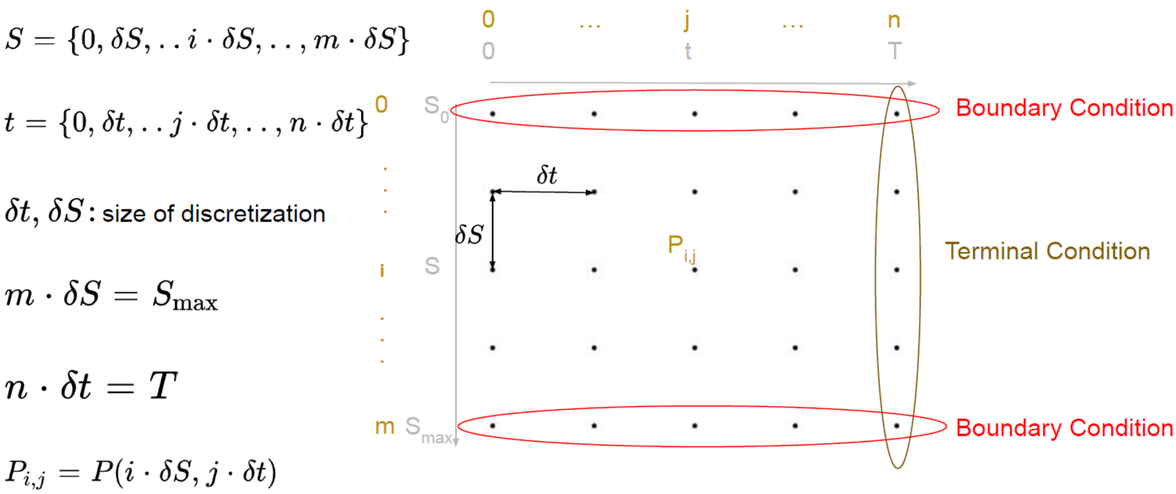
It follows that by grid notation $P_{i,j} = P(i \cdot \delta S, j \cdot \delta t$. S_{max} is a suitable large asset price that cannot be reached by the maturity time T. Thus δS and dt are intervals between each node in the grid, incremented price and time respectively. The terminal condition at expiration time T for every value of S is $max(S-K,0)$ for a call option and $max(K-S,0)$ for put option. The grid traverses backward from the terminal conditions, complying with the PDE while adhering to the boundary conditions of the grid, such as the payoff from an earlier exercise.

The boundary conditions are defined valued at the extreme ends of the nodes, where $i = 0$ and $i = N$ for every time at t. Values at the boundaries are used to calculate the values of all other lattice nodes ly using PDE.

A visual representation of the grid is given below. As i and j increase from the top-left corner of the grid, the price S tends towards S_max (the maximum price possible) at the bottom-right corner of the grid

Dimensional Grid

Figure 6.4: Dimensional grid



1st derivative

$$f(x+h) = f(x) + h \cdot f'(x) + \frac{h^2}{2} \cdot f''(x) + ... + \frac{h^n}{n!} \cdot f^n(x) + O(h^{n+1})$$

$O(h^{n+1})$ is the order of the error

$$f'(x) = \frac{f(x+h) - f(x)}{h} + O(h)$$

This is the forward difference approximation of the first derivative

Similarly the backward difference approximation of the first derivative is as follow:

$$f(x+h) = f(x) - h \cdot f'(x) + \frac{h^2}{2} \cdot f''(x) + ... + \frac{h^n}{n!} \cdot f^n(x) + O(h^{n+1})$$

$$f'(x) = \frac{f(x) - f(x-h)}{h} + O(h)$$

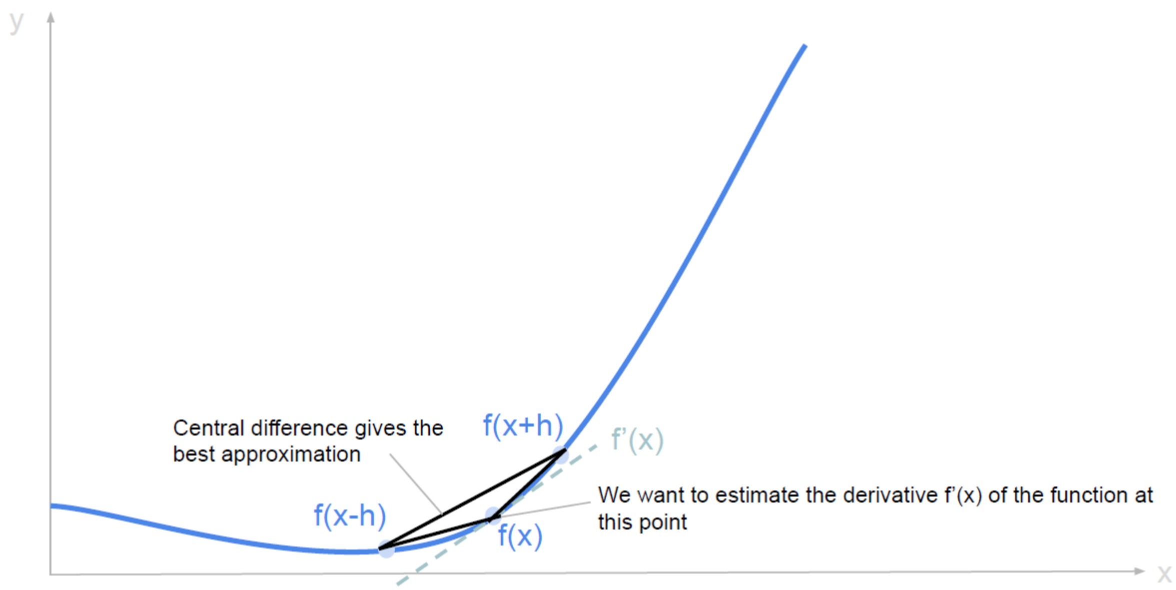
By subtracting the two developments and dividing by 2 we get:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2)$$

This is the central difference approximation for the first derivative

It is the of three approximation with an error in h^2 for the two others

Figure 6.5: Central difference



2nd derivative

We also get the approximation of the second derivatives by adding the two developments and dividing by h^2.

$$f''(x) = \frac{f(x+h) - 2 \cdot f(x) + f(x-h)}{h^2} + O(h^2)$$

This is the second-order central difference approximation for the second derivative. The error is in h^2.

Summary

By taking the notation of the dimensional grid we have:

- **Forward difference:**

$$\frac{dP}{dS} = \frac{P_{i+1,j} - P_{i,j}}{dS}, \frac{dP}{dt} = \frac{P_{i,j+1} - P_{i,j}}{dt}$$

- **Backward difference:**

$$\frac{dP}{dS} = \frac{P_{i,j} - P_{i-1,j}}{dS}, \frac{dP}{dt} = \frac{P_{i,j} - P_{i,j-1}}{dt}$$

- **Central difference:**

$$\frac{dP}{dS} = \frac{P_{i+1,j} - P_{i-1,j}}{2dS}, \frac{dP}{dt} = \frac{P_{i,j+1} - P_{i,j-1}}{2dt}$$

- **2nd derivative:**

$$\frac{d^2P}{dS^2} = \frac{P_{i+1,j} - 2P_{i,j} + P_{i-1,j}}{dS^2}$$

6.4.1 Fininite Difference to approximate Black-Scholes

Figure 6.6: Central difference for B-S

From the Black Scholes PDE, we need approximations for:

$$\frac{\partial P}{\partial t} + r \cdot \frac{\partial P}{\partial S} \cdot S + \frac{1}{2} \cdot \sigma^2 \cdot S^2 \cdot \frac{\partial^2 P}{\partial S^2} = r \cdot P$$

Forward Difference	$\frac{\partial P}{\partial t} = \frac{P_{i,j+1} - P_{i,j}}{\delta t}$	$\frac{\partial P}{\partial S} = \frac{P_{i+1,j} - P_{i,j}}{\delta S}$		
Backward Difference	$\frac{\partial P}{\partial t} = \frac{P_{i,j} - P_{i,j-1}}{\delta t}$	$\frac{\partial P}{\partial S} = \frac{P_{i,j} - P_{i-1,j}}{\delta S}$		
Central Difference	$\frac{\partial P}{\partial t} = \frac{P_{i,j+1} - P_{i,j-1}}{2 \cdot \delta t}$	$\frac{\partial P}{\partial S} = \frac{P_{i+1,j} - P_{i-1,j}}{2 \cdot \delta S}$	Second-order Central Difference	$\frac{\partial^2 P}{\partial S^2} = \frac{P_{i+1,j} - 2 \cdot P_{i,j} + P_{i-1,j}}{\delta S^2}$

Using grid notations

Figure 6.7: Terminal conditions

At its expiration, the price of an option is equal to its final payoff:

	Continuous	**Discrete**
	$P(S,T) = f(S_T)$	$P_{i,n} = f(i \cdot \delta S)$
Call Option	$P(S,T) = (S_T - K)^+$	$P_{i,n} = (i \cdot \delta S - K)^+$
Put Option	$P(S,T) = (K - S_T)^+$	$P_{i,n} = (K - i \cdot \delta S)^+$

Figure 6.8: Boundary conditions

Dirichlet Boundary Condition (Type I)

Call Option

$$P(0,t) = 0 \qquad P(S_{\max},t) = S_{\max} - K \cdot e^{-r\cdot(T-t)}$$

$$P_{0,j} = 0 \qquad P_{m,j} = m \cdot \delta S - K \cdot e^{-r\cdot(n-j)\cdot\delta t}$$

Put Option

$$P(0,t) = K \cdot e^{-r\cdot(T-t)} \qquad P(S_{\max},t) = 0$$

$$P_{0,j} = K \cdot e^{-r\cdot(n-j)\cdot\delta t} \qquad P_{m,j} = 0$$

Neumann Boundary Condition (Type II)

$$\frac{\partial^2 P}{\partial S^2}(\delta S, t) = 0 \qquad \frac{\partial^2 P}{\partial S^2}((m-1)\cdot\delta S, t) = 0$$

$$P_{0,j} - 2 \cdot P_{1,j} + P_{2,j} = 0 \qquad P_{m-2,j} - 2 \cdot P_{m-1,j} + P_{m,j} = 0$$

Once we have the boundary conditions set up, we can now apply an iterative approach using the explicit, implicit or Cranck-Nolson method.

6.4.2 The explicit method

We will try solve the Balck-Scholes PDE:

$$rf = \frac{df}{dt} + rS\frac{df}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2 f}{dt^2}$$

The explicit method for approximating $f_{i,j}$ is given by the following Black-Scholes equation:

$$rf_{i,j} = \frac{f_{i,j} - f_{i,j-1}}{dt} + ridS\frac{f_{i+1,j} - f_{i-1,j}}{2ds}\frac{1}{2}\sigma^2 j^2 \frac{f_{i+1,j} + f_{i-1,j}}{dS^2}$$

Here, we can see that the first difference is the backward difference with respect to t, the second difference is the central difference with respect to S, and the third difference is the second-order difference with respect to S. When we rearrange the terms, we get the following equation:

$$f_{i,j} = a_i^* f_{i-1,j+1} + b_i^* f_{i,j+1} + c_i^* f_{i+1,j+1}$$

Where:

$$j = N-1, N-2, N=3, ..., 2, 1, 0$$

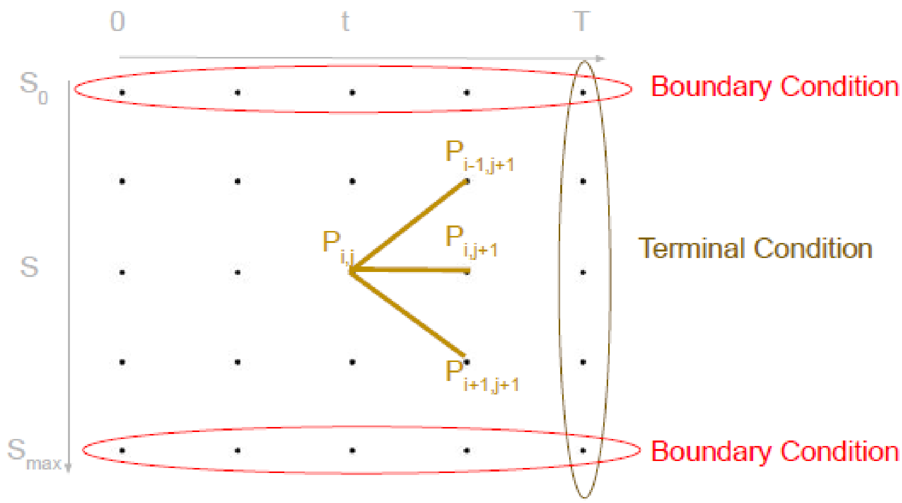
$$i = 1, 2, 3, ..., M-2, M-1$$

Then:

$$a_i^* = \frac{1}{2}dt(\sigma^2 i^2 - ri) b_i^* = 1 - dt(\sigma^2 i^2 - ri) c_i^* = \frac{1}{2}dt(\sigma^2 i^2 + ri)$$

cd

Figure 6.9: Explicit method



6.4.3 The implicit method

The stability problem of the explicit method can be overcome using the forward difference with respect to time. The implicit method for approximating $f_{i,j}$ is given by the following equation

$$rf_{i,j} = \frac{f_{i,j+1} - f_{i,j}}{dt} + ridS\frac{f_{i+1,j} - f_{i-1,j}}{2ds}\frac{1}{2}\sigma^2 j^2 \frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{dS^2}$$

Here, we can see that the first difference is the backward difference with respect to t, the second difference is the central difference with respect to S, and the third difference is the second-order difference with respect to S. When we rearrange the terms, we get the following equation:

$$f_{i,j} = a_j f_{i-1,j} + b_i f_{i,j} + c_i f_{i+1,j+1}$$

Where:

$$j = N-1, N-2, N=3, ..., 2, 1, 0$$

$$i = 1, 2, 3, ..., M-2, M-1$$

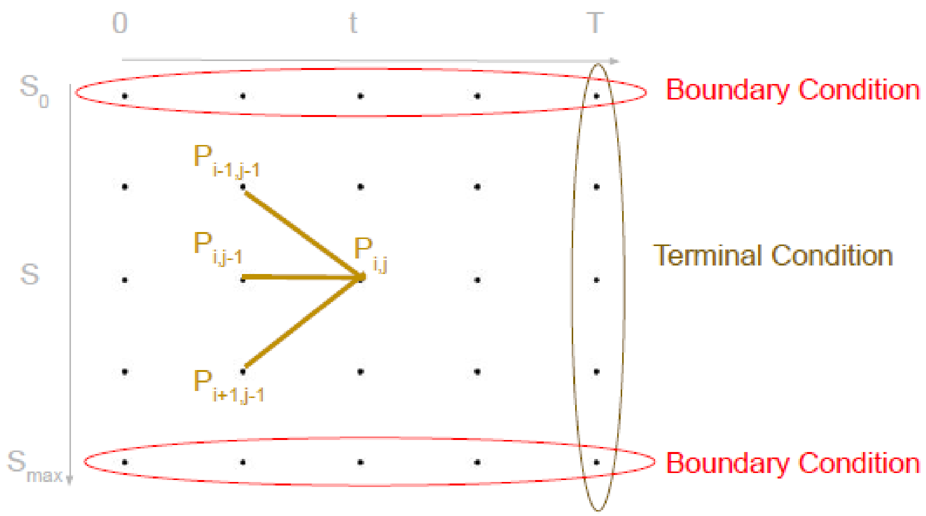
Then:

$$a_i = \frac{1}{2}(ridt - \sigma^2 i^2 dt)$$

$$b_i = 1 + \sigma^2 i^2 dt + rdt$$

$$c_i = \frac{1}{2}(ridt + \sigma^2 i^2 dt)$$

Figure 6.10: Implicit method



6.4.4 Cranck-Nicolson method

Another way of avoiding the instability issue, as seen in the explicit method, is to use the Cranck-Nicolson method. The Cranck-Nicolson method converges much more quickly using a combination of the explicit and implicit methods, taking the avarage both. This leads to the following equation:

$$\frac{1}{2}rf_{i,j-1}+\frac{1}{2}rf_{i,j}=\frac{f_{i,j}-f_{i,j-1}}{dt}\frac{1}{2}ridS(\frac{f_{i+1,j-1}-f_{i-1,j-1}}{2dS})+\frac{1}{2}ridS(\frac{f_{i+1,j}-f_{i-1,j}}{2dS})$$

$$+\frac{1}{4}\sigma^2 i^2 dS^2(\frac{f_{i+1,j-1}-2f_{i,j-1}+f_{i-1,j-1}}{ds^2})$$

$$+\frac{1}{4}\sigma^2 i^2 dS^2(\frac{f_{i+1,j}-2f_{i,j}+f_{i-1,j}}{dS^2})$$

This equation can also be rewritten as follows:

$$-\alpha_i f_{i-1,j-1}+(1-\beta_i)f_{i,j-1}-\gamma_i f_{i+1,j-1)}=\alpha f_{i-1,j}+(1-\beta_i)f_{i,j-1}-\gamma_i f_{i+1,j}$$

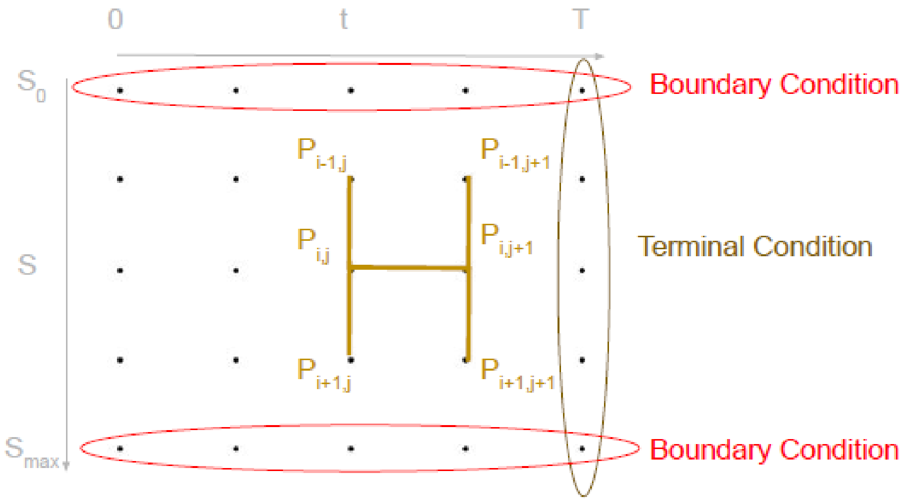
where:

$$\alpha_i=\frac{1}{4}dt(\sigma^2 i^2-ri)$$

$$\beta_i=\frac{1}{2}dt(\sigma^2 i^2+ri)$$

$$\gamma_i=\frac{1}{4}dt(\sigma^2 i^2+ri)$$

Figure 6.11: Cranck-Nicolson method



6.5 FEM vs. FDM vs. FVM

Mesh

- **FEM:** FEM allows for flexible mesh generation, enabling the use of unstructured meshes. This makes it suitable for complex geometries.
- **FDM:** FDM typically uses structured meshes with a uniform grid, which makes it simpler to implement but less flexible for complex geometries.
- **FVM:** FVM can handle both structured and unstructured meshes, providing flexibility for complex geometries while maintaining simplicity for structured grids.

Accuracy and Convergence

- **FEM:** FEM generally provides higher accuracy due to its ability to use higher-order polynomial approximations. It can achieve high convergence rates.
- **FDM:** FDM accuracy depends on the choice of grid spacing, and it typically has lower convergence rates than FEM.
- **FVM:** FVM accuracy depends on the control volume size and the accuracy of flux calculations. It can achieve moderate convergence rates.

As a matter of fact FEM, when properly applied, has advantages over other discretization methods (e.g., the method of finite differences). Some advantages of the FEM (vs.

the standard finite difference method) include its flexibility to effectively accommodate irregular domains (in particular, nonuniform meshes), the ability to directly estimate derivative values (delta and gamma), and the ability of higher order elements to better estimate nonlinear pricing functions.

6.6 Black-Scholes-Merton's Constant Volatility Model (1973)

6.6.1 Black-Scholes assumptions

The Black-Scholes-Merton model serves as the foundational tool for pricing options. Investment banks use this model to determine the fair value of options, including calls and puts, on various underlying assets like stocks, indices, and commodities.

In this section, we present the Black-Scholes differential equation. It's noteworthy that the term 'stochastic' is absent here, indicating that the BS equation lacks randomness. This is advantageous because uncertainty is generally undesirable to obtain closed form formulas.

The assumptions utilized in deriving the BS differential equation, as outlined by Hull (2011), include:

1. European options are under consideration, allowing exercise only at maturity.
2. The stock price adheres to geometric Brownian motion.
3. Short selling of securities with full use of proceeds is permitted.
4. No transactions costs or taxes exist, and all securities are perfectly divisible.
5. No dividends are paid during the life of the derivative.
6. The standard deviation σ of the return of the underlying security is known and constant throughout the option's lifespan.
7. No riskless arbitrage opportunities are present.
8. Security trading occurs continuously.
9. The risk-free rate of interest, r, remains constant and consistent across all maturities.

While acknowledging that some assumptions may not align with reality, they are generally accepted for the purpose of option pricing. Since the inception of the BS model, significant advancements in derivative pricing have been made in academia. Numerous improved models have been developed to rectify various assumptions in the BS model. The subsequent discussion employs the European call option as an example to introduce the BS differential equation.

6.6.2 Black-Scholes PDE derivation

Model Setup :

- **Stock Price (S):** Follows geometric Brownian motion, $dS = \mu S dt + \sigma S dB$, where μ is the drift, σ is the volatility, and B is a Wiener process (or Brownian motion).
- **Risk-Free Bond (B):** Grows exponentially at the risk-free rate, $dB = rBdt$, where r is the risk-free rate.
- **Option Price (V):** Function of stock price and time, $V = V(S, t)$.

Applying Ito's lemma to $V(S, t)$ yields:

- Taylor's expansion : $dV = \frac{V}{\partial t}dt + \frac{V}{\partial S}dS + \frac{{}^2V}{2\partial S^2}dS^2$
- Replacing dS by its formula : $dV = \frac{V}{\partial t}dt + \frac{V}{\partial S}(\mu S dt + \sigma S dB) + \frac{{}^2V}{2\partial S^2}(\mu S dt + \sigma S dB)^2$
- Developing $dV = \frac{V}{\partial t}dt + \frac{V}{\partial S}(\mu S dt + \sigma S dB) + \frac{{}^2V}{2\partial S^2}(\mu^2 S^2 dt^2 + 2\mu\sigma S dt dB + \sigma^2 S^2 dB^2)$
- Removing higher order negligible elements, and replacing $dB^2 = dt$:

  $dV = \frac{V}{\partial t}dt + \frac{V}{\partial S}(\mu S dt + \sigma S dB) + \frac{{}^2V}{2\partial S^2}(\sigma^2 S^2 dt)$
- Rearranging the equation : $dV = \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S \frac{\partial V}{\partial S}dB$

As discussed in the '**Stochastic Process introduction**' of this document, the function f of an Ito process X is also an Ito process. Crucially, the uncertainty associated with both f and X stems from the same underlying Brownian motion. Consequently, the uncertainty in ΔV and ΔS arises from the same ΔB. It implies that a portfolio comprising the stock and this call option can be structured to eliminate this Brownian motion. This Risk-Free portfolio Π is constructed by short selling an option and buying an amount Δ of the stock.

- The portfolio value is : $\Pi = \Delta S - V$

- The change in portfolio value is $d\Pi = \Delta dS - dV$
- The owner of this portfolio is short one derivative and holds a quantity of shares of the stock equal to $\frac{\partial V}{\partial S}$. It is evident that with these weights, the Brownian motion ΔB is entirely offset in this portfolio. This strategy of forming a portfolio is known as Delta hedging.

In order to get the portfolio dynamics, we will substitute dS and dV into $d\Pi$:

$$d\Pi = \Delta(\mu S dt + \sigma S dB) - \left[\left(\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dB\right]$$

In order to make the portfolio risk-free, we will choose Δ such that the term dB is eliminated :

$$d\Pi = \Delta(\mu S dt + \sigma S dB) - \left[\left(\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dB\right]$$

$$d\Pi = \Delta \sigma S dB - \sigma S \frac{\partial V}{\partial S} dB + \Delta \mu S dt - \left(\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt$$

$$d\Pi = \sigma S dB(\Delta - \frac{\partial V}{\partial S}) + \Delta \mu S dt - \left(\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt$$

Hence, in order to eliminate dB, we must choose $\Delta = \frac{\partial V}{\partial S}$.

The Risk-Free portfolio becomes :

$$d\Pi = \frac{\partial V}{\partial S}(\mu S dt + \sigma S dB) - \left[\left(\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dB\right]$$

Which can be simplified as (with stochastic term dB eliminated) :

$$d\Pi = -\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt$$

As dB is absent in the $d\Pi$ expression, it solely depends on time. By shorting one option and taking a long position in $\frac{\partial V}{\partial S}$ shares of stock, we have mitigated the risk within the time interval dt, forming a riskless portfolio.

In an arbitrage-free market, this portfolio is expected to yield the risk-free return of r. hence,

$$\frac{d\Pi}{dt} = r\Pi \iff \frac{d\Pi}{dt} = r(\Delta S - V) \iff \frac{d\Pi}{dt} = r(\frac{\partial V}{\partial S} S - V)$$

$$d\Pi = r(\frac{\partial V}{\partial S} S - V) dt$$

Substituting $d\Pi$ by its value, we get :

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

Rearranging the above equation gives the Black-Scholes PDE :

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

By employing delta hedging, we have eliminated the stochastic term in this equation, rendering it a standard (partial) differential equation instead of a stochastic differential equation. The solution to this differential equation necessitates appropriate boundary conditions. In the case of a European call option, the boundary condition stipulates that at $t = T$ (the option's maturity), C must adhere to the condition $C = max(S(T) - K, 0)$, where K represents the strike price.

To solve this PDE, a first approach is to apply the appropriate boundary conditions for a call or put option :

- Call Option Boundary Conditions :
  - At expiry $t = T$, $C(S, T) = \max(S - K, 0)$.
  - As $S \to \infty$, $C(S, t) \to S$.
  - As $S \to 0$, - $C(S, t) \to 0$.
- Put Option Boundary Conditions :
  - At expiry $t = T$, $P(S, T) = \max(K - S, 0)$.
  - As $S \to \infty$, $P(S, t) \to 0$.
  - As $S \to 0$, $P(S, t) \to K - S$.

Using pure mathematical methods, resolving the PDE with these boundary conditions for a call and a put option leads to the Black-Scholes formulas.

However, we can also find the Black and Scholes formula using the Expectation of the payoff.

6.6.3 Black-Scholes Formula derivation

Re-reading : a complement of the proof can be found here bc now does not make a lot of sense : https://frouah.com/finance%20notes/Black%20Scholes%20Formula.pdf

To calculate the expectation of the price of a European call option under the Black-Scholes model, we use the payoff formula for a call option and then apply the risk-neutral valuation principle. The payoff of a European call option at maturity is given by: $C = \max(S_T - K, 0)$. Where C is the call option payoff, S_T is the stock price at maturity, and K is the strike price of the option. We need to calculate the expected value of this payoff under the risk-neutral probability measure and then discount it back to present value using the risk-free interest rate.

The stock price at maturity S_T in the Black-Scholes model is given by:

$$S_T = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z\right)$$

where S_0 is the current stock price, r is the risk-free interest rate, σ is the volatility, T is the time to maturity, and Z is a standard normally distributed random variable.

Hence, the B&S Payoff can be written :

$$C = \max(S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z\right) - K, 0)$$

To calculate $\mathbb{E}[C]$, we need to integrate this payoff function over the range of possible values of Z, weighted by the probability density of Z. This involves solving an integral where the lower limit corresponds to the value of Z that makes $S_T = K$.

The expected payoff is the average value of the payoff function, weighted by the probability distribution of S_T. This involves integrating the payoff function over the distribution of S_T:

$$\mathbb{E}[C_T] = \int_{-\infty}^{\infty} \max\left(S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z\right) - K, 0\right) f_Z(z)\, dz$$

where $f_Z(z)$ is the probability density function of the standard normal distribution.

This integral can be simplified by recognizing that the payoff is zero for values of Z that result in $S_T \leq K$. Let's denote the critical value of Z, where $S_T = K$, as Z^*. Then, the integral is only over the range where $Z > Z^*$. The payoff function $\max(S_T - K, 0)$ yields a positive value only when $S_T > K$. This means that the option has value only when the stock price at maturity is greater than the strike price. If $S_T \leq K$, the option is not exercised and the payoff is zero.

We need to find the critical value Z^* for which $S_T = K$. Setting $S_T = K$ and solving for Z, we get:

$$K = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z^*\right)$$

Taking the natural logarithm of both sides:

$$\ln(K) = \ln(S_0) + (r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z^*$$

Rearranging for Z^*:

$$Z^* = \frac{\ln(K/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

Given that the payoff is zero for $S_T \leq K$, we only need to integrate over the range where $Z > Z^*$. Therefore, the integral becomes:

$$\mathbb{E}[C_T] = \int_{Z^*}^{\infty} \left(S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z\right) - K\right) f_Z(z)\, dz$$

In the real world, investors require a risk premium, and assets have expected returns greater than the risk-free rate. However, for pricing derivatives like options, we use the risk-neutral measure, which assumes that all investors are indifferent to risk. Under this measure, all assets are assumed to grow at the risk-free interest rate r.

In a risk-neutral world, the expected return on a risky asset (like a stock) is the risk-free rate. This simplification is justified by the argument that any deviation from the risk-free rate can be hedged away in a complete market (as per the Black-Scholes assumptions). Thus, the expected return under the risk-neutral measure is used for pricing derivatives.

Under the risk-neutral measure, the expected return μ in the stock price dynamics is replaced by the risk-free rate r. So the stock price at maturity in the Black-Scholes model becomes:

$$S_T = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z\right)$$

This adjustment is already incorporated in the formula we are using.

Now, we calculate the expected payoff under this risk-neutral probability measure. The integral from Step 5 is evaluated under this measure:

$$\mathbb{E}^Q[C_T] = \int_{Z^*}^{\infty} \left(S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z\right) - K\right) f_Z(z)\, dz$$

Where $\mathbb{E}^Q[\cdot]$ denotes the expectation under the risk-neutral measure Q, and $f_Z(z)$ is still the standard normal density function.

By adopting the risk-neutral valuation approach, we align the option pricing process with the fundamental theorem of asset pricing, which states that in a market without arbitrage opportunities, there exists a risk-neutral probability measure under which all assets' discounted expected future payoffs (under this measure) equal their current prices.

To compute $\mathbb{E}^Q[C_T]$, the expected payoff of a European call option under the risk-neutral measure, we will follow the integral approach that was outlined in the previous steps. Recall that the expected payoff is given by:

$$\mathbb{E}^Q[C_T] = \int_{Z^*}^{\infty} \left(S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z\right) - K\right) f_Z(z)\, dz$$

The integral can be split into two parts, one for each term in the max function. The first part involves the integral of the stock price's exponential term, and the second part involves the integral of the strike price K. However, since K is a constant, its integration over Z simplifies to K multiplied by the cumulative normal distribution of Z^*.

Thus, the expected payoff becomes:

$$\mathbb{E}^Q[C_T] = S_0 \int_{Z^*}^{\infty} \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z\right) f_Z(z)\, dz - K \int_{Z^*}^{\infty} f_Z(z)\, dz$$

The first integral represents the present value of receiving the stock at time T if the option is in the money, and the second integral represents the present value of paying the strike price K if the option is exercised.

The integrals can be solved using properties of the standard normal distribution. The cumulative distribution function of the standard normal distribution, $N(d)$, is used to evaluate these integrals.

- **First Integral :**

$$\int_{Z^*}^{\infty} S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z\right) f_Z(Z)\, dZ$$

This integral is the expected value of receiving the stock S_T if the option is exercised (i.e., if $S_T > K$). The integration is over the range of Z values where this condition holds.

Inside the integral, the exponential term is complex. Let's break it down: $\exp((r - \frac{1}{2}\sigma^2)T)$ is a constant factor for a given r, σ, T.$\exp(\sigma\sqrt{T}Z)$ is the stochastic part, depending on the random variable Z.

Let's simplify the integral by making a change of variable. Define d_1 such that:

$$d_1 = \frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

d_1 is a re-expression of the terms in the exponential function.

Rewriting the integral in terms of d_1 gives us:

$$\int_{d_1}^{\infty} S_0 \exp((r - \frac{1}{2}\sigma^2)T) \exp(\sigma\sqrt{T}(d_1 - Z^*)) f_Z(d_1)\, dd_1$$

The exponential terms can be simplified and combined:

$$\exp((r - \frac{1}{2}\sigma^2)T) \exp(\sigma\sqrt{T}(d_1 - Z^*)) = \exp(\ln(S_0/K) + rT)$$

Now, the integral becomes:

$$S_0 \exp(\ln(S_0/K) + rT) \int_{d_1}^{\infty} f_Z(d_1)\, dd_1$$

This integral evaluates to the cumulative distribution function of the normal distribution:

$$\int_{d_1}^{\infty} f_Z(d_1)\, dd_1 = N(d_1)$$

Combining these steps, we find that the first integral evaluates to:

$$S_0 N(d_1)$$

This is a key part of the Black-Scholes formula and represents the expected value of receiving the stock at maturity if the option is in the money, discounted to present value. The use of the normal cumulative distribution function $N(d_1)$ captures the probability of the option being exercised under the risk-neutral measure.

- **Second Integral :**

$$-\int_{Z^*}^{\infty} K f_Z(Z)\, dZ$$

The integral calculates the expected value of paying the strike price K at maturity, but only in scenarios where the option is exercised (i.e., when $S_T > K$).

The term $f_Z(Z)\, dZ$ represents the infinitesimally small probability of Z being in a differential range around a particular value.

The integral essentially calculates the probability that the option will be in the money at expiration. It does this by integrating the probability density $f_Z(Z)$ from Z^* to infinity.

The cumulative distribution function $N(Z)$ for a standard normal variable gives the probability that a normally distributed random variable is less than or equal to Z. Therefore, the probability that Z is greater than Z^* is $1 - N(Z^*)$.

The integral simplifies to:

$$-K\left(1 - N(Z^*)\right)$$

Recall that d_2 in the Black-Scholes formula is defined as:

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

$$= \frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} - \sigma\sqrt{T}$$

$$= \frac{\ln(K/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

$$= Z^*$$

Therefore, the integral can be rewritten as:

$$-KN(d_2)$$

The second integral evaluates to $-KN(d_2)$, which represents the present value of the strike price K paid at maturity if the option is in the money, adjusted for the probability of this happening. This term is part of the Black-Scholes formula for the pricing of European call options and reflects the risk-neutral expected cost of the option's exercise.

By combining these two integrals, we get:

$$\mathbb{E}^Q[C_T] = S_0N(d_1) - KN(d_2)$$

These calculations involve simplifying the integrals using properties of the normal distribution and applying the change of variables technique to transform the integrals into a form that aligns with the cumulative distribution functions of the normal distribution. This is a fundamental aspect of the Black-Scholes model for option pricing.

- For a European Call Option:

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

- Similarly, for a European Put Option:

$$P(S,t) = Ke^{-r(T-t)}N(-d_2) - SN(-d_1)$$

- Where:

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right]$$

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

- $N(x)$ is the cumulative distribution function of the standard normal distribution.

The Black-Scholes formula calculates the theoretical price of European options using the current stock price, the option's strike price, time to expiration, risk-free rate, and the stock's volatility.

6.6.4 Black-Scholes limitations

- **Limits Usefulness:** As stated previously, the Black-Scholes model is only used to price European options and does not take into account that U.S. options could be exercised before the expiration date.
- **Lacks Cashflow Flexibility:** The model assumes dividends and risk-free rates are constant, but this may not be true in reality. Therefore, the Black-Scholes model may lack the ability to truly reflect the accurate future cashflow of an investment due to model rigidity.
- **Assumes Constant Volatility:** The model also assumes volatility remains constant over the option's life. In reality, this is often not the case because volatility fluctuates with the level of supply and demand.
- **Misleads Other Assumptions:** The Black-Scholes model also leverages other assumptions. These assumptions include that there are no transaction costs or taxes, the risk-free interest rate is constant for all maturities, short selling of securities with use of proceeds is permitted, and there are no risk-less arbitrage opportunities. Each of these assumptions can lead to prices that deviate from actual results.

6.6.5 The volatility surface

The Black-Scholes model, although elegant in its design, faces challenges in practical applications. Notably, real-world observations reveal instances where stock prices exhibit occasional jumps, deviating from the continuous movement predicted by the Geometric

Brownian Motion (GBM) model. Furthermore, stock prices tend to demonstrate fatter tails than anticipated by the GBM. Additionally, if the Black-Scholes model were accurate, the implied volatility surface would be uniformly flat. The volatility surface, defined implicitly as a function of strike (K) and time-to-maturity (T), is represented by the equation: $C(S, K, T) := BS(S, T, r, q, K, \sigma(K, T))$

Here, $C(S, K, T)$ denotes the current market price of a call option with time-to-maturity T and strike K, while $BS()$ represents the Black-Scholes formula for pricing a call option. In this context, $\sigma(K, T)$ represents the volatility that, when substituted into the Black-Scholes formula, yields the market price $C(S, K, T)$. Theoretically, if the Black-Scholes model were correct, the volatility surface would remain flat with $\sigma(K, T) = \sigma$ for all K and T. However, in practical scenarios, the volatility surface not only deviates from flatness but also exhibits significant variations over time.

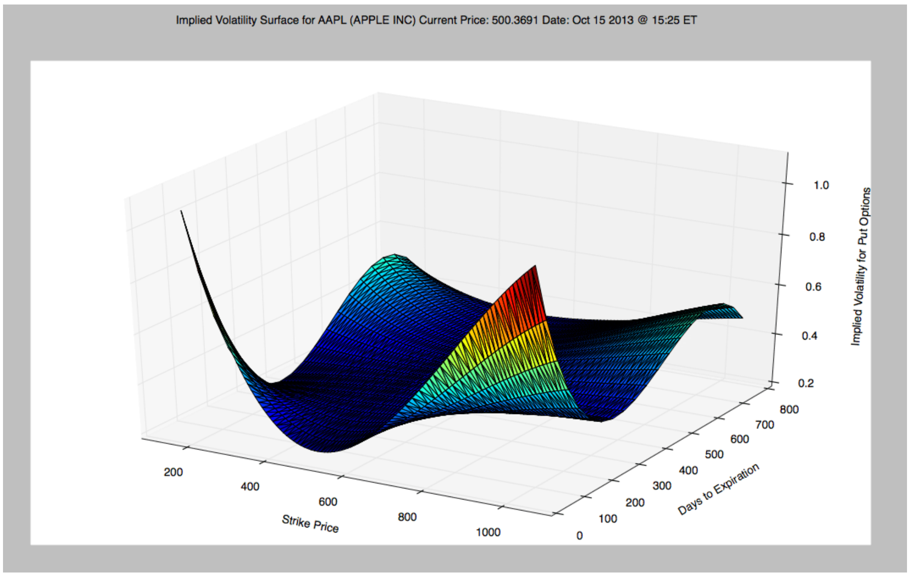


Figure 6.12: Implied Volatility

The volatility surface, a crucial aspect in financial markets, reveals that options with lower strikes tend to have higher implied volatilities, termed the volatility skew or smile. This phenomenon is observed for a given maturity (T) and strike (K). Implied volatility may either increase or decrease with time-to-maturity, stabilizing towards a constant for longer maturities. In certain scenarios, a short-term inverted volatility surface occurs during market stress.

Diverse implementations of the Black-Scholes model can lead to varying implied volatility surfaces. Despite this variability, accurate implementations are expected to result in similar shapes for these surfaces. Single-stock options, often American, may generate distinct volatility surfaces for put and call options, with put-call parity not applying to American options.

Despite acknowledging the Black-Scholes model's limitations, market participants widely adopt its language. Trading desks commonly compute Black-Scholes implied volatility surfaces, and the associated Greeks are referred to as Black-Scholes Greeks.

The reason for the existence of a skew in the volatility surface, especially notable for stocks and stock indices, can be attributed to dynamic factors. The skew manifests as a decline in implied volatility with increasing strike (K) for a fixed maturity (T), with this effect being more pronounced in shorter expirations. Two primary explanations account for the skew:

- Risk Aversion:
  - Stocks don't strictly adhere to Geometric Brownian Motion (GBM) with a constant volatility; instead, they exhibit jumps, with downward jumps being more significant and frequent.
  - Fear in the market tends to increase volatility as markets decline.
  - Supply and demand dynamics contribute, as investors seek protection by buying out-of-the-money puts, leading to higher demand for options with lower strikes.
- Leverage Effect:
  - The total value of a company's assets (debt + equity) is a more natural candidate to follow GBM.
  - Equity volatility is expected to increase as the equity value decreases, based on fundamental accounting equations.
  - The equity value can be viewed as the value of a call option on the total company value, recognizing that equity represents a call option with a strike equal to the debt value.

Expressed mathematically, the relationship between firm value volatility (σV) and equity volatility (σE) is approximated by the ratio of equity value to total value: σE approx =

(V/E)σV. This illustrates the leverage effect, where equity volatility is influenced by the dynamics of the total firm value and debt.

6.7 Binomial Tree Models

The binomial tree model breaks down the time to expiration of an option into potentially many small intervals or steps. At each step, the price of the underlying asset can move up or down by a specific factor, which leads to the creation of a tree-like structure of possible price outcomes.

6.7.1 Step-by-Step Construction of a Binomial Tree

To construct a binomial tree:

1. Divide the option's life into several time intervals or steps.
2. Calculate the up and down movement factors, usually denoted as u and d, using volatility and other market factors.
3. Develop the tree by iterating these steps over the option's life.

The formulae for u and d are typically:

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

$$d = \frac{1}{u},$$

where σ is the volatility of the underlying asset, and Δt is the time step.

6.7.2 Option Valuation using Binomial Trees

Option valuation in the binomial tree framework involves calculating the option value at each final node (at expiration) and then using backward induction to determine the present value at the initial node. This method incorporates the concept of risk-neutral valuation.

6.7.3 Example Calculation

Consider an option with a strike price of \$100, expiring in one year, and the underlying asset priced at \$100. Assuming an up-factor (u) of 1.1 and down-factor (d) of 0.9 per period, and a risk-free rate of 5%, we can calculate the option value at each node using risk-neutral probabilities. Suppose that we want to value an option using a one-period binomial tree. We know at expiration that an option is worth exactly its intrinsic value, the maximum of $[S - X, 0]$ for a call and the maximum of $[X - S, 0]$ for a put. In a one-period binomial tree, the expected value of a call is

$$p \max[S_u - X, 0] + (1 - p) \max[S_d - X, 0]$$

The theoretical value of the call is the present value of the expected value

$$\frac{p \max[S_u - X, 0] + (1 - p) \max[S_d - X, 0]}{1 + rX}$$

Using the same reasoning, the theoretical value of the put is

$$\frac{p \max[X - S_u, 0] + (1 - p) \max[X - S_d, 0]}{1 + rX}$$

Suppose that we expand our binomial tree to two periods each of length $t/2$ and also make the assumption that u and d are multiplicative inverses. Then

$$d = \frac{1}{u} \quad \text{and} \quad u = \frac{1}{d} \quad \text{so} \quad ud = 1$$

This means that an up move followed by a down move or a down move followed by an up move results in the same price. If the magnitudes of the up and down moves u and d are the same at every branch in our tree, then in a risk-neutral world, the probability of an upward move will always be

$$p = \frac{\left(\frac{1+r}{u}\right) - d}{u - d}$$

and the probability of a down move will always be $1 - p$.

Three-Period Binomial Tree for Option Valuation

There are now three possible prices for the underlying at expiration—S_{uuu}, S_{uud}, and S_{udd}. There is only one path that will lead to either S_{uuu} or S_{ddd}. But there are two possible paths to the middle price S_{uud}. The underlying can go up and then down or down and then up. The theoretical value of a call in the two-period example is

$$C = p\max[S_{uu} - X, 0] + 2p(1-p)\max[S_{ud} - X, 0] + (1-p)^2\max[S_{dd} - X, 0]$$

where

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

and the value of a put is

$$P = p\max[X - S_{uu}, 0] + 2p(1-p)\max[X - S_{ud}, 0] + (1-p)^2\max[X - S_{dd}, 0]$$

Using this approach, we can expand our binomial tree to any number of periods.

- n = number of periods in the binomial tree
- t = time to expiration in years
- r = annual interest rate

the possible terminal underlying prices are

$$S_i = S_0 u^i d^{n-i} \quad \text{for} \quad i = 0, 1, 2, \ldots, n$$

The number of paths that will lead to each terminal price is given by the *binomial expansion*

$$\binom{n}{i} = \frac{n!}{i!(n-i)!}$$

The values of a European call and put are

$$\text{Call} = \frac{1}{(1+r\Delta t)^n} \sum_{i=j}^{n} \binom{n}{j} p^j (1-p)^{n-j} \max[S_i - X, 0]$$

$$\text{Put} = \frac{1}{(1+r\Delta t)^n} \sum_{i=j}^{n} \binom{n}{j} p^j (1-p)^{n-j} \max[X - S_i, 0]$$

A Three-Period Example

Suppose that

- $n = 3$
- $S = 100$
- $t = 9$ months (0.75 year)
- $r = 4$ percent (0.04)
- $u = 1.05$
- $d = \frac{1}{u} = 0.9524$

Then the values of p and $1 - p$ are

$$p = \frac{(1 + r\Delta t) - d}{u - d} = \frac{(1 + 0.03) - 0.9524}{1.05 - 0.9524} = 0.59$$

$$1 - p = 1 - 0.59 = 0.41$$

6.8 Heston's Stochastic Volatility Model (1993)

The Heston model is a mathematical model developed by Steven Heston (1993) for pricing options. It is an extension of the Black-Scholes model, which allows for stochastic volatility. In the Black-Scholes model, volatility is a constant, but the Heston model assumes that volatility is a random process, leading to more accurate pricing of options, especially for longer maturities and for options far in or out of the money.

The Heston model describes the dynamics of an asset price and its variance through two stochastic differential equations (SDEs). Let S_t represent the asset price at time t, and V_t represent the variance of the asset's returns at time t.

The Heston Model is defined by two stochastic processes :

1. **Asset Price Dynamics :** The asset price follows a standard geometric Brownian motion with a stochastic volatility:

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_t^S$$

   Here, μ is the rate of return of the asset, and dW_t^S is a Wiener process modeling the random market movements.

2. **Variance Dynamics :** The variance itself follows a mean-reverting square-root process:

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^V$$

   In this equation, κ is the rate of mean reversion, θ is the long-term variance, σ is the volatility of volatility, and dW_t^V is another Wiener process, **correlated with** dW_t^S.

Where :

- S_t is the asset price at time t,
- r is the risk-free interest rate,
- V_t is the instantaneous variance,
- κ is the rate of mean reversion,
- θ is the long-term variance,
- σ is the volatility of volatility,
- dW_t^S and dW_t^V two Wiener process under Q. The correlation between the two Wiener processes dW_t^S and dW_t^V is denoted by ρ, adding a layer of complexity by allowing the model to capture the leverage effect - a phenomenon where asset prices

and their volatility are negatively correlated : $E[dW_t^S dW_t^V] = \rho dt$

The Heston model is widely used in the pricing of bond and currency options, as well as other financial derivatives where stochastic volatility is a significant factor. Its ability to capture the volatility smile and skew, which are observed in real markets, makes it a powerful tool in financial modeling.

In particular, the model is useful for pricing European options, where the random paths of the asset price and its variance can be simulated to compute an option's expected payoff under the risk-neutral measure. The model's ability to capture stochastic volatility makes it more flexible and realistic compared to models assuming constant volatility.

The Heston model represents a significant advancement in financial modeling, providing a more realistic framework for understanding asset price dynamics, particularly in environments characterized by variable volatility. Its introduction has had a profound impact on the fields of derivative pricing and risk management.

6.8.1 Heston Model derivation

a. Deriving the Stochastic Differential Equation (SDE) :

We will now apply Ito's Lemma to a function $F(t, S_t, V_t)$. The general form of Ito's Lemma for a function of two stochastic variables is:

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

Replacing dS_t and dV_t by their value, previously defined :

$$dF = \tfrac{\partial F}{\partial t}dt + \tfrac{\partial F}{\partial S_t}\left(\mu S_t dt + \sqrt{V_t}S_t dW_t^S\right) + \tfrac{\partial F}{\partial V_t}\left(\kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^V\right) + \tfrac{1}{2}\tfrac{\partial^2 F}{\partial S_t^2}\left(\mu S_t dt + \sqrt{V_t}S_t dW_t^S\right)^2 + \tfrac{1}{2}\tfrac{\partial^2 F}{\partial V_t^2}\left(\kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^V\right)^2 + \tfrac{\partial^2 F}{\partial S_t \partial V_t}\left(\mu S_t dt + \sqrt{V_t}S_t dW_t^S\right)\left(\kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^V\right)$$

Removing higher order elements (dt^2, $dWdt$) and replacing $dX_t^2 = dt$:

$$dF = \tfrac{\partial F}{\partial t}dt + \tfrac{\partial F}{\partial S_t}\left(\mu S_t dt + \sqrt{V_t}S_t dW_t^S\right) + \tfrac{\partial F}{\partial V_t}\left(\kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^V\right) + \tfrac{1}{2}\tfrac{\partial^2 F}{\partial S_t^2}\left(V_t S_t^2 dt\right) + \tfrac{1}{2}\tfrac{\partial^2 F}{\partial V_t^2}\left(\sigma^2 V_t dt\right) + \tfrac{\partial^2 F}{\partial S_t \partial V_t}\left(\sigma S_t V_t dt\right)$$

Rearranging the equation gives :

$$\begin{aligned}
dF = &\left(\frac{\partial F}{\partial t} + \mu S_t \frac{\partial F}{\partial S_t} + \kappa(\theta - V_t)\frac{\partial F}{\partial V_t}\right) dt \\
&+ \sqrt{V_t} S_t \frac{\partial F}{\partial S_t} dW_t^S + \sigma\sqrt{V_t}\frac{\partial F}{\partial V_t} dW_t^V \\
&+ \frac{1}{2} V_t S_t^2 \frac{\partial^2 F}{\partial S_t^2} dt + \frac{1}{2}\sigma^2 V_t \frac{\partial^2 F}{\partial V_t^2} dt \\
&+ \frac{\partial^2 F}{\partial S_t \partial V_t}\sqrt{V_t} S_t \sigma \sqrt{V_t} dW_t^S dW_t^V
\end{aligned}$$

This is a more complex equation due to the interaction of S_t and V_t. The last term, involving $dW_t^S dW_t^V$, depends on the correlation between the two Wiener processes, and in practice, it's often expressed in terms of a correlation coefficient ρ, making it $\rho\sqrt{V_t} S_t \sigma \sqrt{V_t} \frac{\partial^2 F}{\partial S_t \partial V_t} dt$.

b. Setting up the Partial Differential Equation (PDE) :

After applying Ito's Lemma and deriving the stochastic differential equation (SDE) for the option price function $F(t, S_t, V_t)$, the next step in the Heston model derivation is to solve the corresponding partial differential equation (PDE). This PDE is derived from the SDE and describes the evolution of the option price over time.

The option price function $F(t, S_t, F_t)$ should satisfy the risk-neutral valuation principle, which states that the discounted expected return of any derivative must be zero. Therefore, the drift term of df should equal the risk-free rate rf times the option price $F(t, S_t, V_t)$ times dt:

$$rfF(t, S_t, V_t)dt = \left(\frac{\partial F}{\partial t} + rfS_t \frac{\partial F}{\partial S_t} + \kappa(\theta - V_t)\frac{\partial F}{\partial V_t} + \frac{1}{2} V_t S_t^2 \frac{\partial^2 F}{\partial S_t^2} + \frac{1}{2}\sigma^2 V_t \frac{\partial^2 V}{\partial V_t^2} + \rho\sigma V_t S_t \frac{\partial^2 F}{\partial S_t \partial V_t}\right) dt$$

Removing the dt and rearranging the terms, we get the PDE:

$$\frac{\partial F}{\partial t} + rS_t \frac{\partial F}{\partial S_t} + \kappa(\theta - V_t)\frac{\partial F}{\partial V_t} + \frac{1}{2} V_t S_t^2 \frac{\partial^2 F}{\partial S_t^2} + \frac{1}{2}\sigma^2 V_t \frac{\partial^2 F}{\partial V_t^2} + \rho\sigma V_t S_t \frac{\partial^2 F}{\partial S_t \partial V_t} - rf = 0$$

The market price of volatility risk could also be introduced, providing the equation

$$dV_t = \kappa(\theta - V_t)dt + \lambda V_t dt + \sigma\sqrt{V_t} dW_t^V$$

, where $\lambda V_t dt$ is the additional term representing the risk premium for holding volatility risk. This term modifies the drift component of the variance process. This adjustment reflects the idea that the market's risk preferences affect the pricing of derivatives through a volatility risk premium. Deriving the model in the same way we did earlier, we get :

$$\frac{\partial F}{\partial t}+\frac{1}{2}S_t^2V_t\frac{\partial^2 F}{\partial S_t^2}+rS_t\frac{\partial F}{\partial S_t}+\left[\kappa(\theta-V_t)-\lambda V_t\right]\frac{\partial F}{\partial V_t}+\frac{1}{2}\sigma^2V_t\frac{\partial^2 F}{\partial V_t^2}+\rho\sigma S_tV_t\frac{\partial^2 F}{\partial S_t\partial V_t}-rfF=0$$

The unspecified $\lambda\,(S_t,v,t)$ is the market price of volatility risk.

c. Solving Partial Differential Equation (PDE) :

The boundary and final conditions depend on the specific option being priced. For example, for a European call option with strike price K and expiration T, the final condition at expiration is: $F(T,S_T,v_T)=\max(S_T-K,0)$

The PDE derived above is a second-order partial differential equation with non-constant coefficients, reflecting the stochastic volatility of the underlying asset. Solving this PDE generally requires numerical methods such as finite difference methods, Monte Carlo simulation, or Fourier transform techniques.

Heston guessed a solution of this form :

$$C(S_t,V_t,t,T)=S_tP_1-Ke^{-r(T-t)}P_2 \tag{6.1}$$

where the first term is the present value of the the spot asset upon optimal exercise, and the second term is the present value of the strike price payment. Both of these term must satisfy the PDE (0.2). writing $x=\ln(S_t)$ for convenience and substituting equation (0.3) in equation (0.2), then P_1 and P_2 are defined by the inverse Fourier transformation

$$P_j(x,V_t,T,K)=\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \Re\left\{\frac{e^{-i\phi\ln(K)}f_j(x,V_t,T,\phi)}{i\phi}\right\}d\phi,\quad j=1,2 \tag{6.2}$$

where :

- $f_j(x,V_t,T,\phi)=e^{C(T-t,\phi)+D(T-t,\phi)V_t+i\phi x}$
- $C(T-t,\phi)=r\phi i\tau+\frac{a}{\sigma^2}\left[(b_j-\rho\sigma\phi i+d)\tau-2\ln\left(\frac{1-ge^{d\tau}}{1-g}\right)\right]$

- $D(T-t,\phi) = \frac{b_j - \rho\sigma\phi i + \frac{d}{\sigma^2}(1-e^{d\tau})}{1-ge^{d\tau}}$
- $g = \frac{b_j - \rho\sigma\phi i - d}{b_j - \rho\sigma\phi i + d}$
- $d = \sqrt{(\rho\sigma\phi i - b_j)^2 - \sigma^2(2u_j\phi i - \phi^2)}$

Also, defined in the paper :

- $a = \kappa\theta$
- $b_1 = \kappa + \lambda - \rho\sigma$
- $b_2 = \kappa + \lambda$

This completes the closed-form approximation in the Heston model.

6.8.2 Discretisation for MC

The Heston model is also a useful model for **simulating** stochastic volatility and its effect on the potential **paths** an asset can take over the life of an option. While volatility is considered constant in BSM model, volatility in Heston model follows stochastic process. Heston model consists of two set of Stochastic Differential Equations (SDE). The first one is to simulate stock price paths using Geometric Brownian Motion (GBM) and the second one is to simulate stock volatility using a stochastic, mean-reverting process.

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_{t,S} \\ dV_t &= \kappa(\theta - V_t)dt + \gamma\sqrt{V_t}dW_{t,V} \end{aligned}$$

such that $dW \sim \mathcal{N}(0, dt)$. Given the correlation ρ between the above two independent Brownian Motions, the covariance between $dW_{t,S}$ and $dW_{t,V}$ is.

$$\mathbb{E}[dW_{t,S} \cdot dW_{t,V}] = \rho\sqrt{dt}\sqrt{dt} = \rho dt$$

Variance: $V = \sigma^2$, and Volatility: σ $W_{t,S}$ Brownian motion of the asset price.
$W_{t,V}$ Brownian motion of the asset variance.

Discretize Stock Price

Under the real work measure, Brownian Motion SDE $dS_t = \mu S_t dt + \sigma S_t dW_t$

Drift $= \mu S_t$, and Diffusion $= \sigma S_t$, μ is the average growth rate of the asset and σ is the associated volatility of the returns on the asset. dW is an increment of a Brownian Motion, known as winner process; such that $dW \sim \mathcal{N}(0, dt)$.

The values of S is log-normally distributed and the returns $\frac{ds}{S}$ are normally distributed.

Under the risk-neutral measure, it changes to $dS_t = rS_t dt + \sigma S_t dW_t$

The good thing is that the closed form solution exists for this SDE. But derivation can be shown separately. Integrating the solution over t to T, resulting into

$$\ln\left(\frac{S_T}{S_t}\right) = \left(\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(W_T - W_t) \tag{6.3}$$

$$\ln\left(\frac{S_T}{S_t}\right) = \left(\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma Z\sqrt{T-t} \tag{6.4}$$

For simplicity, integrate the solution over 0 to t, resulting into

$$\ln\left(\frac{S_t}{S_0}\right) = \left(\mu - \frac{1}{2}\sigma^2\right)(t-0) + \sigma(W_t - W_0) \tag{6.5}$$

Given $W_0 = 0$

$$S_t = S_0 \cdot \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right] \tag{6.6}$$

Simulate it using Euler discretization

$$\widetilde{S}_{t+\delta t} = S_t \cdot \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)\delta t + \sigma \widetilde{W}_{\delta t}\right] \tag{6.7}$$

Since $\widetilde{W}_{\delta t} = Z\sqrt{\delta t}$, where $Z \sim \mathcal{N}(0, 1)$.

$$\widetilde{S}_{t+\delta t} = S_t \cdot \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)\delta t + \sigma Z\sqrt{\delta t}\right] \tag{6.8}$$

It will give us the simulated asset prices.

Discretize Variance

$dV_t = \kappa(\theta - V_t)dt + \gamma\sqrt{V_t}dW_t$

Where κ, θ and γ are all positive. θ is the long-run mean of the instantaneous variance. κ is the rate of mean-reversion; in other words, it is rate at which V_t reverts to θ. γ is the volatility of volatility $\sqrt{V_t}$. ***These parameters should come from calibration.***

Using Euler discretization, $V_{t+\delta t} = V_t + \kappa(\theta - V_t)\delta t + \gamma\sqrt{V_t}\delta t Z$

Simulation Steps

1. Start with the initial values S_0 for the stock price and V_0 for the variance.
2. Simulate Variance (squared of volatility) : $V_{t+\delta t} = V_t + \kappa(\theta - V_t)\delta t + \gamma\sqrt{V_t}\delta t Z_V$
3. Simulate Stock Price : $S_{t+\delta t} = S_t \cdot \exp\left[\left(r - \frac{1}{2}v_t\right)\delta t + \sqrt{v_t}\delta t Z_S\right]$, Where: $E[Z_S, Z_V] = \rho$. The process for V_t is guaranteed to remain positive as long as $\kappa\theta > \frac{1}{2}\gamma^2$. In equity markets, where we often witness increased volatility associated with sharp negative returns, we would expect ρ to be negative. In general, though, the model does not require any particular sign of ρ.
4. Capture the terminal stock prices and compute option values
5. Compute present value (PV) of the mean of the computed option values

Generate Correlated Standard Normal Distribution

Let Z_1 and Z_2 be the two independent $\mathcal{N}(0,1)$ random variables. Now construct two correlated standard normal random variables as $X = Z_1$ and $Y = \rho Z_1 + \sqrt{1-\rho^2}Z_2$

Such that $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ and $\text{Cov}(X,Y) = \rho = E[XY] - E[X]E[Y] = E[XY]$. Its derivation can be shown separately.

6.9 Dupire's Local Volatility Model (1994)

The Dupire model, proposed by Bruno Dupire in 1994, addresses the limitations of constant volatility assumptions in option pricing. This model introduces the concept of local volatility, which allows volatility to vary with both the price of the underlying asset and time. Dupire's local volatility model has been instrumental in refining the theoretical framework for the pricing of European options. His model recognizes the volatility smile, where implied volatility varies with strike price and expiration.

The Dupire model posits that the price of an option can be determined by the local volatility function, denoted as $\sigma(S_t, t)$, where S_t is the underlying asset price at time t. The local volatility function is derived from market prices of European call and put options using the Dupire equation, which is a forward equation derived from the Black-Scholes PDE (partial differential equation). The Dupire equation can be written as:

$$\frac{\partial C}{\partial T} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma(S,T)^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC = 0$$

Here, C is the price of the European call option, r is the risk-free rate, and T is the maturity of the option. The local volatility $\sigma(S,T)$ is the function to be determined.

The Dupire model is primarily used for pricing and hedging European options, and it is also the foundation for more complex derivatives pricing models that incorporate stochastic volatility. The calibration of the local volatility surface is crucial, as it impacts the pricing accuracy for options with different strikes and maturities.

Dupire's original work, "Pricing with a Smile," introduced the local volatility model and provided a new perspective on volatility smile - the pattern observed in the implied volatility of options across different strikes and maturities. It was published in "Risk" magazine in 1994 and has since been a seminal paper in quantitative finance.

The Dupire Local Volatility model is a cornerstone in the field of derivative pricing. By capturing the dynamic nature of volatility, it enhances the theoretical understanding and practical applications of option pricing beyond the constant volatility assumption. The model's ability to match the observed market prices of options has made it a standard tool in the financial industry.

Dupire's model generalizes the Black-Scholes model by allowing volatility to be a function of both the strike price and the time to maturity, i.e., $\sigma = \sigma(K,T)$.

Rearranging the Dupire's formula and solving for $\sigma^2(K,T)$, we obtain:

$$\sigma^2(K,T) = \frac{2\left(\frac{\partial C}{\partial T} + rC - rS\frac{\partial C}{\partial S}\right)}{S^2\frac{\partial^2 C}{\partial S^2}}. \tag{6.9}$$

In practice, Dupire's formula is used to compute the local volatility surface from market prices of European options. The procedure typically involves:

1. **Calibrating to Market Data:** Use market prices of European call options to calibrate the model. This involves numerically estimating the partial derivatives in the formula.

2. **Constructing the Volatility Surface:** For each pair of K and T, compute $\sigma(K,T)$ using the calibrated model. This results in a volatility surface that reflects market variations in implied volatility across different strikes and maturities.

Dupire's local volatility model, therefore, allows for a more accurate representation of market dynamics and is particularly useful for pricing exotic options and managing the risk of complex derivatives portfolios.

To illustrate the application of Dupire's formula, consider a set of market data for European call options on a particular underlying asset. This data includes option prices for various strikes and maturities. Our goal is to use Dupire's formula to construct the local volatility surface.

1. **Data Collection:** Gather market prices of European call options across a range of strike prices and maturities.

2. **Computing Derivatives:** Calculate the partial derivatives $\frac{\partial C}{\partial T}$ and $\frac{\partial^2 C}{\partial K^2}$. This can be done numerically using finite difference methods.

3. **Applying Dupire's Formula:** For each combination of strike price and maturity, apply Dupire's formula to compute the local volatility.

4. **Constructing the Volatility Surface:** Plot the calculated local volatilities against the corresponding strikes and maturities to visualize the local volatility surface.

This surface provides valuable insights into how market participants perceive future volatility and can be used to price and hedge complex derivative products more accurately.

6.10 Volatility

6.10.1 Historical volatility

Historical volatility is also known as stochastic volatility. It measures the past variability of a financial asset's price. It is calculated by analyzing historical data on asset prices over a specified period. This can be done using methods such as calculating the variance or standard deviation of past returns. Stochastic volatility is retrospective, as it is based on data that has already been observed.

6.10.2 Implied volatility

Implied volatility, on the other hand, is a measure of expected future volatility as reflected in the current price of options on a financial asset. It is deduced from the market price of options. Option pricing models, such as the Black-Scholes model, are used to calculate implied volatility by inverting these models to solve for volatility. Implied volatility is forward-looking, as it gives an indication of the volatility expected by market participants for the remaining life of the option. Implied volatility is used to estimate the market's perception of the future volatility of an underlying asset. Implied volatility is expressed as the percentage change in the underlying asset price over one year.

6.10.3 Local volatility

Local volatility is also known as local stochastic volatility. It is used in the context of more advanced models, such as the Heston model, which are designed to better capture the complex dynamics of market volatility. Unlike implied volatility, which is considered constant in the Black-Scholes model, local volatility can vary with time and the price level of the underlying. Local volatility is generally modeled to track market fluctuations more realistically, recognizing that volatility can change over time and at different price levels (volatility smile).

6.10.4 Historical volatility vs. Implied volatility vs. Local volatility

Stochastic volatility is based on historical prices to measure past variability, while implied volatility is derived from current option market prices to represent expected volatility in the future. Traders and investors use these measures to assess risk and make option trading decisions. Implied volatility is an aggregate measure of expected future volatility, often used in simpler models, while local volatility is specific to more sophisticated models and takes into account the variation in volatility at different price levels and times.

6.10.5 Smile volatility

The implied Volatility of options that are heavily out-of-the-money or largely in-the-money is higher than the implied Volatility recalculated from in-the-money options. This phenomenon is known as the Volatility Smile (plot of Volatility as a function of strike price in the form of a smile).

A concept similar to skew is the volatility smile. The volatility smile corresponds to a situation where in-the-money options have a lower implied volatility than out-of-the-money options and in-the-money options. In this situation, the implied volatility of puts and calls increases as the strike price moves further away from the current price of the underlying asset.

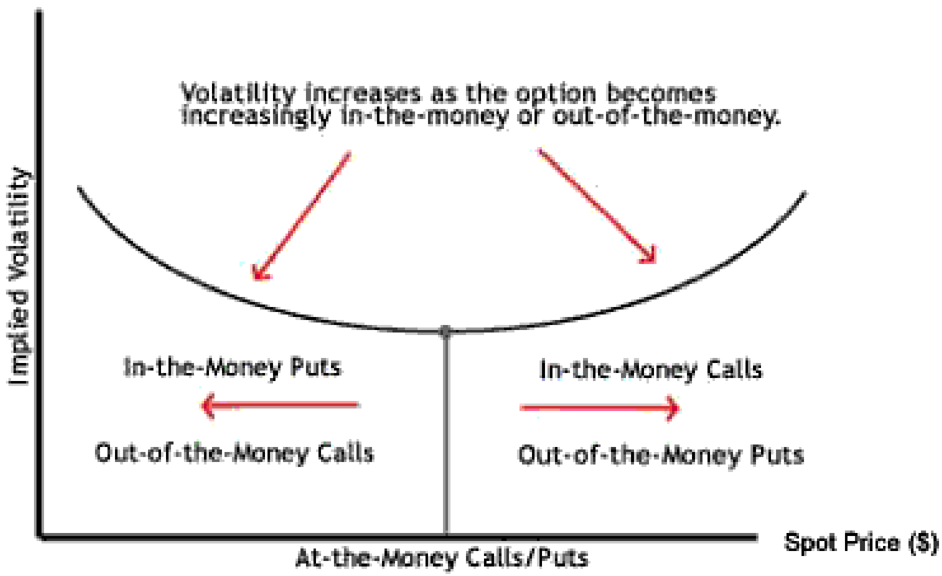


Figure 6.13: Volatility smile

6.11 Greeks and Options Hedging

In the dynamic world of financial markets, managing risks associated with investments and derivatives is a critical aspect of portfolio management. Traders and investors often employ sophisticated strategies to safeguard against market uncertainties and optimize returns. Two fundamental concepts in this domain are the Greeks and hedging.

In options trading, the Greeks are a set of risk measures that help quantify the sensitivity of option prices to various factors. These factors include changes in the underlying asset's price, time decay, implied volatility, and interest rates. Commonly denoted by Greek letters—Delta, Gamma, Theta, Vega, and Rho—these metrics provide valuable insights into how an option's value may fluctuate under different market conditions. Understanding the Greeks is essential for crafting strategies that align with an investor's risk tolerance and market outlook.

Hedging is a risk management strategy aimed at offsetting potential losses in one investment by taking an opposing position in another. In the context of derivatives, such as options or futures, hedging involves creating a strategic combination of positions to reduce exposure to adverse price movements. The goal is to minimize the impact of market fluctuations on the overall portfolio.

6.11.1 Delta-Gamma-Vega Approximations to Option Prices

Using Taylor's Theorem, we can express the price change ($P\&L$) of a derivative security, denoted as $C(S, \sigma)$, due to small changes in the stock price (ΔS) and implied volatility ($\Delta\sigma$) as follows:

$$C(S + \Delta S, \sigma + \Delta\sigma) \approx C(S, \sigma) + \Delta S \frac{\partial C}{\partial S} + \frac{1}{2}(\Delta S)^2 \frac{\partial^2 C}{\partial S^2} + \Delta\sigma \frac{\partial C}{\partial \sigma}$$

This expression can be further simplified as:

$$P\&L = \delta \Delta S + \frac{\Gamma}{2}(\Delta S)^2 + \text{vega}\Delta\sigma$$

Here,

- δ is the delta, representing the sensitivity of the option price to changes in the stock price ($\frac{\partial C}{\partial S}$),
- Γ is the gamma, which measures the rate of change of delta with respect to changes in the stock price ($\frac{\partial^2 C}{\partial S^2}$),
- vega represents the sensitivity of the option price to changes in volatility ($\frac{\partial C}{\partial \sigma}$),
- $P\&L$ is the profit and loss $C(S+\Delta S, \sigma+\Delta\sigma) - C(S,\sigma)$,
- ΔS is the change in stock price,
- $\Delta\sigma$ is the change in volatility.

When $\Delta\sigma = 0$, this simplifies to the delta-gamma approximation, commonly used in calculations such as historical Value-at-Risk (VaR) for portfolios containing options.

In a more intuitive form, the profit and loss can be expressed as:

$$P\&L = \delta S\left(\frac{\Delta S}{S}\right) + \frac{\Gamma S^2}{2}\left(\frac{\Delta S}{S}\right)^2 + \text{vega}\Delta\sigma$$

6.11.2 Delta - Underlying price movements

Investment banks use the concept of Delta, derived from these models, to hedge the risks associated with the directional movement of the underlying asset's price. The delta (Δ) of an option represents the rate of change in the option price concerning a change in the price of the underlying security. For a European call option, the delta is given by:

$$\Delta = \frac{\partial C}{\partial S} = e^{-qT}\Phi(d_1)$$

This formula assumes a volatility surface that remains constant as the underlying security moves (sticky-by-strike). If the volatility of the option changes with the movement of the underlying security, an additional term, vega $\times \frac{\partial\sigma(K,T)}{\partial S}$, would be included. This scenario is known as a sticky-by-delta volatility surface. Equity markets typically employ the sticky-by-strike approach, while foreign exchange markets often use the sticky-by-delta approach. Similar considerations apply to gamma.

Facts :

- Range : For a call option, delta ranges from 0 to 1. For a put option, it ranges from -1 to 0.
- ATM Options: At-the-money options typically have a delta close to 0.5 (for calls) or -0.5 (for puts) as there's an approximately equal chance of finishing in or out of the money.
- Deep In or Out of the Money: Deep in-the-money call options approach a delta of 1, while deep out-of-the-money calls approach 0. For puts, it's the opposite.
- Time to Expiry: Delta changes as the option approaches expiry, becoming more sensitive for at-the-money options.

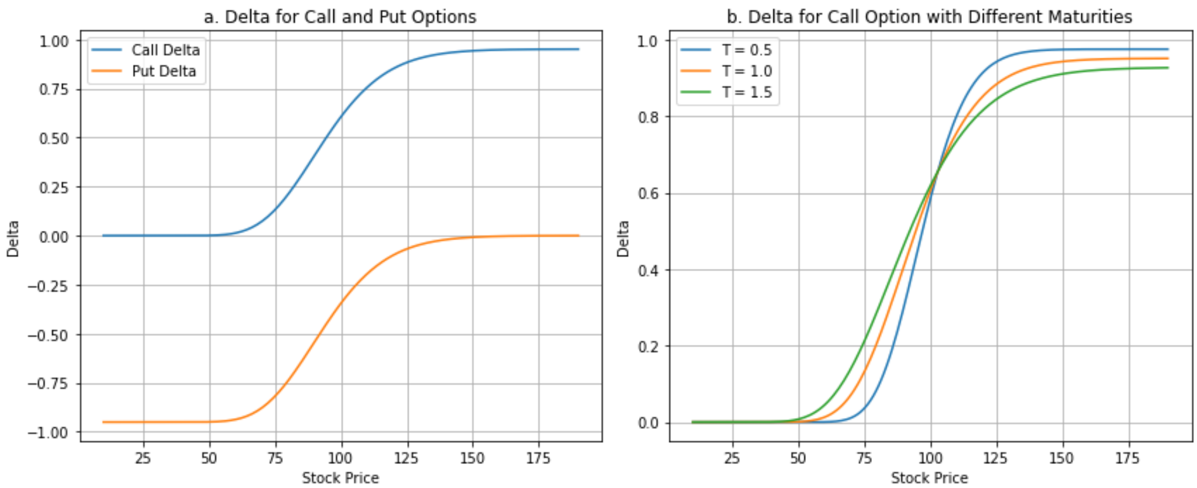


Figure 6.14: Delta for European Options

Through put-call parity, the delta for a put option (Δ_{put}) can be expressed as $\Delta_{\text{call}} - e^{-qT}$. Figure (1.a.) illustrates the delta for call and put options with respect to the underlying stock price, while Figure (1.b.) depicts the delta for a call option across various times-to-maturity. The strike (K) is a key factor, influencing delta steepness around K as time-to-maturity decreases. $\Delta_{\text{call}} = \Phi(d_1)$ is commonly interpreted as the (risk-neutral) probability of the option expiring in the money (although this probability is technically $\Phi(d_2)$).

Figure (2) displays the delta of a call option in relation to time-to-maturity for options with different moneyness. Observations and surprises in this plot prompt consideration of the corresponding delta plot for put options.

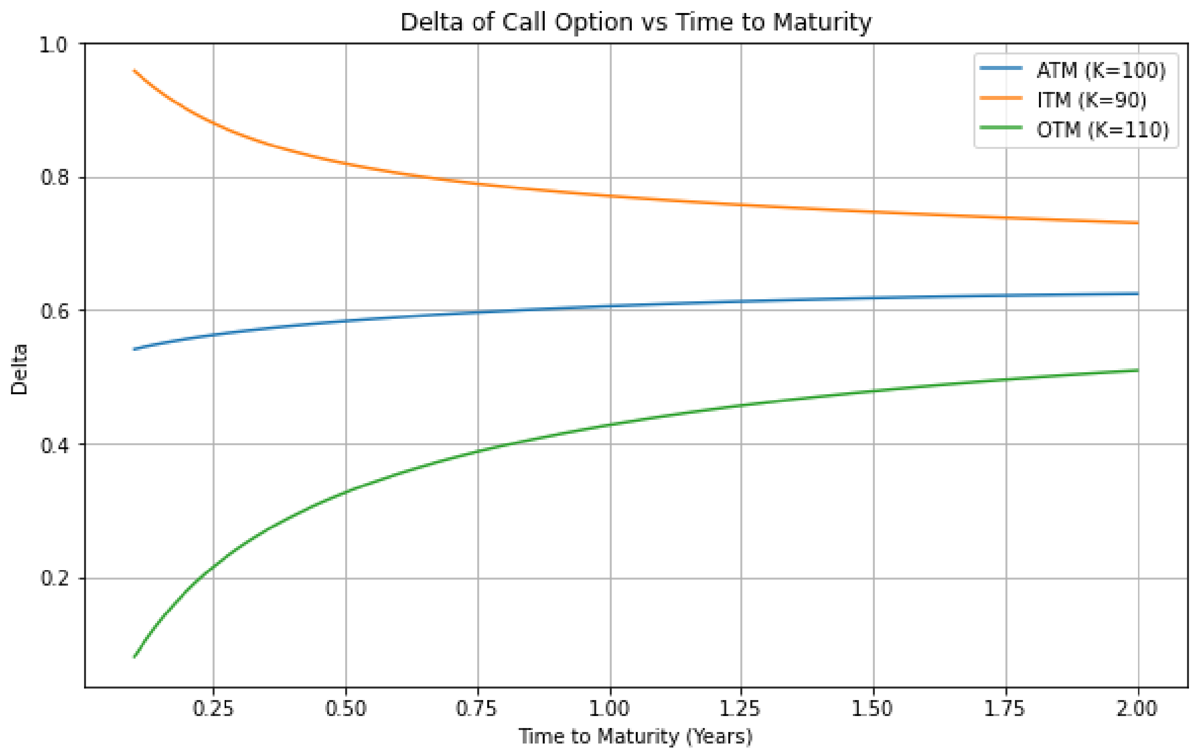


Figure 6.15: Delta for European Call Options as a Function of Time to Maturity

In the Black-Scholes model with Geometric Brownian Motion (GBM), an option can be exactly replicated through a delta-hedging strategy. The Black-Scholes Partial Differential Equation (PDE) was derived based on this delta-hedging and replication concept. The primary idea behind delta-hedging is to continuously adjust the portfolio of the option and the stock to maintain a total delta of zero after each re-balancing. While continuous hedging may not be practical, periodic or discrete hedging is often employed, leading to some replication errors.

Let P_t denote the value of the discrete-time self-financing strategy aiming to replicate the option payoff, and C_0 is the initial option value. The replicating strategy is given by:

$$P_0 = C_0$$

$$P_{t+1} = P_t + (P_t - \delta_t S_t) r \Delta t + \delta_t$$

Selling options when realized volatility (σ) is lower than implied volatility (σ_{imp}) leads to expected losses during delta-hedging.

Selling options when realized volatility (σ) is higher than implied volatility (σ_{imp}) results in expected gains during delta-hedging.

When $\sigma = \sigma_{\text{imp}}$, the total P&L behavior depends on the entire price path taken by the stock over the given time interval.

The payoff from continuously delta-hedging an option is path-dependent, meaning it relies on the price trajectory of the stock throughout the entire time interval. Mathematically, the payoff is expressed as: $P\&L = \int_0^T S_t^2 \frac{\partial^2 V_t}{\partial S^2}\, dt$

Considering this equation, where V_t is the time t value of the option and $_t$ is the realized instantaneous volatility at time t. The term $S_t^2 \frac{\partial^2 V_t}{\partial S^2}$ is commonly referred to as the dollar gamma. This term, representing the second derivative with respect to the stock price, is always positive for both call and put options. However, it approaches zero as the option significantly moves into or out of the money.

There is no restriction to using geometric Brownian motion, and alternative models such as diffusion or jump-diffusion models can be employed. It becomes interesting to simulate these alternative models and observe the effects on the replication error, where the δ_t values are computed, assuming (perhaps incorrectly) a geometric Brownian motion for price dynamics.

Hedging with delta is a crucial concept in options trading and risk management. Delta hedging involves creating a portfolio that is insensitive to small movements in the price of the underlying asset. Here are more details about the key concepts and techniques involved:

- **Delta Neutral Hedging :** Delta neutral hedging aims to offset the overall delta of a portfolio. If a portfolio has a positive delta, it can be hedged by taking a position with a negative delta, and vice versa. The goal is to make the net delta of the portfolio zero, or as close to zero as possible. *Example: If you own a call option with a delta of 0.6 (i.e., the option's price increases by $0.60 for every $1 increase in the underlying asset), you can achieve delta neutrality by shorting 60 shares of the underlying asset for every 100 call options you own.*

- **The Hedge Ratio :** The hedge ratio (also known as the delta ratio) is used to determine the quantity of the underlying asset needed to hedge the option position. It's the ratio of the option delta to the quantity of the underlying asset. Hedge

Ratio $= \Delta \times$ Number of Options

- **Dynamic Delta Hedging :** Delta is not static; it changes as the price of the underlying asset and time to expiry change. Hence, a delta-hedged portfolio needs continuous rebalancing. This process is known as dynamic delta hedging.
- If the underlying asset's price moves significantly, the delta of the option changes, and the hedge will no longer be neutral.
- The portfolio needs to be frequently rebalanced by buying or selling shares of the underlying asset to maintain delta neutrality.

Delta hedging is a sophisticated technique that requires a deep understanding of options and their Greeks, as well as constant monitoring and adjustment of the portfolio. While it can effectively protect against price changes in the underlying asset, it also involves challenges such as transaction costs and gap risk.

6.11.3 Gamma - Delta sensitivity

The gamma (Γ) of an option measures how the option's delta responds to changes in the price of the underlying security. For a call option, the gamma is given by:

$$\Gamma = \frac{\partial^2 C}{\partial S^2} = e^{-qT} \frac{\phi(d_1)}{\sigma S \sqrt{T}}$$

where $\phi(\cdot)$ represents the standard normal probability density function.

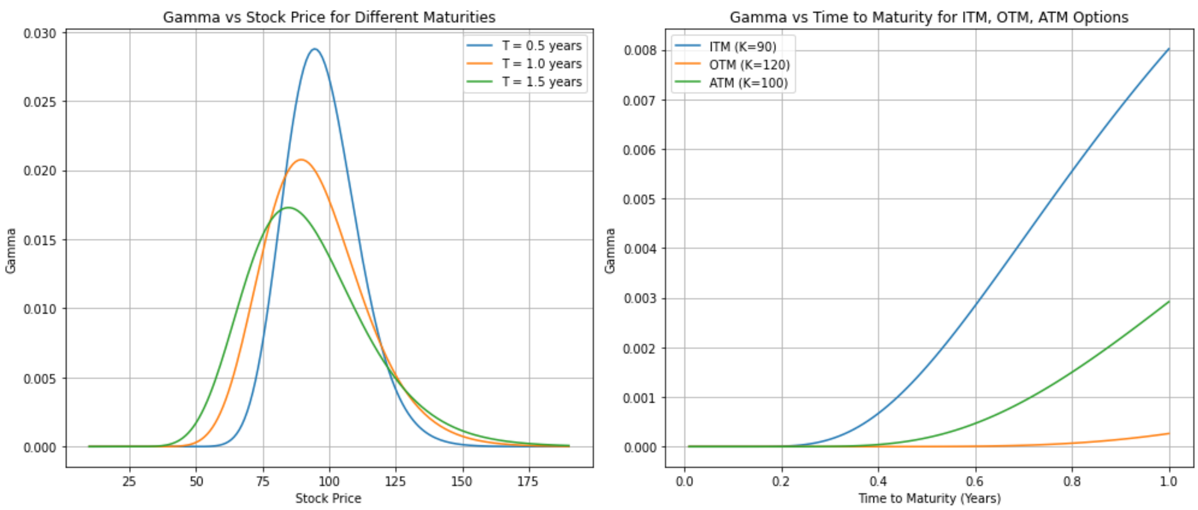


Figure 6.16: Gamma for European Options

In Figure (1), we illustrate the gamma of a European option with respect to stock price for various time-to-maturities. It's important to note that, according to put-call parity, the gamma for European call and put options with the same strike are equal. Gamma is consistently positive, reflecting the convexity of options. Traders holding a long gamma position can profit through gamma scalping, a strategy involving regular portfolio rebalancing to maintain delta-neutrality. However, this strategy requires an upfront payment in the form of the option premium. Figure 4(b) depicts gamma as a function of time-to-maturity, prompting analysis of the observed behavior in the three curves.

Uses in Hedging :

- Dynamic Hedging: High gamma means the delta of the option changes rapidly with movements in the underlying asset. In dynamic hedging, this requires more frequent rebalancing of the portfolio.
- Gamma Neutral Hedging: A portfolio is gamma neutral if the total gamma of all positions is zero. Gamma neutral portfolios are less sensitive to large movements in the price of the underlying asset.
- Managing Risk: Traders use gamma along with delta to manage the risk of an options portfolio. A high gamma position may entail higher risk (and potential reward) because of the greater sensitivity to price changes of the underlying asset.

Hedging with Gamma

- Balancing Delta and Gamma: In practice, hedging with gamma involves balancing both delta and gamma. This can be complex because adjusting a position to be gamma neutral can affect delta neutrality, and vice versa. In a delta-hedged portfolio, gamma becomes important because it affects how often you need to rebalance the portfolio. If gamma is high, delta changes more rapidly, requiring more frequent rebalancing.
- Cost Considerations: Like delta hedging, gamma hedging involves transaction costs. High gamma may necessitate more frequent rebalancing, increasing these costs.
- Importance for Short-Term Traders: Gamma is particularly important for short-

term traders, such as market makers, because of its emphasis on the short-term responsiveness of an option's price to changes in the underlying asset.

Gamma is a key metric for understanding the dynamics of an option's price relative to movements in the underlying asset. It plays a crucial role in sophisticated trading strategies and risk management, particularly when dealing with portfolios that require frequent rebalancing due to rapidly changing market conditions. Understanding gamma, along with delta, is essential for effective options trading and hedging strategies.

6.11.4 Vega - Volatility sensitivity

The vega (V) of an option quantifies the sensitivity of the option price concerning changes in volatility. Specifically, for a call option, the vega is expressed as:

$$V = \frac{\partial C}{\partial \sigma} = e^{-qT}\frac{S\sqrt{T}}{\sqrt{2\pi}}e^{-\frac{1}{2}d_1^2}$$

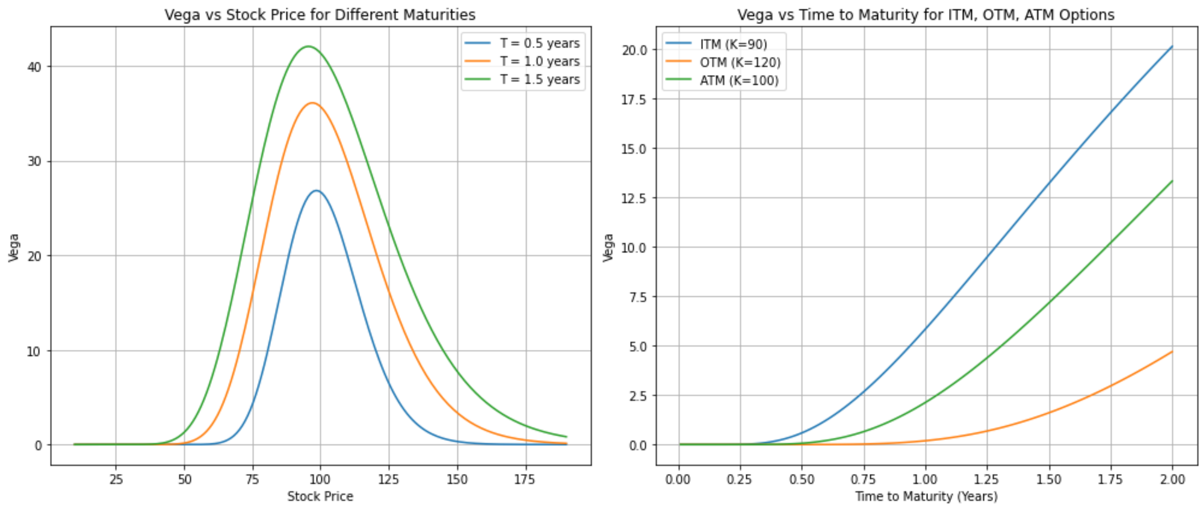


Figure 6.17: Vega for European Options

In Figure 5(b), we graph vega as a function of the underlying stock price, assuming $K = 100$. It's crucial to reiterate that, according to put-call parity, the vega of a call option equals the vega of a put option with an equivalent strike. The increase in vega with time-to-maturity is observed, and near the strike, vega reaches a peak. Analyzing Figure 5(b), one notices that vega approaches 0 as time-to-maturity approaches 0, aligning with the observations in Figure 5(a) and is consistent with the vega expression.

Crépey's Delta Hedging Vega Risk : Since Delta-Hedging does not account for the risk associated with changes in implied volatility.

Crépey's 2004 paper addresses this by proposing a method to hedge both delta and vega risk. The key insight is recognizing the dynamic nature of implied volatility and its impact on option pricing. To achieve a portfolio that is both delta and vega neutral, one must balance the sensitivity of the portfolio to changes in the underlying asset's price (delta) and to changes in its volatility (vega). The delta of a portfolio Π can be expressed as: $\Delta_\Pi = \frac{\partial \Pi}{\partial S}$ and the vega of the portfolio as: $\text{Vega}_\Pi = \frac{\partial \Pi}{\partial \sigma}$

To delta-vega hedge, one must adjust the portfolio such that both Δ_Π and Vega_Π are as close to zero as possible.

Practical Implementation Implementing a delta-vega hedging strategy involves several steps:

1. **Portfolio Analysis:** Determine the current delta and vega of your portfolio.

2. **Identifying Hedging Instruments:** Select appropriate hedging instruments (e.g., options with varying strikes and maturities).

3. **Hedging Strategy:** Calculate the required positions in the chosen instruments to neutralize the portfolio's delta and vega. This often involves solving a system of equations:

$$\begin{aligned} \Delta_\Pi + \sum \Delta_i x_i &= 0, \\ \text{Vega}_\Pi + \sum \text{Vega}_i x_i &= 0, \end{aligned} \tag{6.10}$$

where x_i are the quantities of the hedging instruments.

4. **Continuous Rebalancing:** As the market evolves, continuously monitor and rebalance the portfolio to maintain delta and vega neutrality.

Implementing delta-vega hedging requires sophisticated risk management tools and a deep understanding of the options market. It's also important to consider transaction costs and market liquidity when rebalancing the portfolio.

6.11.5 Theta - time sensitivity

The theta (Θ) of an option characterizes the sensitivity of the option price to a negative change in time-to-maturity. Specifically, for a call option, the theta is given by:

$$\Theta = -\frac{\partial C}{\partial T} = -e^{-qT} S\phi(d_1)\frac{\sigma}{2\sqrt{T}} + qe^{-qT}N(d_1) - rKe^{-rT}N(d_2)$$

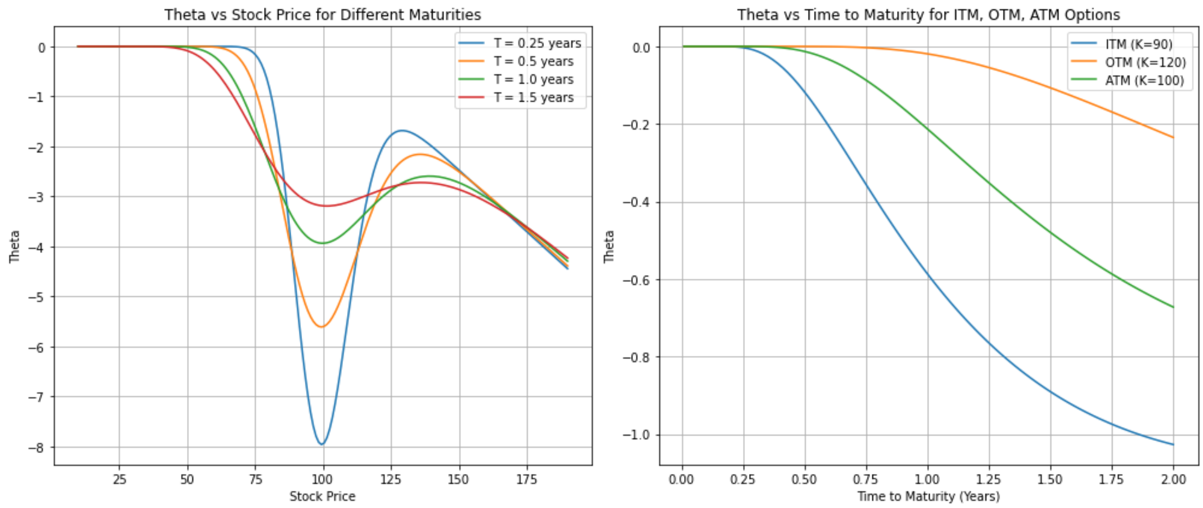


Figure 6.18: Theta for European Options

In Figure (a), we depict theta for four call options with varying times-to-maturity as a function of the underlying stock price. It is notable that the call option's theta is consistently negative. The negative nature of theta is explained, and the increasing negativity as time-to-maturity approaches 0 is examined.

In Figure (b), theta is again illustrated for three call options with different money-ness, this time in relation to time-to-maturity. The at-the-money (ATM) option exhibits the most negative theta, intensifying as time-to-maturity diminishes.

Positive theta is observed for in-the-money put options. Additionally, positive theta for call options can be achieved with a substantial q. However, in typical scenarios, theta for both call and put options is negative.

6.11.6 Rho - Interest rate sensitivity

Rho is one of the less frequently discussed "Greeks" in options trading, but it's important for understanding the sensitivity of an option's price to changes in interest rates. Let's explore Rho in detail:

Rho measures the rate of change of the option's price with respect to changes in the risk-free interest rate:

$$\text{Rho} = \frac{\partial V}{\partial r}$$

For a European call or put option under the Black-Scholes model, the Rho can be expressed as:

- For a Call Option:$\text{Rho}_{\text{call}} = K(T-t)e^{-r(T-t)}N(d_2)$
- For a Put Option:$\text{Rho}_{\text{put}} = -K(T-t)e^{-r(T-t)}N(-d_2)$

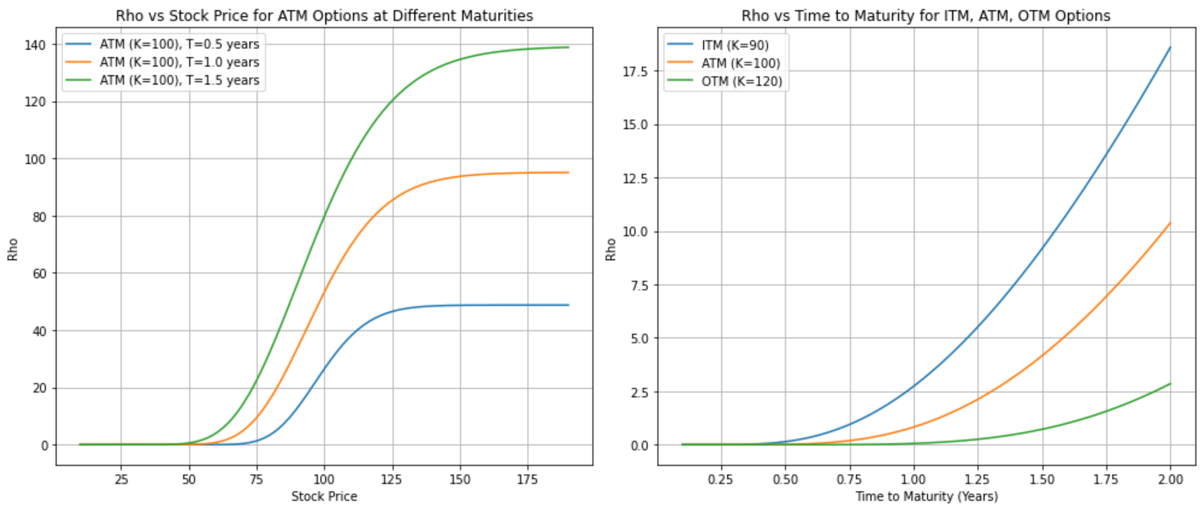


Figure 6.19: Rho for European Options

Rho indicates how much an option's price changes when the interest rate changes by one percentage point. For call options, Rho is typically positive, meaning the option's value increases with rising interest rates. For put options, Rho is usually negative. Rho's impact is more pronounced for options with longer maturities because the present value of the strike price payment is more sensitive to interest rate changes.

Rho is used to gauge and manage the exposure of an options portfolio to changes in interest rates, particularly important in environments where interest rate shifts are expected. Rho can help in structuring a portfolio to diversify or hedge against interest rate risk.

Achieving Rho neutrality means constructing a portfolio where the sensitivity to interest rate changes is minimized. However, this is less commonly targeted compared to Delta and Vega hedging. In a rising interest rate environment, holding positions with positive Rho (like long call options) can be beneficial, while in a declining interest rate environment, positions with negative Rho (like long put options) might be favored.

6.11.7 Vanna - Delta sensitivity to Vega (Skew)

Vanna is a second-order "Greek" in options trading, less commonly discussed than primary Greeks like Delta and Vega, but it's crucial in sophisticated trading and risk management strategies. Vanna measures the sensitivity of an option's delta with respect to changes in volatility, or alternatively, the sensitivity of the option's vega with respect to changes in the underlying asset's price.

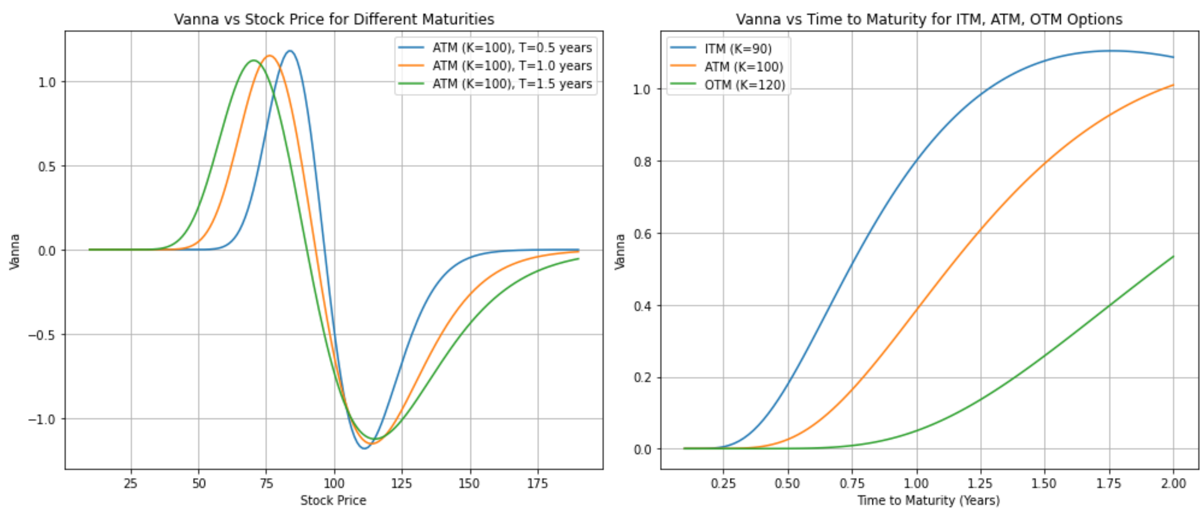


Figure 6.20: Vanna for European Option

Mathematical Definition

Vanna is defined as the second mixed partial derivative of the option's value V with respect to the underlying asset's price S and the volatility σ:

$$\text{Vanna} = \frac{\partial^2 V}{\partial S \partial \sigma}$$

In the Black-Scholes model, for a European call or put option, Vanna can be expressed as:

$$\text{Vanna} = -e^{-r(T-t)} N'(d_1) \left[d_2 / \sigma \right]$$

Vanna indicates how much the delta of an option will change as volatility changes, which is crucial when adjusting delta-hedged positions in response to shifts in market volatility. Alternatively, it shows how much the vega of an option will change as the underlying asset's price changes, which is important for managing vega risk in a portfolio.

Vanna helps in understanding and managing the skewness and curvature of the volatility surface. This is particularly important in environments where volatility is not constant. Traders use Vanna to adjust the delta and vega of their portfolios in response to movements in the underlying asset's price or changes in volatility. In the pricing and hedging of exotic options, Vanna plays a crucial role due to the higher sensitivity of these options to changes in volatility and the underlying asset's price.

Vanna can be used to dynamically hedge a portfolio against small changes in both the underlying asset's price and volatility. Vanna is important when balancing delta and vega risks, as adjustments to manage one can affect the other.

6.11.8 Volga - Vega sensitivity

Volga, also known as "Vomma," is another second-order "Greek" in options trading, which is particularly important for managing the convexity of an option's Vega with respect to changes in volatility. It's a measure used primarily in sophisticated trading strategies and risk management.

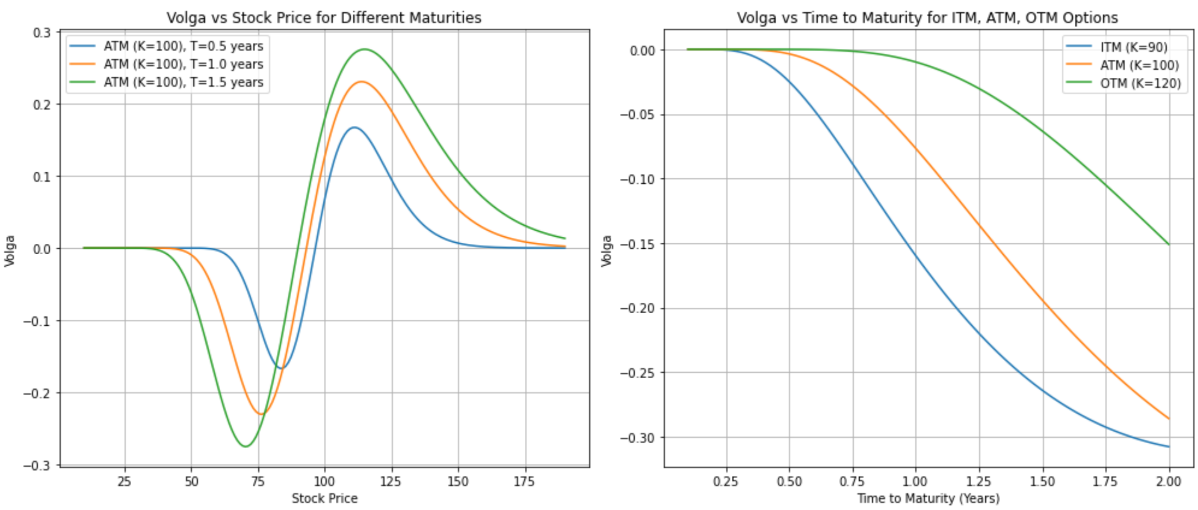


Figure 6.21: Volga for European Option

Volga is defined as the second partial derivative of the option's value V with respect to the volatility σ:

$$\text{Volga} = \frac{\partial^2 V}{\partial \sigma^2}$$

In the Black-Scholes model, for a European call or put option, Volga can be expressed as:

$$\text{Volga} = S\sqrt{T-t}e^{-r(T-t)}N'(d_1)d_1 d_2/\sigma$$

Volga measures the convexity of an option's Vega with respect to changes in the volatility of the underlying asset. It indicates how Vega will change as volatility changes. Options with high Volga are more sensitive to changes in volatility, meaning that their Vega changes more rapidly as volatility changes.

Traders use Volga to understand and manage the risks associated with large movements in volatility, particularly for portfolios with significant vega exposure. Volga is crucial in the pricing and hedging of exotic options, where the exposure to volatility changes can be more complex. In strategies where volatility trading is a key component, understanding Volga helps in making more informed decisions.

Like other second-order Greeks, Volga can be used in dynamic hedging to adjust the portfolio's exposure to changes in volatility. Traders might need to adjust the vega of their portfolios based on their Volga exposure to manage the convexity risk.

6.11.9 Other Greeks

Besides the primary Greeks (Delta, Gamma, Vega, Theta, and Rho) and the second-order Greeks (Vanna and Volga), there are several other Greeks used in more sophisticated options trading and risk management. These include:

1. Lambda (Elasticity or Leverage): Measures the percentage change in an option's value per percentage change in the underlying asset's price. It's similar to Delta but adjusts for the leverage effect.

2. Charm (Delta Decay): Measures the rate of change of Delta over time. It's useful for managing delta-hedging strategies as it indicates how Delta will change as the option

approaches expiration.

3. Veta (Vera or Volga Decay): Measures the change in Volga over time. It's important for assessing how the convexity of Vega changes as the option gets closer to its expiry.

4. Color: Measures the rate of change of Gamma with respect to the passage of time. It helps in understanding the behavior of Gamma as the option nears expiration.

5. Speed: A third-order Greek, measuring the rate of change in Gamma relative to the movement in the underlying asset's price. It's used for fine-tuning Gamma hedging strategies.

6. Ultima: Measures the sensitivity of the option's vomma with respect to changes in volatility. It's a third-order Greek and is used in managing portfolios with large Vega and Volga exposures.

7. Zomma: Measures the sensitivity of Gamma with respect to changes in volatility. It's used to adjust hedging strategies, especially in portfolios sensitive to large movements in volatility.

Each of these Greeks offers a different perspective on the risk characteristics of options and their portfolios. They're particularly useful in advanced trading strategies and for managing large, complex portfolios. These Greeks help traders and risk managers understand and hedge against the nuanced and often complex risks inherent in options trading.

6.12 Greeks summary

The five option Greeks

Δ	**Delta**	Represents the sensitivity of an option's prive to changes in the value of the underlying security.
Θ	**Theta**	Represents the rate of time decay of an option.
Γ	**Gamma**	Represents the rate of change of delta relative to the change of the price of the underlying security.
V	**Vega**	Represents an option's sensitivity to volatility.
ρ	**Rho**	Represents how sensitive the price of an option is relative to interest rates.

Figure 6.22: Greeks Summary

7 Overview of Rates Models

7.1 Introduction to Interest Rates Models

To begin with, the Hull-White, SABR, CIR (Cox-Ingersoll-Ross), and Vasicek models, are key pillars in the domain of financial mathematics for modeling interest rates and derivatives.

The journey through these models begins with the Vasicek model, introduced in 1977 by Oldrich Vasicek. This model was a pioneering effort in the interest rate modeling landscape, utilizing a mean-reverting stochastic differential equation. It's known for its simplicity and mathematical tractability, characterized by the equation $dr_t = a(b - r_t)dt + \sigma dW_t$. While the Vasicek model's assumption of normally distributed rates and its simplicity are appealing, it has its drawbacks, such as the potential for negative interest rates and the unrealistic assumption of constant volatility. Vasicek's seminal work was published in the Journal of Financial Economics in 1977, setting a foundational stone in interest rate theory.

Building on the Vasicek model, the Cox-Ingersoll-Ross (CIR) model, introduced in 1985 by John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross, addressed some of Vasicek's limitations. The CIR model ensures non-negative interest rates by incorporating a square root process, as seen in its equation $dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t$. This model retained the mean-reversion feature and introduced a volatility component dependent on the level of interest rates. Despite being more complex than its predecessor and presenting calibration challenges, the CIR model marked a significant advancement, as detailed in their publication in Econometrica.

In the early 1990s, the financial world witnessed the introduction of the Hull-White model by John Hull and Alan White. This model is an extension of the Vasicek and CIR models, offering more flexibility through time-dependent mean reversion level and volatility. Represented by the equation $dr_t = (\theta(t) - ar_t)dt + \sigma(t)dW_t$, the Hull-White model can accurately fit the initial term structure and is relatively simple to calibrate. However, its increasing complexity with time-dependent parameters is a noted limitation. Hull and White's contribution significantly enhanced the practical application of interest rate models, as discussed in their work in the Review of Financial Studies.

Lastly, the SABR model, introduced in 2002 by Patrick Hagan, Deep Kumar, Andrew Lesniewski, and Diana Woodward, marked a shift towards modeling the volatility smile in derivatives markets. Unlike its predecessors, the SABR model is a stochastic volatility model involving two coupled stochastic differential equations for the asset price and the volatility. It is particularly adept at capturing the volatility smile in derivative markets, making it a popular choice in interest rate and FX markets. Despite its robustness in capturing market phenomena, the SABR model is complex and requires sophisticated numerical methods for pricing, along with challenging calibration processes. The model's intricacies and applications were comprehensively explained in the authors' publication titled "Managing smile risk" in The Best of Wilmott.

7.2 Overview of Vasicek Model

7.2.1 Theoretical aspect

The Vasicek model is described by the following stochastic differential equation:

$$dr_t = a(b - r_t)dt + \sigma dW_t$$

Where:

- r_t is the short-term interest rate at time t.
- a, b, and σ are parameters of the model:
  - $a > 0$ is the speed of reversion.
  - b is the long-term mean level.
  - $\sigma > 0$ is the instantaneous volatility.
- dW_t represents the Wiener process (or Brownian motion).

Its key features are the following :

1. Mean Reversion: The term $a(b - r_t)$ drives the rate r_t towards the long-term mean b with speed a. This reflects the empirical observation that interest rates tend to revert

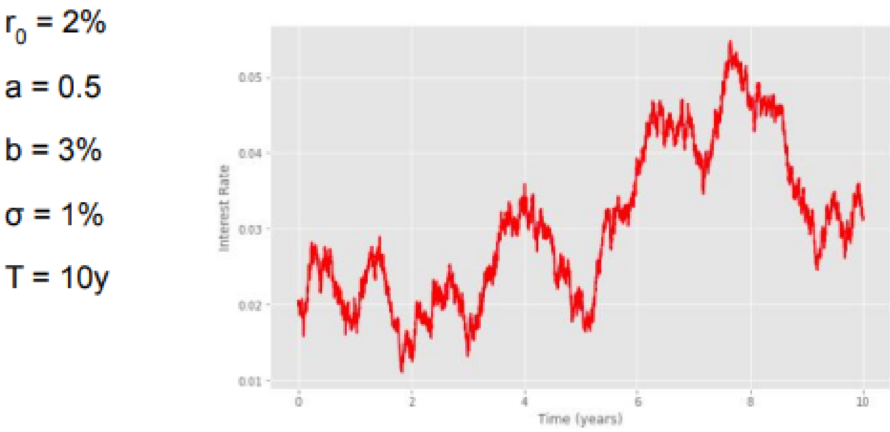


Figure 7.1: Simulation of Interest Rates with Vasicek Model (Source: QuantNext)

to an average level over time.

2. Normal Distribution: The solution to the Vasicek model yields normally distributed rates, which is a simplification and can be a drawback as it allows for negative interest rates, which may not be realistic in all financial contexts.

3. Analytical Tractability: One of the strengths of the Vasicek model is that it allows for an analytical solution for bond prices, which is not possible in many other models of interest rates.

The previous SDE can be solved such as :

Applying Itô's Lemma with $x = r_t \cdot e^{a \cdot t}$

$$dx_t = dr_t \cdot e^{a \cdot t} + r_t \cdot a \cdot e^{a \cdot t} dt$$

With :

$$dr_t = a \cdot (b - r_t) \cdot dt + \sigma \cdot dW_t$$

gives :

$$dx_t = a \cdot b \cdot e^{a \cdot t} \cdot dt + \sigma \cdot e^{a \cdot t} \cdot dW_t$$

$$x_t - x_0 = a \cdot b \cdot \int_0^t e^{a \cdot s} \cdot ds + \sigma \cdot \int_0^t e^{a \cdot s} dW_s$$

$$r_t = r_0 e^{-a \cdot t} + b \cdot (1 - e^{-a \cdot t}) + \sigma \cdot e^{-a \cdot t} \cdot \int_0^t e^{a \cdot s} \cdot dW_s$$

And r_t follows a normal distribution :

$$r_t \sim \mathbf{N}\left(r_0 \cdot e^{-a \cdot t} + b \cdot (1 - e^{-a \cdot t}), \frac{\sigma^2}{2 \cdot a} \cdot (1 - e^{-2a \cdot t})\right)$$

The mean tends to b when t tends to infinite, and the standard deviation tends to $\frac{\sigma^2}{2a}$ when t tends to infinite.

7.2.2 Applications

The Vasicek model, fundamentally a theoretical construct for interest rate dynamics, is widely used in finance for pricing interest rate derivatives, valuing bonds, and managing risk. Its theoretical utility is derived from its closed-form solutions for bond prices and the yield curve, as well as its facility for simulation in risk management and scenario analysis.

Theoretical Usage :

1. Bond Pricing: In the Vasicek model, the price of a zero-coupon bond can be expressed in a closed-form equation, which is rare in interest rate modeling. This is particularly useful for pricing bonds and bond derivatives.

2. Yield Curve Modeling: The model can be used to construct the entire yield curve. Since interest rates are normally distributed in this model, the yield curve can be straightforwardly calculated.

3. Risk Management: The model's parameters can be calibrated to market data, allowing risk managers to simulate various interest rate scenarios and assess the impact on a portfolio.

4. Option Pricing on Bonds: The model can be used to price options on bonds, such as caps, floors, and swaptions.

For zero-coupon bonds, the Vasicek model provides a closed-form solution for bond prices. The formula for a zero-coupon bond price is:

$$P(t, T) = e^{(A(t,T) - B(t,T) r_t)}$$

Where:

- $P(t, T)$ is the price at time t of a zero-coupon bond maturing at time T,
- $A(t, T)$ and $B(t, T)$ are functions of the model parameters and time, given by:

$$A(t, T) = \left[\left(b - \frac{\sigma^2}{2a^2} \right) (B(t, T) - (T - t)) - \frac{\sigma^2 B^2(t, T)}{4a} \right]$$

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a}$$

7.2.3 Discretization and Simulation

To simulate the interest rate path over time, one typically employs the Euler-Maruyama method to discretize the stochastic differential equation. The discrete-time version of the Vasicek model can be written as:

$$r_{t+\Delta t} = r_t + a(b - r_t)\Delta t + \sigma\sqrt{\Delta t} Z_t$$

Where:

- Δt is the time step,
- Z_t is a standard normal random variable.

These simulations allow risk managers to assess the interest rate risk and to perform stress testing under various economic scenarios.

The model parameters a, b, and σ are usually calibrated to fit the current market data. This involves optimizing these parameters so that the model's output (e.g., bond prices or the yield curve) aligns closely with the observed market data. Calibration is typically performed using numerical optimization techniques.

It is used in the applications of :

- Derivative Pricing: Options on bonds and interest rate derivatives like swaptions can be priced using the Vasicek model.
- Asset and Liability Management: Financial institutions use the model to assess interest rate risk on their balance sheets.

7.2.4 Calibration by least squares

The discretisation equation is :

$$r_{t+\Delta t} = r_t + a(b - r_t)\Delta t + \sigma\sqrt{\Delta t}Z_t$$

Which can be rearranged as :

$$r_{t+\Delta t} - r_t = (1 - a\Delta t) \cdot r_t + a \cdot b \cdot \Delta t + \sigma\sqrt{\Delta t}\varepsilon$$

$$r_{t+\Delta t} - r_t = \alpha \cdot r_t + \beta + \xi$$

with ξ iid normally distributed.

$(\hat{\alpha}, \hat{\beta})$ estimated by least squares method (regression of $r_{t+\Delta t}$ on r_t.

Then,

$$(\hat{a}, \hat{b}) = \left(\frac{1 - \hat{\alpha}}{\Delta t}, \frac{\hat{\beta}}{1 - \hat{\alpha}}\right)$$

and,

$$\sigma^2 = \sqrt{\frac{Var(\xi)}{\Delta t}}$$

Maximum Likelihood Estimation : This calibration can also be done using Maximum Likelihood methods.

7.2.5 Limitations

While the Vasicek model is popular due to its simplicity and analytical tractability, it has limitations:

- It can produce negative interest rates, which may not be realistic.
- The assumption of constant volatility may not hold true in all market conditions.

To address some of these limitations, several extensions and variations of the Vasicek model have been proposed, like the Hull-White model and the Cox-Ingersoll-Ross (CIR) model, which modify the dynamics to prevent negative rates and to allow for time-varying parameters.

7.3 Overview of Cox-Ingersoll-Ross (CIR) Model

The Cox-Ingersoll-Ross (CIR) model, developed by John C. Cox, Jonathan E. Ingersoll Jr., and Stephen A. Ross in 1985, is a mathematical model used in financial economics to describe the evolution of interest rates. It is a type of one-factor short-rate model and is primarily used for modeling the dynamics of interest rates and pricing interest rate derivatives.

The model was introduced as an alternative to the Vasicek model. It assumes that the short-term interest rate follows a mean-reverting stochastic process, it does not allow negative interest rates while preserving analytical solution for bond pricing. It is also used in the popular stochastic volatility Heston model to model the stochastic variance.

The CIR model describes the evolution of interest rates using a stochastic differential equation (SDE) that incorporates mean reversion, a characteristic of many economic variables. The SDE for the short rate, r_t, under the CIR model is given by:

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$$

Here:

- r_t is the short rate at time t.
- κ is the speed of mean reversion.
- θ is the long-term mean level of the interest rate.
- σ is the volatility of the interest rate.
- dW_t is a Wiener process, representing the random market risk.

A key feature of the CIR model is that it ensures interest rates remain positive, a property not guaranteed in some other models, like the Vasicek model. If the **Feller condition** $(2 \cdot a \cdot b > \sigma^2)$ is respected, then r_t is strictly positive.

The CIR model is widely used in the valuation of interest rate derivatives, such as bond options and interest rate caps/floors. It is also used in the construction of yield curves and in risk management strategies.

The Cox-Ingersoll-Ross model represents a significant contribution to the field of financial economics, providing a robust framework for understanding and modeling the dynamics of interest rates. Its introduction has been influential in both theoretical finance and practical applications, particularly in the areas of bond valuation and risk management.

The probability density function of the future value r_s at the future time s conditionally to its current value r_t at the current time t is given by :

$f(r_s, s; r_t, t) = c \cdot e^{-(u+v)} \cdot (\frac{v}{u})^{q/2} \cdot I_q(2 \cdot (u \cdot v)^{\frac{1}{2}})$ Where :

- $c = \frac{2\kappa}{\sigma^2(1-e^{-\kappa(s-t)})}$
- $u = c \cdot r_t \cdot e^{-\kappa \cdot (s-t)}$
- $v = c \cdot r_s$
- $q = \frac{2 \cdot \kappa \cdot \theta}{\sigma^2} - 1$
- I_q the modified Bessel function of the first kind of order q.

We can calculate the expected value of r_s :

$E(r_s/r_t) = r_t \cdot e^{-\kappa \cdot (s-t)} + \theta \cdot (1 - e^{-\kappa \cdot (s-t)})$

$E(r_s/r_t) \to \theta \quad \text{as} \quad s \to +\infty$

As well as the variance of r_s :

$Var(r_s/r_t) = r_t \cdot \frac{\sigma^2}{\kappa} \cdot (e^{-\kappa \cdot (s-t)} - e^{-2 \cdot a \cdot (s-t)}) + \theta \cdot \frac{\sigma^2}{2\kappa} \cdot (1 - e^{-\kappa(s-t)})^2$

$Var(r_s/r_t) \to \theta \cdot \frac{\sigma^2}{2\kappa} \quad \text{as} \quad s \to +\infty$

The distribution function is a non-central Chi-squared with $2q + 2$ degrees of freedom and non-centrality parameter of $2u$. The asymptotic function of r_s when s becomes large enough is a Gamma distribution with the following density :

$f(r) = \frac{\omega^\nu}{\Gamma(\nu)} \cdot r^{\nu-1} \cdot e^{-\omega \cdot r}$

Where :

- $\omega = \frac{2\kappa}{\sigma^2}$

- $\nu = \frac{2\cdot\kappa\cdot\theta}{\sigma^2}$

The Half-Life of the Mean-Reversion is the average time that it takes to be half-way back to the mean.

$r_0 \cdot e^{-\kappa\cdot t_{1/2}} + \theta \cdot (1 - e^{-\kappa\cdot t_{1/2}} - r_0 = \frac{\theta - r_0}{2}$

$t_{1/2} = \frac{ln(2}{\kappa}$

The higher the speed of reversion, the smaller the half life.

Zero-coupon bond pricing: The price of a ZCB at the date t with maturity T has the following expression in the CIR model, under the No-Arbitrage assumption :

$P(t,T) = A(t,T) \cdot e^{-B(t,T)\cdot r(t)}$

With :

- $A(t,T) = (\frac{2\cdot\gamma\cdot e^{(\gamma+\kappa)\cdot\frac{T-t}{2}}}{2\cdot\gamma+(\kappa+\gamma)\cdot(e^{\gamma\cdot(T-t)}-1)})^{\frac{2\cdot\kappa\cdot\theta}{\sigma^2}}$
- $B(t,T) = \frac{1\cdot(e^{\gamma\cdot(T-t)}-1)}{2\cdot\gamma+(\kappa+\gamma)\cdot(e^{\gamma\cdot(T-t)}-1)}$
- $\gamma = \sqrt{\kappa^2 + 2\cdot\sigma^2}$

The limits of this model :

Impossibility to have negative values, which was positive to model interest in pre-crisis world but can be problematic in post-crisis one. The flexibility of the model is limited, it is typically unable to reproduce all SC curve shapes observed in the market. It is a single factor model, with a constant volatility parameter and it does not incorporate jumps.

7.4 Overview of Hull-White Model

7.4.1 Theoretical aspect

The Hull-White model, developed by John Hull and Alan White, is an extension of the Vasicek and Cox-Ingersoll-Ross (CIR) models. It's a one-factor interest rate model used extensively in financial engineering for the valuation of interest rate derivatives. The Hull-White model modifies the Vasicek model to allow for a time-varying mean reversion level and volatility, which provides a better fit to the initial term structure of interest rates.

The Hull-White model is often expressed as follows:

$$dr_t = [\theta(t) - ar_t]dt + \sigma(t)dW_t$$

Where:

- r_t is the short-term interest rate at time t.
- a is the speed of mean reversion.
- $\theta(t)$ is a time-dependent function chosen to fit the current term structure of interest rates.
- $\sigma(t)$ is the time-dependent volatility.
- dW_t is the Wiener process (or Brownian motion).

The key features of the model are the following :

1. Time-Varying Parameters: Unlike the Vasicek model, the Hull-White model includes time-dependent parameters $\theta(t)$ and $\sigma(t)$, allowing for a more flexible fit to the term structure of interest rates observed in the market.

2. No Arbitrage: The model is calibrated to the current yield curve, ensuring no-arbitrage conditions.

3. Mean Reversion: Similar to the Vasicek model, it includes a mean reversion term, but the mean reversion level can change over time.

4. Flexibility and Tractability: The model retains the analytical tractability of the Vasicek model while providing more flexibility to fit market data.

The function $\theta(t)$ is determined in such a way that the model is consistent with the initial term structure of interest rates. This is achieved by calibrating $\theta(t)$ so that the theoretical zero-coupon bond prices match the market prices.

7.4.2 Applications

The model's strength lies in its ability to incorporate time-varying parameters while still allowing for analytical solutions for certain financial instruments, like zero-coupon bonds and European options on bonds.

In the Hull-White model, the price of a zero-coupon bond can be derived analytically. The formula is given by:

$$P(t,T) = A(t,T)e^{-B(t,T)r_t}$$

Here, $A(t,T)$ and $B(t,T)$ are functions defined as:

$$A(t,T) = \exp\left(\left(\int_t^T f(0,s)ds - \frac{1}{4a}\int_t^T B^2(s,T)\sigma^2(s)ds\right)\right)$$
$$B(t,T) = \frac{1-e^{-a(T-t)}}{a}$$

In these equations:

- $f(0,s)$ represents the instantaneous forward rate observable at time 0 for time s.
- $\sigma(s)$ is the time-dependent volatility.
- a is the speed of mean reversion.
- r_t is the short-term interest rate at time t.

The Hull-White model also allows for a closed-form solution for the price of European options on bonds. For a European call option on a zero-coupon bond, the price can be calculated using the Black formula, which is adapted for interest rate options.

For Monte Carlo simulations or other numerical methods, the Hull-White model can be discretized. The discrete-time version of the model is often represented as:

$$r_{t+\Delta t} = r_t + [\theta(t) - ar_t]\Delta t + \sigma(t)\sqrt{\Delta t}Z_t$$

Where:

- Δt is the time step.
- Z_t is a standard normal random variable.

The calibration process involves adjusting the time-dependent function $\theta(t)$ and the volatility function $\sigma(t)$. This is done so that the model's output matches the current market data, particularly the current term structure of interest rates.

The Hull-White model's ability to provide closed-form solutions for certain financial instruments, despite its time-varying parameters, makes it a powerful tool in financial engineering. It is particularly useful in the valuation of interest rate derivatives and in risk management applications. The model's flexibility in fitting the initial term structure and its analytical tractability in pricing complex derivatives are key reasons for its widespread use in the financial industry.

7.4.3 Limitations

The Hull-White model, while versatile and widely used, has several limitations. Firstly, its assumption of a normally distributed interest rate process can lead to negative interest rates, which may be unrealistic, especially in low-interest-rate environments. Secondly, the model's simplicity, though a strength in terms of tractability, can be a drawback when dealing with more complex term structure dynamics observed in real markets. The time-varying parameters, while adding flexibility, also introduce complexity in calibration, requiring more sophisticated techniques and accurate market data. Additionally, the model's one-factor nature limits its ability to capture the multi-dimensional nature of interest rate movements, particularly the different behaviors of short-term and long-term rates. These limitations mean that while the Hull-White model is useful for many practical applications, it may not fully capture the nuances of real-world interest rate dynamics, necessitating the use of more complex models or complementary approaches for certain advanced financial applications.

7.5 Overview of SABR Model

7.5.1 Theoretical aspect

The SABR (Stochastic Alpha, Beta, Rho) model is a sophisticated tool in financial mathematics, primarily used for modeling and estimating the volatility of financial instruments, particularly in the interest rate derivatives market. Developed by Patrick S. Hagan, Deep Kumar, Andrew S. Lesniewski, and Diana E. Woodward in 2002, the SABR model has become a standard in the industry due to its ability to handle varying market conditions and its adaptability in fitting market data.

The SABR model describes the behavior of a forward rate F and its corresponding volatility σ. It's a two-factor model where one stochastic process represents the forward rate and another represents the volatility of the forward rate. The dynamics are governed by the following stochastic differential equations:

$$dF_t = \sigma_t F_t^{\beta} dW_t^1$$

$$d\sigma_t = \alpha \sigma_t dW_t^2$$

Here:

- F_t is the forward rate.
- σ_t is the stochastic volatility.
- β is the elasticity parameter (often set to 0 for lognormal behavior or 1 for normal behavior).
- α is the volatility of volatility.
- dW_t^1 and dW_t^2 are two Wiener processes with correlation ρ.

The key features of the model are the following :

1. Stochastic Volatility: One of the primary features of the SABR model is its stochastic volatility component, allowing it to capture the volatility smile seen in market data.

2. Correlation: The model incorporates the correlation ρ between the forward rate and its volatility, which is crucial for accurately pricing interest rate derivatives.

3. Flexibility: The SABR model can adapt to various market conditions due to its flexible parameterization.

While the SABR model does not yield exact closed-form solutions for option prices, Hagan et al. provided an approximation formula for European options, which is widely used in practice:

$$\begin{aligned}\sigma_{\mathrm{BS}}(F,K,T) = {} & \frac{\alpha}{(FK)^{(1-\beta)/2}\left(1+\frac{(1-\beta)^2}{24}\log^2\frac{F}{K}+\frac{(1-\beta)^4}{1920}\log^4\frac{F}{K}\right)} \\ & \cdot\left(1+\left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(FK)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu}{(FK)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2\right)T\right)\end{aligned} \tag{7.1}$$

This formula provides an implied volatility (σ_{BS}) to be used in the Black-Scholes formula for option pricing.

7.5.2 Applications

The SABR model is a sophisticated tool in the field of financial engineering, especially valued for its ability to accurately capture the complex dynamics of volatility, including the volatility smile in interest rate markets. Its theoretical foundation and practical applications are grounded in advanced stochastic calculus and numerical methods.

1. Volatility Smile Modeling: The SABR model is extensively used to model the volatility smile in interest rate derivatives markets. The volatility smile is a pattern observed in markets where options' implied volatility varies with strike price and maturity, deviating from the constant volatility assumption of the Black-Scholes model.

2. Pricing Interest Rate Derivatives: It is particularly effective in pricing European-style interest rate derivatives like swaptions, caps, and floors. The model provides a more realistic framework for understanding the behavior of volatility across different strikes and maturities.

3. Risk Management: In risk management, the SABR model helps in understanding the behavior of the volatility surface over time, which is crucial for managing the risk of portfolios containing interest rate derivatives.

The SABR model itself does not offer a closed-form solution for option pricing. However, a widely-used approximation formula for the implied volatility (σ_{BS}) in the Black-Scholes model was derived by Hagan and his colleagues:

$$\begin{aligned}\sigma_{\mathrm{BS}}(F,K,T) = {} & \frac{\alpha}{(FK)^{(1-\beta)/2}\left(1+\frac{(1-\beta)^2}{24}\log^2\frac{F}{K}+\frac{(1-\beta)^4}{1920}\log^4\frac{F}{K}\right)} \\ & \cdot\left(1+\left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(FK)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu}{(FK)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2\right)T\right)\end{aligned} \tag{7.2}$$

Where:

- F is the forward rate.
- K is the strike price.
- T is the time to maturity.
- α, β, ρ, and ν are the SABR model parameters.

Calibrating the SABR model involves estimating the parameters α, β, ρ, and ν from market data. This process typically requires sophisticated optimization techniques and a deep understanding of the model's behavior under different market conditions.

7.5.3 Simulations

The SABR model is defined by a pair of coupled stochastic differential equations (SDEs) that describe the dynamics of both the forward rate F_t and its volatility σ_t. These SDEs are:

$$dF_t = \sigma_t F_t^{\beta} dW_t^1$$

$$d\sigma_t = \alpha \sigma_t dW_t^2$$

In these equations:

- F_t is the forward rate at time t.
- σ_t is the stochastic volatility at time t.
- β is the elasticity parameter, affecting the behavior of the forward rate.
- α is the volatility of volatility, a measure of how erratic the volatility is.
- dW_t^1 and dW_t^2 are two Wiener processes with a correlation coefficient ρ.

To simulate paths using the SABR model, these SDEs are discretized using numerical methods, such as the Euler-Maruyama method. The discretized equations over a small time interval Δt are:

$$F_{t+\Delta t} = F_t + \sigma_t F_t^{\beta} \sqrt{\Delta t} Z_t^1$$

$$\sigma_{t+\Delta t} = \sigma_t + \alpha \sigma_t \sqrt{\Delta t} Z_t^2$$

Here:

- Z_t^1 and Z_t^2 are normally distributed random variables with correlation ρ.
- Δt is the time step in the discretization.

These discretized equations are used in Monte Carlo simulations to generate possible future paths for the forward rate and its volatility. By simulating a large number of paths, one can compute the expected value or distribution of future payoffs for various financial instruments, such as options.

7.5.4 Limitations

The SABR model, while powerful in capturing the volatility smile and stochastic dynamics of interest rates, has certain limitations. First, its complexity can be a barrier, as accurate calibration and implementation require advanced mathematical and computational expertise. Second, the model's reliance on numerical methods for option pricing and risk analysis can lead to computational intensity, particularly in Monte Carlo simulations. This can be resource-intensive and time-consuming, especially for large-scale applications or real-time analytics. Third, while the model excels at short-term forecasting, its long-term predictions may be less reliable due to the inherent difficulty in accurately estimating parameters over extended periods. Additionally, the SABR model assumes constant parameters α, β, and ρ over time, which might not always align with real market dynamics where these parameters can evolve. Finally, like many models in finance, it relies on historical data for calibration, making it potentially vulnerable to changes in market conditions and structure that are not captured in past data. Despite these limitations, the SABR model remains a widely used and respected tool in quantitative finance for its robust handling of complex market phenomena.

8 Critique of model

8.1 Limitations of Existing Models

Black-Scholes-Merton equation falls short of accurately modeling price movements in financial markets primarily because of the inherent assumptions it makes. The model assumes constant volatility, risk-free interest rates, and a log-normal distribution of asset prices. However, these assumptions often diverge from the complex and dynamic reality of financial markets. Market conditions can change, volatility is not constant, interest rates fluctuate, and the distribution of asset prices may not always conform to the assumed log-normal distribution. As a result, the model's simplicity and reliance on these assumptions limit its effectiveness in capturing the intricacies and nuances of real-world financial markets.

References

[1] Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 81(3), 637-654.

[2] Boyle, P. P. (1977). Options: A Monte Carlo Approach. *Journal of Financial Economics*, 4(3), 323-338.

[3] Boyle, P., Broadie, M., & Glasserman, P. (1997). Monte Carlo Methods for Security Pricing. *Journal of Economic Dynamics and Control*, 21(8-9), 1267-1321.

[4] Cameron, P. (2017). Equity Derivatives: Corporate and Institutional Applications. Palgrave Macmillan.

[5] Chance, D. M., & Brooks, R. (2015). An Introduction to Derivatives and Risk Management. Cengage Learning.

[6] Chi, S. (2019, June 8th). Brownian motion, Ito's lemma, and the Black-Scholes formula (Part I and II). LinkedIn. https://www.linkedin.com/pulse/brownian-motion-itos-lemma-black-scholes-formula-part-chuan-shi

[7] Chriss, N. A. (n.d.). Black-Scholes and Beyond: Option Pricing Models. Columbia University. Retrieved [date you accessed the resource], from https://www.columbia.edu/~mh2078/FoundationsFE/BlackScholes.pdf

[8] Clewlow, L., & Strickland, C. (1998). Implementing Derivatives Models. Wiley.

[9] Cofnas, A. (2008). The Forex Options Course: A Self-Study Guide to Trading Currency Options. Wiley.

[10] Cox, J.C., Ingersoll, J.E., & Ross, S.A. (1985). "A Theory of the Term Structure of Interest Rates." *Econometrica*, 53(2), 385-407.

[11] Dash, S. (2004). Derivatives: The Tools That Changed Finance. Risk Books.

[12] Dupire, B. (1994). Pricing with a Smile. *Risk*, 7(1), 18-20.

[13] Fabozzi, F. J., Mann, S. V., & Choudhry, M. (2012). The Handbook of Fixed Income Securities. McGraw-Hill Education.

[14] Geman, H. (2005). Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals, and Energy. Wiley.

[15] Glasserman, P. (2004). Monte Carlo Methods in Financial Engineering. Springer.

[16] Hagan, P., Kumar, D., Lesniewski, A., & Woodward, D. (2002). "Managing Smile Risk." *The Best of Wilmott*, 1, 249-296.

[17] Hardy, M. (2009). Credit Derivatives: Understanding and Working with the 2007 ISDA Definitions. ISDA.

[18] Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2), 327-343.

[19] Hull, J. C. (2017). Options, Futures, and Other Derivatives. Pearson Education Limited.

[20] Hull, J. C. (2017). Risk Management and Financial Institutions. Wiley.

[21] Hull, J., & White, A. (1990). "Pricing Interest-Rate-Derivative Securities." *The Review of Financial Studies*, 3(4), 573-592.

[22] Joshi, M. (2003). The Concepts and Practice of Mathematical Finance. Cambridge University Press.

[23] Joshi, M. S. (2008). The Concepts and Practice of Mathematical Finance (Mathematics, Finance and Risk) (2nd ed.). Cambridge University Press.

[24] Kemeny, J. G., & Snell, J. L. (1960). Finite Markov Chains. Van Nostrand.

[25] Kolb, R. W., & Overdahl, J. A. (2007). Futures, Options, and Swaps. Wiley-Blackwell.

[26] Markov, A. A. (1913). An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains. *Science in Context*.

[27] McDonald, R. L. (2006). Derivatives Markets. Pearson Education.

[28] Norris, J. R. (1997). Markov Chains. Cambridge University Press.

[29] Shiller, R. J. (2003). The New Financial Order.

[30] Smith, C. W., Smithson, C. W., & Wilford, D. S. (1995). Managing Financial Risk. HarperBusiness.

[31] Tett, G. (2009). Fool's Gold: How the Bold Dream of a Small Tribe at J.P. Morgan Was Corrupted by Wall Street Greed and Unleashed a Catastrophe. Free Press.

[32] Tuckman, B., & Serrat, A. (2011). Fixed Income Securities: Tools for Today's Markets. Wiley.

[33] Vasicek, O. (1977). "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics*, 5(2), 177-188.

[34] Wilmott, P. (2007). Paul Wilmott Introduces Quantitative Finance. Wiley.