

How Become a Quantitative Analyst

May 13, 2019

Contents

1	Introduction.	9
1.1	Basis notions.	9
1.1.1	Bonds vs Stocks.	11
1.1.2	Derivatives.	12
2	Calculus	13
2.1	Taylor Series	13
2.2	Numerical methods for one dimensional nonlinear problems.	14
2.2.1	Bisection method.	14
2.2.2	Newton's method.	14
3	Probability.	17
3.1	Conditional expectation.	17
3.1.1	Properties of Conditional expectation.	18
3.2	Log-normal distribution	19
3.3	Central limit theorem.	19
4	Stochastic process.	21
4.1	Ito Lemma	21
4.2	Ito integral	22
4.3	Martingale.	22
4.4	Radon-Nikodym theorem.	23
4.5	Girsanov theorem.	23
4.6	Brownian motion.	24
4.6.1	Basic Properties of Brownian Motion.	24
4.7	Poisson process.	25
4.7.1	Characteristic of Poisson process.	25
4.8	Compound Poisson process.	25
4.9	Options	25
5	Black-Scholes model	27
5.1	Induction	27
5.2	Assumptions	28
5.3	Derivation the Black-Scholes equation for a stock.	29
5.4	Facts about the Black-Scholes equation	29
5.5	Valuation formula.	29
5.5.1	Black Scholes 76 (BS 76)	30
5.5.2	The Black-Scholes PDE	31
5.5.3	Limitations of Black-Scholes model.	31
5.6	Put-call parity.	31
5.7	No-Arbitrage bounds on option prices.	32

5.7.1	Bonds on call options	32
5.7.2	Bounds on put options.	34
5.7.3	Early exercise	34
5.8	Greeks.	35
5.8.1	Properties of Greeks.	35
5.8.2	Greeks for forward contract	36
5.8.3	Useful formulas for Greeks.	37
5.8.4	Methods of Calculating Greek Parameters.	37
5.9	Combination	38
5.10	Hedging and replication.	38
6	Building interest rate curves.	39
6.1	Introduction.	39
6.1.1	The classical single-curve approach.	39
6.1.2	Interpolation.	40
6.1.3	Cubic spline interpolation.	40
7	Implied volatility	41
7.1	Summary about volatility.	41
8	Local volatility models.	43
8.1	Displaced diffusion model.	43
8.2	Constant elasticity of variance model (CEV)	43
8.3	Duprie formula	43
9	Jump diffusion model.	45
10	Stochastic volatility models.	47
10.1	Induction.	47
10.2	Heston model.	47
10.2.1	Influence of parameters.	48
10.2.2	Advantages and disadvantages of the Heston model.	48
10.3	SABR model.	48
10.4	Discretization.	50
10.5	Option pricing via Monte Carlo simulation.	50
10.5.1	Option price proprieties	51
10.5.2	Stock price.	51
10.5.3	Strike price.	51
10.5.4	volatility	51
10.5.5	Time to maturity	51
10.5.6	Interest rate	51
10.6	Forward contract and future contract.	51
10.7	Currency option valuation formula.	52
10.8	Exotic options	53
10.8.1	Chooser options	53
10.8.2	Asian options	53
10.9	Stochastic models of bond prices.	54

11 Interest rates	55
11.0.1 Forward interest rate.	55
11.0.2 Short-term interest rate	56
11.1 Short-term rate models.	56
11.1.1 Notation.	56
11.1.2 Mean reversion process	56
11.1.3 Ornstein-Uhlenbeck process.	57
11.1.4 Merton model.	57
11.1.5 Vasicek model.	58
11.1.6 Cox Ingersoll Ross model.	58
11.1.7 Hull and White model.	58
11.1.8 General remarks.	58
11.2 Delta-hedging	59
11.2.1 Primary market.	59
11.2.2 Single period case	60
11.2.3 Fundamental theorem.	60
11.3 Bonds.	61
11.3.1 Price of bonds	61
11.3.2 Duration and convexity.	62
11.3.3 Convexity.	63
11.4 Hypothesis testing an overview.	63
11.5 Assumptions of model	63
11.6 Classic model of linear regression	64
11.7 Analysis of Variance	65
11.8 Confidence interval.	66
11.9 Hypothesis testing.	66
11.10 Another tests	66
11.11 How to measure a risk.	67
11.11.1 Introduction.	67
11.11.2 Definition and interpretation	67
11.11.3 Value at risk method.	69
11.11.4 Advantages and disadvantages of VaR.	69
11.11.5 Expected shortfall.	69
11.12 Back testing of VaR.	70
11.13 Exchange and Clearing House	70
11.14 The Margin System	70
11.15 Interest rate swaps	70
11.15.1 Idea and definition.	70
11.15.2 Pricing interest rate swaps (IRS).	71
11.15.3 Overnight indexed swap.	71
11.15.4 Interest basis swap.	71
11.16 Default	71
11.17 Spreads and default risk.	72
11.17.1 Credit default swap.	72
11.18 Credit valuation adjustment (CVA)	73
11.19 Estimation probability of default	73
11.20 Merton's approach to corporate debt.	74
11.21 Basic.	74
11.22 Linear Processes.	76
11.23 Autoregressive series.	77
11.23.1 Why is stationarity important in the historical VaR model?	77
11.23.2 How are we testing stationarity?	77

11.24EWMA and GARCH	78
11.24.1EWMA	78
11.24.2GARCH(1,1).	78
11.25Risk	78
12 Credit Risk.	81
12.1 induction	81
12.2 Popular questions.	81
13 Cracking the Coding Python	83
13.1 Numbers	83
13.2 Numerical analysis	84
13.3 Sorting	85
13.4 Searching	85
13.5 LeetCode Problems	86
13.6 Useful tricks	87
14 SQL	89
15 Coding in C++	91
15.1 C++.	91
15.2 Python.	91
15.3 Useful phrases.	91
15.4 EXCEL	93
16 SQL	95
16.1 Popular Interview Questions.	95
16.2 Data definition	95
16.3 Data manipulation	95

- Ω sample space
- \mathcal{F} sigma field
- \mathbb{P}, \mathbb{Q} probability measure
- $(\Omega, \mathcal{F}, \mathbb{P})$ probability space
- T^* time horizon for all market activities
- $(F_t)_{t \in [0, T^*]}$

Chapter 1

Introduction.

1.1 Basis notions.

Definition 1.1.1 (*Plain vanilla*)

Plain vanilla signifies the most basic or standard version of a financial instrument, usually options, bonds, futures and swaps. *Plain vanilla* is the opposite of an exotic instrument, which alters the components of a traditional financial instrument, resulting in a more complex security.

Definition 1.1.2 (*OTC*)

Financial securities instruments are traded in one of two ways, either on a exchange such as New York Stock Exchange (NYSE) or over the counter. The phrase *over-the-counter* can be used to refer to stocks that trade via a dealer network as opposed to on a centralized exchange. The idea is that over the counter privately negotiated contracts can be designed. *OTC* contract tend to be less regulated. Hence it does of course mean more risk.

Definition 1.1.3 (*LIBOR*)

LIBOR (London Interbank Offered Rate) is the interest rate at which banks borrow large amounts of money from each other. It is a widely used benchmark for short term (overnight to 1 year) interest rate.

Definition 1.1.4 (*Portfolio*)

Portfolio is a combination of stocks, bonds and financial derivatives.

Definition 1.1.5 (*Derivative*)

A *derivative* is a security whose value depends on the value of some underlying security.

Definition 1.1.6 (*Bond*)

A *bond* is a certificate issued by a government or a public company promising to repay borrowed money at a fixed rate of interest at a specified time.

Definition 1.1.7 (*Spread*)

A *spread* is the difference between the bid and the ask price of a security or asset.

Definition 1.1.8 (*Arbitrage*)

Arbitrage is the simultaneous purchase and sale of an asset in order to profit from a difference in the price. Arbitrage exists as a result of market inefficiencies. Arbitrage basically expresses the concept that one cannot make money for nothing. It is sometimes called the "no free lunch principle". Arbitrage is probably simplest to explain in the context of foreign exchange.

Definition 1.1.9 (Mutual fund)

A mutual fund is nothing more than a collection of stocks and bonds. We can think of a mutual fund as a company that brings together a group of people and invest their money in stock, bonds and other securities. Each investor owns shares which represent a portion of the holdings of the fund.

Definition 1.1.10 (Return on an investment)

It is a revenue expressed as fraction of the initial investment.

Definition 1.1.11 (Basel)

Basel II is a **set of banking regulations** put forth by the Basel Committee on Bank Supervision, which regulates finance and banking internationally. Basel II attempts to integrate Basel capital standards with national regulations, by setting the minimum capital requirements of financial institutions with the goal of ensuring institution liquidity.

Basel II is the second of the Basel Committee on Bank Supervision's recommendations, and unlike the first accord, Basel I, where focus was mainly on credit risk, the purpose of Basel II was to create standards and regulations on how much capital financial institutions must have put aside. Banks need to put aside capital to reduce the risks associated with its investing and lending practices.

Definition 1.1.12 (Fair value adjustment)

A fair value adjustment is a type of accounting process that makes it possible to reassess the fair value when there is a considerable difference between that figure and the current book value of an asset. Managing this type of adjustment requires taking some time to engage in what is known as revaluing in order to bring the two figures into closer harmony. There are a number of reasons why a fair value adjustment may be necessary, including significant shifts in the market value of the assets involved, or when the assets are involved in a business acquisition.

Definition 1.1.13 (Net present value.)

Net Present Value (NPV) is the difference between the present value of cash inflows and the present value of cash outflows. NPV is used in capital budgeting to analyze the profitability of a projected investment or project.

$$NPV(r) := \sum_{t=1}^T \frac{C_t}{(1+r)^t} - C_0$$

A positive net present value indicates that the projected earnings generated by a project or investment (in present dollars) exceeds the anticipated costs (also in present dollars). Generally, an investment with a positive NPV will be a profitable one and one with a negative NPV will result in a net loss.

Definition 1.1.14 (Internal rate of return.)

The value of r for which $NPV(i) = 0$ is called the **yield rate** or **internal rate of return**.

Definition 1.1.15 (Stress testing.)

Stress testing is a simulation technique used on asset and liability portfolios to determine their reactions to different financial situations. Stress tests are also used to gauge how certain stressors will affect a company or industry. They are usually computer-generated simulation models that test hypothetical scenarios. The Monte Carlo simulation is one of the most widely used methods of stress testing.

Definition 1.1.16 (Back testing.)

Back testing is a technique for simulating a model or strategy on past data to gauge its accuracy and effectiveness. Back-testing in value at risk is used to compare the predicted losses from the calculated value at risk with the actual losses realized at the end of the specified time horizon. This comparison identifies the periods where the value at risk is underestimated or where the portfolio losses are greater than the original expected value at risk. The value at risk predictions can be recalculated if the back-testing values are not accurate, thereby reducing the risk of unexpected losses.

Definition 1.1.17 (Mark-to-market.)

Mark-to-market it is measuring the value of financial asset at their market value and calculating unrealised gains or losses resulting from changing exchange rates. It is calculated everyday between trade date and settlement date. It is calculated by comparing current and contracted value of foreign exchange (FX).

Definition 1.1.18 (Yield to maturity.)

There are such a variety of fixed-income products, with different coupon structures, amortization, fixed and/or floating rates that it is necessary be able to consistently compare different products. One way to to this is through measure of how much each contract earns. There are several measures of this all coming under the name yield.

Definition 1.1.19 Term structure of interest rates. *A term structure of interest rates is a set of interest rates sorted by time to maturity. The curve shows the relation between the (level of) interest rate (or cost of borrowing) and the time to maturity.*

Definition 1.1.20 (Yield Curve.)

- *Yield refers to periodic return calculated as coupon/bond price.*
- *Yield curve depicts the relationship between yield and maturity.*
- *Slope of the yield curve help us in predicting future interest rate changes.*

Definition 1.1.21 (Forward rate.)**1.1.1 Bonds vs Stocks.**

Corporations have two options when it comes to raising money without taking out a loan. You can issue corporate bonds or sell shares of stock. There is no limit as to how many bonds or stock shares you can sell or how often you can sell them. However, issuing bonds or stock shares affect your corporation in substantially different ways. With a bond issue, you have the benefit of reducing your tax liability without sacrificing control of your corporation.

1. Ownership protection.

Issuing bonds instead of selling stock does not change your ownership percentage in your corporation. When you issue bonds, you are asking investors to loan your company money. Bondholders are your corporation's creditors but not the owners. Each stock share, by comparison, represents a potential ownership claim in your corporation.

2. Interest deduction.

Bond debt works like every other type of corporate debt. The principal amount is a loan and does not increase your taxable income. You can deduct the interest payments make to bondholders on your corporate income tax return. This in turn reduces your taxable income and your tax liability. When you sell stock, though, the proceeds are taxed as income. Your corporation is further taxed on the dividends you pay to your stockholders. This increases your taxable income and your tax liability.

3. Fixed maturity date.

Bonds have a finite life span. This lets you know in advance how much the corporation will spend paying interest and repaying the original loan. When the bonds reach their maturity date, your obligation to make interest payments ends. The bond issue is retired and removed from the corporate books. When you issue stock, the shareholders can keep their shares forever. The shares can be given away or left to a beneficiary under a will or trust. If you want to buy the shares back, you must offer your shareholders a premium over the original share price.

4. Callable future.

When you issue callable bonds, you have the right to call the bonds in before their maturity date. The interest stops accumulating as of the callable date. The bond is paid off based on its face value. This is useful if interest rates fall and you want to call in old bonds so you can sell new bonds at the lower interest rate. If you sell stock and shareholders have the right to receive dividends, your corporation is committed to paying dividends indefinitely. If you miss a dividend payment or stop making payments, the shareholders can file a lawsuit against your corporation.

1.1.2 Derivatives.

Below are some of the motives for someone to use derivatives:

- to reduce (or hedge) exposure to risk. For example, a wheat farmer and a wheat miller could enter into a futures contract to exchange cash for wheat in the future. Both parties have reduced a future risk: for the wheat farmer, the uncertainty of the price, and for the wheat miller, the availability of wheat.
- to speculate expected changes in future prices with the hope of making profit. In this case, speculation increases the exposure to risk. The potential gain or loss can be leveraged relative to the initial investment.
- to reduce transaction costs, such as commissions and other trading costs.
- to maximize return on investments through asset management activities, tax loopholes, and regulatory restrictions. For example, a company can use derivatives to produce temporary losses to lower its taxes. We refer to this motive as **regulatory arbitrage**.

Chapter 2

Calculus

2.1 Taylor Series

Taylor's series is an expansion of a function into an infinite series of a variable x or into a finite series plus a remainder term.

Let $f: I \rightarrow \mathbb{R}$ smooth ($f \in C^\infty$), $I \subset \mathbb{R}$ open, $0 \in I$.

Definition 2.1.1 Taylor's series of f in fixed point a define as

$$T_a^n f(x) := \sum_{k=0}^n \frac{1}{k!} f^{(k)}(a)(x-a)^k$$

Fix $x_0 \in I$. We may also define reminder by the formula:

$$R_n(x) = f(x) - T_{x_0}^n f(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k$$

Theorem 2.1.1 Let $f: I \rightarrow \mathbb{R}$, $f \in C^{n+1}([a, b])$, and $a < x_0 < x < b$. Then exist $\xi \in \mathbb{R}$ such that $x_0 < \xi < x$ and

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{n+1} (x-x_0)^{n+1}$$

Theorem 2.1.2 Let $I \subset \mathbb{R}$ open interval. Let $f: I \rightarrow \mathbb{R}$ $f \in C^{n+1}$ and we have

$$|f^{(n+1)}(x)| \leq c, \quad x \in I.$$

Then $\forall x_0, x \in I$ we have

$$\left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-x_0)^k \right| \leq \frac{c}{(n+1)!} |x-x_0|^{n+1}$$

Several important Maclaurin series expansions follows:

1. $f(x) = \frac{1}{1-x}$

$$T_0^n f(x) = \sum_{k=0}^n x^k$$

2. $f(x) = e^x$

$$T_0^n f(x) = \sum_{k=0}^n \frac{x^k}{k!}.$$

3. $f(x) = \log(1+x)$

$$T_0^n f(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{n+1} x^{n+1}$$

Applications

1. Some functions have no anti-derivative which can be expressed in terms of familiar functions. This makes evaluating definite integrals of these functions difficult because the fundamental theorem of calculus cannot be used.

$$\int_0^1 \sin(x^2) dx$$

2. It is useful in many proofs for instance Ito formula, Jensen inequality
3. approximation for example to calculate \sqrt{e}

2.2 Numerical methods for one dimensional nonlinear problems.

Many problems in finance require solving equations of the form

$$f(x) = 0,$$

where $f(x)$ is not a linear function. Anyone can easily list examples of nonlinear problems arising in the mathematical finance:

- computing the yield of a bond
- computing the implied volatility

2.2.1 Bisection method.

Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous functions such that $f(a)$ and $f(b)$ have different signs. From the intermediate value theorem, it follows that exists one or more points x that function $f(x)$ is equal 0. (There might be more points where $f(x) = 0$, but the bisection method will find only one such zero.)

The idea of the bisection method is to divide the interval $[a, b]$ into two equal intervals $[a, c]$ and $[c, b]$ with $c = \frac{a+b}{2}$. Since $f(a)$ and $f(b)$ have different signs either $f(a)$ and $f(c)$ have different signs, or $f(c)$ and $f(b)$ have different signs. Unless $f(c) = 0$, which means that a solution for the problem $f(x) = 0$ has already been found.

2.2.2 Newton's method.

One way to derive the recursion formula for Newton's method for one dimensional problems is as follows: Let $x_k \dots$

Definition 2.2.1 Let $f : [a, b] \rightarrow \mathbb{R}$. The tangent line to the graph of $f(x)$ passing through the point $(x_0, f(x_0))$ where $x_0 \in [a, b]$ is

$$y - f(x_0) = f'(x_0)(x - x_0)$$

Put $(x_1, 0)$ we get formula

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Keep this logic we obtain:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad \forall k \geq 0. \quad (2.1)$$

Remark 2.2.1 *A more insightful way of deriving (2.1), which can easily be extended to N -dimensional problems is to use the Taylor expansion of the function $f(x)$ around the point x_k .*

Unlike the bisection method, Newton's method does not necessarily converge for any function $f(x)$ and any initial guess x_0 . To obtain convergence in Newton's method, and, in particular, to obtain fast convergence, a good choice of the initial approximation x_0 is important.

Chapter 3

Probability.

There are 52 basic cards in a deck. Playing cards we have following: Jack, Queen, King, Ace. Usual signs are

1. spades
2. hearts
3. diamonds
4. clubs

3.1 Conditional expectation.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be probability space and $\mathcal{G} \subset \mathcal{F}$ is sub-field.

Definition 3.1.1 *Expectation of random variable X such that $\int_{\Omega} |X| d\mathbb{P} < \infty$ is defined as integral of random variable X with respect to \mathbb{P} and is denoted by $\mathbb{E}^{\mathbb{P}}[X]$. Thus*

$$\mathbb{E}_{\mathbb{P}}[X] = \int_{\Omega} X d\mathbb{P}.$$

Theorem 3.1.1 *Let \mathcal{G} be σ -algebra contained \mathcal{F} . Then if X jest \mathcal{G} -measurable random variable such that $\int_{\Omega} |X| d\mathbf{P} < \infty$ and*

$$\int_A X d\mathbf{P} = 0 \quad \forall A \in \mathcal{G},$$

to $X = 0$ p.w. $[\mathbf{P}]$.

Proof.

Wystarczy pokazać, że $\mathbf{P}(\{X = 0\}) = 1$. Zauważmy, że X przyjmuje dowolnie małe wartości z prawdopodobieństwem 1. Istotnie ustalamy dowolne $\varepsilon > 0$ oraz definiujemy zbiór $A = \{X \geq \varepsilon\}$. Wówczas

$$0 \leq \varepsilon \mathbf{P}(A) = \varepsilon \int_A d\mathbf{P} = \int_A \varepsilon d\mathbf{P} \leq \int_A X d\mathbf{P} = 0.$$

Stąd $\mathbf{P}(\{X \geq \varepsilon\}) = 0$. Stosując powyższe do zmiennej losowej $-X$ otrzymujemy $\mathbf{P}(\{-X \leq \varepsilon\}) = 0$. Stąd

$$\mathbf{P}(\{|X| < \varepsilon\}) = 1 - \mathbf{P}(\{X \geq \varepsilon\}) - \mathbf{P}(\{-X \leq \varepsilon\}) = 1.$$

Zatem zmierzając z ε do zera otrzymamy tezę. Bardziej formalnie, niech $\varepsilon = \frac{1}{n}$, $n \in \mathbb{N}$. Rozważmy ciąg zbiorów $B_n = \{|X| < \frac{1}{n}\}$. Jest to ciąg zbiorów zstępujących takich, że $\mathbf{P}(B_n) = 1$. Zatem dzięki ciągłości miary mamy

$$\mathbf{P}(\{X = 0\}) = \mathbf{P}(\{|X| = 0\}) = \mathbf{P}\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow \infty} \mathbf{P}(B_n) = 1.$$

□

Theorem 3.1.2 Niech X i Y będą \mathcal{G} -mierzalnymi zmiennymi losowymi takimi że $\int_{\Omega} |X| d\mathbf{P} < \infty$ oraz $\int_{\Omega} |Y| d\mathbf{P} < \infty$. Jeśli

$$\int_A X d\mathbf{P} = \int_A Y d\mathbf{P} \quad \forall A \in \mathcal{G}, \quad (3.1)$$

to wówczas $X = Y$, p.w. $[\mathbf{P}]$.

Proof.

Ponieważ $\int_A (X - Y) d\mathbf{P} = 0 \quad \forall A \in \mathcal{G}$ z powyższego twierdzenia wynika, że $X - Y = 0$ p.w. $[\mathbf{P}]$ a stąd $X = Y$ p.w. $[\mathbf{P}]$.

□

Definition 3.1.2 Conditional expectation of X with respect to $\mathcal{G} \subset \mathcal{F}$ is random variable Y such that:

1. Y is \mathcal{G} -measurable,
2. for any $A \in \mathcal{G}$

$$\int_A X d\mathbf{P} = \int_A Y d\mathbf{P}.$$

Remark 3.1.1 Random variable Y is usually denoted by $\mathbb{E}^{\mathbf{P}}[X|\mathcal{G}]$.

3.1.1 Properties of Conditional expectation.

Theorem 3.1.3 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be measurable space. Further let \mathcal{G} be another σ -algebra contained in \mathcal{F} . Consider random integrable random variable X independent from \mathcal{G} it means $\sigma(X)$ and \mathcal{G} are independent with respect to \mathbf{P} . Then:

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

Proof. Niech $\mathcal{F} = \sigma(X)$. Po pierwsze zauważmy, że $\mathbb{E}[X]$ jest \mathcal{G} -mierzalna. Istotnie $\mathbb{E}[X]$ jest stała a zatem mierzalna względem dowolnej σ -algebry. Ponieważ X i \mathcal{G} są niezależne zatem zmienne losowe X oraz $\mathbb{1}_B$ są niezależne dla dowolnego $B \in \mathcal{G}$. Wówczas

$$\int_B \mathbb{E}[X] d\mathbf{P} = \mathbb{E}[X] \int_B d\mathbf{P} = \mathbb{E}[X] \int_{\Omega} \mathbb{1}_B d\mathbf{P} = \mathbb{E}[X] \mathbb{E}[\mathbb{1}_B].$$

Ponieważ zmienne losowe niezależne są nieskorelowane, zatem

$$\mathbb{E}[X] \mathbb{E}[\mathbb{1}_B] = \mathbb{E}[X \mathbb{1}_B] = \int_B X d\mathbf{P}.$$

Zatem otrzymaliśmy równość

$$\int_B \mathbb{E}[X] d\mathbf{P} = \int_B X d\mathbf{P}, \quad B \in \mathcal{G}.$$

Zatem z definicji warunkowej wartości oczekiwanej i powyższej równości otrzymujemy

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

□

3.2 Log-normal distribution

Let $Y \sim \mathcal{N}(\mu, \sigma)$. Then random variable $X = e^Y$ is said to be log-normal distribution $X \sim \mathcal{LN}(\mu, \sigma)$. The probability density function of log-normal distribution is

$$f_X(x) = \frac{1}{x\sqrt{2\pi}\sigma} e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}.$$

Then

$$\mathbb{E}[X] = e^{\mu + \frac{1}{2}\sigma^2},$$

$$\text{Var}[X] = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1).$$

3.3 Central limit theorem.

There is actually several forms of the central limit theorem, which formally is just a weak convergence results in probability. Central Limit theorem involves the average of sum of independent identically distributed random variables with finite second moment. The theorem states that as the number of random variables increases, the average approaches a normal distribution with known parameters.

Theorem 3.3.1 (*Linderber-Levy*)

Let X_1, X_2, \dots, X_n be a sequence of n independent and identically distributed random variables with finite mean $\mathbb{E}[X_i] = \mu$ and variance $D^2[X_i] = \sigma^2$, for $i = 1, \dots, n$. Let define $S_n := X_1 + X_2 + \dots + X_n$. Then for all $t \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} P\left(\left\{\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq t\right\}\right) = \Phi(t), \quad \forall t \in \mathbb{R}. \quad (3.2)$$

Remark 3.3.1 *Linderberg proved a central limit theorem for independent random variables which are not identically distributed.*

If additionally we assume that sequence of random variables X_1, X_2, \dots, X_n has Bernoulli distribution then we have following version of central limit theorem

Theorem 3.3.2 *Moiivre'a-Laplace'a*

Let X_1, X_2, \dots, X_n be a sequence of n independent and identically distributed random variables comes from Bernoulli distribution with common parameter $p \in (1, 0)$. Then

$$\lim_{n \rightarrow \infty} P\left(\left\{\frac{S_n - np}{\sqrt{np(1-p)}} \leq t\right\}\right) = \Phi(t), \quad \forall t \in \mathbb{R}. \quad (3.3)$$

Chapter 4

Stochastic process.

Definition 4.0.1 Let us consider probability space $(\Omega, \mathcal{F}, \mathbb{P})$. An increasing family of sigma fields that is $\mathcal{F}_u \subset \mathcal{F}_t$ for any $0 \leq u \leq t \leq T$ where T is a finite strictly positive numbers is called filtration and we shall write $(\mathcal{F}_t)_{t \in [0, T]}$.

Generally in finance we work in probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with \mathbb{P} being market probability measure equipped with a filtration $(\mathcal{F}_t)_{t \in [0, T]}$ which captures the information available up to and including time $t \geq 0$.

Definition 4.0.2 Set $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbf{T}}, \mathbb{P})$, where $(\mathcal{F}_t)_{t \in [0, T]}$ is filtration defined above is called filtered probability space.

Definition 4.0.3 Given a positive integer d , a d -dimensional stochastic process on the given probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbf{T}}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t \in [0, T]}$ is a collection $(X_t)_{t \in [0, T]}$, where each X_t is a d -dimensional random vector; i.e., a function $X_t: \Omega \rightarrow \mathbb{R}^n$ such that $X_t^{-1}(B) = \{\omega \in \Omega: X_t(\omega) \in B\} \in \mathcal{F}$ for each Borel set B of \mathbb{R}^n

4.1 Ito Lemma

Definition 4.1.1 An adapted continuous process X is called an **Ito process** if it admits a representation

$$X_t = X_0 + \int_0^t \mu(u, X_u) du + \int_0^t \sigma(u, X_u) dW_u, \quad t \in [0, T], \quad (4.1)$$

with $\mu(t, X_t)$ being the drift term and $\sigma(t, X_t)$ the diffusion term for some adapted processes μ, σ defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbf{T}}, \mathbb{P})$ and integrable in a suitable sense.

Equation (4.1) may be written in differential form as

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

Theorem 4.1.1 Suppose that $g: [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ is function of class $C^{1,2}([0, T] \times \mathbb{R}, \mathbb{R})$. Then for any Ito process X , the process $Y_t = g(t, X_t), t \in [0, T]$, is an Ito process. Moreover, its canonical decomposition is given by the Ito formula

$$dY_t = \left(\frac{\partial g}{\partial t}(t, X_t) + \mu(t, X_t) \frac{\partial g}{\partial x}(t, X_t) + \frac{1}{2} \sigma^2(t, X_t) \frac{\partial^2 g}{\partial x^2}(t, X_t) \right) dt + \sigma(t, X_t) \frac{\partial g}{\partial x}(t, X_t) dW_t.$$

Remark 4.1.1 Ito lemma shows how to switch between processes. With a well chosen transformation, a simpler SDE can be derived and in some cases an analytical solution of SDE can be found. Another applications are calculating stochastic integral and checking if stochastic process is a martingale.

4.2 Ito integral

Theorem 4.2.1 For any continuous function $f: [0, T] \rightarrow \mathbb{R}$ the Ito integral $\int_0^T f(t) dW_t$ has Gaussian distribution with mean zero and variance $\int_0^T f^2(t) dt$.

Let

$$\mathcal{M} = \left\{ f: [0, T] \times \Omega \rightarrow \mathbb{R} : f \text{ is adapted, } \mathbb{E} \left[\int_0^T f^2(t) dt \right] < \infty \right\}$$

Theorem 4.2.2 For $f \in \mathcal{M}$

$$\mathbb{E} \left[\int_0^T f_t dW_t \right] = 0$$

Theorem 4.2.3 For $f \in \mathcal{M}$

$$\mathbb{E} \left[\int_0^T f_t dW_t \right]^2 = \mathbb{E} \left[\int_0^T f_t^2 dt \right]$$

Several important stochastic integral

1. $\int_0^t 1 dW_s$
2. $\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}$
3. $\int_0^t \sqrt{W_s} dW_s =$

4.3 Martingale.

Definition 4.3.1 Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbf{T}}, \mathbb{P})$ be filtered probability space. A stochastic process $(X_t)_{t \in [0, T]}$ is a martingale if

1. $\mathbb{E}[X_t | \mathcal{F}_s] = X_s$ for all $0 \leq s \leq t$,
2. $\mathbb{E}[|X_t|] < \infty$,
3. X_t is \mathcal{F}_t -adapted.

The key property of martingale is that the mean of the future is equal to the present regardless of the past. This is only a statement about the mean not about the entire distribution. The martingale property states that the future expectation of stochastic process is equal to the current value, given all information about the prior events. The martingale property ensures that in a "fair game", knowledge of the past will be of no use in the predicting future winnings.

Below we present several important martingales :

1. $X_t = W_t^2 - t$
2. $Y_t = W_t^3 - 3tW_t$
3. $Z_t = \exp(\sigma W_t - \frac{1}{2}\sigma^2 t)$

Theorem 4.3.1 Let $(W_t)_{t \in [0, T]}$ be a Wiener process. Then

$$\mathbb{E}[(W_t - W_s)^2 | \mathcal{F}_s] = \mathbb{E}[(W_t - W_s)^2] = t - s, \quad t > s,$$

$$\mathbb{E}[W_s^2 | \mathcal{F}_s] = W_s^2, \quad s > 0.$$

Question 4.3.1 Let $(W_t)_{t \in [0, T]}$ be a Wiener process. Define

$$X_t = \int_0^t W_u du$$

What is the distribution of X_t ? Is the process X_t martingale?

beginquestion Let $(W_t)_{t \in [0, T]}$ be a Wiener process. Calculate

$$\mathbb{E} \left[\int_s^t W_u^2 | \mathcal{F}_s \right]$$

Question 4.3.2 Check if $X_t = W_t^3$ is a martingale

$$dX_t = 3W_t dW_t + W_t^3 (dW_t)^2 =$$

Question 4.3.3 What is the difference and relation between a Markov process and a martingale process?

4.4 Radon-Nikodym theorem.

Definition 4.4.1 Consider two measures \mathbb{P} and \mathbb{Q} defined on (Ω, \mathcal{F}) . Measure \mathbb{Q} is absolutely continuous with respect \mathbb{P} and we denote this fact by $\mathbb{Q} \ll \mathbb{P}$ if

$$\mu(A) = 0 \implies \nu(A) = 0, \quad \forall A \in \mathcal{F}$$

Definition 4.4.2 The measures \mathbb{P} and \mathbb{Q} are said to be equivalent if \mathbb{P} and \mathbb{Q} are absolutely continuous with respect to each other.

Theorem 4.4.1 (Radona-Nikodyma)

Let (Ω, \mathcal{F}) be a measure space. \mathbb{P} and \mathbb{Q} are σ -finite, such that $\mathbb{Q} \ll \mathbb{P}$. Then exists \mathcal{F} -measurable function (random variable) $f : \Omega \rightarrow [0, \infty]$ such that

$$\mathbb{Q}(A) = \int_{\Omega} f d\mathbb{P}, \quad \forall A \in \mathcal{F}.$$

Remark 4.4.1 The random variable f is called the Radon-Nikodym derivative and is often written $\frac{\mathbb{Q}}{\mathbb{P}}$. The..

4.5 Girsanov theorem.

The biggest use of measure changes in derivatives pricing in Girsanov's theorem, which is used to change the drift of Brownian Motion. Assume that we have probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variable $X : \Omega \rightarrow \mathbb{R}$ such that $\mathbb{E}[X] = 1$. Then we define new measure $\mathbb{Q} : \mathcal{F} \rightarrow [0, 1]$ by

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_A X] = \int_A X d\mathbb{P}, \quad \forall A \in \mathcal{F}.$$

Theorem 4.5.1 (Girsanov theorem.)

Let γ be an adapted process satisfying certain conditions and define

$$X_t = \exp \left(-\frac{1}{2} \int_0^t |\gamma_u|^2 du + \int_0^t \gamma_u dW_u \right).$$

If $\mathbb{E}^\mathbb{P}[X_T] = 1$, then X_t is martingale and defines a new measure \mathbb{Q} where

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = X_T.$$

Under \mathbb{Q} , the process

$$W_t^\mathbb{Q} = W_t - \int_0^t \gamma_u du$$

is a standard Brownian motion.

Hence Girsanov's theorem has defined a new measure under which the Brownian motion with drift $W_t^\mathbb{Q}$ is a driftless Brownian motion

Remark 4.5.1 Every time we change a measure we must remember about two golden rules:

- The new measure must assign probabilities to the same subsets
- subset that has measure zero under the old measure must have the zero measure under new measure

4.6 Brownian motion.

Definition 4.6.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Stochastic process $B_t = (B_t)_{t \in \mathbf{T}}$ is defined to be a Brownian motion if:

1. $B_0 = 0$,
2. B_t has continuous trajectory,
3. for each $t > 0$ and $s > 0$, $B_{t+s} - B_t \sim \mathcal{N}(0, s)$,
4. for each $t > 0$ and $s > 0$, $B_{t+s} - B_t$ and B_t are independent.

4.6.1 Basic Properties of Brownian Motion.

Theorem 4.6.1 Set $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbf{T}}, \mathbb{P})$ is filtered probability space. If $B_t = (B_t)_{t \in \mathbf{T}}$ is a Brownian motion then we have following:

1. $B_{t+s} - B_t$ is independent from \mathcal{F}_t for $s > 0$
2. $W_t := B_{t+s} - B_t$ for $u > 0$ is also a standard Brownian motion
3. $W_t :=$

Fix time horizon $T > 0$ and $M > 0$.

Theorem 4.6.2 Let $0 = t_0 < t_1 < t_2 < \dots < t_M = T$ be subdivision of the interval $[0, T]$ and $Z_1, Z_2, \dots, Z_M \sim \mathcal{N}(0, 1)$ to be independent. Define

$$Y_0 = 0, \quad Y_i = Y_{i-1} + \sqrt{t_i - t_{i-1}} Z_i, \quad (i = 1, 2, \dots, M)$$

Then $(Y_0, Y_1, Y_2, \dots, Y_M)$ has the same distribution as $(B_{t_0}, B_{t_1}, B_{t_2}, \dots, B_{t_M})$ where B is a standard Brownian motion.

4.7 Poisson process.

Poisson processes are continuous time stochastic processes defined on the interval $[0, \infty)$ with the space $\mathcal{S} = \{0, 1, 2, \dots\}$. It is a counting process whose state at time t represents the total of certain events that have occurred up to time t .

Definition 4.7.1 *A counting process $N = (N_t)_{t \in [0, \infty)}$ is said to be a Poisson process with intensity λ if:*

- $N_0 = 0$
- *the process N has independent increments. That is for any time points $t_0 < t_1 < \dots < t_n$ the process increments*

$$N_{t_1} - N_{t_0}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$$

are independent

- *for all $s \geq 0$ and $t > 0$ the random variable $N_{t+s} - N_s$ is Poisson distributed with parameter λt that is*

$$\mathbb{P}(\{N_{t+s} - N_s = n\}) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, \dots$$

4.7.1 Characteristic of Poisson process.

Let $N = (N_t)_{t \in [0, \infty)}$ is a Poisson process with rate λ . Then

- $\mathbb{E}[N_t] =$
- $\text{var}(N_t) =$
- $\text{Cov}(N_t, N_{t+h})$

4.8 Compound Poisson process.

4.9 Options

Definition 4.9.1 *(call option)*

*A **European call option** written on a common stock is a financial security that gives its holder the right (but not the obligation) to buy the underlying stock on a prespecified date and for a prespecified price.*

The act of making this transaction is referred to as exercising the option. If an option is not exercised, we say it is abandoned. Another class of options comprises so-called American options. These may be exercised at any time on or before the prespecified date. The prespecified fixed price, say K , is termed the strike or exercise price; the terminal date, denoted by T in what follows, is called the expiry date or maturity. It should be emphasized that an option gives the holder the right to do something; however, the holder is not obliged to exercise this right. In order to purchase an option contract, an investor needs to pay an option's price (or premium) to a second party at the initial date when the contract is entered into.

Chapter 5

Black-Scholes model

5.1 Induction

Let's take a geometric (or exponential) Brownian motion as a stochastic process which models the stock price. The evolution of the stock price process S is assumed to be described by the following linear stochastic differential (SDE)

$$dS_t = \mu S_t dt + \sigma S_t dW_t \quad \mu \in \mathbb{R}, \sigma \in (0, \infty), S_0 \in (0, \infty). \quad (5.1)$$

It is shorthand notation for the following Ito integral equation

$$S_t = S_0 + \int_0^t \mu S_u du + \int_0^t \sigma S_u dW_u \quad t \in [0, T].$$

SDE with Lipschitz continuous coefficients has a unique solutions.

It is elementary to check using Ito formula that the process S which equals

$$S_t = S_0 \exp(\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t), \quad \forall t \in [0, T], \quad (5.2)$$

is indeed a solutions of (5.1), starting from S_0 at time 0. It's apparent from (5.2) that the stock returns are log-normal, meaning that the random variable $\log \frac{S_t}{S_u}$ has under \mathbb{P} a Gaussian probability distribution for any choice of dates $u \leq t \leq T^*$.

In the Black-Scholes setting, the martingale measure for the discounted stock price process is unique, and is explicitly known, as the following result shows.

Theorem 5.1.1 *The unique martingale measure \mathbf{Q} for discounted stock price process S^* is given by the Radon-Nikodym derivative*

$$\frac{d\mathbf{Q}}{d\mathbb{P}} = \exp\left(\frac{r - \mu}{\sigma} W_T - \frac{1}{2} \frac{(r - \mu)^2}{\sigma^2} T\right)$$

The dynamics of the discounted stock price S^ under the martingale measure \mathbf{Q} are*

$$dS_t^* = \sigma S_t^* dW_t^* \quad (5.3)$$

and the process W^ which equals*

$$W^* = W_t - \frac{r - \mu}{\sigma} t, \quad \forall t \in [0, T],$$

follows a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbf{Q})$.

The discounted stock price S^* follows under \mathbf{Q} follows a strictly positive martingale, since

$$S_t^* = S_0 \exp(\sigma W_t^* - \frac{1}{2}\sigma^2 t), \quad \forall t \in [0, T] \quad (5.4)$$

Notice also that the dynamics of the stock price S under \mathbf{Q} are

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t = r S_t dt - r S_t dt + \mu S_t dt + \sigma S_t dW_t = \\ &= r S_t dt + \sigma S_t dW_t^* \end{aligned} \quad (5.5)$$

We are in a position to introduce the class of admissible trading strategies. An unconstrained Black-Scholes market model would involve arbitrage opportunities, so that reliable valuation of derivative instruments would not be possible.

Definition 5.1.1 *A trading strategy $\varphi \in \Phi$ is called \mathbf{P}^* -admissible if the discounted wealth process*

$$V_t^*(\varphi) = B_t^{-1} V_t(\varphi), \quad t \in [0, T],$$

follows a martingale under \mathbf{P}^ .*

It is not hard to check that by restricting our attention to the class of \mathbf{P}^* admissible strategies, we have guaranteed the absence of arbitrage opportunities in the Black-Scholes market. Consequently, given a contingent claim X which settles at time $T \leq T^*$ and is attainable (i.e., can be replicated by means of a \mathbf{P}^* -admissible strategy) we can uniquely define its arbitrage price, $\pi_t(X)$, as the wealth φ_t at time t of any \mathbf{P}^* -admissible trading strategy φ which replicates X – that is, satisfies $V_T(\varphi) = X$. If no replicating \mathbf{P}^* -admissible strategy exists, the arbitrage price of such a claim is not defined.

Conforming with the definition of an arbitrage price, to value a derivative security we will usually search first for its replicating strategy. Another approach to the valuation problem is also possible, as the following simple result shows.

Theorem 5.1.2 *Let X be a attainable European contingent claim which settles at time T . The arbitrage price $\pi_t(X)$ at time $t \in [0, T]$ is given by the risk-neutral valuation formula*

$$\pi_t(X) = B_t \mathbb{E}^{\mathbf{P}^*} [B_T^{-1} X | \mathcal{F}_t], \quad t \in [0, T].$$

5.2 Assumptions

As we know in the Black-Scholes world the evolution of the stock price S_t is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad \mu \in \mathbb{R}, \sigma \in (0, \infty).$$

We also assume

1. stock does not pay dividends
2. there are no transaction cost
3. the continuously compounding interest rate $r > 0$
4. volatility is constant
5. there is no jumps

The latter assumption implies the evolution of the risk-free asset $B = (B_t)_{t \in [0, T]}$ is given by

$$dB_t = rB_t dt$$

We are interested in pricing option which is a function of the stock price at time $T > 0$, S_T . One possible example is a call option with a strike $K > 0$, that is derivative which at time T pays

$$X = \max\{S_T - K, 0\}.$$

While the form of the payoff is not particularly important, that it is a function of the stock price at time T , and only at time T is important. Under this condition we can show that the call option price is a function of current time t and current price S_t only. Thus we denote by $C(t, S_t)$.

To price derivative in the Black-Scholes world we must do so under a measure does not allow arbitrage. Such a measure is called a risk-neutral measure. It's not hard to prove that under this measure, the drift term of the stock price changes so that

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

We are now ready to proceed with our derivation. In the risk-neutral world, $\frac{C(t, S_t)}{B_t}$ is a martingale and hence if we calculate its differential we know it must have zero drift. Applying Ito's lemma to $C(t, S_t)$ gives ...

5.3 Derivation the Black-Scholes equation for a stock.

5.4 Facts about the Black-Scholes equation

Let $V = V(S_t, t)$ denote the option value as a function of asset price

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

- It is a partial differential equation because it has more than one independent variable, here S and t .
- The equation is linear and homogeneous (there is no right-hand side) so that you can value a portfolio of derivatives by summing the values of the individual contracts

5.5 Valuation formula.

- S_t price at t (at valuation date)
- K strike
- r spot rate
- σ volatility
- T expiry
- t valuation date
- $\tau = T - t$ time to maturity

Define

$$d_1 := \frac{\log \frac{S_t}{K} + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

$$d_2 := \frac{\log \frac{S_t}{K} + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

Price of call option:

$$c^{BS}(S_t, K, \sigma, \tau, r) = S_t \mathcal{N}(d_1) - K e^{-r\tau} \mathcal{N}(d_2)$$

Price of put option:

$$p^{BS}(S_t, K, \sigma, \tau, r) = -S_t \mathcal{N}(-d_1) + K e^{-r\tau} \mathcal{N}(-d_2)$$

Approximation formula for at-the-money options. It states that option's price is approximately

$$0.4S_0\sqrt{T}$$

We say that at given t before or at expiry, a call option is **in-the-money** if $S_t > K$ and **out-the-money** $S_t < K$, respectively. Similarly, a put option is said to be **out-the-money** and **in-the-money** respectively.

We know that $\mathcal{N}(d_2)$ is the probability of the call ending up in the money. But what does $\mathcal{N}(d_1)$ represent? The answer is that $\mathcal{N}(d_1)$ and $\mathcal{N}(d_2)$ are both probabilities of the call ending up in the money but under different measures.

Undoubtedly, the most striking feature of the Black-Scholes result is the fact that the drift μ does not enter the valuation formula. This is not surprising, however, as expression which describes the dynamics of the stock price under the martingale measure P^* , does not involve the stock the drift rate μ .

5.5.1 Black Scholes 76 (BS 76)

Black in 1976 made some slight modifications to his and his colleagues preceding work. Instead of using the spot price of an underlying asset S_0 , the model now discounts a forward price F_t . This model is more suitable to price interest rate derivatives (bond options, interest caps/floors and swaptions.) Under BS 76 model, the theoretical value of call and put

$$c_{76}(t) = e^{-r(T-t)}[F_t \mathcal{N}(d_1) - K \mathcal{N}(d_2)]$$

$$p_{76}(t) = e^{-r(T-t)}[-F_t \mathcal{N}(-d_1) + K \mathcal{N}(-d_2)]$$

where

$$d_1 := \frac{\log \frac{F_t}{K} + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{(T-t)}},$$

$$d_2 := \frac{\log \frac{F_t}{K} - \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{(T-t)}}.$$

Problem 5.5.1 What is the benefit of pricing an option on the forward instead of the spot??

The forward price F_t at time t is martingale under the risk neutral probability \mathbf{Q} . We get simpler expressions if, say, we want to price an at-the-money European call on a forward, where the strike $M = F_0$. The parameter d_1 in Black-Scholes becomes $d_1 = \frac{1}{2}\sigma\sqrt{T}$. In addition, note that the payoff for a European-style call on the forward $\max(F_T - K, 0)$ is the same as the payoff for a call on the spot $\max(F_T - K, 0)$, since the forward price converges to the spot price at time T .

5.5.2 The Black-Scholes PDE

Theorem 5.5.1 *Let $g: \mathbb{R} \rightarrow \mathbb{R}$ be a Borel-measurable function, such that the random variable $X = g(S_T)$ is integrable under \mathbf{P}^* . Then the arbitrage price of the claim X which settles at time T is given by the equality $\pi_t(X) = v(S_t, t)$, where the function $v: \mathbb{R}_+ \times [0, T] \rightarrow \mathbb{R}$ solves Black-Scholes partial differential equation*

$$\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} + rs \frac{\partial v}{\partial s} - rv = 0, \quad \forall (s, t) \in (0, \infty) \times (0, T),$$

subject to the terminal condition $v(s, T) = g(s)$.

Using s we may note

$$\theta + \frac{1}{2}\sigma^2 s^2 \gamma + rs\delta = rv$$

5.5.3 Limitations of Black-Scholes model.

Where Black-Scholes really shines is a common language between options traders. It's the oldest, simplest and the most intuitive option pricing around. Every option trader or quantitative analyst understands it and it is easy to calculate. In practise the modified models are simply extensions of the original Black-Scholes model. The model is intuitive, simple and works surprisingly well under normal market conditions

5.6 Put-call parity.

Put-call parity states that European call option minus European put option with the same underlying asset, strike price and time is equal to forward contract.

Put-Call parity is very useful concept in options pricing and it is model independent, so it must hold no matter which model one is working with. Put-call parity could be used for example, to price a put option instead of the required call option, due to the fact a put option is sometimes easier to price as its payoff is bounded. It can also be used to prove that the drift of the underlying asset does not affect the price of an option and is very helpful to find boundaries of options.

Theorem 5.6.1 *For continuous compounded interest rate r the following put-call parity relationship is valid*

$$c_t - p_t = S_t - Ke^{-rt}.$$

Proof. The proof begins with the following true expression

$$(S_T - K)^+ - (K - S_T)^+ = S_T - K,$$

where $t < T$ is the expiration time of options. Re-arranging we get:

$$(S_T - K)^+ - K = (K - S_T)^+ + S_T.$$

Now we multiply each side by the discount factor $e^{-r(T-t)}$ with $t < T$. Then

$$e^{-r(T-t)}(S_T - K)^+ + e^{-r(T-t)}K = e^{-r(T-t)}(K - S_T)^+ - e^{-r(T-t)}S_T$$

Take conditional expectations under the risk neutral measure with respect to \mathcal{F}_t at some time $t < T$

$$\begin{aligned} \mathbb{E}^{\mathbf{Q}}[e^{-r(T-t)}(S_T - K)^+ | \mathcal{F}_t] + \mathbb{E}^{\mathbf{Q}}[e^{-r(T-t)}K | \mathcal{F}_t] &= \\ = \mathbb{E}^{\mathbf{Q}}[e^{-r(T-t)}(K - S_T)^+ | \mathcal{F}_t] + \mathbb{E}^{\mathbf{Q}}[e^{-r(T-t)}S_T | \mathcal{F}_t] \end{aligned}$$

Now recall that risk neutral pricing theory tells us that the discounted value of risky asset is a martingale. this immediately gives us that

$$c_t + e^{-r(T-t)}K = p_t + S_t, \quad \forall t \in [0, T].$$

Thus

$$c_t - p_t = S_t - e^{-r(T-t)}K, \quad \forall t \in [0, T].$$

□

There is another proof of put-call parity for European option. First we need to clarify one thing.

Remark 5.6.1 *The value of portfolio is equal to the cash amount generated if the portfolio is liquidated, and not to the cash amount needed to set up the portfolio.*

Theorem 5.6.2 *For continuous compounded interest rate r the following put-call parity states that*

$$p_t + S_t - c_t = Ke^{-r(T-t)}.$$

Proof. Consider a portfolio made o the following assets

- long 1 put option
- long 1 share
- short 1 call option

The value of portfolio (defined in remark above) at time $t \in [0, T]$ is

$$V_t = p_t + S_t - c_t$$

It is easy to see that

$$V_T = p_T + S_T - c_T = K$$

regardless of the value S_T . We will see it by analysing what happens if $S_T < K$ or if $S_T \geq K$. We know that $V_t = V_T e^{-r(T-t)} = Ke^{-r(T-t)}$

□

5.7 No-Arbitrage bounds on option prices.

5.7.1 Bonds on call options

Let denote $P_{Eur}(K, T)$, $(P_{Ame}(K, T))$ and $C_{Eur}(K, T)$, $(C_{Ame}(K, T))$ be the value of European (American) put and call options. Let's assume no stock pay dividends. Since an American option can duplicate a European option by exercising the American option at the maturity date, it follows that an American option is always worth at least as much as a European option on the same asset with the same strike price and maturity date. In symbols

$$C_{Ame}(K, T) \geq C_{Eur}(K, T)$$

$$P_{Ame}(K, T) \geq P_{Eur}(K, T)$$

Theorem 5.7.1 *For a non-dividend-paying stock one has*

$$C_{Ame}(K, T) = C_{Eur}(K, T).$$

Proof.

Suppose that $C_{Ame}(K, T) > C_{Eur}(K, T)$. We will show that this creates an arbitrage opportunity. Consider the position of selling the American call and buying the European call. The net cash flow $C_{Ame}(K, T) - C_{Eur}(K, T)$ would be invested at the risk-free rate r . If the owner of the American call chooses to exercise the option at some time $t \leq T$ sell short a share of the security for amount K and add the proceeds to the amount invested at the risk-free rate. At time T close out the short position in the security by exercising the European option. The amount due is

$$(C_{Ame}(K, T) - C_{Eur}(K, T))e^{rT} + K(e^{r(T-t)} - 1) > 0$$

If the American option is not exercised, the European option can be allowed to expire and the amount due is

$$(C_{Ame}(K, T) - C_{Eur}(K, T))e^{rT} > 0$$

In either case, an arbitrage opportunity occurs. □

Since the best one can do with a call stock option is to own the stock so the call price cannot exceed the current stock price.

Theorem 5.7.2

$$S_0 \geq C_{Ame}(K, T) \geq C_{Eur}(K, T)$$

Proof.

Suppose $S_0 < C_{Ame}(K, T)$.

transactions	time = 0	$S_t \leq K$	$S_t > K$
buy a stock	$-S_0$	S_t	S_t
sell a call option	C_{Amer}	0	$K - S_t$
Total	$C_{Amer} - S_0 > 0$	S_t	K

As we can see in the last row every entry is non-negative. Thus an arbitrage occurs. □

Let's have a look at bounds of call options.

- The price of a call or a put option has to be non-negative because with these options you are offered the possibility for a future gain with no liability. In symbols, we have

$$0 \leq C_{Eur}(K, T) \leq C_{Ame}(K, T).$$

- The price of a European call option must satisfy the put-call parity. Thus, for a non-dividend-paying stock we have

$$C_{Eur}(K, T) = P_{Eur}(K, T) + S_0 - Ke^{rT} \geq S_0 - Ke^{rT}.$$

- Combining all of the above results we can write

$$S_0 \geq C_{Ame}(K, T) = C_{Eur}(K, T) \geq \max \{0, S_0 - Ke^{rT}\}.$$

- For a continuous dividend paying stock we have

$$S_0 \geq C_{Ame}(K, T) \geq C_{Eur}(K, T) \geq \max \{0, S_0 e^{-\delta T} - Ke^{rT}\}.$$

5.7.2 Bounds on put options.

Next, we consider bounds on put options

- Put options are non-negative.

$$0 \leq P_{Eur}(K, T) \leq P_{Ame}(K, T).$$

- The best one can do with a European put option is to get the strike price K at the maturity date T . So a European put cannot be worth more than the present value of the strike price. That is,

$$P_{Eur}(K, T) \leq Ke^{-rT}.$$

- The best one can do with an American put option is to exercise it immediately after time zero and receive the strike price K . So an American put cannot be worth more than the strike price. In symbols

$$P_{Ame}(K, T) \leq K.$$

- The price of a European put option must satisfy the put-call parity. Thus, for a non-dividend-paying stock we have

$$P_{Eur}(K, T) = C_{Eur}(K, T) + Ke^{-rT} - S_0 \geq Ke^{-rT} - S_0.$$

- Combining all of the above results we can write

$$K \geq P_{Ame}(K, T) = P_{Eur}(K, T) \geq \max \{0, Ke^{-rT} - S_0\}.$$

- For continuous dividend paying stock we have

$$K \geq P_{Ame}(K, T) = P_{Eur}(K, T) \geq \max \{0, Ke^{-rT} - S_0e^{-\delta T}\}.$$

5.7.3 Early exercise

Theorem 5.7.3 *It is NEVER optimal to early exercise an American call option on a non-dividend-paying stock.*

Proof. An American call option can be exercised early if the exercise is higher than the option value. Using generalized put-call parity for European options and the fact that $P_{Eur}(S_t, K, T-t) \geq 0$ we can write

$$\begin{aligned} C_{Eur}(S_t, K, T-t) &= P_{Eur}(S_t, K, T-t) + S_t - Ke^{-r(T-t)} \\ &\geq S_t - Ke^{-r(T-t)} \geq S_t - K. \end{aligned}$$

Now since $C_{Eur}(S_t, K, T-t) \leq C_{Ame}(S_t, K, T-t)$ we have

$$C_{Ame}(S_t, K, T-t) \geq S_t - K$$

This shows that the value $S_t - K$ can never be higher than the option value $C_{Ame}(S_t, K, T-t)$ for all $0 \leq t \leq T$. Thus if you sell the call option you receive $C_{Ame}(S_t, K, T-t)$ while you exercise the option you will receive $S_t - K$.

□

5.8 Greeks.

To measure quantitatively the influence of an option's position on given portfolio of financial assets we will have to examine the dependence of its price on the fluctuations of the current stock price, time to expiry, strike price volatility and other relevant parameters. For fixed expiry date T and arbitrary $t \leq T$ we denote τ time to option expiry, that is we put $\tau = T - t$. We write $c^{BS}(S_t, K, \sigma, \tau, r)$ and $p^{BS}(S_t, K, \sigma, \tau, r)$ to denote the price call and put option respectively.

Given closed-form solutions for European options, we can derive all partial derivatives.

5.8.1 Properties of Greeks.

Consider the value of a European option $V = V(S_t, t)$, without paying a dividend satisfying the following Black-Scholes equation

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

Below we present some basic Greek parameters.

Definition 5.8.1 (*Delta, Δ*).

The rate of change of the option price with respect to the price of underlying asset. Mathematically it is partial derivation of call price with respect to price of underlying asset.

$$\Delta := \frac{\partial V}{\partial S_t}.$$

- Delta is used for hedging purposes. It tells the option seller how to edge his exposure by holding $\frac{\partial V}{\partial S}$ units of stock S , at any time.
- When S is close to the strike K , Δ of call option is close to 0.5, and the option behaves like a position of 0.5 in the underlying asset.
- Delta for call option is always positive and belongs to interval $(0, 1)$. Thus, an increase in stock price increases the call option value.
- Delta for put option is always negative and belongs to interval $(-1, 0)$. Thus, an increase in stock price decreases the put option value.

Definition 5.8.2 (*Gamma, Γ*).

The rate of change in Delta with respect to the price of the underlying asset,

$$\Gamma = \frac{\partial^2 V}{\partial S_t^2}.$$

The key features are:

- A small (large) gamma means delta change slowly (rapidly)
- Gamma is important in that it express how much hedging cost in small time interval. In particular if the Gamma of our portfolio is positive we will make money by delta-hedging and if negative we will lose money
- Calls and puts have the same Gamma by the put-call parity since forward contract has $\Gamma = 0$.
- The positivity of Gamma is due to call and put options being convex functions of spot as the second derivative for any convex function is positive.

- At-the-money options have the highest gamma, which indicates that Δ changes very fast as S changes. In contrast, both in-the-money options and out-of-the-money options have low gammas because their delta is constant, close to one or zero, respectively.
- As maturity nears, the option gamma increases
- Gamma is similar to the concept of convexity developed for bonds. Fixed coupon bonds, however, always have positive convexity, whereas options can create negative convexity.

Definition 5.8.3 (*Vega, \mathcal{V}*) The rate of change of the option price with respect to the volatility of underlying asset. Mathematically it is partial derivation of call price with respect to volatility underlying asset.

$$\mathcal{V} = \frac{\partial V}{\partial \sigma}$$

The key features are:

- The Vega expresses the position the trader is taking on volatility: a positive Vega expresses the opinion that volatilities will go up, and a negative Vega the opinion that they will go down.
- As a forward is insensitive to volatility, it follows from put-call parity that the Vega of a put will equal the Vega of the call with the same strike.
- Note that for a call or put, the Vega is always positive. An immediate consequence is that the map from volatilities to prices is injective - if two volatilities give the same price then they are equal. This is part of the reason that the practice of quoting volatilities instead of prices is popular. This observation also allow us to find implied volatility.
- Vega is highest for long-term at-the-money options

Definition 5.8.4 (*Theta, θ*) The rate of change of option price with respect to the passage of time.

$$\Theta = \frac{\partial V}{\partial t}$$

Definition 5.8.5 (*Rho, ρ*) The rate of change of option price with respect to the risk-free rate.

$$\Theta = \frac{\partial V}{\partial t}$$

Unlike linear contracts, options are exposed not only to movements in the direction of the spot price, but also in its volatility. Options therefore can be viewed as "volatility bets."

5.8.2 Greeks for forward contract

Consider now a forward contract, with strike price K and maturity T . Then

$$X = F_T = S_T - K$$

Denote by $F = F_t = S_t - Ke^{-r(T-t)} = S_t - Ke^{-r\tau}$. Then

- delta

$$\delta_F = \frac{\partial F}{\partial s} = 1$$

- gamma

$$\gamma_F = \frac{\partial^2 F}{\partial s^2} = 0$$

- theta

$$\theta_F = \frac{\partial F}{\partial t} = -\tau K e^{-r\tau}$$

- rho

$$\rho_F = \frac{\partial F}{\partial r} = \tau K e^{-r\tau}$$

- vega

$$\nu_F = \frac{\partial F}{\partial \sigma} = 0$$

5.8.3 Useful formulas for Greeks.

To prove this we should use the following facts:

-

$$S_t N'(d_1) - K e^{-r\tau} N'(d_2) = 0$$

-

$$\frac{\partial d_1}{\partial s} = \frac{\partial d_2}{\partial s} = \frac{1}{s\sigma\sqrt{\tau}}$$

-

$$\frac{\partial d_1}{\partial r} = \frac{\partial d_2}{\partial r} = \frac{\sqrt{\tau}}{\sigma}$$

-

$$\frac{\partial d_2}{\partial t} - \frac{\partial d_1}{\partial t} = \frac{\sigma}{2\sqrt{\tau}}$$

-

$$\frac{\partial d_1}{\partial \sigma} - \frac{\partial d_2}{\partial \sigma} = \sqrt{\tau}$$

Below we present the table of Greek parameters split by option type

name	symbol	derivative	call option	put option
Delta	Δ	$\frac{\partial V}{\partial S_t}$	$N(d_1)$	$-N(-d_1)$
Gamma	Γ	$\frac{\partial^2 V}{\partial S_t^2}$	call formula	put formula
Vega	\mathcal{V}	$\frac{\partial V}{\partial \sigma}$	call formula	put formula
Rho	ρ	$\frac{\partial V}{\partial r}$	call formula	put formula
Theta	θ	$\frac{\partial V}{\partial t}$	call formula	put formula

5.8.4 Methods of Calculating Greek Parameters.

There are several ways to compute the Greeks. Three the most popular methods:

1. finite difference approximations
2. path-wise method
3. likelihood ratio method.

Finite-difference approximations involve calculating the price for a given value of parameter, say spot, then changing the parameter value slightly, by ε and recalculating the price. If we let f be the payoff and θ the parameter we are interested in, an estimate of the sensitivity will be

$$\Delta = \frac{f(\theta + \varepsilon) - f(\theta)}{\varepsilon}.$$

This method is easy to implement does not require much thought, apart from choosing an appropriate ε . It is however biased.

5.9 Combination

One of the most appealing features of options (apart from the obvious chance of making extraordinary returns) is the possibility of easy speculation on the future behaviour of a stock price. Usually this is done by means of so called combinations – that is, combined positions in several options, and possibly the underlying asset. For instance, a bull spread is portfolio created by buying a call option on a stock with a certain strike price and selling a call option on the same stock with a higher strike price (both options have the same expiry date). Equivalently, bull spreads can be created by buying a put with a low strike price and selling a put with a high strike price. An investor entering a bull spread is hoping that the stock price will increase. Like a bull spread, a bear spread can be created by buying a call with one strike price and selling a call with another strike price. The strike price of the option purchased is now greater than the strike price of the option sold, however. An investor who enters a bear spread is hoping that the stock price will decline.

A butterfly spread involves positions in options with three different strike prices. It can be created by buying a call option with a relatively low strike price, buying another call option with a relatively high strike price, and selling two call options with a strike price halfway between the other two strike prices. The butterfly spread leads to a profit if the stock price stays close to the strike price of the call options sold, but gives rise to a small loss if there is a significant stock price move in either direction. A portfolio created by selling a call option with a certain strike price and buying a longer-maturity call option with the same strike price is commonly known as a calendar spread. A straddle involves buying a call and put with the same strike price and expiry date. If the stock price is close to this strike price at expiry of the option, the straddle leads to a loss. A straddle is appropriate when an investor is expecting a large move in stock price but does not know in which direction the move will be. Related types of trading strategies are commonly known as strips, straps and strangles.

5.10 Hedging and replication.

1. What is pricing by replication?

Pricing by replication involves trading in financial instruments in such way as to match the payoff of another instrument that is most likely exotic. For example, a barrier option can be priced using a portfolio of vanilla options whose payoff matches that of the barrier option. One of the main benefits of pricing exotic options by (static) replication with vanilla options is that exotic will be priced in a way that is consistent with current volatility smile.

Chapter 6

Building interest rate curves.

6.1 Introduction.

The area of interest rates, specially interest rate options, usually requires the most complicated quant work. The main reason for this is that an interest rate derivative depends on the yield curve which is a one dimensional object, as opposed to an equity option which depends on zero dimensional object: the stock price. Pricing complex interest rate derivatives requires modelling the future dynamics of the yield curve term structure. Hence, most of the literature assumes the existence of the current yield curve as given, and its construction is often neglected, or even obscured, as it is considered more a question of art than a science. Financial institutions, and practitioners have developed their own methodologies in order to extract the current yield curve term structure from quoted prices of a finite number of market instruments. Broadly speaking there are two classes of yield curve constructions algorithms:

- best fit,
- exact fit.

Best-fit algorithms assume a functional form for the term structure and calibrate its parameters using a selection of calibration instruments quoted on the market, such that to minimize the repricing error. Usually it is based on the Nelson-Siegel model or some kind of its extension. Such approach is popular due to the smoothness of the curve, calibration easiness, intuitive financial interpretation of functional form parameters (level slope curvature). On the other side, the fit quality is typically not good enough for trading purposes in liquid interest rate markets, where several basis points can make the difference.

In practice exact-fit algorithms are often preferred: they fix the yield curve on a time grid of N points (pillars) in order to exactly reprice N pre-selected market instruments. The implementation of such algorithms is often incremental, extending the yield curve step-by-step with the increasing maturity of the ordered instruments, in a so called bootstrap approach. Intermediate yield curve values are obtained by interpolation on the bootstrapping grid. Different interpolation algorithms are available, but little attention has been devoted in the literature to the fact that interpolation is often used already during bootstrapping, not just after that, and that the interaction between bootstrapping and interpolation can be subtle if not nasty.

6.1.1 The classical single-curve approach.

The pre-crisis standard market practice for the construction of the single interest rate yield curve can be summarized in the following procedure:

- Interbank credit/liquidity issues do not matter for pricing, Libors are a good proxy for risk free rates, Basis Swap spreads are negligible.
- The collateral do not matter for pricing, Libor discounting is adopted.
- Select one finite set of the most convenient (i.e. liquid) vanilla interest rate instruments traded in real time on the market, with increasing maturities. For instance a very common choice in the EUR market is a combination of short-term EUR Deposit, medium-term Futures on Euribor3M and medium-long-term Swaps on Euribor6M.
- Build one yield curve using the selected instruments plus a set of bootstrapping rules (e.g. pillars, priorities, interpolation, etc.).
- Compute on the same curve FRA rates, cash flows, discount factors, and work out the prices by summing up the discounted cash flows.

Definition 6.1.1 (*Term structure of interest rates.*)

A term structure of interest rates is a set of interest rates sorted by time to maturity. The curve shows the relation between the (level of) interest rate (or cost of borrowing) and the time to maturity.

We can build several types of curves using rates of a different nature, for example zero coupon yield curve, forward rates curve, instantaneous forward curve. Below we provide basic mathematics formulae to deal with these term structures of interest rates.

6.1.2 Interpolation.

Interpolation is a method of constructing new data points within the range of a discrete set of known data points. In other words, it is a process of approximating the value of function $y(t)$ between two points at which it has prescribed values.

Interpolation is very important in yield curve construction and determines some characteristics of the curve. A lot of risk and money can be hidden behind the interpolation method. There are many choices of interpolation function, however we are interested in interpolations that preserve arbitrage-free conditions, localness (a change in an input, affects the shape of the curve only locally), smoothness, positivity and stability (a change in an input, does not affect the entire shape of the curve) of forward rates.

6.1.3 Cubic spline interpolation.

In order to generate a stochastic term structure with a no-arbitrage model that needs the current yield curve as an input parameter, it was necessary to include appropriate interpolation scheme in the computer simulation program.

Chapter 7

Implied volatility

All potential practical applications of Black-Scholes formula hinge on knowledge of the volatility parameter of the return of the stock prices.

Definition 7.0.1 *Given an observed European call option price C^{obs} for a contract with strike price K and expiration date T , the implied volatility is defined to be the value of the volatility parameter that must go into the Black-Scholes formula such that*

$$C_{BS}(S, K, r, T, t; \sigma_{iv}) = C^{obs}.$$

Indirect calculation show that

$$\frac{\partial C_{BS}}{\partial \sigma} > 0$$

Given that C_{BS} is a monotonic increasing function of σ , it means that for every option price, there must be a corresponding implied volatility.

If one calculates the implied volatility of options with the same expiry date but different strike prices and plots the volatilities, there is often a smile or skew shape.

It is convincingly believed that the constant volatility assumption of the Black-Scholes model is rejected by any empirical facts. The phenomenon of volatility smile shows the variation of the Black-Scholes implied volatility with respect to the strike price. Smile is common among currency options, where both tails are thicker than log-normal. For equity options common is skew. One possible explanation is that the market is more likely to move up or down by a large amount than is assumed within the Black-Scholes model. Hence the smile reflects the market's view of the imperfections in the Black Scholes model.

7.1 Summary about volatility.

There are many possible explanations for why some markets exhibit volatility smiles. One possible explanation is that the market is more likely to move up or down by a large amount than is assumed within Black-Scholes model. Hence the smile reflects the market's view of the imperfections in the black Scholes model.

In general, if implied volatility is higher than historical volatility it gives some indication that option prices may be high. If implied volatility is below historical volatility, this may mean option prices are discounted.

Many option traders focus specifically on volatility. When the implied volatility is low, it is generally a good time to buy an option. When implied volatility is high, it can be a good time to sell an option or use a spread strategy.

A volatility number of 0.55, whether implied or historical, means that in one year stock has a 68% chance (one standard deviation) of being 55% higher or lower.

In order to fix the issue raised by the smile effect, that is for estimating and fitting such volatility smiles, two majors approaches have been developed: local volatility and stochastic volatility.

Merton (1971) suggested to make the volatility a deterministic function of time. This would indeed explain the different volatility for different tenors, but would not explain the smile effect for different strikes.

Other local volatility model (Dupire, 1994), suggests that we should consider the volatility model to be a deterministic function of the underlying stock and the time. However, it cannot explain the persistent smile shape which does not vanish over time with longer maturities. The SABR model has a unique feature that allows you to compute the implied volatility directly for a given strike

Chapter 8

Local volatility models.

Local volatility models are self-consistent, arbitrage-free, and can be calibrated to roughly match some observed market smiles and skews. Currently these models are the most popular way of managing smile and skew risk.

However, it has recently been observed (Hagan et al., 2002) that the dynamic behaviour of smiles and skews predicted by local volatility models is exactly opposite to the behaviour observed in the marketplace: local volatility models predict that the skew moves in the opposite direction to the market level, in reality, it moves in the same direction. This leads to extremely poor hedging results within these models, and the hedges are often worse than the naive Black model hedges, because these naive hedges are in fact consistent with the smile moving in the same direction as the market.

8.1 Displaced diffusion model.

8.2 Constant elasticity of variance model (CEV)

This model have been developed to fit the volatility smiles that are observed in practice. The constant elasticity of variance model is an example for a diffusion model where the risk neutral process for stock price is

$$dS_t = \mu S_t dt + \sigma S_t^\alpha dW_t, \quad \mu \in \mathbb{R}, \quad \sigma \in (0, \infty).$$

When $\alpha = 1$, the CEV model is the geometric Brownian motion model. When $\alpha < 1$ the volatility increases as the stock price decreases. This creates a probability distribution similar to that observed for equities with a heavy left tail and a less heavy right tail. When $\alpha > 1$ the volatility increases as the stock price increases. This creates a probability distribution with a heavy right tail and less heavy left tail.

The CEV model is more suitable for catching the volatility smile similar to that observed for equity options, while jump-diffusion models are better for currency options.

8.3 Duprie formula

Chapter 9

Jump diffusion model.

We have mentioned the imperfections of the Black-Scholes model of stock-price evolution. In this chapter, we look at one method of improving it. Our improvement is to add the possibility of the stock price jumping discontinuously. The motivation for this model is that stock markets do crash and during a crash there is no opportunity to carry out a continuously-changing Delta hedge. One consequence of this will be the impossibility of perfect hedging; at any given time the stock price can increase slightly or decrease slightly or fall a lot. It is not possible to be hedged against all of these simultaneously. The impossibility of perfect hedging means that the market is incomplete, that is not every option can be replicated by a self-financing portfolio. The price of a non-replicable option can then only be bounded rather than fixed using no-arbitrage methods.

One explanation of this smile is that it is caused by strong demand for slightly out-of-the-money put options. A fund manager has his performance reviewed every three months. He wants to be protected against the possibility of a market crash in the mean time. He therefore buys put options which guarantee that his portfolio's value can only fall by a small amount even if the market crashes. Thus he is buying the put option as insurance. Since there are many fund managers doing the same thing, hedging is impossible, and no one wants extra exposure to crashes, the market is all one way, and the price of the put option is bid up.

Suppose we have a stock moving under geometric Brownian motion, with the added possibility of crashes. What properties should crashes have? They should occur instantaneously: the probability of one occurring in a given small time interval should be roughly proportional to the length of a time interval. We can achieve these properties by modelling crashes with a Poisson process.

We recall the characteristics of a Poisson process. The main parameter is the intensity λ and the probability of an event in a given small time interval, Δt , is $\lambda\Delta t$ plus a smaller error. We denote the number of events up to time t by $N(t)$. We therefore have that $N(t)$ is an integer-valued function which is constant for a while and then jumps up by 1. Another important property is that the probability of an event is independent of the number of events that have already occurred: that is for $t > s$, we have that $N(t) - N(s)$ is independent of the value of $N(s)$. We have

$$\mathbb{P}(N(t) = j) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}.$$

that is, the number of jumps up to time t is Poisson distributed with parameter λt .

Returning to our stock, it moves according to a geometric Brownian motion with superimposed jumps. We suppose that at a jump the stock price is multiplied by a random variable J . We could take J to be log-normal, or to be a single constant number, or anything we want. Then

$$dS_t = \mu S_t dt + \sigma S_t dW_t + S_t J(n-1) dN(t).$$

Note that when a jump occurs, S changes by $S(J-1)$, which is equivalent to multiplying S by J .

Remark 9.0.1 *It is worth to remember*

- *Jump diffusion models have the foreign exchange rate moving as geometric Brownian motion with an added random jump component*
- *One disadvantage of jump diffusion model is that the smiles produced flatten with maturity quickly.*

Chapter 10

Stochastic volatility models.

10.1 Induction.

A stochastic volatility model takes volatility to be driven by the random process, thus those models have two sources of randomness, the stock price and the volatility. One of the parameter in the stochastic models family is a correlation between the two sources of randomness. This correlation is typically negative so that a fall in the stock price is often accompanied by rise in volatility. This results in a negative skew for implied volatility.

It is then possible that the volatility will become large causing large movements in the underlying asset price. These large moves give the distribution of the underlying asset fatter tails than in the Black-Scholes model. This means that options away from the money will be priced more expensively than in Black-Scholes leading to a volatility smile. Two the most popular stochastic volatility models

10.2 Heston model.

There are many popular forms for the volatility process, the Heston model for example uses

$$\begin{aligned} dS_t &= \mu S_t dt + S_t \sqrt{V_t} dW_t^{(1)}, \quad S_0 > 0 \\ dV_t &= \kappa(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_t^{(2)}, \quad V_0 = \sigma_V^2 = 0 \end{aligned}$$

where θ is the long term average volatility, κ is the rate at which the process reverts to its mean, σ_V is volatility of volatility and $W_t^{(1)}$ $W_t^{(2)}$ are Brownian motions with correlation ρ .

The variance process $(V_t)_{t \geq 0}$ is called Feller diffusion or CIR process and the Yamada-Watanabe conditions ensure that a non negative unique strong solution exists.

Shape of smiles can be easily changed by tweaking the parameters. Having the flexibility of changing the smile shape has its disadvantages in that all these parameters need to be fitted in a stable and consistent way with market which is not a straightforward task.

Heston model is incomplete. Therefore, it is not possible to obtain an unique price for any contingent claim using only the underlying asset and a bank account, which is normally the case for complete models such as the Black-Scholes model. To complete the market in the Heston model, one has to add an European call option for example.

The Heston is quite often a benchmark especially in terms of equity pricing because it has following futures:

- The variance process is mean reverting with θ as long term average volatility and κ is the rate at which the process reverts to its mean.
- It has a quasi closed form solution.

- it takes into account leverage effect, it means that $dW_t^{(1)}$ and $dW_t^{(2)}$ are negative correlated.

10.2.1 Influence of parameters.

It is important to understand the meaning of the Heston parameters, since the effective use of the stochastic volatility model depends on the initial parameters and calibration parameters. A brief summary of parameters of Heston model.

- initial variance V_0
- long run variance
- mean reversion
- correlation

10.2.2 Advantages and disadvantages of the Heston model.

Both academia and practitioners have recognized the significance of the Heston model nevertheless it is not a model without any drawbacks. Advantages of the Heston model:

-

Disadvantages of the Heston model

10.3 SABR model.

The SABR model is a volatility model, which attempts to capture both the correct shape of the smile, as well as the correct dynamics of the volatility smile. The model does not provide option prices exactly, instead, it gives an estimate of the implied volatility curve, which is then considered as an input in Black's model to price swaptions, caps, and other interest rate derivatives.

SABR is an acronym of Stochastic, Alpha, Beta, Rho. The SABR model was developed by Patrick Hagan, Deep Kumar, Andrew Lesniewski, and Diana Woodward. Widely used by practitioners, the model is used to model forward swap rates, forward stock prices, forward Libor rates or any others forward rate.

Definition 10.3.1 *Under T -forward measure, the SABR model is described by the following system of stochastic differential equations:*

$$dF = \hat{\alpha} F^\beta dW_1, \quad F(0) = f,$$

$$d\hat{\alpha} = \nu \hat{\alpha} dW_2, \quad \alpha(0) = \alpha,$$

$$\mathbb{E}[dW_1 dW_2] = \rho dt,$$

where F is any forward rate, $\hat{\alpha}$ is the volatility, and W_1 and W_2 are two correlated Wiener processes.

We have the following parameters:

1. α is the initial variance and satisfies the condition $\alpha \geq 0$.
2. β is the exponent for the forward rate and satisfies the condition $0 \geq \beta \leq 1$
3. ν is the volatility of variance.
4. ρ is the correlation between the two Wiener processes and satisfies $0 < \rho < 1$.

We shall point out the following particular cases of SABR model. When

1. $\beta = 0$, the SABR model is reduced to the normal model
2. $\beta = \frac{1}{2}$, the SABR model is reduced to stochastic Cox-Ingersoll-Ross model
3. $\beta = 1$, the SABR model is reduced to the stochastic log-normal model.
4. $\mu = 0$, the SABR model is reduced to the CEV (Constant Elasticity of Variance) model.

No closed form expression for this probability distribution is known, only for the special cases: $\beta = 0$ and $\beta = 1$.

The prices of European call options in the SABR model are forced into the form of the Black model valuation formula. The implied volatility $\sigma_B(f, K)$ is approximately given by:

$$\sigma_B(f, K) \simeq U \left(\frac{z}{X(z)} V \right)$$

where

$$U = \alpha \left\{ (fK)^{\frac{(1-\beta)}{2}} \left[1 + \frac{(1-\beta)^2}{24} \log^2 \left(\frac{f}{K} \right) + \frac{(1-\beta)^4}{1920} \log^4 \left(\frac{f}{K} \right) \right] \right\}^{-1},$$

$$V = \left\{ 1 + \left[\frac{(1-\beta)^2}{24} \frac{\alpha^2}{(fK)^{1-\beta}} + \frac{\rho\beta\nu\alpha}{4(fK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24} \nu \right] t_{ex} + \dots \right\}.$$

Here z is defined as

$$z = \frac{\nu}{\alpha} (fK)^{(1-\beta)/2} \log \left(\frac{f}{K} \right)$$

and $X(z)$ is defined by

$$X(z) = \log \left(\frac{\sqrt[2]{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right).$$

We note it easily the particular case of at-the-money options, options on struck at $K = f$ formula of volatility reduces to

$$\sigma_{ATM} = \left\{ 1 + \left[\frac{(1-\beta)^2}{24} \frac{\alpha^2}{(f)^{2-2\beta}} + \frac{\rho\beta\nu\alpha}{4(fK)^{(1-\beta)}} + \frac{2-3\rho^2}{24} \nu \right] t_{ex} + \dots \right\}.$$

Calibration of parameters.

At a fixed f , the implied volatility σ_B is reduced to a function of K when the parameters α , β , ρ and ν are estimated. Here, we examine how these parameters are estimated.

Usually, the parameter β is estimated first and from experience it is observed that market smiles can be fitted equally well with any specific value of β . Once β is known, there are two ways of estimating α , ρ and ν : we can calibrate α , ρ and ν directly (in one step) or we can calibrate ρ and ν directly (first step) and deduce α from ρ , ν and σ_{ATM} (second step).

- Estimating β

Parameter β can be chosen using our predictions about the model. Supporters of choice $\beta = 0$ believe that a normal model is a powerful tool for managing risks should be used for markets where $f < 0$ or $f \simeq 0$. Supporters of choice $\beta = 0.5$ are usually US interest rate desks that have developed trust in CIR models. Finally, followers of the choice $\beta = 1$ include desks trading foreign exchange options, they believe that log-normal models are more natural.

- Estimating α , ρ and ν

Once β is known estimating α , ρ and ν consists of minimizing the errors between the market volatilities σ_j^{mkt} and the model volatilities σ_B for the same maturity t_{ex} . That is

$$(\hat{\alpha}, \hat{\rho}, \hat{\nu}) = \underset{\alpha, \rho, \nu}{\operatorname{argmin}} \sum_j [\sigma_B(f_j, K_j; \alpha, \rho, \nu) - \sigma_j^{mkt}]^2$$

Dynamics of SABR model.

The three parameters α , β and ν have different effects on the curve. We will adjust each parameter while keeping the others constant and observe the result in the smile's shape.

•

10.4 Discretization.

Typically the dynamics of these stock prices and interest rates are assumed to be driven by a continuous-time stochastic process. Simulation, however, is done at discrete time steps. Hence, the first step in any simulation scheme is to find a way to discretize a continuous-time process into a discrete time process. Let's denote:

- S_t price over the time interval $[0, T]$,
- $0 = t_0 < t_1 < \dots < t_m = T$, t_i time steps for $i = 0, 1, 2, \dots, m$,
- $\Delta t = t_{i+1} - t_i$,
- $\mu \in \mathbb{R}$, $\sigma > 0$,

As we mentioned in previous paragraph the stock prices S_t is driven by the stochastic differential equation (SDE)

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

It is very important that $W_{t+\Delta t} - W_t$ and $\sqrt{\Delta t}Z$ are identical in distribution, where $Z \sim \mathcal{N}(0, 1)$. Hence we get

$$S_{t+\Delta t} = S_t + \mu S_t \Delta t + \sigma S_t \sqrt{\Delta t} Z.$$

From definition of Brownian motion it's obvious that it may takes negative values. Of course price of equities are always positive. So for financial application we need discretize geometric Brownian motion. To simulate the time paths of daily stock prices, from the day of an option investment to the expiry date of an option, we need an explicit expression of the stock price on each day in terms of the stock price a day earlier. Specifically, if we use t and $t + \Delta t$ to indicate two successive points in time, we can note

$$S_{t+\Delta t} = S_t \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) \Delta t + \sigma \sqrt{\Delta t} Z \right). \quad (10.1)$$

10.5 Option pricing via Monte Carlo simulation.

In general the technique is to generate several thousand possible (but random) price paths for the underlying asset via simulation, and then calculate the associated payoff of the option for each path. These payoffs are then averaged and discounted to today, and this result is the value of the option today. We use this method to calculate the price of an option with complicated features. Consider an option on a stock with strike price K and expiration time T . Let $S_T^1, S_T^2, \dots, S_T^n$ be n randomly drawn stock prices at time T . Let $V(S, T)$ denote the option payoff at time T . We define the time at 0 Monte Carlo price of the option by

$$V(S_0, 0) = \frac{1}{n} e^{-rT} \sum_{i=1}^n V(S_T^i, T).$$

For the case of call option we have

$$V(S_T^i, T) = \max\{0, S^i - K\}.$$

For a put, we have

$$V(S_T^i, T) = \max\{0, K - S^i\}.$$

Remark 10.5.1 *Monte Carlo simulations, are central to financial engineering and risk management. They allow financial engineers to price complex financial instruments. They allow also risk managers to build the distribution of portfolios that are too complex to model analytically.*

Remark 10.5.2 *Use binomial trees for low dimensional problems involving early exercise and Monte Carlo method for high dimensional problems involving path dependence.*

Problem 10.5.1 *One of nice example using of Monte Carlo methods is the estimation of Π .*

10.5.1 Option price proprieties

10.5.2 Stock price.

All other things equal, as the stock price increases, so does the call value. A call with an exercise price of 60 would be worth more if the current market price were 50 than if it were 45, all else equal. The probability that the market price will eventually exceed the exercise price by any given amount is higher in the former case than in the latter, and so, consequently, is the payoff expected from the holding call. Similarly, the lower the stock price, the more a put option is worth. If you want to have the right to sell stock at 60, you would pay more for that put option when the stock is 59 than when it is 65. The lower the current stock price, the more a put is worth.

10.5.3 Strike price.

Wouldn't you pay more for the right to buy stock at say 380 than for the right to buy stock at 410. Of course you would always prefer the right to buy stock at a lower price any day of the week! Thus, calls become more expensive as the strike price moves lower. Likewise, puts become more expensive in value as the strike price increases.

10.5.4 volatility

Stocks that are volatile go through strike price more frequent than the non-volatile stocks. With these big moves, you have a higher chance of making money.

10.5.5 Time to maturity

Options have a definitive life because of expiration. Therefore, an option will increase in value with more time. Why? Well, the more the time until expiration, the greater the probability of a profitable move.

10.5.6 Interest rate

This is really a small factor in determining an option's price. When interest rates are on the rise, the value of call options rise as well. If a trader decides to buy a call option instead of stock, then the extra cash they have should theoretically earn interest for them. While this doesn't necessarily work so easily in the real world the theory behind it does make sense.

Theorem 10.5.1 *Price of call option is convex function of the strike price*

10.6 Forward contract and future contract.

Definition 10.6.1 (*Future contract*)

A future contract is an agreement to buy or sell an asset at a certain date in the future for a certain price.

The important feature of these contracts is that are traded on exchanges.

Definition 10.6.2 (Forward contract)

Forward contract is an agreement established at time t to pay or receive on a settlement date T a preassigned payoff, say X at an agreed forward price.

Let us emphasize that there is no cash flow at contract's initiation and the contract is not marked to market. In contrast to stock options and futures contracts, forward contracts are not traded on exchanges. By convention, the party who agrees to buy the underlying asset at time T for the delivery price K is said to assume a long position in a given contract. Consequently, the other party, who is obliged to sell the asset at the same date for the price K , is said to assume a short to buy an asset worth S_T at maturity for K . It is clear that the payoff from the long position (from the short position, respectively) in a given forward contract with a stock S being the underlying asset corresponds to the time T contingent claim X where

$$X = S_T - K,$$

$$X = K - S_T,$$

As we can see value of forward contract (payoff) depends on the relationship between delivery price and underlying price.

Since the final value (at maturity) of a forward position depends on the spot price which will then be prevailing, this contract can be viewed, from a purely financial point of view, as a bet on the future spot price.

To value forward contract we need to remember simple formula for the forward price at time T of non-dividend paying stock,

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

When the stock pays it dividend the price will drop by the dividend amount, so we need to take this into account in forward price. We simply just subtract the suitable accumulated dividend amount, giving

$$F_T = S_0 e^{rT} - de^{T-T_d}$$

10.7 Currency option valuation formula.

Definition 10.7.1 Set $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbf{T}}, \mathbb{P})$, where $(\mathcal{F}_t)_{t \in [0, T]}$ is filtration defined above is called filtered probability space.

The exchange rate process Q which is used to convert foreign payoffs into domestic currency is modelled by the following stochastic differential equation

$$dQ_t = \mu Q_t dt + \sigma Q_t dW_t \quad \mu \in \mathbb{R}, \sigma \in (0, \infty), S_0 \in (0, \infty)$$

Example of currency option, we consider a standard European call option, whose at the expiry date T equals

$$C_T^Q := (Q_T - K)^+$$

where Q_T is the spot price of the deliverable currency.

Theorem 10.7.1 The arbitrage price, in units of domestic currency, of a currency European call option is given by the risk-neutral valuation formula

$$C_t^Q = e^{-r_d(T-t)} \mathbb{E}^{\mathbb{P}^*} [(Q_T - K)^+ | \mathcal{F}_t]$$

Moreover, the price C_t^Q is given by the following expression

$$C_t^Q = Q_t e^{-r_f(T-t)} \mathcal{N}(h_1) - K e^{-r_d(T-t)} \mathcal{N}(h_2)$$

where

$$h_1 := \frac{\log \frac{q}{K} + (r_d - r_f + \frac{1}{2}\sigma_Q^2)\tau}{\sigma\sqrt{\tau}}$$

$$h_2 := \frac{\log \frac{q}{K} + (r_d - r_f - \frac{1}{2}\sigma_Q^2)\tau}{\sigma\sqrt{\tau}}$$

10.8 Exotic options

The aim of this chapter is to study examples of more sophisticated option contracts. For convenience, we give the generic name exotic option to any option contract which is not a standard European or American option. It should be made clear that we shall restrict our attention to the case of exotic spot options. We find it convenient to classify the large family of exotic options as follows:

- **packages** – options that are equivalent to a portfolio of standard European options, cash and the underlying asset (stock, say);
- **chooser options** – option contracts that are chosen by their holders to be call or put at a prescribed future date;
- **binary options** – contracts whose payoff is defined by means of some binary function;
- **barrier options** – options whose payoff depends on whether the underlying asset price reaches some barrier during the option's lifetime;
- **Asian options** – options whose payoff depends on the average price of the underlying asset during a prespecified period;
- **basket options** – options with a payoff depending on the average of prices of several assets;
- **lookback options** – options whose payoff depends, in particular, on the minimum or maximum price of the underlying asset during options' lifetimes;

10.8.1 Chooser options

A chooser option is an agreement in which one party has the right to choose at some future date whether the option is to be a call or put option with a common exercise price K and remaining time to expiry $T - T_0$. Therefore, the payoff at T_0 of chooser option is

$$CH_{T_0} := \max \{C(S_{T_0}, T - T_0, K), P(S_{T_0}, T - T_0, K)\}$$

10.8.2 Asian options

An Asian option (or an average option) is a generic name for the class of options (of European or American style) whose terminal payoff is based on average asset values during some period within the options' lifetimes. Due to their averaging feature, Asian options are particularly suitable for thinly traded assets (or commodities). Actually, in contrast to standard options, Asian options are more robust with respect to manipulations near their expiry dates. Typically, they are also less expensive than standard options. Let T be the exercise date, and let $0 \leq T_0 \leq T$ stand for the beginning date of the averaging period. Then the payoff at expiry of an Asian call option equals

$$C_T^A := (A_S(T_0, T) - K)^+$$

where

$$A_S(T_0, T) = \frac{1}{T - T_0} \int_{T_0}^T S_u du$$

The main difficulty in pricing and hedging Asian options is due to the fact that the random variable $A_S(T_0, T)$ does not have a log-normal distribution. This feature makes the task of finding an explicit formula for the price of an Asian option surprisingly involved. For this reason, early studies of Asian options were based either on approximations or on the direct application of the Monte Carlo method.

10.9 Stochastic models of bond prices.

There is no doubt that management of interest rate risk, by which we mean the control of changes in value of a stream of future cash flows resulting from changes in interest rates, or more specifically the pricing and hedging of interest rate products, is an important and complex issue. It creates a demand for mathematical models capable of covering all sorts of interest rate risks. A **bond** is a certificate issued by a government or a public company promising to repay borrowed money at a fixed rate of interest at a specified time. Let $T^* > 0$ be a fixed horizon date for all market activities. By a zero-coupon bond (a discount bond) of maturity T we mean a financial security paying to its holder one unit of cash at a prespecified date T in the future. We assume that bonds are *default-free*. The price of zero-coupon bond of maturity T at any instant $t \leq T$ will be denoted $B(t, T)$.

Chapter 11

Interest rates

Let us consider a zero-coupon bond with maturity date $T \leq T^*$. We assume that for any fixed maturity $T \leq T^*$, the bond price $B(t, T)$, $t \in [0, T]$ is a strictly positive and adapted process (could be numeraire) on filtered probability space.

In finance we can point out spot rate, short rate and forward rate. It is important to distinguish them

Definition 11.0.1 *The simple rate of return from holding the bond over the time $[t, T]$ equals:*

$$\frac{1 - B(t, T)}{B(t, T)}.$$

Definition 11.0.2 *An adapted process $Y(t, T)$ defined by the formula*

$$Y(t, T) = -\frac{1}{T - t} \log B(t, T)$$

*is called the **yield-to-maturity** on a zero-coupon bond maturing at time T .*

Definition 11.0.3 *The **term structure of interest rates**, know also as the **yield curve** is the function that relates the yields $Y(t, T)$ to maturity T . It is obvious that, for arbitrary fixed maturity date T , there is a one-to-one correspondence between the bond price process $B(t, T)$ and its yield-to-maturity process $Y(t, T)$. Given the yield-to-maturity process $Y(t, T)$, the corresponding bond price process $B(t, T)$ is uniquely determined by the formula.*

In practice, the term structure of interest rates is derived from the prices of several actively traded interest rate instruments, such as Treasury bills, Treasury bonds, swaps and futures. Note that the yield curve at any given day is determined exclusively by market prices quoted on that day. The shape of an historically observed yield curve varies over time; the observed yield curve may be upward sloping, flat, descending, or humped.

11.0.1 Forward interest rate.

We have

$$f(t_1, t_2) = \frac{\frac{P(t_1)}{P(t_2)} - 1}{t_2 - t_1}$$

where $f(t_1, t_2)$ represents the forward rate from time t_1 to t_2 $P(t_1)$ is the value today of zero coupon bonds expiring at time t_1 .

11.0.2 Short-term interest rate

Most traditional stochastic interest rate models are based on the exogenous specification of a short-term rate of interest. We write r_t to denote the instantaneous interest rate (also referred to as a short-term interest rate, or spot interest rate) for risk-free borrowing or lending prevailing at time t over the infinitesimal time interval $[t, t + dt]$.

In a stochastic set-up, the short-term interest rate is modelled as an adapted process, say r , defined on a filtered probability space for some $T^* > 0$. We assume throughout that r is a stochastic process with all sample paths integrable on $[0, T^*]$ with respect to Lebesgue measure. We may then introduce an adapted process A of finite variation and with continuous sample paths, given by the formula

$$A_t = \exp \left(\int_0^t r_u du \right), \quad \forall t \in [0, T^*]$$

Equivalently, for almost all $\omega \in \Omega$, the function $A_t = A_t(\omega)$ solves the differential equation $dA_t = r_t A_t dt$, with $A_0 = 1$

11.1 Short-term rate models.

11.1.1 Notation.

In order to introduce the basic terminology, a Zero bond is considered that pays 1\$ at maturity. Today the time point is $t = 0$ and the Zero is assumed to mature at time point $t = T \geq 0$. The unit for the time axis is always years throughout this chapter. In finance crucial question is how much would someone be willing to pay for this bond at time t , $0 \leq t \leq T$. Let's introduce following notation:

- $P(t, T) :=$ price at time t of a Zero bond that matures at time T ,
- $R(t, T) :=$ spot rate, i.e., continuously compounded yield at time t on the Zero bond that matures at time T ,
- $r_t :=$ short-term interest rate at time t (also called short rate),
- $f(t, T) :=$ instantaneous forward rate at time t for time T ,

The pricing relationship for a Zero bond is:

$$P(t, T) = e^{-R(t, T)(T-t)}, \quad \forall t \in [0, T].$$

If the bond price is given, the spot rate $R(t, T)$ can be calculated as

$$R(t, T) = -\frac{1}{T-t} \log P(t, T), \quad \forall t \in [0, T].$$

Definition 11.1.1 *The instantaneous forward rate $f(t, T)$ is defined as*

$$f(t, T) := -\frac{\partial}{\partial T} \ln P(t, T).$$

11.1.2 Mean reversion process

A possible mean reversion process for a stock is

$$dS_t = \alpha(\mu - S_t)dt + \sigma S_t dW_t$$

where α is the mean reversion rate, and μ is the long term mean. These are the real world dynamics of the stock process, but to price an option we are only interested in the risk-neutral dynamics.

When we move to risk-neutral dynamics the drift of the stock becomes $r dt$. Therefore the real world drift, be it mean reverting or not, does not change the price of the option.

There is strong economic theory and evidence that interest rates follow a mean reversion process. If interest rates are high the economy slows down and the demand for money decreases, pushing interest rates back down. As interest rates decrease the economy becomes more active and demand for money increases, pushing interest rates back up. A rule of thumb regarding whether a quantity is mean-reverting is whether you expect it to be the same range roughly 100 years from now, if you do, then it must have some sort of mean reversion. Financial models are designed to capture market behaviour, so if the market behaves in a mean reverting way, then a good model will capture this feature.

A big difference between the interest rate models and models for the equity market is that in a short rate model the drift matters. The reason for this is that the short rate is NOT a tradeable asset so if the drift is mean reverting it will affect derivative prices.

11.1.3 Ornstein-Uhlenbeck process.

The Ornstein-Uhlenbeck process is a one-dimensional Ito diffusion that is given as the solution of the following linear SDE in the stronger sense:

$$dX_t = \alpha(\mu - X_t)dt + \sigma dW_t, \quad t \geq 0, \quad X_0 = x_0 \in \mathbb{R}, \quad \alpha \in \mathbb{R}^+, \quad \sigma \in \mathbb{R}^+. \quad (11.1)$$

This process mean reverts around μ . The SDE is solved by introducing the integrating factor

$$\Phi(t) = e^{-\alpha t}, \quad t \geq 0,$$

and evaluating the differential $d(e^{-\alpha t} X_t)$. Solving equation (11.1) we get

$$X_t = x_0 e^{-\alpha t} + \mu(1 - e^{-\alpha t}) + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW_s, \quad t \geq 0.$$

11.1.4 Merton model.

Merton proposed to model the short-term rate process r through the formula

$$r_t = r_0 + at + \sigma W_t^*, \quad t \in [0, T], \quad (11.2)$$

where W^* is one dimensional standard Brownian motion under the spot martingale measure W^* and a, r_0, σ are positive constants

Theorem 11.1.1 *The arbitrage price $B(t, T)$ of T -maturity zero-coupon bond in Merton's model equals*

$$B(t, T) = e^{-r_t(T-t) - \frac{1}{2}a(T-t)^2 + \frac{1}{6}\sigma^2(T-t)^3}. \quad (11.3)$$

the dynamics of the bond price under \mathbb{P}^ are*

$$dB(t, T) = B(t, T)(r_t dt - \sigma(T-t)dW_t^*).$$

Proof. Let us evaluate the price $B(0, T)$ using the formula

$$B(0, T) = \mathbb{E}^{\mathbf{P}^*} \left[e^{-\int_0^T r_u du} \right], \quad T \in [0, T^*]$$

We need to find the probability distribution of the integral $\xi_T = \int_0^T r_u du$. Let us consider $Y_t = r_t(T-t)$. Using Ito formula we obtain

□

11.1.5 Vasicek model.

The short-term interest rate r is defined as the unique strong solution of SDE

$$dr_t = (a - br_t)dt + \sigma dW_t^* \quad (11.4)$$

a, b, σ are strictly positive constants.

Theorem 11.1.2 *The unique solution of (11.4) is given by the formula*

$$r_t = r_s e^{-b(t-s)} + \frac{a}{b} \left(1 - e^{-b(t-s)}\right) + \sigma \int_s^t e^{-b(t-u)} dW_u^*.$$

Proof. Let us fix $s > 0$ and let us consider the process $Y_t = r_t e^{-b(t-s)}$, where $t \geq s$. Then using Ito formula.

□

11.1.6 Cox Ingersoll Ross model.

The Cox Ingersoll short rate model is defined as

$$dr_t = a(b - r_t)dt + \sigma \sqrt{r_t} dW_t,$$

where a is the rate at which the process mean reverts (a larger number results in a faster mean reverting process), b is the long run average interest rate. The model has been a benchmark for many years because of its analytical tractability and the fact that contrary to the Vasicek model the instantaneous short rate is always positive.

11.1.7 Hull and White model.

The Hull-White Short Rate Model is defined as

$$dr_t = a(b_t - r_t)dt + \sigma \sqrt{r_t} dW_t$$

11.1.8 General remarks.

When you chose model you should consider the following questions:

- Does the dynamics imply positive rates, $r_t > 0$ for each $t \in [0, T]$?
- What distribution does the dynamics imply for the short rate r_t (normal, fat tails)?
- Are bond prices $P(t, T) = \mathbb{E} \left[e^{-\int_t^T r_s ds} | \mathcal{F}_t \right]$ explicit computable from dynamics?
- Are option price explicit computable from the dynamics?
- Is the model mean-reverting?
- How suited is the model for Monte Carlo simulation?

11.2 Delta-hedging

The model requires the option seller to hedge his exposure by holding $\frac{\partial Q}{\partial S}$ a units of the stock, S , at any time. This quantity is known as the Delta of the option and the hedging strategy is known as Delta-hedging. Therefore, as the stock moves up and down, the option seller has to continuously change his holding to remain Delta-hedged. In the real world this is not practical, as it takes time to execute a trade. Thus it is impossible to re hedge truly continuously. As a consequence, the seller will never be perfectly hedged. Another problem is that executing a trade costs money: transaction costs may be low but they will never be non-existent. The more trades made, the more transaction costs mount up and increase the costs of hedging the option.

11.2.1 Primary market.

Definition 11.2.1 *A trading strategy (in the primary market) is collection of pairs of random variables*

$$\varphi = (\alpha_t, \beta_t)_{t=1}^T$$

where the random variable α_t represents the number of shares to be held over the time interval $(t-1, t]$ and the random variable β_t represents the number of the units of the bond to be held over the time interval $(t-1, t]$.

For simplicity we allow α_t, β_t to take any values in \mathbb{R} . In particular, we do not restrict to integer numbers of shares of stock or units of the bond, and we allow short selling of shares ($\alpha_t < 0$) and borrowing $\beta_t < 0$. To avoid strategies that anticipate the future, it is assumed that α_t, β_t are predictable (\mathcal{F}_{t-1}) measurable random variables for $t = 1, 2, \dots, T$. In this discrete model setting, this simply means that α_t and β_t can be expressed as real-valued functions of S_0, S_1, \dots, S_{t-1} . We will restrict attention to **self-financing** trading strategies, namely those trading strategies φ such that

$$\alpha_t S_t + \beta_t B_t = \alpha_{t+1} S_t + \beta_{t+1} B_t, \quad t = 1, \dots, T-1,$$

and the investor's initial wealth is equal to

$$\alpha_1 S_0 + \beta_1 B_0.$$

Definition 11.2.2 *We say that a trading strategy φ represents a portfolio whose value at time t is given by process $V(\varphi) = (V_t(\varphi))_{t=0}^T$, where*

$$V_0(\varphi) = \alpha_1 S_0 + \beta_1 B_0,$$

$$V_t(\varphi) = \alpha_1 S_t + \beta_1 B_t, \quad t = 1, \dots, T.$$

Definition 11.2.3 *An **arbitrage opportunity** (in the primary market) is a trading strategy φ such that*

1. $V_0(\varphi) = 0$,
2. $V_T(\varphi) \geq 0$,
3. $\mathbb{P}(\{V_T(\varphi) > 0\}) > 0$.

Definition 11.2.4 *Contingent claim for primary market is any \mathcal{F}_T -measurable random variable $X = f(S_T)$*

Definition 11.2.5 *A replicating (or hedging) strategy for European contingent claim is a trading strategy φ such that $V_T(\varphi) = X$.*

11.2.2 Single period case

What if we want to find the initial arbitrage free price of a European contingent claim. Existence of such a price depends on the existence of a so-called risk neutral probability, and uniqueness depends on there being a replication strategy for the contingent claim.

We begin by examining the single period case. Let $T=1$. For this case, we first show that there is a replicating strategy for any European contingent claim X . Given X we seek a trading strategy $\varphi = (\alpha_1, \beta_1)$, where α_1 and β_1 are constants such that

$$V_1(\varphi) = \alpha_1 S_1 + \beta_1 B_1 = X. \quad (11.5)$$

Now S_1 has two possible values, $S_0 u, S_0 d$. X is a function of S_1 , since it is \mathcal{F}_1 -measurable ($\mathcal{F}_1 = \sigma(S_0, S_1)$). Let X^u denote the value of X when $S_1 = S_0 u$ and X^d denote the value of X when $S_1 = S_0 d$. Then considering these two possible outcomes, (11.5) yields two equations for the two deterministic unknowns α_1, β_1 :

$$\alpha_1 S_0 u + \beta_1 (1+r) = X^u \quad (11.6)$$

$$\alpha_1 S_0 d + \beta_1 (1+r) = X^d. \quad (11.7)$$

Solving for α_1, β_1 yields

$$\alpha_1 = \frac{X^u - X^d}{(u-d)S_0} \quad (11.8)$$

$$\beta_1 = \frac{1}{1+r} \frac{uX^d - dX^u}{u-d}. \quad (11.9)$$

The initial wealth needed to finance this strategy (manufacturing cost of the contingent claim) is

$$\begin{aligned} V_0 &= \alpha_1 S_0 + \beta_1 B_0 = \\ &= \frac{1}{(1+r)(u-d)} ((1+r-d)X^u + (u-(1+r))X^d) \\ &= \frac{1}{1+r} (p^* X^u + (1-p^*) X^d) = \mathbb{E}^{p^*}[X^*], \end{aligned}$$

where $X^* = \frac{X}{(1+r)}$, $p^* = \frac{1+r-d}{u-d}$. Note that $p^* \in (0, 1)$ and so, just as for p , each of the two possible outcomes for (S_0, S_1) has positive probability under p^* . Furthermore discounted process $S^* = (S_0, \frac{1}{(1+r)S_1})$ is martingale under p^* relative to the filtration $(\mathcal{F}_t)_{t=0}^1$. Thus, under p^* the average rate of return of the risky asset is the same as that of the riskless asset. For this reason p^* is called the **risk neutral probability**.

It's important to realize that computing expectations under p^* is a mathematical device. We are not assuming the the stock price actually moves according to this probability. That is p^* may be unrelated to the subjective probability p that we associate with the binomial model for movements in the stock price.

We hedge or replicate financial derivatives using underlying assets. That is a combination of shares and bonds that is worth exactly same value as the derivative, regardless of what happens to the stock price in the future.

11.2.3 Fundamental theorem.

Definition 11.2.6 Set $K \subset L^0(\mathbb{P})$ defined by

$$K := \{(\varphi^* S^*)_T : \varphi^* \text{ jest strategi\u0105 o warto\u015bciach w } \mathbb{R}^d\}$$

is called the set of contingent claims attainable at price 0.

Definition 11.2.7 We call the convex cone $C \subset L^\infty(\mathbb{P})$ defined by

$$C := \{g \in L^\infty(\mathbb{P}) : \text{istnieje } f \in K \text{ takie że } f \geq g\}.$$

the set of contingent claims super-repliable at price 0.

Definition 11.2.8 Financial market satisfies the no-arbitrage condition if

$$K \cap L_+^0(\mathbb{P}) = \{0\}.$$

Economically speaking, the no arbitrage condition demands that any terminal payoff that precludes loss and is attainable at zero initial investment should not allow any chance of making money no matter how small.

Theorem 11.2.1 (Fundamental theorem of asset pricing.)

For a financial market $S^* = (S_t)_{t=1}^T$ modelled on a finite filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$ the following statements are equivalent:

- (i) process S^* satisfies the no-arbitrage condition.
- (ii) $\mathcal{M}^e(S^*) \neq \emptyset$,

where $\mathcal{M}^e(S^*)$ is set of all equivalent martingale measures.

11.3 Bonds.

11.3.1 Price of bonds

Definition 11.3.1 Price of the bond, P , is defined as sum of all future coupons C_i and principal A returned at the end, suitably discounted

$$P = \sum_{i=1}^n \frac{c_i}{(1+y_n)^i} + \frac{A}{(1+y_n)^n}$$

Remark 11.3.1 Consider semi-annual coupon bond, with face value F , coupon rate C , and maturity T , pays the holder of the bond a coupon payment equal to $\frac{C}{2}F$ every six months, except at maturity.

Definition 11.3.2 Let B be the value of a bond with future cash flows c_i to be paid to the holder of the bond at times t_i , $i = 1, \dots, n$. Let $r(0, t_i)$ be the continuously compounded zero rates corresponding to t_i , $i = 1, \dots, n$. Then

$$B = \sum_{i=1}^n c_i e^{-r(0, t_i)t_i}.$$

Remark 11.3.2 Note that $e^{-r(0, t_i)t_i}$ is the discount factor corresponding to time t_i , $i = 1, \dots, n$.

Definition 11.3.3 The yield of a bond is internal rate of return of the bond, i.e., the constant rate at which the sum of discounted future cash flows of the bond is equal to the price of the bond. If B is the price of a bond with cash flows c_i at time t_i , $i = 1, \dots, n$. and if y is the yield of the bond, then

$$B = \sum_{i=1}^n c_i e^{-yt_i}. \quad (11.10)$$

Remark 11.3.3 As expressed above, the price of the bond B can be regarded as a function of the yield. Therefore, whenever needed, we may think of B as being a function of the yield, i.e., $B = B(y)$

Remark 11.3.4 *It is easy to see that the price of the bond goes down if the yield goes up, and it goes up if the yield goes down.*

Remark 11.3.5 *To compute the yield of bond with a known price B , we must solve (11.10) for y . this can be written as equation in y , i.e.,*

$$f(y) = 0, \quad \text{where} \quad f(y) = \sum_{i=1}^n c_i e^{-yt_i} - B.$$

which is then solved numerically.

Definition 11.3.4 (Par yield) *Par yield is the coupon rate that makes the value of the bond equal to its face value.*

In other words par, yield is the value C of the coupon rate such that $B = F$. For the semi-annual coupon bond we substitute $B = F$. Then, the par yield of the bond can be obtained by solving the following linear equation for C

$$1 = \sum_{i=1}^{n-1} \frac{C}{2} e^{-r(0,t_i)t_i} + (1 + \frac{C}{2}) e^{-r(0,T)T}. \quad (11.11)$$

11.3.2 Duration and convexity.

Duration and convexity are two of the most important parameters to estimate when investing in a bond, other than its yield.

Duration provides the sensitivity of the bond price with respect to small changes in the yield, while convexity distinguishes between two bond portfolios with the same duration. (the portfolio with higher convexity is more desirable.)

Duration of a bond is the weighted time average of the future cash flows of the bond discounted with respect to the yield of the bond, and normalized by dividing by the price of the bond.

Definition 11.3.5 *The duration D of the bond with price B and yield y , with cash flows c_i at time t_i , $i = 1, \dots, n$, is*

$$D = \frac{\sum_{i=1}^n t_i c_i e^{-yt_i}}{B}. \quad (11.12)$$

From (11.10) and (11.12) it is easy to see that

$$\frac{\partial B}{\partial y} = - \sum_{i=1}^n t_i c_i e^{-yt_i} = -BD,$$

and therefore

$$D = -\frac{1}{B} \frac{\partial B}{\partial y}. \quad (11.13)$$

The duration of bond gives the relative change in the price of a bond for small changes Δy . Duration of bond which is a measure of how long on average the bond owner needs to wait to receive cash flows.

Remark 11.3.6 *The discretized version of (11.13) is*

$$D \approx -\frac{1}{B} \frac{B(y + \Delta y) - B(y)}{\Delta y} = -\frac{\Delta B}{B \Delta y}$$

which is equivalent to

$$\frac{\Delta B}{B} \approx -\Delta y D. \quad (11.14)$$

In other words, the percentage change in the price of the bond can be approximated by the duration of the bond multiplied by the parallel shift in the yield curve with opposite sign.

Remark 11.3.7 • Zero-Coupon Bond – Duration is equal to its time to maturity

- Vanilla Bond - Duration will always be less than its time to maturity.

11.3.3 Convexity.

For very small parallel shift in the yield curve, the approximation formula (11.14) is accurate. For larger parallel shifts, convexity is used to better capture the effect of the changes in the yield curve on the price of the bond.

11.4 Hypothesis testing an overview.

Hypothesis testing allows us to carry out inferences about population parameters using data from a sample. In order to test a hypothesis in statistics, we must perform the following steps:

1. Formulate a null hypothesis and an alternative hypothesis on population parameters.
2. Build a statistic to test the hypothesis made.
3. Create critical region
4. Make a decision to reject or not to reject the null hypothesis.

11.5 Assumptions of model

Regression is one of the most widely used of all statistical methods. The goals of regression modeling include the investigation of how Y is related to X_1, \dots, X_k , estimation conditional expectation of Y given X_1, \dots, X_k and prediction of future Y values when the corresponding values of X_1, \dots, X_k are already available. These goals are closely connected.

- t -number of observation
- T -number of observations
- k -number of predictors
- y_t - t -th observation of response variable
- x_{ti} - t -th observation of k -th response
- ε_t noise or error
- X' -transpose of matrix X

The multiple linear regression model relating Y to the predictor variables is

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t, \quad t = 1, 2, \dots, T$$

Using matrix notation

$$\mathbf{y}_{[T \times 1]} = \mathbf{X}_{[T \times k]} \boldsymbol{\beta}_{[k \times 1]} + \boldsymbol{\varepsilon}_{[T \times 1]} \quad (11.15)$$

In this paragraph vector means column.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}$$

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \ddots & \vdots & \vdots \\ x_{T1} & x_{T2} & \dots & x_{Tk} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_T \end{bmatrix}$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

Assumptions are:

1. Matrix X is known and not random
2. $r(X) = k$ rank of matrix
3. $E(\boldsymbol{\varepsilon}) = \mathbf{0}_{[T \times 1]}$
4. $V(\boldsymbol{\varepsilon}) = \sigma^2 I_T$, $\sigma^2 > 0$

11.6 Classic model of linear regression

Definition 11.6.1 A vector of regression coefficient defined as

$$\hat{\boldsymbol{\beta}} := (X^T X)^{-1} X^T \mathbf{y}$$

can be estimated by the method of least squares. Geometrically we are minimizing the sum of the squared lengths of the vertical lines. (here we may put some figure in the future).

Theorem 11.6.1 In Classic model of linear regression the most effective estimator of vector $\boldsymbol{\beta}$ in a class of all linear and unbiased estimators is estimator $\hat{\boldsymbol{\beta}}$ is estimator defined above. Covariance matrix of vector $\hat{\boldsymbol{\beta}}$ is equal

$$V(\hat{\boldsymbol{\beta}}) = \sigma^2 (X^T X)^{-1}.$$

Proof. Later

□

Definition 11.6.2 Let $\hat{\boldsymbol{\beta}}$ is a vector of parameters defined as above. Then vector defined as $\hat{\mathbf{y}} = X \hat{\boldsymbol{\beta}}$ is called theoretical value vector.

Theorem 11.6.2 *Let*

$$SSE := \sum_{t=1}^T (Y_t - \bar{Y})^2 = \varepsilon' \varepsilon = \sum_{t=1}^T \varepsilon_t^2.$$

The estimate for σ^2 is the estimate

$$s^2 = \frac{1}{T-k} \sum_{t=1}^T (Y_t - \hat{Y}_t)^2$$

Theorem 11.6.3 *Unbiased estimate of covariance matrix in linear regression model is matrix*

$$\hat{V}(\hat{\beta}) := s^2 (X^T X)^{-1}$$

Definition 11.6.3 *Standard errors are squared elements of diagonal covariance matrix $\hat{V}(\hat{\beta})$. They inform how we are wrong if we change unknown parameter β_i by $\hat{\beta}_i$*

Interpretation of vector of parameters β .

β_1 represents the difference in the predicted value of Y for each one-unit difference in X_1 , if X_2, \dots, X_k remains constant.

11.7 Analysis of Variance

Definition 11.7.1 (SSE) *The total variation in Y can be partitioned into two parts: the variation that can be predicted by X_1, \dots, X_k and the variation that cannot be predicted. The variation that can be predicted is measured by the regression sum of squares SSR which is*

$$SSR := \sum_{t=1}^T (\hat{Y}_t - \bar{Y})^2$$

Definition 11.7.2 (SSE) *The amount of variation in Y that cannot be predicted by a linear function of X_1, \dots, X_k is measured by the residual error sum of squares, which is the sum of squared residuals; i.e.,*

$$SSE := \sum_{t=1}^T (Y_t - \hat{Y}_t)^2 \quad (11.16)$$

Definition 11.7.3 (SST) *The total variation is measured by the total sum of squares which is the sum of the squared deviations of Y from its mean; that is*

$$SST := \sum_{t=1}^T (Y_t - \bar{Y})^2$$

It can be shown algebraically following theorem

Theorem 11.7.1 *Let SSE and SSR will be defined as above. Then*

$$SST = SSR + SSE$$

It is useful to measure the *strength* of linear relationship. Such a measure is provided by the **coefficient of determination** R^2 . In R it is called R -squared.

Definition 11.7.4 *R -squared, denoted by R^2 is*

$$R^2 := \frac{SSR}{SST} = 1 - \frac{SSE}{SST}.$$

It measures the proportion of the total variation of Y that can be linearly predicted by X . $R^2 \in [0, 1]$. The closer R^2 is to 1, the stronger relationship between X and Y

11.8 Confidence interval.

We put one more assumption. Now additional we assume that vector $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T]'$ has multi-normal distribution. Hence $\boldsymbol{\varepsilon} \sim \mathcal{N}_T(\mathbf{0}_{[T \times 1]}, \sigma^2 \mathbf{I}_T)$

Theorem 11.8.1 *In classical normal linear regression estimator has normal with mean $\boldsymbol{\beta}$ and covariance matrix $\sigma^2(X'X)^{-1}$. Then*

$$\hat{\boldsymbol{\beta}} \sim \mathcal{N}_k(\boldsymbol{\beta}, \sigma^2(X'X)^{-1})$$

For fixed i in CNLRM (classical normal linear regression model) we have

$$\frac{\hat{\beta}_i - \beta_i}{\hat{D}(\hat{\beta}_i)} \sim t_{T-k}$$

Confidence interval for one parameter. Fixed confidence level $1 - \alpha$. Confidence level for coefficient β_i is equal

$$[\beta_i - t_{T-k}(1 - \frac{\alpha}{2})\hat{D}(\hat{\beta}_i); \beta_i + t_{T-k}(1 - \frac{\alpha}{2})\hat{D}(\hat{\beta}_i)]$$

11.9 Hypothesis testing.

- **One coefficient**

Hypothesis $H_0 : \beta_i = 0$

$H_1 : \beta_i \neq 0$

Test statistic:

$$\frac{\hat{\beta}_i}{\hat{D}(\hat{\beta}_i)} \sim t_{T-k}$$

- **Joint testing for all coefficient**

Let

$$y_t = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t, \quad t = 1, 2, \dots, T$$

We assume that there is an intercept in our model.

Hypothesis:

$H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0$ (all coefficients are zero, there are no linear impact on response variable)

$H_1 : \beta_2 \neq 0 \vee \beta_3 \neq 0 \vee \dots \vee \beta_k \neq 0$ (at least one coefficient has linear impact on response variable) Statistic

$$F = \frac{\frac{R^2}{k-1}}{\frac{1-R^2}{T-k}} \sim F_{(k-1, T-k)}$$

One side critical region always.

11.10 Another tests

Jarque and Bera test

H_0 : sample comes from normal distribution. Here $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$

H_1 : another distribution.

$$JB = T \cdot \left(\frac{skewness^2}{4} + \frac{(kurtosis)^2}{24} \right)$$

Kurtosis

$$kurtosis = \frac{\mu_4}{(\sigma^2)^2}$$

where μ_4 fourth central moment.

Durbin and Watson test

In statistics, the Durbin–Watson statistic is a test statistic used to detect the presence of autocorrelation (a relationship between values separated from each other by a given time lag) in the residuals (prediction errors) from a regression analysis. A similar assessment can be also carried out with the Breusch–Godfrey test and the Ljung–Box test. Assumptions:

1. $\varepsilon_t = \rho_1 \varepsilon_{t-1} + \eta_t$, $|\rho_1| < 1$, $\eta_t \sim i.i.d N(0, \sigma_\eta^2)$
2. there is an interceptor in model

It is worth to emphasis, that, DW is not appropriate for testing for serial correlation of residuals in autoregressive (AR) models. The CFA curriculum doesn't specify exactly why DW is inappropriate for AR models; suffice it to say that it isn't. That's a good point to remember.

Tests

$H_0 : \rho = 0$ (there is no autocorrelation)

$H_1 : \rho > 0$ if $\hat{\rho} > 0$ (residuals create autoregressive process order 1)

$H_1 : \rho < 0$ if $\hat{\rho} < 0$ (residuals create autoregressive process order 1)

where,

$$\hat{\rho} = \frac{\sum_{t=2}^T \varepsilon_t \varepsilon_{t-1}}{\sum_{t=1}^T \varepsilon_t^2}$$

it is estimator first-order serial correlation coefficient

Statistic

$$d = \frac{\sum_{t=2}^T (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^T \varepsilon_t^2}$$

11.11 How to measure a risk.

11.11.1 Introduction.

Measures such as delta, gamma, vega for describing different aspects of the risk in a portfolio of derivatives. A financial institution usually calculates each of these measures each day for every market variables. A delta-gamma-vega analysis, therefore, leads to a very large number of different risk measures being produced each day.

Value at risk (VaR) is an attempt to provide a single number summarizing the total risk in a portfolio of financial assets.

11.11.2 Definition and interpretation

Value-at-risk is defined as the loss level that will not be exceeded with a certain confidence level during a certain period of time. For example, if a bank's 10-day 99 *VAR* is 3 million, there is considered to be only a 1 chance that losses will exceed 3 million in 10 days. If we are going to use one measure to describe the risk in a particular situation, is *VAR* the best choice? One problem with *VAR* is that, when used in an attempt to limit the risks taken by a trader, it can lead to undesirable results.

Suppose a bank tells a trader that the one-day 99 *VAR* of the trader's portfolio must be kept at less than 10 million. There is a danger that the trader will construct a portfolio where there is a 99

chance that the daily loss is less than 10 million and a 1 chance that it is 500 million. The trader is satisfying the risk limits imposed by the bank, but is clearly taking unacceptable risks. Most traders would, of course, not behave in this way - but some might.

A measure that produces better incentives for traders than VAR is expected shortfall. This is also sometimes referred to as conditional VAR , or tail loss. Where VAR asks the question 'how bad can things get?', expected shortfall asks 'if things do get bad, what is our expected loss?'.

Expected shortfall, like VAR , is a function of two parameters: N (the time horizon in days) and α (the confidence level). It is the expected loss during an N -day period, conditional that the loss is greater than the α percentile of the loss distribution. For example, with $\alpha = 99$ and $N = 10$, the expected shortfall is the average amount that is lost over a 10-day period, assuming that the loss is greater than the 99-th percentile of the loss distribution.

Definition 11.11.1 *Given a loss L and the confidence level $\alpha \in (0, 1)$. $VaR_\alpha(L)$ is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $1 - \alpha$ i.e (a) i)*

$$\begin{aligned} VaR_\alpha(L) &= \inf \{l \in \mathbb{R} : \mathbb{P}(L > l) \leq 1 - \alpha\} \\ &= \inf \{l \in \mathbb{R} : 1 - \mathbb{P}(L \leq l) \leq 1 - \alpha\} = \inf \{l \in \mathbb{R} : 1 - F_L(l) \leq 1 - \alpha\} \\ &= \inf \{l \in \mathbb{R} : F_L(l) \geq \alpha\} \end{aligned}$$

We might think of L as the (potential) loss resulting from holding a portfolio over some fixed time horizon. In market risk the time is typically one day or ten days. In credit risk the portfolio may consist of loans and the time horizon is often one year.

Remark 11.11.1 *Fix $\alpha = 0.05$. Then about 5% of time we expect the return to be worse than -1.64.*

Remark 11.11.2 $VaR_\alpha(L) = q_\alpha(F)$, where F is a loss distribution.

Remark 11.11.3 *Suppose the loss distribution is normal $L \sim N(\mu, \sigma^2)$. This means*

$$L \stackrel{d}{=} \mu + \sigma L', \quad L' \sim N(0, 1)$$

Then

$$VaR_\alpha(L) = \mu + \sigma \Phi^{-1}(\alpha)$$

Remark 11.11.4 *Value at risk does not give any information about how bad losses may be when things go wrong. In other words, what is the size of an average loss given that the loss exceeds the 95%.*

Let $t = 1, \dots, T$ be individual time periods. When the portfolio is linear and created by N assets, $R_{j,t}$ is return of j -th asset in the time period t and $w_{j,t}$ is current weight of the j -th asset in this portfolio, then the return of the portfolio, $R_{p,j}$ can be calculated as

$$R_{p,j} = \sum_{j=1}^N w_{j,t} R_{j,t}.$$

To calculate historical value at risk we must

1. Rank values calculated

$$R_{p,(j)}, \dots, R_{p,(T)}$$

2. For fixed α we must estimate the empirical quantile q_α

$$q_\alpha =$$

3. Calculate VAR as

$$VaR_\alpha = q_\alpha P_0$$

where P_0 is the initial portfolio value.

11.11.3 Value at risk method.

There is a question which of these approaches or methods is the best for calculating Value at Risk. Unfortunately there is no easy answer. The various methods differ in complex risk measurement capabilities of market factors, in difficulty of implementation. The choice of method depends on the parameters which the risk manager considers more important.

Historical simulation.

Variance-Covariance method

Using the properties of the multivariate normal distribution...

$$VaR_{\alpha}(L) =$$

Analytic solutions can be obtained: no simulations required. Easy to implement. Parametric approach to calculating VaR is used at the academy rather than in practice.

Monte Carlo method

11.11.4 Advantages and disadvantages of VaR.

Advantages:

- VaR is easy to understand - VaR is measured in price units (Dollars, Euros) or as a percentage of portfolio value. This makes VaR very easy to interpret and also to further use in analyses.
- Comparing VaR of different assets and portfolios - We can measure and compare VaR of different types of assets and various portfolios (stocks, bonds, currencies, derivatives, or any other assets with price).
- Risk factors always change over time. Sensitivity measures and scenario analysis don't take into account differences in volatilities across markets, correlations across risk factors, as well as the probability of adverse moves in the risk factors.

Disadvantages:

- VaR gets difficult to calculate with large portfolios - When you are calculating VaR of a portfolio, we need to measure or estimate not only the return of individual assets, but also the correlations between them. With growing number and diversity of positions in the portfolio, the difficulty of this task grows exponentially.
- VaR is not additive - Correlations between individual risk factors enter the VaR calculation is also the reason why VaR is not simply additive. The portfolio VaR containing assets A and B does not equal to the sum of asset A VaR and asset B VaR.
- Different VaR methods lead to different results - Different approaches (Historical simulation, Analytical simulation, MC simulation,...) can also lead to very different results with the same portfolio, so the representativeness of VaR can be questioned.

11.11.5 Expected shortfall.

Definition 11.11.2 For a loss L with continuous loss distribution function F_L the **expected shortfall** at confidence level $\alpha \in (0, 1)$ is given by

$$\mathbb{E}[S_{\alpha}(L)] := \mathbb{E}[L | L \geq VaR_{\alpha}(L)]$$

11.12 Back testing of VaR.

One of VaR validation methods is backtesting, which verifies the actual losses are in line with the VaR measurement. The first concern of backtesting is the number of exceptions. An exception occurs if its losses exceed the VaR number. Therefore, the expected number of exceptions N in a total of T observations is $T(1 - c)$. Certainly, the number of exceptions will not exactly be $T(1 - c)$. Instead it could swing in a small range.

11.13 Exchange and Clearing House

The exchange specifies the conditions for the trading. It defines - among other things which contracts are traded and their specifications.

The settlement of the deals is carried out by the clearing house of exchange. The clearing house has the following main functions:

- It is counterparty for both the buyer and the seller in all traded contracts. Placing the clearing house between buyer and seller reduces the credit risk. To reduce this risk to a minimum, the clearing house deals solely with registered clearing members who for their part offer the services as brokers or clearers. In order to protect against default risk of exchange members, so called initial margins and variation margins are calculated.
- Daily revaluation and accounting of variation margins for all open deals
- Fixing of the initial margin. The initial margin depends on the market's volatility. Therefore it is adjusted regularly to the actual market conditions

11.14 The Margin System

11.15 Interest rate swaps

11.15.1 Idea and definition.

Generally speaking, a **swap contract** (or a **swap**) is an agreement between two parties to exchange cash flows at some future dates according to a prearranged formula. The two most popular kinds of swap agreements are standard **interest rate swaps** and **cross-currency swaps** (known also as **differential swaps**).

In a **plain vanilla interest rate swap**, one party, say A, agrees to pay to the other party, say B, amounts determined by a fixed interest rate on a notional principal at each of the payment dates. At the same time, the party B agrees to pay to the party A interest at a floating reference rate on the same notional principal for the same period of time. Thus an interest rate swap can be used to transform a floating-rate loan into a fixed-rate loan or vice versa. In essence, a swap is a long position in a fixed-rate coupon bond combined with short positions in floating-rate notes (alternatively, it can be seen as a portfolio of specific forward contracts).

Remark 11.15.1 *The notional principal amount, in an interest rate swap, is the predetermined dollar amounts on which the exchanged interest payments are based. The notional principal never changes hands in the transaction, which is why it is considered notional, or theoretical.*

Remark 11.15.2 *Swaps are now so liquid and exist for enormous range of maturities that their price determine the yield curve and not vice versa. The swaps market is big. The notional principal amount in the US dollars, currently comfortably in 14 numbers.*

11.15.2 Pricing interest rate swaps (IRS).

Interest rate swaps IRS are OTC contracts in which, in general, two counter-parties agree to exchange two streams of cash flows in the same currency, typically tied to a floating Libor rate $L_x(T_{i-1}, T_i)$ versus a fixed rate K . These payment streams are called fixed and floating leg of the swap, respectively.

The IRS is characterised by the following schedule

- $\mathbf{T} = \{T_0, T_1, \dots, T_n\}$, floating leg schedule,
- $\mathbf{S} = \{S_0, S_1, \dots, S_m\}$, fixed leg schedule

with $T_0 = S_0$, $T_n = S_m$. Coupon payoffs are

$$IRS_{float}(T_j; T_j, T_j, L) = NL(T_{j-1}, T_j) \delta_L(T_{j-1}, T_j), j = 1, \dots, n,$$

$$IRS_{fix}(S_i; S_i, S_j, K) = NK \delta_K(S_{i-1}, S_i), i = 1, \dots, m$$

At the day t , the present value of coupons at payments dates of T_i are:

Floating:

$$IRS_{float}(t, T_j; T_j, T_j, L) = NP(t; T_j) F(t; T_{j-1}, T_j) \delta_L(T_{j-1}, T_j).$$

Fixed:

$$IRS_{fix}(t, S_i; S_i, S_j, K) = NKP(t; T_j) \delta_K(S_{i-1}, S_i).$$

At $T_0 = S_0$ valuation

11.15.3 Overnight indexed swap.

An overnight indexed swap is an agreement between two counterparties to exchange at each payment date or at maturity the difference between fixed rate and floating rate on the nominal amount. The periodic floating rate is equal to the geometric average of an overnight rate over every day of the payment period. For the EUR market the fixing is named EONIA for USD we have Fed Funds.

11.15.4 Interest basis swap.

Interest rate basis swaps (IRBS) are OTC contract in which two counterparties agree to exchange two streams of cash flows in the same currency, ties to two floating Libor rates with different tenors x and y . There are two ways of building (IRBS) instruments depending on how the two floating legs are packed together: as a portfolio of two fixed vs floating IRS, or a single IRS floating vs floating plus spread.

11.16 Default

Definition 11.16.1 *Default occurs when a debtor is unable to meet the legal obligation of debt repayment.*

Banks and other financial institutions set out loans to obligors. These loans are subject to credit risk, i.e. risk induced by the fact that some of the obligors may not (fully) meet their financial obligations, for example by not repaying the loan. When this happens, we say that the obligor defaults, or is in default.

In order to prevent the event the bank needs some sort of insurance. The basic idea behind insurance is always the same. For example, in health insurance the costs of a few sick customers are covered by the total sum of revenues from the fees paid to the insurance company by all customers.

For bank loans one can argue exactly the same way. Banks charge an appropriate risk premium

for every loan and collecting these risk premiums in an internal bank account called expected loss reserve. In this way bank creates a capital cushion for covering losses arising from defaulted loans. The basic idea is as follows: The bank assigns to every customer:

1. a **default probability (DP)**
2. **exposure at default (EAD)** the amount of money subject to be lost in the considered time period.
3. a loss fraction called the **loss given default (LGD)**, describing the fraction of the loan's exposure expected to be lost in case of default

The loss of any obligor is then defined by a loss variable

$$L := EAD \times LGD \times \mathbf{1}_D, \quad \mathbf{1}_D \sim \text{Bin}(n, PD)$$

Next, consider a portfolio of n obligors, each bearing its own risk. We attach a subscript i to the variables, indicating that they belong to obligor i in the portfolio. We can then define the portfolio loss as

$$\tilde{L}_n = \sum_{i=1}^n EAD_i \times LGD_i \times \mathbf{1}_{D_i},$$

which is just the sum over all individual loss variables.

The distribution of \tilde{L}_n is of course of great interest, since it contains all the information about the credit risk. Several characteristics of this distribution have special names. Some of them are

- Expected Loss (EL) defined as $EL := \mathbb{E}[\tilde{L}_n]$
- Unexpected Loss (UL), defined as the standard deviation of the portfolio loss $UL := \sqrt{\text{Var}[\tilde{L}_n]}$
- Value-at-Risk
- Economic Capital, defined as $EC_\alpha = VaR_\alpha - EL$

What makes the analysis of the distribution of \tilde{L}_n interesting, is that variables $\mathbf{1}_{D_i}$ are usually correlated. This means, for example, when they are positively correlated, knowing that one of the obligors defaults, there is a higher probability of other obligors defaulting as well.

11.17 Spreads and default risk.

11.17.1 Credit default swap.

Definition 11.17.1 *Credit default swaps shifts the risk onto an insurance company or other CDS seller in exchange for certain premium.*

Well simply put it is a risk transfer.

One of the ways that financial institution can make money is by issuing loans and credit. Imagine that we manage pension fund of teachers. And there is a company which wants to take a loan from us. But we have in our status that we may invest only in very very safe instruments. We can only invest in triple AAA companies. Such company has a very low chance of default. But assume that is a company which wants to raise money. Well the company issues out a debt in form of bonds. Of course if we decide to buy these bonds we are expecting two things: principal and interest. But there is a problem because this company for instance (SONY) is only double BB. So what we can do?? This guy needs money I have the money but actually we can buy bonds which the company

issue because of to low rating which Sony has. And this is where credit default swaps come in. Insurance company or a bank guarantee that Sony not will default or if the do the highly rated company will come and pay the full value of the bond. In return for covering the risk of this 10 years bonds the investment manager mus pay for insurance company 10 dollars for every year the bond is good. People in financial world don't call CDS insurance so that CDS could be exempt from insurance law regulation.

11.18 Credit valuation adjustment (CVA)

It is the difference between the risk-free portfolio value and the true portfolio value that takes into account the possibility of a counterpart's default. In other words, CVA is the market value of counterpart credit risk. This price depends on counterparty credit spreads as well as on the market risk factors that drive derivatives values.

$$CVA = \mathbb{E}^{\mathbb{Q}}[L^*] = (1 - R) \int_0^T \mathbb{E}^{\mathbb{Q}} \left[\frac{B_0}{B_t} E(t) | \tau = t \right]$$

where T is the maturity of the longest transaction in the portfolio, B_t is the future value of one unit of the base currency invested today at the prevailing interest rate for maturity t , R is the fraction of the portfolio value that can be recovered in case of a default, τ is the time of default, $E(t)$ is the exposure at time t , and $PD(s, t)$ is the risk neutral probability of counterparty default between times s and t . These probabilities can be obtained from the term structure of credit default swap (CDS) spreads.

11.19 Estimation probability of default

There are several alternatives for estimating the probability of default. Default probabilities may be estimated from a historical data base of actual defaults using modern techniques like logistic regression. Default probabilities may also be estimated from the observable prices of credit default swaps, bonds, and options on common stock. The simplest approach, taken by many banks, is to use external ratings agencies such as Standard and Poors, Fitch or Moody's. For small business default probability estimation, logistic regression is again the most common technique. These models are both developed internally and supplied by third parties.

A binary response Y can take only two values 0, and 1 which code two possible outcomes for example, that a company goes into default or it does not. Binary regression models of predictors $X_{i,1}, \dots, X_{i,p}$. Since a probability is constrained to lie between 0 and 1, a linear model is not appropriate for a binary response. However, linear models are so convenient that one would like a model that has many of the features of linear model. This has motivated the development of generalized linear models (GLMs). Generalized linear models for binary responses are of the form

$$\mathbb{P}(Y_i = 1 | X_{i,1}, \dots, X_{i,p}) = H(\mathbf{x}_i^T \boldsymbol{\beta}).$$

The most common generalized linear models for binary responses are probit regression where $H(x) = \Phi(x)$ and logistic regression where

$$H(x) = \frac{1}{(1 + e^{-x})}$$

The parameter vector $\boldsymbol{\beta}$ can b estimated by maximum likelihood. Assume that conditional on $\mathbf{x}_1, \dots, \mathbf{x}_n$ the binary response Y_1, \dots, Y_n are mutually independent. Then the likelihood (conditional on $\mathbf{x}_1, \dots, \mathbf{x}_n$) is

$$\prod_{i=1}^n H(\mathbf{x}_i^T \boldsymbol{\beta})^{Y_i} (1 - H(\mathbf{x}_i^T \boldsymbol{\beta}))^{1-Y_i}.$$

The following variables are included in the data set

1. Yearly income
2. does the individual own home
3. is the individual self-employment
4. Ratio of monthly credit card expenditure to yearly income.

11.20 Merton's approach to corporate debt.

Let's denote L promised deterministic payoff which comes from liability. Let us recall the most important assumptions

- trading takes place continuously in time,
- an unrestricted borrowing and lending of funds is possible at the same interest rate,
- all traded assets are infinitely divisible,
- no restrictions on the short-selling of traded securities are present,
- the bankruptcy and/or reorganization costs in case of default are negligible,
- there is no transaction costs and taxes.

Consider $B(t, T)$ the price of a T -maturity zero-coupon bond equals

$$B(t, T) =$$

Denote by $E(V_t)$ and $D(V_t)$ the value of the firm's equity and debt respectively. The total value of firm's assets satisfies

$$V_t = E(V_t) + D(V_t)$$

We postulate that the firm's value process V follows a geometric Brownian motion under the spot martingale measure \mathbf{P}^*

$$dV_t = V_t((r - \kappa)dt + \sigma_V dW_t^*),$$

where κ reflects an inflow of capital to the firm. We postulate that default may only occur at the debt's maturity date T . Specifically, if at the maturity T the total value V_T of the firm's assets is less than the notional value L of the firm's debt, the firm defaults and the bondholders receive the amount V_T . The fixed amount L may be interpreted as the face value (or par value) of a corporate zero-coupon bond maturing at time T . Formally, the value of the firm's debt at time t thus equals

$$D(V_t) = LB(t, T) - P_t$$

where P_t is the price of the **put-to-default**.

11.21 Basic.

Let $t \in \mathbb{Z}$.

Definition 11.21.1 Let us consider probability space $(\Omega, \mathcal{F}, \mathbb{P})$. An increasing family of sigma fields that is $\mathcal{F}_u \subset \mathcal{F}_t$ for any $u, t \in \mathbb{Z}$ is called filtration and we shall write $(\mathcal{F}_t)_{t \in \mathbb{Z}}$.

Definition 11.21.2 Set $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{Z}}, \mathbb{P})$, where $(\mathcal{F}_t)_{t \in \mathbb{Z}}$ is filtration defined above is called filtered probability space.

Definition 11.21.3 Given a positive integer d , a d -dimensional stochastic process on the given probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{Z}}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t \in \mathbb{Z}}$ is a collection $(X_t) = (X_t)_{t \in \mathbb{Z}}$, where each X_t is a d -dimensional random vector; i.e., a function $X_t: \Omega \rightarrow \mathbb{R}^n$ such that $X_t^{-1}(B) = \{\omega \in \Omega: X_t(\omega) \in B\} \in \mathcal{F}$ for each Borel set B of \mathbb{R}^n

Definition 11.21.4 Time series

Let (X_t) be stochastic process and $T \in \mathbb{Z}$ and $T < \infty$. We define time series as the realisations of random variables X_1, \dots, X_T . Then we write $(X_t)_{t=1}^T$.

As we can see time series can be considered a sample from, the stochastic process. The mean and the variance of random variables have a special place in the theory of statistics. In time series analysis, the analogues of these are the mean function and the autocovariance function.

Definition 11.21.5 Mean function

The mean function $\mu(t)$ of a stochastic process X_t is defined as $\mu(t) := \mathbb{E}[X_t]$

In general $\mu(t)$ depends on time t , as, for example, processes with a seasonal or periodical structure or processes with a deterministic trend.

Definition 11.21.6 Autocovariance function

The autocovariance function $\gamma(s, t)$ of a stochastic process X_t is defined as

$$\gamma(s, t) := \text{cov}(X_s, X_t) = \mathbb{E}[(X_s - \mu(s))(X_t - \mu(t))]$$

The autocovariance function is symmetric, i.e. $\gamma(t, u) = \gamma(t - u, -u)$. For the special case $u = 0$ the result is the variance function $\gamma(t, 0) = \text{var}(X_t)$. In general $\gamma(t, u)$ is dependent on t as well as on u .

Stochastic stationary processes are probability models for time series with time invariant behaviour.

Definition 11.21.7 Strict stationarity

Process X_t is said to be strictly stationary if for any $k > 0$ any $t_1, \dots, t_k \in \mathbb{Z}$ distribution

$$X_{t_1}, \dots, X_{t_k}$$

is the same as distribution

$$X_{t_1+u}, \dots, X_{t_k+u}$$

for every value of u

This is a very strong condition that is hard to verify empirically. A weaker version of stationarity is often assumed.

Definition 11.21.8 Weak stationarity.

Process X_t is said to be strictly stationary (covariance stationarity) if

- $\mathbb{E}|X_t|^2 < \infty, \quad \forall t \in \mathbb{Z},$
- $\mu(t) = \mu, \quad \forall t \in \mathbb{Z},$
- $\text{cov}(X_t, X_{t-u}) = \gamma(t, t-u) = \gamma(u), \quad \forall t, u \in \mathbb{Z}.$

Remark 11.21.1 In the case of Gaussian time series, the two definitions of stationarity are equivalent

When time series are stationary it is possible to simplify the parametrisation of the mean and autocovariance functions. The weak stationarity implies that the time plot of the data would show that the T values fluctuate with constant variation around a fixed level.

Remark 11.21.2 The covariance $\text{cov}(X_t, X_{t-u}) = \gamma(u)$ is called u -lag autocovariance function of X_t .

White noise is the simplest example of stationary process. We will define several types of white noise with increasingly restrictive assumptions.

Definition 11.21.9 Autocorrelation function

The correlation coefficient between X_t and X_{t-u} is called the u -lag autocorrelation of X_t and is commonly denoted by $\rho(u)$, which under the weak stationarity assumption is a function of u only. Specifically, we define

$$\rho(u) := \frac{\text{cov}(X_t, X_{t-u})}{\sqrt{\text{var}(X_t)\text{var}(X_{t-u})}} = \frac{\gamma(u)}{\gamma(0)}$$

Often the ACF is plotted as a function of u , the so called correlogram. This is an important graphical instrument to illustrate linear dependency structures of the process. In practice, if all sample ACFs are close to zero, then the series is a white noise series.

Suppose that we observe time series $(X_t)_{t=1}^T$ from stationary process (X_t) . To estimate mean μ and variance σ^2 of the process we may use the sample mean \bar{X} and sample variance S . To estimate the autocovariance function, we use sample autocovariance function

$$\hat{\gamma}(u) := \frac{1}{T} \sum_{t=1+u}^T ((X_{t-u} - \bar{X})(X_t - \bar{X})), \quad 0 \leq u \leq T-1.$$

To estimate ρ we use autocorrelation function (ACF)

$$\hat{\rho}(u) := \frac{\hat{\gamma}(u)}{\hat{\gamma}(0)} = \frac{\sum_{t=1+u}^T ((X_{t-u} - \bar{X})(X_t - \bar{X}))}{\sum_{t=1}^T (X_t - \bar{X})^2}, \quad 0 \leq u \leq T-1.$$

Definition 11.21.10 Weak white noise

The stochastic process (X_t) is white noise if the following holds:

- $\mathbb{E}[X_t] = \mu, \quad t \in \mathbb{Z},$
- $\text{var}(X_t) = \sigma^2, \quad t \in \mathbb{Z},$
-

$$\gamma(u) = \begin{cases} \sigma^2 & u = 0, \\ 0 & u \neq 0. \end{cases}$$

Remark 11.21.3 If stochastic process (X_t) is i.i.d. then we call it i.i.d. white noise. An i.i.d white noise process is also white noise process, but not vice versa.

Definition 11.21.11 Lag operator.

Let (X_t) is stochastic process. Lag operator L is defined as

$$LX_t := X_{t-1}$$

Remark 11.21.4 The lag operator is linear

11.22 Linear Processes.

We will now turn to an examination of the large class of useful time series models. Consider autoregressive model defined by

$$X_t = \varphi X_{t-1} + \varepsilon_t.$$

where $|\varphi| < 1$ and ε_t is sequence of uncorrelated random variables, each with mean 0 and variance σ^2 . We should check if such process existed. Using lag operator we have

$$(1 - \varphi L)X_t = \varepsilon_t$$

Formally inverting the operator $(1 - \varphi L)$ leads to

$$X_t = \frac{1}{(1 - \varphi L)} \varepsilon_t = \sum_{u=0}^{\infty} \varphi^u L^u \varepsilon_t = \sum_{u=0}^{\infty} \varphi^u \varepsilon_{t-u}.$$

Definition 11.22.1 linear process

Process $(X_t)_{t=1}^T$

$$X_t = \sum_{u=-\infty}^{\infty} \varphi^u \varepsilon_{t-u},$$

where ε_t is a white-noise and $\sum_{u=-\infty}^{\infty} |\varphi|^u$

Remark 11.22.1 The general linear process depends on both past and future values of ε_t . A linear process which depends only on the past and present values of ε_t is said to be causal.

11.23 Autoregressive series.

Time series $(X_t)_{t=1}^T$ is called autoregressive of order p if satisfies

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t,$$

where ε_t is a white-noise and the ϕ_1, \dots, ϕ_p are constants. Then we denote $AR(p)$ Autoregressive series are important because:

- They have a natural interpretation - the next value observed is a slight perturbation of a simple function of the most recent observations
- It's easy to estimate their parameters. It can be done with standard regression software.
- They are easy to forecast.

11.23.1 Why is stationarity important in the historical VaR model?

In margining, the return should be relatively easy to predict over the time. In situation when returns alone are not stationary, returns over volatility $\frac{r_t}{\sigma_t}$ whose statistical properties suggest much more stationarity are examined.

11.23.2 How are we testing stationarity?

- **Augmented Dickey-Fuller test (ADF).**

The null hypothesis for this test is the presence of a unit root. A unit root means that the unconditional expected value of the mean is undefined, and therefore the long term expected value of the return cannot be constant (stationarity).

A unit-root (null hypothesis) guarantees non-stationarity, so we hope to reject this test. If the p-value is between 0 and 0.05, we can reject the null hypothesis and do not have a unit root. Our desire is to reject the null hypothesis.

11.24 EWMA and GARCH

Volatility tends to happen in clusters. The assumption that volatility remains constant at all times can be fatal. In order to forecast volatility in stock market, there must be methodology to measure and monitor volatility modelling. Recently, EWMA (exponentially weighted moving average) and GARCH models have become critical tools for time series analysis in financial applications. These two methods enhance the quality of *VaR* models.

For *VaR* calculations EWMA and GARCH models assume returns on financial assets have serial correlations. Both models give more weight to the latest returns than the old ones.

11.24.1 EWMA

RiskMetrics approach was developed in 1994 and it uses exponential weighted moving average (EWMA) to forecast the time-varying risk.

Formally, the forecast variance for time t is a weighted average of the previous forecast, using weight λ , and of the latest squared innovation, using weight $(1 - \lambda)$. Thus the variance can be computed from

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2.$$

Here, the parameter λ is called the decay factor and must be less than unity.

Vanilla historical VaR (where returns are not scaled) is not quick to changing volatilities, and thus fails to provide the desired margin profile interest rate swaps portfolios:

- the margins are too low in high volatility environment,
- This methodology exposes a company to larger shortfalls.

11.24.2 GARCH(1,1).

There is substantial empirical evidence that the conditional volatility models successfully forecast risk. The general assumption is that the conditional returns have a normal distribution, although this could be extended to other distributions such as the t-student's.

The GARCH model assumes that the conditional variance depends on the latest innovation, and on the previous conditional variance. Define $h_t := \sigma^2$ as the conditional variance, using information up to time $t-1$, and r_{t-1} as the previous day's return. The simplest such model is the GARCH(1,1) process,

$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta h_{t-1}$$

which involves one lag of the innovation and one lag of the previous forecast. The β term is important because it allows persistence in the shock, which is realistic future of the data.

The unconditional variance is found by setting $\mathbb{E}[r_{t-1}^2] = h_t = h_{t-1} = h$. Solving for h , we find

$$h = \frac{\alpha_0}{1 - \alpha_1 - \beta}.$$

This model will be stationary when the sum of parameters $\gamma = \alpha_1 + \beta < 1$

11.25 Risk

Risk means uncertainty in the future returns from an investment. In particular that the investment could earn less than the expected return. Risk is often measured as standard deviation. Recently there has been a trend toward measuring risk by the value at risk (VaR) or by expected shortfall. Because risk depends upon the probability distribution of a return, probability and statistic are fundamental tools for finance. There is several types of risk.

The most popular types of risk are:

- Market risk is due to changes in prices.
- Credit risk.
- liquidity risk is the potential extra cost of liquidating a position because buyers are difficult to locate.
- Operational risk is due to fraud mismanagement, human errors.

Measure risk such as duration analysis are used to estimate the market of fixed rate securities. In contrast value at risk and expected shortfall are widely used because they can be applied to all types of risks and securities, including complex portfolios.

Chapter 12

Credit Risk.

12.1 induction

Definition 12.1.1

Price is what you pay, value is what you get. One of the main drivers for trading is the all-in price. This requires calculation of credit, debt, margin, funding and liquidity value adjustments — known as CVA, DVA, MVA, FVA, LVA respectively. Or collectively: XVA. **What does quantative analyst**

Quants are typically expected to get involved with the mathematics and implementation of models for pricing of derivative instruments. The focus of what a quant does really depends on the nature of his or her role, but invariably the need for a deep fundamental understanding of maths is the key.

invariably=always

How would you describe yourself

I would like to avoid repeating information included in my CV. But I would say that my background to date has been centred around preparing myself to work in finance. I'm motivated and put a lot of effort into everything I do whatever I'm studying at university or even I'm playing sport. I work well under pressure and learn quick new things. I have a good knowledge of mathematics especially finance mathematics stochastic process, measure theory probability and theory of risk.

12.2 Popular questions.

Exercise 12.2.1 *Your friend has two children. He told you that he has at least one daughter. What is the probability that his second child is also a girl*

Exercise 12.2.2 *How observation may be use to simulate random variables.*

Exercise 12.2.3 *Find approximate value \sqrt{e}*

Chapter 13

Cracking the Coding Python

13.1 Numbers

Problem 13.1.1 (*Finding Π value using Monte Carlo method.*)

```
def find_pi_value_fun(runs):
    u1=uniform.rvs(0,1,runs)
    u2=uniform.rvs(0,1,runs)
    hit_or_miss=zeros(len(u1))
    for i in range(len(u1)):
        if (u1[i]**2+u2[i]**2<=1):
            hit_or_miss[i]=1
        else:
            hit_or_miss[i]=0
    pi=average(hit_or_miss)*4
    return pi
```

Problem 13.1.2 (*Factorial using recursion*)

```
def factorial_recursion_fun(n):
    if n==0:
        factorial=1
    else:
        factorial=n*factorial_recursion_fun(n-1)
    return factorial
```

Problem 13.1.3 (*Factorial using loop*)

```
def factorial_loop_fun(n):
    factorial=[0]*n
    factorial[0]=1
    for i in range(1,n):
        factorial[i]=factorial[i-1]*(i+1)
    return factorial[-1]
```

Problem 13.1.4 (*Prime number checking*)

```
def prime_number(n):
    if n == 1:
        return False
```

```

    elif n == 2:
        return True
    else:
        for i in range(2, n):
            if (n % i == 0):
                return False
        return True

#####

def prime_number_ver2(n):
    if (n<2):
        return False
    for i in range(2,int(n**0.5)+1):
        if n%i==0:
            return False
    return True

x=int(input('give the number>>'))
print(prime_number(n=x))

```

13.2 Numerical analysis

Problem 13.2.1 (*Mid point rule.*)

```

def mid_point_rule_fun(left,right,n):
    def f(x):
        return 1/((x+1)**2)
    h=(right-left)/n
    nots=linspace(left,right,n+1)
    midpoints=zeros((n+1))
    midpoints[0]=left
    for i in range(1,9):
        midpoints[i]=(nots[i]+nots[i-1])/2
    integral=h*sum(f(midpoints[1:]))
    return integral

```

Problem 13.2.2 (*Trapezoidal rule*)

```

def trapeizodal_rule_fun(left,right,n):
    def f(x):
        return 1/((x+1)**2)
    h=(right-left)/n
    nots=linspace(left,right,n+1)
    values=f(nots)
    integral=0.5*h*(values[0]+2*sum(values[1:(len(values)-1)])+values[-1])
    return integral

```

Problem 13.2.3 (*Root finding*)

```

class RootFinding():

```

```

def __init__(self,a,b,epsilon):
    self.a=a
    self.b=b
    self.epsilon=epsilon

def bisection(self):
    #here we can function to find root
    def f(x):
        return x**3+x-1
    c=(self.a+self.b)/2
    while (self.b-self.a)/2>self.epsilon:
        if f(c)==0:
            return c
        elif f(self.a)*f(c)<0:
            self.b=c
        else:
            self.a=c
        c=(self.a+self.b)/2
    return c

```

13.3 Sorting

A quick sort first selects a value, which is called the pivot value. There are many different ways to choose the pivot value, we will simply use the first item in the list. The role of the pivot value is to assist with splitting the list. The actual position where the pivot value belongs in the final sorted list, commonly called the split point, will be used to divide the list for subsequent calls to the quick sort.

13.4 Searching

Constructor for various searching .

```

class SearchingAlgoorythm():
    def __init__(self,a_list,item):
        self.a_list=a_list
        self.item=item

```

Below we can define method for binary search.

```

def binary_search(self):
    first=0
    last=len(self.a_list)-1
    found =False
    while first<=last and not found:
        midpoint=(first+last)//2
        if self.a_list[midpoint]==self.item:
            found=True
        elif self.item<self.a_list[midpoint]:
            last=midpoint-1
        else:
            first=midpoint+1

```

```
return found
```

Problem 13.4.1 (*Root finding*)

```
class RootFinding():
    def __init__(self,a,b,epsilon):
        self.a=a
        self.b=b
        self.epsilon=epsilon

    def bisection(self):
        #here we can function to find root
        def f(x):
            return x**3+x-1
        c=(self.a+self.b)/2
        while (self.b-self.a)/2>self.epsilon:
            if f(c)==0:
                return c
            elif f(self.a)*f(c)<0:
                self.b=c
            else:
                self.a=c
            c=(self.a+self.b)/2
        return c
```

13.5 LeetCode Problems

Problem 13.5.1 *Given an array of integers, return indices of the two numbers such that they add up to a specific target. You may assume that each input would have exactly one solution, and you may not use the same element twice.*

```
class Solution:
    def twoSum(self,nums,target):
        """
            :type nums: List[int]
            :type target: int
            :rtype: List[int]
        """
        for i in range(0,len(nums)):
            for j in range(i+1,len(nums)):
                temp_sum=nums[i]+nums[j]
                if (temp_sum==target):
                    return [i,j]

object_solution=Solution()
print(object_solution.twoSum(nums=[2,5,5,11],target=10))
OUTPUT
[1, 2]
```

Problem 13.5.2 *Given an array of integers, return indices of the two numbers such that they add up to a specific target. You may assume that each input would have exactly one solution, and you may not use the same element twice.*

```
def myPower(self,x,n):
    """
        :type x: float
        :type n: int
        :rtype: float
    """

    if n==0:
        return 1
    else:
        return x*self.myPower(x,n-1)
```

Problem 13.5.3 # Pascal triangle.

```
def generate_pascal_triangle(self,rows_num):
    if(rows_num==0):
        return []
    else:
        result=[[1]]
        for i in range(1,rows_num):
            result.append([0]*(i+1))
            for k in range(i+1):
                if k==0 or k==i:
                    result[i][k]=1
                else:
                    result[i][k]=result[i-1][k-1]+result[i-1][k]
        return result
```

13.6 Useful tricks

Problem 13.6.1 (*enumerate*)

```
my_list = ['apple', 'banana', 'grapes', 'pear']
counter_list = list(enumerate(my_list, 1))
print(counter_list)
OUTPUT
[(1, 'apple'), (2, 'banana'), (3, 'grapes'), (4, 'pear')]
```


Chapter 14

SQL

Remark 14.0.1 *Format of inputs is following: **tablename**: col1[type], col2[type], ...*

Problem 14.0.1 (*Remove duplicates*)

Write a SQL query to find all duplicate emails in a table named Person.

SOLUTION.

```
SELECT email
FROM Person
GROUP BY email
HAVING COUNT(email)>1
```

Problem 14.0.2 *left join* *We have two tables:*

1. **Person**: *PersonId[int], FirstName[varchar], LastName[varchar]*
2. **Address**: *AddressId[int], PersonId[int], City[varchar], State[varchar]*

Write a SQL query for a report that provides the following information for each person in the Person table, regardless if there is an address for each of those people:

```
SELECT FirstName,LastName,City,State
FROM Person
LEFT JOIN Address
ON Person.PersonId =Address.PersonId
```


Chapter 15

Coding in C++

15.1 C++.

Problem 15.1.1 (*Normal distribution sample.*)

```
vector<double> normal_sample_fun(double mu, double sigma,int n)
{
    default_random_engine generator(time(0));
    vector<double>result;
    normal_distribution<double> distribution(mu, sigma);
    for (int i = 0; i < n; i++)
    {
        result.push_back(distribution(generator));
    }
    return result;
}
```

15.2 Python.

15.3 Useful phrases.

- I'm a conscientious student and usually the last person to leave the library.
- Cracov is full of financial whiz kids.
- All else being equal, the option is worth more the more volatile the stock is.
- Information structure of the model is based on observation of the stock price process only.
- We will restrict attention...
- There are two main uses of derivatives, namely speculation and hedging
- notional=existing only as idea, not as something real
- to offset=to balance one influence against an opposing influence, so that there is no great difference as a result.
- to exchange cash flows at some future dates according to a prearranged formula
- calculate the associated payoff of the option for each path

- We use this method to calculate the price of an option with complicated features
- if it admits a representation
- integrable in a suitable sense
- exposure to risk
- Risk means uncertainty in the future returns from an investment
- Recently there has been a trend toward
- volatile (adjective)=likely to change suddenly
- Well, you see
- to enhance = to improve the quality, amount or strength of something
- be subject to =to have or experience a particular thing, especially something unpleasant
- But you mean..., right?
- One of the most appealing features of options it is possibility.
- Another class of options comprises so-called American options.
- consecutive numbers 12,13,14,
- prevailing *prewejling* = existing in a particular place or at a particular time.
- proceeds = the amount of money received from particular event or activity or when something is sold
- We adopt the following general framework. Let ...
- using ... we can work out ...
- we are interested in pricing
- Existence arbitrage in our model is cause for concern.
- I need a new computer and training course to get the best out of me.
- explore two possible scenarios
- What's so interesting about this model ...
- I guess that's why ...
- Dropping the arguments
- the only term affecting the sign
- symbol $n!$ is called factorial
- inflection point
- to push a student
- I found this theorem difficult

15.4 EXCEL

Definition 15.4.1 EXCEL

Microsoft Excel is an electronic spreadsheet program, created by multiple highly skilled engineers from Microsoft. It enables users to organize, format, and calculate data with formulas using a spreadsheet system broken up by rows and column.

We also use this tool for storing, organizing and manipulating the data. In addition, it also offers programming that supports VBA, and we can use external database to make dynamic reports, analysis etc. Smart use of this program saves a lot of time and helps in creating our own applications too.

- **SUMIF**

The SUMIF function will look for a certain criteria and if it finds it, then it will Sum up related cells.

- **COUNTIF**

We use COUNTIF function to count the specified cells, with a given condition or criterion.

- **VLOOKUP**

This function will look for a piece of information in a large table of data and pull in any field from that table into your new table.

- **MAX**

This function will simply return the largest and smallest result from a range of numbers.

- **IF**

The IF function is used to determine whether a statement is True or False and then performs an action based on the result. We may nest IF function seven times.

- **PIVOT TABLES**

Excel Pivot Tables are tables that summarise large amounts of data in an Excel spreadsheet. Another feature of Excel pivot tables is the ability to quickly extract the data from any part of the pivot table.

- **WEEKDAY**

return to the day of the week of a particular date.

- **COUNTIF**

We use COUNTIF function to count the specified cells, with a given condition or criterion.

Chapter 16

SQL

Structured Query Language is a standard Database language which is used to create, maintain and retrieve the database.

16.1 Popular Interview Questions.

Question 16.1.1 *What is subquery?*

Question 16.1.2 *What is the GROUP BY clause used for?*

Question 16.1.3 *What is the difference between WHERE and HAVING*

16.2 Data definition

- **CREATE** Creates a new table, a view of a table, or other object in database
- **ALTER** Modifies an existing database object, such as a table.
- **DROP** Deletes an entire table, a view of a table or other object in the database.

16.3 Data manipulation

- **INSERT** Creates a record
- **UPDATE** Modifies records
- **DELETE** Deletes records

Remember that **record** = row and **fields** = column