1   AN INTRODUCTION TO PROBABILITY & STATISTICS

This introductory chapter gives you a brief overview of things to come.

1.1   DATA AND STATISTICS

We are surrounded by an enormous amount of data. We encounter it in different forms such as data on crime, on popularity ratings of political leaders, voting numbers, sports teams, etc.

Organizations take decisions based on information or data, either internally generated or externally acquired (through media outlets, the government, data vendors etc.). However, this data needs to be properly understood and analysed in order to derive theoretical and actionable insights from it. Statistics is the discipline which makes the data speak!

1.2   WHAT IS STATISTICS?

Formally, statistics is the science which deals with the collection, organization, analysis, interpretation and presentation of data. It helps decision makers in making informed decisions. The larger field of statistics can further be divided into:
- ▪ Descriptive statistics: deals with summarizing and visualizing the data
- ▪ Inferential statistics: deals with making inferences about a population using samples drawn from it.

We will cover the basic elements of each of these in this primer.

1.3   DIFFERENCE BETWEEN PROBABILITY AND STATISTICS

Probability and statistics are often mentioned in the same breath and studied together. However, there are differences between them we must make note of.

Probability is concerned with calculating the likelihood of an event occurring and is generally calculated before the said event takes place. On the other hand, statistics comes into play once the event has taken place and we have some data to analyze.

We will cover the basics of probability theory in this primer.

1.4   PRACTICAL APPLICATIONS OF STATISTICS

We see statistics in its applied avatar across many fields. We enlist some of them below.

a) Medicine: To predict the presence of a disease in a patient, medical data analysis etc.
b) Economics: To analyze the relationship between economic measures such as GDP and employment, etc.

c) Business: To study past trends of customer behavior and use it to advertise and sell high demand products, etc.
d) Policy formulation: To formulate budget allocations, based on estimated revenue and expenditure from different sources etc.
e) Disaster and risk mapping: To identify hot-spots, disaster prone areas, to predict the occurrence of hazardous events etc.
f) Insurance: To calculate the risk presented by the insured client/property (insurance premium is directly proportional to the risk posed by the insured) etc.
g) Finance: To evaluate risks and returns for stocks, portfolio optimization, risk management, credit risk modelling, trading strategy formulation, etc.

1.5 STATISTICS IN QUANT FINANCE & ALGORITHMIC TRADING

Statistics is widely used in finance. Even more so in its subdomains, quantitative finance and algorithmic trading. Any trading strategy that you encounter here will involve statistical analysis.