Algorithmic Trading
Summer of Science 2024

Geet Sethi

May 2024 - July 2024

Contents

Chapter 1

Fundamentals of Stock Markets

1.1 Types of Assets (Asset Classes)

Definition 1.1.1: Asset Class

An asset class is a category of investment with particular risk and return characteristics. The following are some of the popular asset classes:

- Fixed income instruments
- Equity
- Real estate
- Commodities (e.g. precious metals)

1.1.1 Fixed Income Instruments

These carry very limited risk to the principle and the return is paid as an interest based on the particular fixed income instrument.
Interest paid could be at quarterly, semi-annual or annual intervals. The capital is returned to the investor at the end of the term of the deposit.
Typical fixed income investments include:

- Fixed deposits
- Bonds issued by the government and its agencies
- Bonds issued by corporates

1.1.2 Equity

Investment in equities involve buying shares of publicly listed companies.
The shares are traded both on the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE).
Unlike fixed income securities, these offer no guarantee against the capital. However, as a trade off, these investments can yield very attractive returns. Indian equities have generated returns close to 14%-15% CAGR (compound annual growth rate).

1.1.3 Real Estate

Real estate investment involves transacting commercial and non-commercial land (e.g. sites, apartments, commercial buildings, etc.).

> **Note 1.1** There is no official metric to measure the returns generated by real estate.

1.1.4 Commodity

Investments in precious metals such as gold and silver is considered one of the most popular investment avenues. Gold and silver over a long-term period have yielded a CAGR return of approximately 8% over the last 20 years.

1.2 Financial Intermediaries

1.2.1 The Regulator

In India, the stock market is regulated by the **The Securities and Exchange Board of India** also referred to as **SEBI**. The main objective of SEBI is to promote the development of stock exchanges, protect the interests of retail investors, regulate the activates of market participants and financial intermediaries.

1.2.2 The Stock Broker

Definition 1.2.1: Stock Broker

A stock broker is a private entity, registers as a trading member with the stock exchange and a stock broking license. They are a gateway to the stock exchanges. An individual must go through a stock broker to buy/sell stocks.

An individual who wishes to trade in the stock market must open a ”Trading Account” with a broker through which they can then trade at the stock exchanges.

1.2.3 Depository and Depository Participants

Earlier, when you bought stocks, the only way to identify that you owned the stock was a piece of paper called the share certificate. Hence it became extremely important to store the property papers in a safe and secure place.
Seeing the obvious problem of storing the certificates, after 1996, the share certificates were converted to a digital format.

Definition 1.2.2: Dematerialization

The process of converting paper format share certificates to digital format is called dematerialization and is often abbreviated as **DEMAT**

The share certificate, though in digital form, still needs to be stored securely. This is done through DEMAT accounts.

Definition 1.2.3: Depository

A depository is a financial intermediary which offers the service of a DEMAT account.

DEMAT accounts act as a digital vault for your shares.

1.2.4 Banks

Banks help in facilitating the fund transfer from your bank account to your trading account.

1.2.5 Clearing Corporations

National Security Clearing Corporation Ltd (NSCCL) and Indian Clearing Corporation Ltd (ICCL) are wholly owned subsidiaries of NSE and BSE respectively.
The job of the clearing corporation is to ensure guaranteed settlement of your trades/transactions.
The typical roles of the clearing corporation is to ensure the following:

- Identify the buyer and seller and match the debit and credit process
- Ensure no defaults - The clearing corporation also ensures there are no defaults by either party.

1.3 Calculating Returns

1.3.1 Absolute Return

This is the return that your investment has generated in absolute terms.

$$\left(\frac{\text{Ending Period Value}}{\text{Starting Period Value}} - 1\right) \times 100$$

Example 1.3.1 (Calculating absolute return)
Let's say that you bought a stock at 3030 and sold it at 3550. What absolute return did you generate?

$$\begin{aligned}\text{absolute return} &= \left(\frac{3550}{3030} - 1\right) \times 100 \\ &= 0.1716 \times 100 \\ &= 17.16\%\end{aligned}$$

1.3.2 Compound Annual Growth Rate (CAGR)

The formula to calculate CAGR is:

$$\text{CAGR} = \left(\frac{\text{Ending Value}}{\text{Beginning Value}}\right)^{\left(\frac{1}{\text{No. of years}}\right)} - 1$$

Example 1.3.2 (Calculating CAGR)
Let's say that you bought a stock at 3030 and sold it at 3550. What CAGR did you generate?

$$\begin{aligned}\text{CAGR} &= \left(\frac{3550}{3030}\right)^{\frac{1}{2}} - 1 \\ &= 9.2 - 1 \\ &= 8.2\%\end{aligned}$$

1.4 Index

There are two main market indices in India. The **S&P Sensex** representing the BSE and **CNX Nifty** representing the NSE.

1.4.1 Practical Uses of the Index

- **Information**: The index reflects the general market trend for a period of time.
- **Benchmarking**: The index can be used to judge whether the returns you got over a period of time by comparing the returns to the increase in the index.
- **Trading**: Majority of the traders in the market trade the index.
- **Portfolio Hedging**: Investors usually build their own portfolio which typically contains 10-12 stocks which they would have bought from a long term perspective. If they can foresee a prolonged adverse movement in the market (such as in 2008) which could potentially erode the capital in the portfolio, the investors can use the index to hedge the portfolio.

1.4.2 Index Construction Methodology

Every stock in the index is assigned a certain weightage. There are many ways to calculate these weights but the Indian stock exchange follows a method called **free float market capitalization**. In the method, the larger the market capitalization of the company, higher its weight.

$$\text{Free float market capitalization} = \text{total number of shares outstanding in the market} \times \text{price of the stock}$$

1.4.3 Sector Specific Indices

While the Sensex and Nifty represent the broader markets, there are certain indices that represent specific sectors. These are called **sectoral indices**. For example, Bank Nifty on NSE represents mood specific to the banking industry.

1.5 Clearing & Settlement Process

1.5.1 What happens when you buy a stock?

Day 1 - The trade (T Day)

The day one makes a trade is referred to as the trade date (represented as 'T Day').
By the end of the day, your broker will debit the amount required for the trade and other applicable charges towards the purchase.
An important point to note is that the money is debited from your account but the stock does not come into your DEMAT account yet.
The same day, the broker generates a 'contract note'. A contract note typically shows the break up of all transactions done during the day along with the trade reference number as well as the charges charged by the broker.

Day 2 - Trade Day + 1 (T+ Day)

One can sell the stock that they bought on the trade day on this day.

> **Important 1.1** This does involve a slight risk since you do not own the stock that you bought the previous day yet i.e. that stock hasn't been deposited into your DEMAT account.

From the point of view of the user, nothing happens on this day. However, in the background the money required to purchase the shares is collected by the exchange along with other charges.

Day 3 - Trade Day + 2 (T+2 Day)

On this day, around 11 AM, the shares are debited from the person who sold the shares and credited to the brokerage with whom the person is trading, who will in turn credit it to your DEMAT account by end of day. Similarly, money which was debited from your account is credited to the person who sold the shares.

1.5.2 What happens when you sell a stock?

Similar to the process which takes play when you buy a stock, the day when you sold the stock is called the 'trade day'. The moment you sell the stock, the stock gets blocked in your DEMAT account. On the T+1 day, the blocked shares are given to the exchange. On T+2 day, you would receive the funds from the sale which will be credited to your trading account.

1.6 Orders in the Market

1.6.1 Types of Orders

- **Market Order**: A market order is *buying or selling a stock at the best price available.* Generally, this type of order will be *executed immediately.* However, the price at which the market order will be executed is not guaranteed. The *last traded price* (LTP) need not be the price at which the order is executed.
- **Limit Order**: A limit order is an *order to buy or sell a stock at a specific price or better.* A limit order is not guaranteed to be executed. But they do help ensure the investor does not pay more than a predetermined price for a stock.
- **Stop-Loss Order**: A stop order, also referred to as stop-loss order, is an *order to buy or sell a stock once the price reaches a specified price*, also known as the stop price. When the stop price is reached, a stop order becomes a market order. A buy-stop order is entered at a stop price which is above the current market price. Investors generally use a buy-stop order to limit a loss or to protect a profit on a stock that they have sold short. A sell-stop order is entered at a stop price below the current market price. Investors generally use a sell-stop order to limit a loss or to protect a profit on a stock that they own.
- **Stop-Limit Order**: A stop-limit order is an *order to buy or sell a stock that combines the features of a stop order and a limit order.* Once the stop price is reached, a stop-limit order becomes a limit order that will be executed at specified price or better. The benefit of a stop-limit order is that the investor can control the price at which the order can be executed.
- **Take Profit Order**: A take-profit order (sometimes called a profit target) is *intended to close out the trade at a profit once it has reached a certain level.* Execution of a take-profit order closes the position. This type of order is always connected to an open position of a pending order.

1.6.2 Slippage

Definition 1.6.1: Slippage

Slippage refers to the difference between the expected price of a trade and the price at which the trade is executed.

Slippage can occur at any time but is most prevalent during periods of higher volatility when market orders are used. It can also occur when a large order is executed but there isn't enough volume at the chosen price to maintain the current bid/ask spread.
An $x\%$ of slippage means the order was executed $x\%$ below or above the expected price.
Disadvantages of high slippage include:

- Increased trading costs
- Reduced profitability
- Inaccurate risk management
- Difficulty in entering and exiting positions

1.6.3 Risk-Reward Ratio

This ratio is used to assess the potential profitability and risk of a trade or investment opportunity. It is a way to evaluate the relationship between the potential reward of a trade and the amount of risk taken.

The risk-reward ratio is calculated by dividing the potential reward (or profit) of a trade by the potential risk (or loss). The resulting ratio provides an indication of how much profit is expected for each unit of risk assumed.

Chapter 2

Fundamental Analysis

2.1 Overview

Fundamental Analysis (FA) is a holistic approach to study a business. When an investor wishes to invest in a business for the long term(say 3 - 5 years) it becomes extremely essential to understand the business from various perspectives. It is essential for the investor to separate the daily short term noise in the stock prices and concentrate on the underlying business performance. Over the long term, the stock prices of a fundamentally string companies tend to appreciate.

2.2 Investible Grade Attributes

An investible grade company has a few distinguishable characteristics. These characteristics can be classified user two heads namely the 'Qualitative aspect' and the 'Quantitative aspects'.

2.2.1 Qualitative Aspects

These mainly involve understanding the non-numeric aspects of the business. It includes many factors such as:

- Management's background
- Business ethics - is the managements involved in scams, bribery, unfair business practices, etc.
- Corporate governance - appointments of directors, organization structure, transparency, etc.
- Minority shareholders - how does the management treat its minority shareholders, do they consider their interests while taking corporate actions
- Related party transactions - Is the company tendering financial favors to known entities such as promoter's relatives, friends, vendors, etc. at the cost of the shareholders funds?
- Salaries paid to promoters - Is the management paying themselves a hefty salary, usually a percentage of profits
- Operator activity in stocks - Does the stock price display unusual price behavior especially at the time when the promoter is transacting in the shares.
- Shareholders - Who are the significant shareholders in the firm
- Political affiliation
- Promoter lifestyle - Are the promoters too flamboyant and loud about their lifestyle?

2.2.2 Quantitative Aspects

These are matters related to financial numbers. These include many things, to name a few:

- Profitability and its growth
- Margins and its growth
- Earnings and its growth
- Matters related to expenses
- Operating efficiency
- Pricing power
- Matters related to taxes
- Dividends payout
- Cash flow from various activities
- Debt – both short term and long term
- Working capital management
- Asset growth
- Investments
- Financial Ratios

> **Note 2.1** The list is actually endless. In fact, each sector has many different metrics which are relevant to that sector only.
>
> For example, in the retail industry, the following metrics are used:
>
> - Total number of stores
> - Average sales per store
> - Total sales per square foot
> - Merchandise margins
> - Owned store to franchisee ratio

2.3 Annual Report

The annual report (AR) is a yearly publication by the company and is sent to the shareholders and other interested parties. The annual report is published by the end of the Financial Year, and all the data made available in the annual report is dated to 31st March.

> **Important 2.1** No two annual reports are the same; they are all made to suite the company's requirement keeping in perspective the industry they operate in. However, some sections are common across all annual reports.

The annual report also contains three important financial statements namely:

- Profit and Loss Statement
- Balance Sheet
- Cash Flow Statement

2.4 The Profit and Loss Statement

2.4.1 Overview

The Profit and Loss statement is also popularly referred to as the P&L statement, Income Statement, Statement of Operations, and Statement of Earnings.
The P&L statement reports information on:

- The revenue of the company for the given period (yearly or quarterly)
- The expenses incurred to generate the revenues
- The tax and depreciation
- The earnings per share number

2.4.2 Some Jargon Used in these Statements

- **Top Line**: The top line of the company is the revenue generated by the company.
- **Net Sales**: The revenue adjusted after the excise duty is the net sales of the company.
- **Total Operating Revenue**: Revenue from sales of products + sale if services + other operating revenues sums up to give the total operating revenue of the company.
- **Bottom Line**:
- **Tangible asset**: This asset is one which has a physical form and provides an economic value to the company.
- **Intangible asset**: This asset is one which does not have any physical form but still provides an economic value to the company such as brand value, trademarks, etc.
- **Deprecation and Amortization**: An asset (tangible or intangible) has to be depreciated over its useful life. Useful life is defined as the period during which the asset can provide economic benefit to the company. Since an asset would continue to provide economic benefits over its useful life, it makes sense to spread the cost of acquiring the asset over its useful life. This is called depreciation. The deprecation equivalent for intangible assets is called amortization.
- **Profit After Tax (PAT)**
- **Earnings Per Share (EPS)**: EPS reflects the earning capacity of a company on a per share basis. $\text{EPS} = \frac{\text{PAT}}{\text{Total number of outstanding ordinary shares}}$

2.5 The Balance Sheet

Definition 2.5.1: Asset

An asset is a resource controlled by the company, and is expected to have an economic value in the future. Typical examples of assets include plants, machinery, cash, brands, patents etc. Assets are of two types, current and non-current.

Definition 2.5.2: Liability

A liability represents the company's obligation. The obligation is taken up by the company because the company believes these obligations will provide economic value in the long run. Liability in simple words is the loan that the company has taken and it is therefore obligated to repay back.

Note 2.2 In a typical balance sheet, the total assets of a company should be equal to the total liabilities of the company.

Definition 2.5.3: Owners Capital

It is the difference between the assets and the liabilities.
It is also called the 'Shareholders Equity' or the 'Net Worth'

2.5.1 Types of Assets & Liabilities

Current Assets

Current assets are assets that can be easily converted to cash and the company foresees a situation of consuming these assets within 365 days. Current assets are the assets that a company uses to fund its day to day operations and ongoing expenses.

Non-Current Assets (Fixed Assets)

Non-current assets are the assets that the company owns, the economic benefit of which is enjoyed over a long period (beyond 365 days)

Current Liabilities

Current liabilities are a company's obligations which are expected to be settled within 365 days (less than 1 year). The term 'Current' is used to indicate that the obligation is going to be settled soon, within a year.

Non-Current Liabilities

Non-current liabilities represent the long term obligations, which the company intends to settle/pay off not within 365 days/ 12 months of the balance sheet date. These obligations stay on the books for few years. Non-current liabilities are generally settled after 12 months after the reporting period.

2.6 The Cash Flow Statement

Overview

The cash flow statement provides information to the users of the financial statements about the entity's ability to generate cash and cash equivalents as well as indicates the cash needs of a company.

2.6.1 Activities Undertaken by a Company

Any legitimate company has three main activities:

- **Operational activities (OA)**: Activities that are directly related to the daily core business operations are called operational activities. Typical operating activities include sales, marketing, manufacturing, technology upgrade, resource hiring etc.
- **Investing activities (IA)**: Activities pertaining to investments that the company makes with an intention of reaping benefits at a later stage. Examples include parking money in interest bearing instruments, investing in equity shares, investing in land, property, plant and equipment, intangibles and other non current assets etc.
- **Financing activities (FA)**: Activities pertaining to all financial transactions of the company such as distributing dividends, paying interest to service debt, raising fresh debt, issuing corporate bonds etc.

2.7 Financial Ratios

2.7.1 Profitability Ratios

The Profitability ratios help the analyst measure the profitability of the company. The ratios convey how well the company is able to perform in terms of generating profits.

EBITDA Margin

The Earnings before Interest Tax Depreciation & Amortization (EBITDA) Margin tells us how profitable (in percentage terms)the company is at an operating level.

$$\text{EBITDA} = \text{Operating Revenues} - \text{Operating Expenses}$$
$$\text{Operating Revenues} = \text{Total Revenue} - \text{Other Income}$$
$$\text{Operating Expenses} = \text{Total Expense} - \text{Finance Cost} - \text{Depreciation \& Amortization}$$
$$\text{EBITDA Margin} = \frac{\text{EBITDA}}{\text{Total Revenue} - \text{Other Income}}$$

PAT Margin

While the EBITDA margin is calculated at the operating level, the Profit After Tax (PAT) margin is calculated at the final profitability level. At the operating level we consider only the operating expenses however there are other expenses such as depreciation and finance costs which are not considered. Along with these expenses there are tax expenses as well. When we calculate the PAT margin, all expenses are deducted from the Total Revenues of the company to identify the overall profitability of the company.

$$\text{PAT Margin} = \frac{\text{PAT}}{\text{Total Revenue}}$$

Return on Equity (RoE)

It is the return the shareholder earns for every unit of capital invested. RoE measures the entity's ability to generate profits from the shareholders investments.

$$\text{RoE} = \frac{\text{Net Profit}}{\text{Shareholders Equity}} \times 100$$

> **Important 2.2** DuPont Model
>
> Inspecting the RoE closely is very important because as the company takes on more debt instead of investment through equity, the RoE shoots up but that debt is not good for the company!
> To combat this, *DuPont* came up with another way of writing the RoE:
>
> $$\text{RoE} = \frac{\text{Net Profit}}{\text{Net Sales}} \times \frac{\text{Net Sales}}{\text{Avg. Total Assets}} \times \frac{\text{Avg. Total Assets}}{\text{Shareholder Equity}}$$
>
> If you notice, the terms cancel out to give the original formula back. However, in this process of decomposing the formula, one gained insights into three distinct aspects of the company's business.

2.7.2 Leverage Ratios

The Leverage ratios also referred to as solvency ratios/ gearing ratios measures the company's ability (in the long term) to sustain its day to day operations. Leverage ratios measure the extent to which the company uses the debt to finance growth.

Interest Coverage Ration

The interest coverage ratio, also referred to as the debt service ratio or the debt service coverage ratio, helps us understand how much the company is earning relative to the interest burden on the company. Hence, it helps us interpret how easily can the company pay its interest payments.

$$\text{Interest Coverage Ration} = \frac{\text{Earnings before interest and tax}}{\text{Interest Payment}}$$

Debt to Equity Ratio

It measures the amount of total debt capital with respect to the total equity capital.

$$\text{Debt to Equity Ratio} = \frac{\text{Total Debt}}{\text{Total Equity}}$$

Debt to Asset Ratio

It measures the amount of total debt capital with respect to the total equity capital.

$$\text{Debt to Asset Ratio} = \frac{\text{Total Debt}}{\text{Total Assets}}$$

Financial Leverage Ratio

The financial leverage ratio gives us an indication to what extent the assets are supported by equity.

$$\text{Financial Leverage Ratio} = \frac{\text{Average Total Asset}}{\text{Average Total Equity}}$$

2.7.3 Valuation Ratios

The Valuation ratios compare the stock price of the company with either the profitability of the company or the overall value of company to get a sense of how cheap or expensive the stock is trading.

Price to Sales Ratio

In many cases, investors may use sales instead of earnings to value their investments. The earnings figure may not be true as some companies might be experiencing a cyclical low in their earning cycle.

$$\text{P/S Ratio} = \frac{\text{Current Share Price}}{\text{Sales per share}}$$

Price to Book Value Ratio

Definition 2.7.1: Book Value

It is simply the amount of money left on the table after the company pays off all of its obligations. This is the amount of money the company can expect to receive after it sells all of its assets and settles its debts.

$$\text{BV} = \frac{\text{Share Capital} + \text{Reserves}}{\text{Total Number of Shares}}$$

Many investors may choose to value a company based on how much its book value is because that is a guaranteed amount that the company will get for sure even in the worst case scenarios.

$$\text{P/BV Ratio} = \frac{\text{Current Share Price}}{\text{Book Value per share}}$$

Price to Earning Ratio

We know that the *EPS* measures the profitability of a company on a per share basis. Dividing the current market price by the *EPS* gives us the *Price to Earnings ratio* of the firm. It measures the willingness of the market participants to pay for the stock, for every rupee of profit that the company generates.

$$\text{P/E Ratio} = \frac{\text{Current Share Price}}{\text{Earnings per share}}$$

2.7.4 Operating Ratios

The Operating Ratios, also called the 'Activity Ratios' measures the efficiency at which a business can convert its assets (both current and non-current) into revenues. This ratio helps us understand how efficient the management of the company is.

Fixed Assets Turnover

The ratio measures the extent of the revenue generated in comparison to its investment in fixed assets. It tells us how effectively the company uses its assets. Fixed assets include the property, plant and equipment. Higher the ratio, it means the company is effectively and efficiently managing its fixed assets.

$$\text{Fixed Assets Turnover} = \frac{\text{Operating Revenues}}{\text{Total Average Asset}}$$

Working Capital Turnover

Definition 2.7.2: Working Capital

Working capital refers to the capital required by the firm to run its day to day operations.

$$\text{Working Capital} = \text{Current Assets} - \text{Current Liabilities}$$

The working capital turnover, also referred to as net sales to working capital, indicates how much revenue the company generates for every unit of working capital. Higher the number, the better it is.

$$\text{Working Capital Turnover} = \frac{\text{Revenue}}{\text{Average Working Capital}}$$

Total Assets Turnover

The ratio measures the extent of the revenue generated in comparison to its investment in fixed assets. It tells us how effectively the company uses its assets. Fixed assets include the property, plant and equipment. Higher the ratio, it means the company is effectively and efficiently managing its fixed assets.

$$\text{Fixed Assets Turnover} = \frac{\text{Operating Revenues}}{\text{Total Average Asset}}$$

Receivables Turnover Ratio

The receivable turnover ratio indicates how many times in a given period the company receives money/cash from its debtors and customers. Naturally a high number indicates that the company collects cash more frequently.

$$\text{Receivables Turnover Ratio} = \frac{\text{Revenue}}{\text{Average Receivables}}$$

Days Sales Outstanding (DSO) / Average Collection Period

The days sales outstanding ratio illustrates the average cash collection period i.e the time lag between billing and collection.

$$\text{DSO} = \frac{365}{\text{Receivables Turnover Ratio}}$$

2.8 Valuation of a Company using Discounted Cash Flow (DCF) Method

Valuation per say helps the individual determine the 'intrinsic value' of the company. We will now look at a valuation technique called the **Discounted Cash Flow (DCF)** analysis to calculate the intrinsic value of the company.

2.8.1 Time Value of Money

If we have to evaluate, what would be the value of money that we have today sometime in the future, then we need to move the 'money today' through the future. This is called the **Future Value (FV)** of the money. Likewise, if we have to evaluate the value of money that we are expected to receive in the future in today's terms, then we have to move the future money back to today's terms. This is called the **Present Value (PV)** of money.

Definition 2.8.1: Compounding and Discounting

This process of adjusting the money we have today to calculate its future value is called **Compounding** and when we have to calculate its present value of some money we are to receive in the future is called **Discounting**

Future value can be calculated using:

$$\text{Future Value} = \text{Amount} \times (1 + \text{opportunity cost rate})^{\text{Number of Years}}$$

Present value can be calculated using:

$$\text{Present Value} = \frac{\text{Amount}}{(1 + \text{Discount Rate})^{\text{Number of Years}}}$$

Definition 2.8.2: Net Present Value

The sum of all present values of the future cash flow is called the **Net Present Value (NPV)**.

2.8.2 The Free Cash Flow (FCF)

Definition 2.8.3: Free Cash Flow (FCF)

The free cash flow is the excess operating cash that the company generates after accounting for capital expenditures such as buying land, building and equipment.

This is the cash that shareholders enjoy after accounting for the capital expenditures. The mark of a healthy business eventually depends on how much free cash it can generate.

Thus, the free cash is the amount of cash the company is left with after it has paid all its expenses including investments.

$$\text{FCF} = \text{Cash from Operating Activities} - \text{Capital Expenditures}$$

2.8.3 Key Steps of DCF Analysis

1. Estimate the average free cash flow
2. Identify the growth rate
3. Estimate the future cash flows

2.8.4 The Terminal Value

Definition 2.8.4: Terminal Growth Rate

The rate at which the company generates free cash flow grows beyond 10 years is called the terminal growth rate.

Note 2.3 Usually, the terminal growth rate is considered to be less than 5%

Definition 2.8.5: Terminal Value

The terminal value is the sum of all the future cash flow, beyond the 10th year, also called the terminal year.

The terminal value can be calculated by taking the cash flow of the 10th year and grow it at the terminal growth rate.

$$\text{Terminal Value} = \text{FCF} \times \frac{(1 + \text{Terminal Growth Rate})}{(\text{Discount Rate} - \text{Terminal Growth Rate})}$$

Note 2.4 The FCF used in the terminal value calculation is that of the 10th year.

2.8.5 The Share Price

The share price we will be talking about here is not the actual share price on the market but the 'intrinsic value' of the share that we wish to find (the valuation of the company).

Net Debt

Definition 2.8.6: Net Debt

$$\text{Net Debt} = \text{Current Year Total Debt} - \text{Cash \& Cash Balance}$$

Note 2.5 A negative sign indicates that the company has more cash than debt.

This value must be subtracted from the free cash flow to yield the **total present value of the free cash flow**.

$$\text{Share Price} = \frac{\text{Total Present Value of Free Cash Flow}}{\text{Total Number of Shares}}$$

2.8.6 Modelling Error & The Intrinsic Value Band

Though quite scientific, the DCF model makes a bunch of assumptions and hence would most likely lead to errors. Hence, we should accommodate for modelling errors.

One may allow ±10% leeway in the price.

Chapter 3

Technical Analysis

3.1 Overview

Technical Analysis is a research technique to identify trading opportunities in market based on the actions of market participants. The actions of markets participants can be visualized by means of a stock chart. Over time, patterns are formed within these charts and each pattern conveys a certain message.

> **Note 3.1** Technical Analysis (TA) is best used to identify short term trades. Do not use TA to identify long term investment opportunities.

3.2 Assumptions in Technical Analysis

1. **Markets discount everything**: All known and unknown information in the public domain is reflected in the latest stock price.

2. **Prices move in trends**: All major moves in the market are outcomes of trends.

3. **History tends to repeat itself**: In the TA context, the price trend tends to repeat itself. This happens because market participants consistently react to price movements in a remarkable similar way, each and every time the price moves in a certain direction.

3.3 Graphs

3.3.1 Line Chart

This is the most basic chart type and it uses only one data point to form the chart.

When it comes to technical analysis, a line chart is formed by plotting the **closing prices** of a stock or an index.

3.3.2 Bar Chart

This is a slightly more comprehensive chart as it display all the four price variables namely open, high, low, and close.

It has three main components:

1. **The central line**: The top of this line indicated the highest price the stock had reached while the bottom indicates the lowest price in the same period.

2. **The left mark/tick**: It indicates the open price

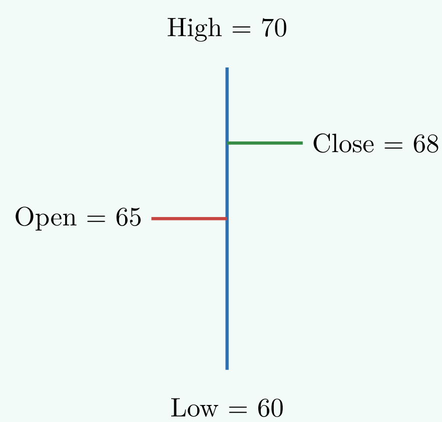
3. **The right mark/tick**: It indicates the close price

Example 3.3.1 (Bar Chart)

For example, assume the OHLC data for a stock is as follows:

Open	65
High	70
Low	60
Close	68

The bar chart for this will look like the following:



3.3.3 Candlestick Chart

This chart is very similar to the bar chart with the difference being that the open and close prices are shown using a rectangular body instead of wicks on the left and right.

A candlestick chart is classified as a bullish or bearish candle usually represented by blue/green/white and red/black candles respectively.

It has three main components:

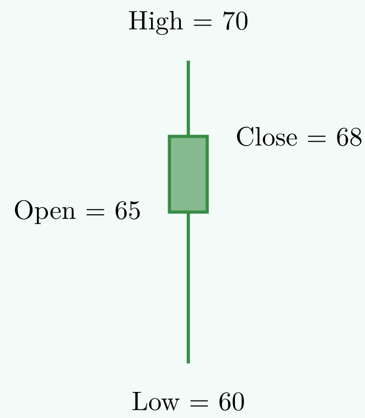
1. **The central real body**: The real body, rectangular in shape, connects the opening and closing price.
2. **Upper shadow**: Connects the high point to the close.
3. **Lower shadow**: Connects the low point to the open.

Example 3.3.2 (Candlestick Chart)

For example, assume the OHLC data for a stock is as follows:

Open	65
High	70
Low	60
Close	68

The candlestick for this will look like the following:



Note 3.2 One needs to pay some attention to the length of the candle while trading based on candlestick patterns. One should avoid trading based on subdued short candles.

3.4 Single Candlestick Patterns

3.4.1 The Marubuzo

Definition 3.4.1: The Marubuzo

It is defined as a candlestick with no upper and lower shadow. It has just the real body.

1. Bullish Marubuzo: Open = Low and Close = High
2. Bearish Marubuzo: Open = High and Close = Low

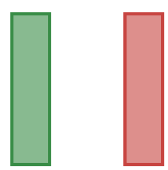


Figure 3.1: Bullish and Bearish Marubuzo

A **risk taker** would buy the stock in the same time interval in which the marubuzo occurred. Obviously, one needs to validate if the stick will be a marubuzo. This can be easily done by checking if the close price is approximately equal to the high price and the opening price is approximately equal to the low price just a few moments before the interval of the candle stick ends.

While a **risk averse** trader would buy the stock in the next interval right after the marubuzo occurs.

3.4.2 The Spinning Top

The spinning top candlestick can be described as follows:

- These candles have a small real body.
- The upper and lower shadows are almost equal.

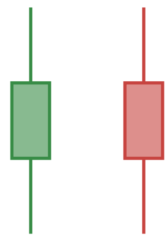


Figure 3.2: Bullish and Bearish Spinning Tops

Significance of the upper and lower shadow

1. The presence of the upper shadow tells us that the bulls did attempt to take the market higher. However, they were not really successful.
2. The presence of the lower shadow tells us that the bears did attempt to take the market lower. However, they were not really successful.

> **Note 3.3** Looking at a spinning top in isolation does not mean much. It just conveys indecision as both bulls and bears were not able to influence the markets. However when you see the spinning top with respect to the trend in the chart it gives out a really powerful message based on which you can position your stance in the markets.

Spinning tops in a downtrend

In a down trend, the bears are in control as they manage to take the price lower. However, with the spinning top in a down trend, the bears cold be consolidating their position before resuming another round of selling or the bulls could have arrested the price fall and have tried to hold onto their position.

Spinning tops in a uptrend

This case is very similar to what happens in a downtrend but reversed. That is, the bulls are in control, however, now they might be going for another buying round or the bears might have entered and are trying to make the prices fall but are unsuccessful.

3.4.3 Paper Umbrella

Definition 3.4.2: Paper Umbrella Candlestick

To qualify a candle as a paper umbrella, the length of the lower shadow should be at least **twice the length of the real body**.

Different types of paper umbrella candlesticks:

1. If the paper umbrella appears at the bottom end of a downward rally, it is called the **Hammer**.
2. If the paper umbrella appears at the top end of a uptrend rally, it is called the **Hanging man**.

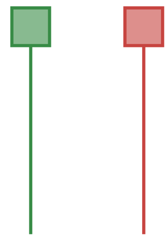


Figure 3.3: Bullish and Bearish Paper Umbrellas

Note 3.4 The hammer or hanging man can be of any color as it does not really matter as long as it qualifies 'the shadow to real body' ratio. However, it is slightly more comfortable to see a green and red colored real body respectively.

3.4.4 The Shooting Star

Definition 3.4.3: Shooting Star

The shooting star looks just like an inverted paper umbrella. Hence, the shooting star does not a long lower shadow. Instead, it has a long upper shadow where the length of the upper shadow is at least twice the length of the real body.

Note 3.5 Just like the paper umbrellas, the color of the shooting star does not matter though the pattern is slightly more reliable if the real body is red.

- The longer the upper wick, the more bearish the pattern is.
- The shooting star is a bearish patter; hence the prior trend should be bullish.

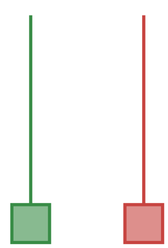


Figure 3.4: Bullish and Bearish Shooting Stars

3.5 Multiple Candlestick Patterns

3.5.1 The Engulfing Pattern

The engulfing pattern need two candlesticks, In a typical engulfing pattern, you will find a small candle followed by a relatively long candle which appears to engulf the smaller one.

- If the engulfing pattern appears at the bottom of the trend, it is called the **Bullish Engulfing**.
- If the engulfing pattern appears at the top of the trend, it is called the **Bearish Engulfing**.

Bullish Engulfing Pattern

The prerequisites for this pattern are as follows:

1. The prior trend must be a downtrend.
2. The first stick of the pattern (P1) should be a red candle reconfirming the bearishness in the market.
3. The 2nd candle of the pattern (P2) should be a green candle, long enough to engulf the red candle.

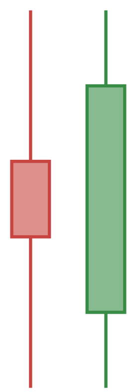


Figure 3.5: Bullish Engulfing Pattern

Bearish Engulfing Pattern

The prerequisites for this pattern are as follows:

1. The prior trend must be a uptrend.
2. The first stick of the pattern (P1) should be a green candle reconfirming the bullishness in the market.
3. The 2nd candle of the pattern (P2) should be a red candle, long enough to engulf the green candle.

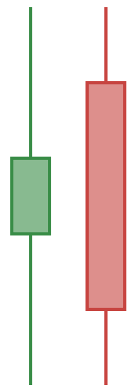


Figure 3.6: Bearish Engulfing Pattern

> **Note 3.6** The bearish engulfing pattern suggests a short trade.

3.5.2 The Piercing Pattern

The piercing pattern is very similar to the bullish engulfing pattern with a very minor variation. In a bullish engulfing pattern the P2's blue candle engulfs P1's red candle completely. However in a piercing pattern P2's blue candle partially engulfs P1's red candle, however the engulfing should be between 50% and less than 100%.

3.5.3 The Dark Cloud Cover

The dark cloud cover is very similar to the bearish engulfing pattern with a minor variation. In a bearish engulfing pattern the red candle on P2 engulfs P1's blue candle completely. However in a dark cloud cover, the red candle on P2 engulfs about 50 to 100% of P1's blue candle. The trade set up is exactly the same as the bearish engulfing pattern.

> **Tip 3.1** Think about the dark cloud cover as the inverse of a piercing pattern.

3.5.4 The Harami Pattern

The Bullish Harami

It is a bullish pattern appearing at the bottom end of the chart. It is similar to the engulfing pattern.

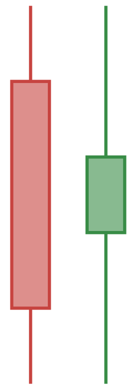


Figure 3.7: Bullish Harami Pattern

The Bearish Harami

It is a bearish pattern appearing at the top end of the chart. It is similar to the engulfing pattern. It presents the trader with a opportunity to initiate a short trade.

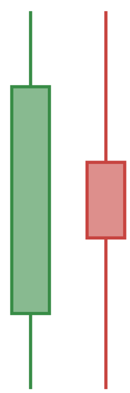


Figure 3.8: Bearish Harami Pattern

3.5.5 The Gaps

Gap Up Opening

It indicated buyer's enthusiasm. Buyers are willing to buy stocks at a price higher than the previous day's close. Hence, because of enthusiastic buyer's outlook, the stock (or the index) opens directly above the previous day's close.

Gap Down Opening

Similar to gap up opening, a gap down opening shows the enthusiasm of the bears. The bears are so eager to sell, that they are willing to sell at a price lower than the previous day's close.

3.5.6 The Morning Star

The morning star is a bullish candlestick pattern which evolves over three periods (i.e. a three candlestick pattern). It is a downtrend reversal pattern. The morning star appears at the bottom of a downtrend.

The conditions for a morning star are:

1. P1 should be a long red candle.
2. With a gap down opening, P2 should be either a doji or a spinning top.
3. P3 opening should be gap up, plus the closing price should be higher than the opening of P1.

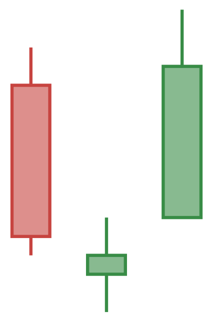


Figure 3.9: The Morning Star Pattern

> **Note 3.7** A stop loss order for a trade made on the basis of a morning star should be put at the value equal to the lowest price among P1, P2 & P3

3.5.7 The Evening Star

The evening star is a bearish equivalent of the morning star. The evening star appears at the top end of an uptrend. Like the morning star, the evening star is a three candle formation and evolves over three trading sessions.

The conditions for an evening star are:

1. P1 should be a long blue candle.
2. With a up down opening, P2 should be either a doji or a spinning top.
3. P3 opening should be gap down, plus the closing price should be lower than the opening of P1.

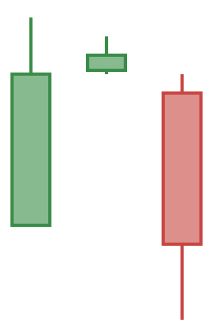


Figure 3.10: The Morning Star Pattern

Note 3.8 A stop loss order for a trade made on the basis of an evening star should be put at the value equal to the highest price among P1, P2 & P3

3.6 The Support and Resistance

Definition 3.6.1: The Resistance

As the name suggests, resistance is something which stops the price from rising further. The resistance level is a price point on the chart where traders expect maximum supply (in terms of selling) for the stock/index. The resistance level is **always above the current market price**.

Definition 3.6.2: The Support

As the name suggests, the support is something that prevents the price from falling further. The support level is a price point on the chart where the trader expects maximum demand (in terms of buying) coming into the stock/index. Whenever the price falls to the support line, it is likely to bounce back. The support level is **always below the current market price**.

3.7 Volume

Definition 3.7.1: Volume

Volumes indicate how many shares are bought and sold over a given period of time. The more active the share, higher would be its volume.

3.7.1 Volume Trends

Price	Volume	Expectation
Increases	Increases	Bullish
Increases	Decreases	Caution - weak hands buying
Decreases	Increases	Bearish
Decreases	Decreases	Caution - weak hands selling

Important 3.1 As a practice, traders usually compare the current session's volume over the average of the last 10 trading sessions.

$$\text{current volume} > \text{average volume} \Rightarrow \text{high volume}$$
$$\text{current volume} = \text{average volume} \Rightarrow \text{average volume}$$
$$\text{current volume} < \text{average volume} \Rightarrow \text{low volume}$$

3.7.2 Thought Process behind Volume Trends

When institutional investors buy or sell they obviously do not transact in small chunks. They buy very huge chunks. Now, if they were to buy a lot of shares from the open market, it will start reflecting in volumes. Besides, because they are buying a large chunk of shares, the share price also tends to go up.

Usually institutional money is referred to as the *smart money*. It is perceived that **smart money always makes wiser moves in the market compared to retail traders**. Hence following the smart money seems like a wise idea.

3.8 Moving Averages

3.8.1 Simple Moving Averages

In this method, we give equal importance to all the data points being considered.

$$\text{SMA} = \sum_{i=0}^{N} \text{value}$$

where N is the number of data points being considered.

3.8.2 Exponential Moving Averages (EMA)

In this method, we give higher importance to the newer data points and lesser to the older ones.

$$\text{EMA}_i = (\text{value} \times K) + (\text{EMA}_{i-1} \times (1 - K))$$

where EMA_0 which is the EMA for the first period is taken to be equal to the SMA for that period and K is multiplier constant which is used to smoothen the curve. It can usually be calculated by the following formula:

$$K = \frac{2}{\text{number of observations} + 1}$$

Example 3.8.1 (Using the SMA/EMA to make trades)

1. Buy (go long) when the current market price turns greater than the 50 day SMA/EMA. Once you go long, you should stay invested till the necessary sell condition is satisfied.
2. Exit the long position (square off) when the current market price turns lesser than the 50 day SMA/EMA.

3.8.3 Moving Average Crossover System

In this system, instead of the usual single moving average, we combine two moving averages. This is referred to as *smoothing*.

Example 3.8.2 (50 Day EMA + 100 Day EMA)

A typical example of this would be to combine a 50 day EMA, with a 100 day EMA. The shorter moving average (50 days in this case) is also referred to as the *faster moving average*. The longer moving average (100 days moving average) is referred to as the *slower moving average*.

Entry and Exit Rules for the Crossover System:

1. Buy long when the short term moving average turns greater than the long term moving average. Stay in the trade as long as this condition is satisfied.
2. Exit the long position when the short term moving average turns lesser than the longer term moving average.

3.9 Indicators

3.9.1 Overview

A technical indicator helps a trader analyze the price movement of a security. Indicators are built on preset logic using which traders can supplement their technical study (candlesticks, volumes, S&R) to arrive at a trading decision. Indicators help in buying, selling, confirming trends, and sometimes predicting trends.

Indicators are of two types namely **leading** and **lagging**.

Leading Indicators

A leading indicator leads the price, meaning it usually signals the occurrence of a reversal or a new trend in advance.

> **Important 3.2** Leading indicators are notorious for giving false signals. Therefore, the trader should be highly alert while using leading indicators.

A majority of leading indicators are called oscillators as they oscillate within a bounded range.

Lagging Indicators

A lagging indicator on the other hand lags the price; meaning it usually signals the occurrence of a reversal or a new trend after it has occurred.

One of the most popular indicators is the moving averages.

3.9.2 Momentum

Definition 3.9.1: Momentum

Momentum is the rate at which the price changes.
For example if stock price is Rs.100 today and it moves to Rs.105 the next day, and Rs.115, the day after, we say the momentum is high as the stock price has changed by 15% in just 3 days. However if the same 15% change happened over let us say 3 months, we can conclude the momentum is low. So the more rapidly the price changes, the higher the momentum.

3.9.3 Relative Strength Index (RSI)

RSI is a leading momentum indicator which helps in **identifying a trend reversal**.

> **Note 3.9** The term *Relative Strength Index* can be a bit misleading as it does not compare the relative strength of two securities, but instead shows the internal strength of the security.

The objective of using RSI is to help the trader identify over sold and overbought price areas. Overbought implies that the positive momentum in the stock is so high that it may not be sustainable for long and hence there could be a correction. Likewise, an oversold position indicates that the negative momentum is high leading to a possible reversal. RSI gives out the strongest signals during the periods of sideways and non-trending ranges.

$$\mathrm{RSI} = 100 - \frac{100}{1 + \mathrm{RS}}$$

$$\mathrm{RS} = \frac{\text{Average Gain}}{\text{Average Loss}}$$

Definition 3.9.2: Look-back Period

The data points used for calculating the RSI determines the *look-back period*.
For example, if one is using daily price data and uses 14 data points for calculating the averages in the RS formula, the look-back period would be 14 days.

Classical Interpretation of RSI

- When the RSI is between 0 and 30, the security is supposed to be oversold and is ready for an upward correction.
- When the RSI is between 70 and 100, it is supposed to be heavily bought and is ready for a downward correction.

Modern Interpretation of RSI

- If the RSI is fixed in an overbought region (0 to 30) for a prolonged period, look for buying opportunities instead of shorting.
- > **Note 3.10** The RSI stays in the overbought region for a prolonged period because of an excess positive momentum.

  If the RSI is fixed in an oversold region for a prolonged period, look for selling opportunities rather than buying.

  > **Note 3.11** The RSI stays in the oversold region for a prolonged period because of an excess negative momentum.

  If the RSI value starts moving away from the oversold value after a prolonged period, look for buying opportunities.

  > **Example 3.9.1**
  > The RSI moving above 30 after a long time may mean that the stock may have bottomed out, hence a case for going long

- If the RSI value starts moving away from the overbought value after a prolonged period, look for selling opportunities.

  > **Example 3.9.2**
  > The RSI moving below 70 after a long time may mean the stock has topped out, hence a case for shorting.

3.9.4 Moving Average Convergence and Divergence (MACD)

MACD is all about convergence and divergence of two moving averages. Convergence occurs when the two moving averages move towards each other, and a divergence occurs when the moving averages move away from each other.

A standard MACD is calculated using a 12 day EMA and a 26 day EMA. We subtract the 26 day EMA from the 12 day EMA, to estimate the convergence and divergence (CD) value. A simple line of this graph is often referred to as the *MACD line*.

Significance of Values of MACD

- The sign of the MACD just indicates the direction of the stock's movement.
- The higher the magnitude of the MACD, the higher is the momentum.

Example 3.9.3

For example if the 12 Day EMA is 6380, and 26 Day EMA is 6220 then the MACD value is +160. We also know that the shorter term average will generally be higher than the longer term only when the stock price is trending upwards. Hence **a positive MACD value indicates that the price is moving upwards!**

Point of Convergence and Divergence

- When the MACD Line crosses the center line from the negative territory to positive territory, it means there is divergence between the two averages. This is a sign of increasing bullish momentum; therefore one should look at buying opportunities.
- When the MACD line crosses the center line from positive territory to the negative territory it means there is convergence between the two averages. This is a sign of increasing bearish momentum; therefore one should look at selling opportunities.

The Signal Line

Traders generally argue that while waiting for the MACD line to crossover the center line a bulk of the move would already be done and perhaps it would be late to enter a trade. To overcome this, there is an improvisation over this basic MACD line. The improvisation comes in the form of an additional MACD component which is the *9 day signal line*. **A 9 day signal line is a exponential moving average (EMA) of the MACD line.**

With these two lines (the MACD line and the signal line), a trade can follow a simple 2 line crossover strategy similar to the crossing over of 2 different moving averages and no longer wait for the center line cross over.

- The sentiment is bullish when the MACD line crosses the 9 day EMA wherein MACD line is greater than the 9 day EMA. When this happens, the trader should look at buying opportunities.
- The sentiment is bearish when the MACD line crosses below the 9 day EMA wherein the MACD line is lesser than the 9 day EMA. When this happens, the trader should look at selling opportunities.

3.9.5 The Bollinger Bands (BB)

BBs are used to determine overbought and oversold levels, where a trader can try to sell when the price reaches the top of the band and try to buy when the price reached the bottom of the band.

The BB has 3 components:

1. **The Middle Line**: It is a 20 day SMA of the closing averages.
2. **The Upper Band**: It is a +2 standard deviation of the middle line.
3. **The Lower Band**: It is a -2 standard deviation of the middle line.

Note 3.12 The standard deviation (SD) is a statistical concept; which measures the variance of a particular variable from its average. In finance, the standard deviation of the stock price represents the volatility of a stock.

For example, if the standard deviation of a stock is 12%, it is as good as saying that the volatility of the stock is 12%.

Chapter 4

Risk Management & Portfolio Optimization

4.1 Risk

4.1.1 Unsystematic Risk

Definition 4.1.1: Unsystematic Risk

At any given point, the drop in the stock price can only be attributed to company specific factors or internal factors. The risk of losing money owning to these factors is termed as *Unsystematic risk*.

Unsystematic risk can be diversified, meaning instead of investing all the money in one company, one can choose to invest in 2-3 different companies (preferably from different sectors).

4.1.2 Diversification

Definition 4.1.2: Diversification

When one chooses to invest their money in multiple companies instead of a single one, they are said to be diversifying their portfolio. The different companies may be from the same or different sectors.

Tip 4.1 Unsystematic risk can be drastically reduced by diversifying investments especially when diversified into various sectors.

The higher the number of stocks in your portfolio, higher the diversification, and therefore lesser the unsystematic risk. This leads to a very important question - how many stocks should a good portfolio have so that the unsystematic risk is completely diversified.

Research has it that up to 21 stocks in the portfolio will have the required necessary diversification effect and anything beyond 21 stocks may not help much in diversification.

The risk vs number of stocks graphs follows an exponential decay pattern i.e. at the start as one adds more stocks, the risk reduces drastically but after about 20 stocks, the risk barely reduces on adding more stocks.

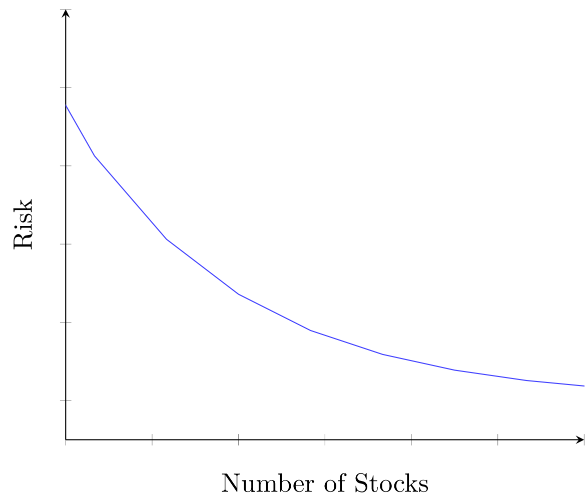


Figure 4.1: Risk vs Number of Stocks

4.1.3 Systematic Risk

Definition 4.1.3: Systematic Risk

Systematic risk is the risk that is common to all stocks in the markets. Systematic risk arises out of common market factors such as the macroeconomic landscape, political situation, geographical stability, monetary framework etc.

A few specific systematic risks which can drag the stock prices down are: –

1. De-growth in GDP
2. Interest rate tightening
3. Inflation
4. Fiscal deficit
5. Geopolitical risk

Systematic risk is inherent in the system and it **cannot really be diversified**. However, systematic risk can be *hedged*.

4.2 Expected Return

Expecting a realistic return plays a pivotal role in investment management. The expected return can be calculated using the following formula:

$$\mathrm{E} = \sum_{i=1}^{n} W_i R_i$$

where E is the expected return of the portfolio, W_i is the weightage of the investment in the entire portfolio and R_i is the expected return of that individual asset.

Example 4.2.1 (Calculating expected return)

Let's say that you invested 25,000 in stock A and expect a return of 20% and invest 25,000 in stock B while expecting 15% return.
Therefore, the expected return of the entire portfolio would be:

$$E = \frac{25,000}{25,000 + 25,000} \times 20 + \frac{25,000}{25,000 + 25,000} \times 15 \qquad = \frac{1}{2} \times 20 + \frac{1}{2} \times 15 = 17.5\%$$

4.3 Variance & Covariance

4.3.1 Variance

Definition 4.3.1: Variance

The variance of stock returns is a measure of how much a stock's return varies with respect to its average daily returns.

The formula to calculate variance is:

$$\sigma^2 = \sum \frac{(x - \mu)^2}{N}$$

where σ^2 is the variance, X is the daily return, μ is the average of the daily return and N is the total number of observations.

As an investor, one should look into the variance to determine the riskiness of the investment. A large variance indicates that the stock could be quite risky while a small variance can indicate lesser risk.

4.3.2 Covariance

Definition 4.3.2: Covariance

Covariance indicates how two (or more) variables move together. It tells us whether the two variables move together (in which case they share a positive covariance) or they move in the opposite direction (negatively covariance).

Covariance in the context of stock market measures how the stock prices of two stocks (or more) move together. The two stocks prices are likely to move in the same direction if they have a positive covariance; likewise, a negative covariance indicates that they two stocks move in opposite direction.

Note 4.1 *Covariance* may sound similar to *correlation*, however, the two are different.

Covariance of two stocks S_1 and S_2 can be calculated as follows:

$$\text{Covariance} = \sum \frac{\left((x_{S_1} - \mu_{S_1}) \times (x_{S_2} - \mu_{S_2})\right)}{n - 1}$$

where X_{S_i} is the daily return of stock S_i, μ_{S_i} is the average return of stock S_i and n is the total number of days.

Important 4.1 Portfolio managers strive to select stocks which share a negative covariance. The reason is quite simple – they want stocks in the portfolio which can hold up. Meaning if one stock goes down, they want, at least the other to hold up. This kind of counter balances the portfolio and reduces the overall risk.

4.3.3 Variance & Covariance Matrix

Covariance can only be calculated between 2 entities. In any ordinary portfolio, there would most likely be more that 2 stocks. In such a case, one would have to calculate the covariance between all of the stocks taking two at a time.

For example, in a portfolio with 4 stocks, one would have to calculate 6 covariances. (Covariance between A & B is the same as between B & A).

To organize this information properly, we generally calculate and tabulate them in a *Variance - Covariance Matrix*.

However, as we will see, this variance-covariance matrix alone does not convey much information and we will develop another matrix called the *correlation matrix* to help us out. Using this correlation matrix, we will calculate the *portfolio variance*.

The size of the variance-covariance matrix for a portfolio with k stocks will be $k \times k$. The formula for the elements of a variance-covariance matrix is:

$$\sum_{k \times k} = \frac{1}{n} X^T X$$

where n is the number of observations, X is the excess return matrix $(n \times k)$.

Definition 4.3.3: Excess Return Matrix (X)

Excess return matrix is defined as the difference between stock's daily return over its average return i.e. $x - \mu$

Note 4.2 This variance-covariance matrix is a symmetric matrix because the covariance between A & B is the same as between B & A.

Important 4.2 The principle diagonal values are the covariance of the entity with itself. This is the same as the variance of that entity. Therefore, the principle diagonal of the variance-covariance matrix contains the variances of each of the entities while the other entries contain the covariances between the two entities according to the row and column!

4.3.4 Correlation Matrix

The correlation of two stocks S_1 and S_2 is expressed in terms of their covariance and their standard deviations ($\sqrt{\text{Variance}}$) as follows:

$$\text{Correlation} = \frac{Cov(S_1, S_2)}{\sigma_{S_1} \times \sigma_{S_2}}$$

Using this formula, we can convert a variance-covariance matrix into a correlation matrix!

Note 4.3 The diagonal entries of a correlation matrix are all equal to 1!

4.3.5 Portfolio Variance

4.3.6 Overview

The first step in calculation portfolio variance is to assign weights to the stocks in the portfolio according to the amount of money invested in each stock.

The next step is to calculate the *weighted standard deviation*.

Definition 4.3.4: Weighted Standard Deviation

It is simply the weight of a stock multiplied by its respective standard deviation.

Note 4.4 The total weight should add up to 100% i.e. the sum of the individual weights of the stocks will add up to 100% but the weighted standard deviation will not!

Finally, we can calculate the *Portfolio Variance* using the following formula:

$$\text{Portfolio Variance} = \sqrt{\text{Wt. SD}^T \times \text{Correlation Matrix} \times \text{Wt. SD}}$$

where Wt. SD is the array of weighted standard deviations.

Note 4.5 Risk or variance or volatility is like a coin with two faces. Any price movement below our entry price is called risk while at the same time, the same price movement above our entry price is called return.

4.3.7 Equity Curve

In a very lose sense, a typical equity curve helps you visualize the performance of the portfolio on a normalized scale of 100. In other words, it will help you understand how Rs.100/- invested in this portfolio would have performed over the given period.

Basically, an *equity curve* can be developed if we plot the portfolio value at a fixed interval (for example, daily).

Tip 4.2 The variance of the portfolio can hence be also calculated as the standard deviation of these values of the equity curve in a fixed interval!

4.4 Expected Returns

Definition 4.4.1: Expected Return

The expected return of the portfolio is the sum of the average return of each stock multiplied bu its weight and further multiplied by 252 (the number of trading days in a year). In simple terms, we are scaling the daily returns to its annual returns and according to the investment made (the weights).

Note 4.6 We can also scale the portfolio variance to represent annual variance by multiplying it by $\sqrt{252}$.

Portfolio returns are *normally* distributed. If we plot the distribution of a portfolio, we are likely to get a normally distributed portfolio. And if the portfolio is normally distributed, then we can estimate the likely return of this portfolio over the next year with a certain degree of confidence.

$$\text{Expected return} = \text{Expected annualized return (calculated before)} \pm \text{annualized portfolio variance}$$

The accuracy of this broadly varies across three levels:

1. Level 1 - one standard deviation away: 68% confidence
2. Level 2 - two standard deviation away: 95% confidence
3. Level 3 - three standard deviation away: 99% confidence

Important 4.3 Variance is measured in terms of standard deviation. So it is important to note that the annualized portfolio variance calculated above is just 1 standard deviation!

4.5 Portfolio Optimization

Generally speaking, portfolio optimization is finding the best set of weights for a given set of stocks in the portfolio which maximizes returns while also minimizing risk!

Definition 4.5.1: Minimum Variance Portfolio

There will be a combination of weights possible such that the variance of the portfolio is minimum. This particular portfolio is also referred to as the "Minimum Variance Portfolio". The minimum variance portfolio represents the least amount of risk you can take.

Definition 4.5.2: Maximum Return Portfolio

Similar to the minimum variance portfolio, this portfolio is made using weights so that the returns are maximized!

Note 4.7 For a fixed level of risk/variance of a portfolio, we can create **at least two unique** portfolios. One such portfolio will yield the highest return among them and one will yield the lowest.

4.5.1 Efficient Frontier

The scatter plot of the return vs risk data where for each risk percentage, maximum and minimum returns are calculated is known as the *efficient frontier*.

4.6 Value at Risk

Value at Risk (VaR) is a metric which gives an individual a sense of the worst case loss they would incur, if the most unimaginable were to occur the next morning.

At the core of this approach lies the concept of normal distribution. We will discuss a 'quick and dirty' method to calculate *VaR* now.

The steps involved in calculating the portfolio VaR are:

1. Identify the distribution of the portfolio returns.
2. Map the distribution (check if it is a normal distribution).
3. Arrange portfolio returns in ascending order.
4. The least value within the last 95% entries is the portfolio VaR while the average of the last 5% is the cumulative VaR or CVaR

Note 4.8 We have taken only 95% of the data because the distribution being a normal distribution, we can expect 95% of the data to be less than 2 SD away from the expected returns!

Example 4.6.1

If we calculated the VaR to be -1.48%, we can expect that the worst case loss of our portfolio would be -1.48% with a confidence of 95%!

4.7 Position Sizing

Position sizing is all about answering how much capital you will expose to a particular trade given that you have 'x' amount of trading capital.

One classic position sizing strategy which most people employ is the standard 5% rule. The 5% rule does not permit you to risk more than 5% of the capital on a given trade.

4.7.1 Estimating Equity Capital

Essentially there are three main techniques or models out there to estimate the current equity capital one has.

1. **Core Equity Model**: This model requires you to deduct the capital allocated to a trade from the existing capital. This way, the exposure to a trade goes on reducing as you ladder up more and more positions.
2. **Total Equity Model**: This model aggregates all the positions in the market along with its respective P&L and cash balance to estimate the equity. That is, free cash along with the margins blocked and the P&L per position is taken into consideration.
3. **Reduced Total Equity Model**: This model is a combination of both the others. It takes into account not only the remaining cash, but also the locked in profits.

Example 4.7.1 (Reduced Total Equity Model)

Let us take an initial corpus of 1,00,000/- and let us set the maximum allowed position of 10% of my corpus. For the first trade, I am allowed to spend a maximum of 10,000/-. Let's say I spend all of that. I am now having a corpus of 90,000/- which gives me a limit of 9,000/-. This is exactly like the *Core Equity Model*. Now let's say that the stock I had bought is up by 5%. What I can do now is put a stop-loss order at this price which ensures that I will get a minimum of this amount and then add this to my corpus while calculating the maximum position for my next trade!

4.7.2 Methods for Position Sizing

Unit per fixed amount

This is one of the simplest models to calculate position size. It requires you to simply state how many shares or lots (in case of futures) you will trade for a given amount.

Example 4.7.2 (Unit per fixed amount position sizing model)

You may setup a position sizing strategy as buying up to 1 lot for every 1,00,000 of capital.

Note 4.9 This **does not factor in the risk** of the investments.

Percentage Margin

This technique requires the trader to position size based on the margins. The maximum amount of margins required for a trade should be set as a percentage of the total capital (calculated using one of the equity estimation models).

This ensures that the trader pays roughly equal margin to all the position.

However, volatility from each position could vary and the trader can end up with risky bets and therefore alters the entire risk profile of the portfolio.

Percentage Volatility

The percentage volatility model accounts for volatility of the underlying asset. The volatility as per this technique is not really the *standard deviation*, but rather the daily expected movement in the underlying.

Example 4.7.3 (Volatility i.e. daily expected movement)

For example, if SBI's OHLC is 276, 279, 274, and 278, then the volatility for the day is simply the difference between low and high i.e $279 - 274 = 5$.

To get a sense of the generic volatility measured this way, one can look at the difference between low and high for last n days and take an average. However, the only problem here would be that we would be ignoring the gap up and gap down openings. For this reason, Van Tharp suggest the use of **Average True Range** to measure the stock's volatility.

The *Percentage Volatility* method of position sizing requires us to define the maximum amount of volatility exposure one can assume for the given equity capital.

> **Example 4.7.4** (Percentage volatility position sizing model)
> For example, if the equity capital is Rs.500,000/- then I could make a rule saying that I do not want to expose more than 2% of the capital to volatility.

Percentage Risk

The percentage risk method, relies upon your own assessment of *loss* that you are willing to bear for a given trade. This, as you may know is also called the *Stop loss* for the trade. The stop loss for a trade is the price at which you decide to close the trade and take a hit. The percentage risk technique controls the position size as a function of risk defined by stop loss.

As a thumb rule, professional traders do not risk more than 1 to 3% of their capital on any single trade.

> **Example 4.7.5** (Percentage risk position sizing model)
> Let us take an example of a future of stock. Let the total capital be $5,00,000$, the lot size be 1500, the trade price be 393.65, the target price be 400 and the stop loss price be 390. Therefore, the target value comes out to be 6.35 and the stop loss value comes to be 3.65. Let the margin required to trade one lot be $73,500$.
> Technically, I can buy $5,00,000/73,500 = 6.8$ lots. But that would mean I could lose up to $3.65 \times 1500 \times 6 = 32,850$ on this trade. This is $\frac{32,850}{5,00,000} \times 100 = 6.57\%$ of my entire capital on one trade. This is not a good thing.
> So let me setup a maximum risk per trade as a percentage of my overall capital - say 1.5%. This means that I can take a maximum loss of $1.5\% \times 5,00,000 = 7,500$ on one trade.
> Since I can lose $3.65 \times 1,500 = 5,475$ per lot of my selected stock, I can buy a maximum of $7,500/5,475 = 1.36$ lots of that stock to stay in my limit.

Kelly's Criterion

The Kelly's Criterion essentially helps us estimate the optimal bet size (or the fraction of our trading capital) considering –

- We have a certain information on the bet we are about to take.
- We have an edge taking that particular bet.

The Kelly's Criterion is an equation, the output of which is a percentage, also known as the Kelly's percent. The equation is as below:

$$\text{Kelly }\% = \left(W - \frac{1-W}{R}\right) \times 100$$

where W is the winning probability and R is the win-loss ratio.

> **Definition 4.7.1: Winning Probability**
> The winning probability is defined as the total number of winning trades divided over the total number of trades.

> **Definition 4.7.2: Win-Loss Ratio**
> The win-loss ratio is the average gain of winning trades divided over average loss of the negative trades.

This Kelly's percentage is the maximum percentage of the capital you should allow on that trade.

However, the Kelly's formula can spit out large percentages such as 70%. We would obviously not want to risk such a large chunk of capital on a single trade.

To mitigate this problem, we can mix this strategy with the percentage risk strategy. We can setup a higher maximum risk percentage such as 5%. Then the actual percentage risk we use in the trade would be the max percentage risk multiplied by the Kelly's percentage.

Example 4.7.6 (Kelly's percentage + percentage risk model)

Let us setup the maximum percentage risk as 5%. For a particular trade, let us say that the Kelly's percentage comes out to be 70%. Then we can calculate the position size by taking the percentage risk as 70% of 5% = 1.5%!

Chapter 5

Time Series Analysis

5.1 Introduction

5.1.1 How Should We Not Predict Stocks

There are many YT videos predicting future stock prices from past data. Many of them also use concepts such as RNNs, LSTMs, and CNNs. Unfortunately, most of them are useless due to what kind of prediction they are doing.

5.1.2 What We Want

Given the past entries, we want to predict the future values. Note the plural form. Let's say, our model is a black-box for now, which takes the entries as the last 'n' samples, and predicts the next 5 samples.
Input: $y_0, y_1, y_2, \ldots, y_{n-1}$
Output: $y_n, y_{n+1}, y_{n+2}, \ldots$

5.1.3 What happens in the YT videos

There, the input is the same, however, the output is generally just one value, the prediction at the next time step.
Input: $y_0, y_1, y_2, \ldots, y_{n-1}$
Output: y_n

However, you must have seen in their thumbnails, that their predictions and the actual values are really close. However, this is SCAM. In fact, what is happening here, is that they divide their dataset into train and validation sets. The validation set, also includes values from the future as inputs! This is exactly why these are useless. Instead of predicting based on the 'estimate'; of the future value, the actual future value is used to predict the next value.

In fact, on a closer look, you'll notice that what these models are actually doing is just copying the previous timestep's value and outputting that as the prediction. This is why the 'predictions' look so close to the real values(they are the same, just delayed by one timestep).

Now, you might wonder what if we use the predicted values again as inputs to predict further values in the future. Sure, this can be done. Unfortunately, the answer is always almost a straight horizontal line, which indicates the last known price in your train set.

This is a glimpse into why predicting stock price movements is hard, because of the Random Walk Hypothesis.

5.1.4 Random Walk Hypothesis

According to this hypothesis, the stock prices follow a random walk.

$$y_t = y_{t-1} + \epsilon$$

The ϵ is a noise random variable, generally assumed to be Gaussian. It is 0 mean.

Based on this, we can never predict the next value in the series, as we cannot predict the noise of the term.

Furthermore, when thinking of stock prices as random variables (with appropriate probability spaces), it can be seen that the expected value of the future value is the same as the current value. This is called the ***naive forecast**.* Thus, this completely destroys any prediction capabilities.

Thankfully, this is just a model and is not completely accurate. The stock prices don't follow a pure random walk, as seen by **volatility clustering**.

Definition 5.1.1: Volatility Clustering

It's a common observation that the future stock price also reflects that volatility, i.e., if the stock is volatile today, it'll probably be volatile tomorrow as well. Similarly, if it's relatively stable today, it'll probably probably be stable tomorrow as well. This is called volatility clustering, as the volatilities tend to cluster together.

5.1.5 What Should We Predict

A common misconception would be that we should obviously predict stock prices. But the problem with stock prices is that they are not stationary, as in, they do not come from the same distribution.

This is a problem. For our model to function properly, we must test it on data that comes from the same distribution it is trained on!

For example, if a model is trained to predict the prices of apples, we cannot use it to predict the prices of cars.

This is why, in most of finance, we work with returns instead of prices. In fact, we work with *log* returns.

Another reason to use returns is that they're scale invariant. It makes more sense to talk in terms of percentage change rather than the actual difference.

$$R_t = P_t/P_{t-1} - 1$$

R_t here is the return. Log returns are given by:

$$r_t = log(1 + R_t)$$

They are more useful than simple returns because we can take the simple 'average log return' to get the final price of the stock, which cannot be done using simple returns R_t.

5.1.6 Efficient Market Hypothesis

In simple words, this means that the current stock prices reflect all the available information in the market. Thus, one cannot beat the market regularly. There are various forms of this hypothesis, namely the Weak Form, the Semi-Strong Form, and the Strong Form.

5.2 Exploring Trends

One of the ways to explore trends is moving averages which we have already heavily discussed earlier. Another method is to use *Holt's Linear Trend Model*.

5.2.1 Holt's Linear Trend Model

First, let's revisit simple exponential smoothing. The basic model is:

$$\hat{y}_{t+1|t} = l_t l_t = \alpha y_t + (1 - \alpha) l_{t-1} \tag{5.1}$$

where l_t denotes the level, an estimate to capture the average, smoothened version of the series.

If we also add a trend term to this equation, we get Holt's linear trend model.

The model consists of three main parts:

$$\begin{aligned}
\text{Forecast equation: } & \hat{y}_{t+h|t} = l_t + hb + t \\
\text{Level equation: } & l_t = \alpha y_t + (1-\alpha)(t_{t-1} + b_{t-1}) \\
\text{Trend equation: } & b_t = \beta(l_t - l_{t-1} + (1-\beta)b_{t-1}
\end{aligned}$$

where h is the number of steps in the forecast.

The first equation is similar to that of a straight line. The second equation $(l_{t-1} + b_{t-1})$ is just $\hat{y}_{t+1|t}$. The last equation is the slope part of the first equation. We are tracking differences between the current and the previous level. This level l_t is an estimate of the smoothened version of the time series.

$$\text{Error} = \sum_{t=1}^{T} (y_t - \hat{y}_{t|t-1})^2$$

$$\alpha, \beta = \text{argmin Error}$$

5.2.2 Holt-Winter's Model

On adding yet another term to the equation, the seasonality of data, we get the Holt-Winter's Model:

$$\begin{aligned}
\text{Forecast equation: } & {}_{t+h|t} = l_t + hb_t + s_{t+h-m} \\
\text{Level equation: } & l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\
\text{Trend equation: } & b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\
\text{Seasonality equation: } & s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m};
\end{aligned}$$

The seasonality terms captures, well, the seasonality of data. If there is a repeating pattern in the time series, this model aims to capture it.

5.3 Statistical Models

5.3.1 AR(p)

AR stands for **Auto-Regressive**, which means that the current value is dependent on p past values. In a way, this is similar to linear regression.

$$\hat{y}_t = b + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \cdots + a_p y_{t-p}$$

Assuming this is similar to linear regression, where the noise in predictions is considered gaussian, we can write the above equivalently as:

$$y_t = b + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \cdots + a_p y_{t-p} + \epsilon_t$$

ϵ_t is a 0 mean, σ standard deviation Gaussian.

5.3.2 MA(q)

This stands for Moving Average. However, this is not to be confused with the Simple Moving Average or the Exponential Moving Average discussed earlier.

This too is a linear model, however, it's a linear combination of past error terms. This is not like any other model we've seen before.

$$y_t = c + \epsilon_t + b_1 \epsilon_{t-1} + b_2 \epsilon_{t-2} + \cdots + b_q \epsilon_{t-q}$$

> **Note 5.1** An *MA* model is said to be **invertible** if it is algebraically equivalent to a converging infinite order AR model. By converging, we mean that the AR coefficients decrease to 0 as we move back in time.

5.3.3 ARMA(p, q)

ARMA is simply a combination of AR(p) and MA(q) models.

$$y_t = b + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \cdots + a_p y_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + b_2 \epsilon_{t-2} + \cdots + b_q \epsilon_{t-q}$$

This is a pretty straightforward model, however, this wouldn't work if the time series we're working with isn't stationary! This is why, we introduce the ARIMA model.

5.3.4 ARIMAp, d, q

The *I* stands for *integrated*.

We had already seen what stationarity is. Now, if there was a way to make the time series stationary, we could simply use *ARMA*. This is exactly the philosophy behind *ARIMA*.

Time differencing a series, as it turns out, helps to make it stationary. Thus, 'd' is the number of times the series is differenced, before applying *ARMA*.

The *ARIMA* model works similar to an ML model. The parameters a and b are learnt as the solution to some optimization problem, generally minimizing the *RMSE*.

For finding the best values of the hyper-parameters p, q, and d, we will use some tests.

> **Note 5.2** We are talking about weak-sense stationarity here. In fact, strong sense stationarity isn't used all that often. All this means, is that the mean and the auto-covariance of the series, should be the same over all equal time intervals. If not, then the whole notion of 'the' mean or 'the' variance is meaningless, as they'd be functions of time!

5.3.5 ADF Test

The Augmented Dickey-Fuller test is designed to check if a time series is stationary or not.

This is a statistical test, in which a null hypothesis is either accepted or rejected, depending on the p-value obtained. This test can be thought of like an API, or a black-box. You input a time series, and a sufficiently low p-value (less than 0.05) would indicate that the series is stationary.

This can be used to to find the parameter 'd' of ARIMA.

We simply keep on differencing until the series becomes stationary!

Whenever we want to fit such models, we want our series to be stationary.

5.3.6 Auto Correlation Function (ACF)

Auto-correlation function is a method used to choose the parameter q of *ARIMA*.

In a nutshell, we input our time series into this function, and we get a plot. The plot consists of points and a margin. The number of values, starting from the origin, which lie above this margin, would be our q. Note, occasionally, due to pure chance, it's possible that a bunch of values are inside this margin and suddenly one jumps out, however, these are insignificant, as this can be attributed to rare events.

5.3.7 Partial Auto Correlation Function (PACF)

The *Partial Auto Correlation Function* is used to select the AR parameter p.

> **Tip 5.1** For a lot of stock data, the best ARIMA model you'll get is ARIMA(0,1,0). This is nothing but a random walk. Hence, this again proves that stock price prediction is not an easy task!

> **Note 5.3** There are other models out there as well, which are a variation on ARIMA, such as seasonal ARIMA and Auto ARIMA.

5.4 GARCH

5.4.1 Overview

GARCH stands for Generalized Auto-Regressive Conditional Heteroskedasticity.

Definition 5.4.1: Heteroskedasticity

In statistics, a sequence of random variables is homoscedastic if all its random variables have the same finite variance; this is also known as homogeneity of variance. The complementary notion is called heteroscedasticity, also known as heterogeneity of variance

GARCH aims to predict the future volatility of a time series. Highly volatile stock data tends to appear together, and so does low volatile data.

This means that at least some aspect of stock data might be predictable!

5.4.2 ARCH

Overview

Till now, our models have been of the form:

$$y_t = f(x_t, x_{t-1}, \ldots, x_{t-p}) + \epsilon_t$$

Our primary focus till now was to try and model the 'mean' part, that is, the $f(x_t, x_{t-1}, \ldots, x_{t-p})$, while waving off ϵ_t as unpredictable noise.

Now, we aim to try and model this noise. Note: the error is still unpredictable. However, we're trying to model the parameters of the random variable representing the error, i.e., the variance.

Modelling the Error

$$\epsilon_t = z_t \sigma_t$$

Here z_t is the stochastic term, while σ_t is the time-dependent variance of the residuals. z_t often represents a noise process, which is strong sense stationary.

Now, σ_t is modeled by:

$$\sigma_t^2 = \omega + a_1 \epsilon_{t-1}^2$$

This is the *ARCH(1)* model. Including past p terms makes it an *ARCH(p)* model.

This now is an auto-regressive model for the variance, with the only subtlety being that the past terms are errors instead of variances. z_t is often considered to be a standard normal.

Therefore, by using this model:

- Given previous error terms, $E[\epsilon_t^2 | \epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-p}] = \sigma_t^2$
- The unconditional expectation, $E[\epsilon_t^2] = \frac{\omega}{1-a_1}$ for *ARCH(1)*
- The ACF and PACF for squared errors contain significant lags, which decay geometrically.

These points make the ARCH model a good fit for financial data. The residuals on stock returns, when squared, also contain significant lags. This and the zero expected values are why ARCH seems like a good fit.

5.4.3 GARCH (Generalized ARCH)

The problem with ARCH is that the error terms are random, which can make the variance swing around quite quickly. However, we want a persistence of variance. Thus, instead of just using past error terms, we also use past variances.

Modelling the Error

$$\sigma_t^2 = \omega + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2$$

This way, a high variance yesterday also causes a high variance today, so we have achieved the persistence we wanted. This now models volatility clustering.

The above formula is for *GARCH(1, 1)*. The general form *GARCH(p, q)* looks like:

$$\sigma_t^2 = \omega + \sigma a_i\epsilon_{t-i}^2 + b_i\sigma_{t-i}^2$$

A common mistake people make is thinking of σ_t to be random. The randomness is caused by z_t.

5.4.4 Uses of GARCH

It is very popular to model the volatility of a stock, which comes in handy when one wants to price financial instruments such as Options. In fact, the *Black-Scholes* option pricing equation uses the volatility of the underlying company's stock over the expiry date.

Another idea which I encourage you to think about, is the combination of *ARIMA* and *GARCH*. This is quite a popular approach, where the *ARIMA* part models the mean and the *GARCH* part models the residuals.