Probability Theory

1.1 Introduction

The theory of probability mainly originated through gambling and games of chance. On that time, the mathematicians had little or no interest in the development of any theory instead they looked only at the combinatorial reasoning involved in each problem. Laplace was the first to develop some theory based on probability theory in his seminal work '*Theorie analytique des probabilités* (1812)'. He introduced the classical definition of the probability of an event that can occur only in a finite number of ways as the proportion of the number of favorable outcomes to the total number of all possible outcomes, provided that all the outcomes are equally likely. According to this definition, the computation of the probability of events was reduced to combinatorial counting problems. It is easy to understand but have limitations due its assumptions. Consider the situations: *i*) A loaded die is rolled, *ii*) A coin is tossed until a head appear, *iii*) Throw a point object in a square and we want to find the probability that it lies in the upper half of the square. Here all three situations can't be handled through the concept of classical probability although last situation can be explained through geometric probability.

An extension of the classical definition of Laplace was used to evaluate the probabilities of sets of events with infinite outcomes. The notion of equal likelihood of certain events played a key role in this development. According to this extension, if Ω is some region with a well-defined measure (length, area, volume, etc.), the probability that a point chosen at random lies in a subregion A of $\Omega = \frac{measure(A)}{measure(\Omega)}$. Many problems of geometric probability were solved using this extension. Although there is still some problems in choosing a random point. Joseph Bertrand (*Calculdes probabilités*, 1889) discussed some problems in geometric probability where the result depended on the method of solution. Interested students can study the famous 'Bertrand paradox' and observe that it resolved through the proper defining of 'probability spaces'.

Later, in 1933, A.N. Kolmogorov axiomatized probability in his fundamental work 'Foundations of the Theory of Probability'. Here, random events are represented by sets and probability is just a normed measure defined on these sets. This measure-theoretic development not only provided a logically consistent foundation for probability theory but also, at the same time, joined it to the mainstream of modern mathematics. However, we will go through some basic definitions e.g. random experiment, sample space etc.

► The term probability defines the measure of chance of occurrence of a phenomena.

1.2 Some Basic Definitions

1.2.1 Random Experiment

A random experiment or, statistical experiment is an experiment with the following three criteria:

1. All outcomes of the experiment are known in advance.
2. Any performance of the experiment results in an outcome that is not known in advance.
3. The experiment can be repeated under identical conditions.

■ Suppose we tossed a coin. Assuming that the coin does not land on the side, there are two possible outcomes of the experiment: heads and tails. On any performance of this experiment one does not know what the outcome will be. The coin can be tossed as many times as desired. This is an example of a random experiment.

[Do It Yourself] 1.1. *Give two examples of random experiment.*

▶ In probability theory, we associate with each such experiment a set Ω, the set of all possible outcomes of the experiment. To engage in any meaningful discussion about the experiment, we associate with Ω a σ-field $\mathbb{S}$, of subsets of Ω.
■ Trial: A trial refers to a special type of experiment in which there are two possible outcomes: success with probability p and failure with probability $q = 1 - p$.
■ Sample: It is a part of the population and is suppose to represent the characteristic of the population.

1.2.2 Classical Definition

If a random experiment have N finite, mutually exclusive, exhaustive and equally likely cases and $N(A)$ of them are favorable to the occurrence of the event A, then the probability of the occurrence of A is $P(A) = \frac{N(A)}{N}$.
▶ It is not applicable if the outcomes are not equally likely.
▶ It is not applicable if the all possible outcomes are not finite.

1.2.3 Statistical or, Empirical or, Frequency Definition

If a trial is repeated a number of times under essentially homogeneous and identical conditions, then the limiting value of the ratio of the number of times the event happens to the number of trials, as the number of trials become indefinitely large, is called the probability of happening of the event (It is assumed that the limit is finite and unique). Symbolically, if in n trials an event E happens f_n times, then the probability of the happening of E is given by $P(E) = \lim_{n\to\infty} \frac{f_n}{n}$.

▶ Statistical probability is an improvement over classical probability in the sense that it can be defined for the experiments does not have equally probably outcomes.

■ Statistically, the term probability means the long term relative frequency of any particular outcome of the experiment.

■ When a large number of repetitions are considered together, a kind of regularity is found to underlie the apparent chaos. This is called statistical regularity. For example if we toss a fair coin a large number of times (say 1000 times), we can say that the proportion of head is nearly 0.5.

1.2.4 Subjective Probability

A type of probability based on personal believes, judgment, or experience about the occurrence of a specific outcome in the future. The calculation of subjective probability contains no formal computations (of any formula) and reflects the opinion of a person based on his/her past experience. The subjective probability differs from subject to subject and it may contain a high degree of personal biasness.

This kind of probability usually based on person's experience, understanding, knowledge, and his intelligence who determining the probability of some specific event (situation). It is usually applied in real-life situations, especially, related to the decision in business, job interviews, promotions of the employee, awarding incentives, and daily life situations such as buying and/or selling of a product. As a matter of fact, an individual may use their own expertise, opinion, past experiences, or intuition to assign the degrees of probability to a specific situation.

▶ One may think that there are 80% chances that your best friend will call you today because his/her car broke down yesterday and he/she will probably need a ride.

▶ You think you have a 50% chances of getting a certain job you applied for as the other applicant is also qualified.

1.2.5 Basics of Set Theory

▶ A set is a well defined collection of distinct objects. The individual objects of the set are called elements of the set. We generally use capital letters to denote a set, small letters for its elements and '$\in$' sign to access its elements.
▶ We can define a set by listing the elements or, roster method in which the elements of the set are listed within the pair of brackets { } and are separated by commas. Sometimes it is not always possible to define a set by roster method and it can be defined by describing the elements property or, set-builder notation in which all the elements of the set follow a certain property.
▶ For example: $A = \{1, 2, 3, 4, 5\}$ is a set formed by roster method whereas we can define $A = \{x : x\ is\ first\ five\ natural\ numbers\}$ by set-builder method.
▶ Also $B = \{x : x\ is\ a\ real\ number\ on\ (0, 1)\}$ is a set which can only be constructed through set-builder method.

[Do It Yourself] 1.2. *Suppose A be the set of all integer points lies within the circle $x^2 + y^2 = 9$. How you will describe A by roster method and set builder method?*

(A) Let, $A = \{2, 1, 4, 6, 7\}$, here A is a set with elements $2, 1, 4, 6, 7$. Now $x \in A \Rightarrow x$ is an element of A.

(B) Let, $B = \{\alpha, 1, \beta, 2, \gamma, 3\}$, here B is a set with elements $\alpha, 1, \beta, 2, \gamma, 3$. Now $x \in A \Rightarrow x$ is an element of A.

(C) Let, $A = \{x : x\ is\ an\ even\ number\} = \{2, 4, 6, \cdots\}$, here A is a set with elements $2, 4, 6, \cdots$. Now $y \in A \Rightarrow y$ is an element of A.

(D) Let, $A = \{x : x\ is\ a\ positive\ root\ of\ x^2 - x - 2 = 0\}$, here A is a set with only element $\{2\}$. Now $x \in A \Rightarrow x = 2$.

(E) Set of natural numbers : $\mathbb{N} = \{1, 2, 3, \cdots\}$.

(F) Set of integers : $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \cdots\}$.

(G) Set of real numbers : $\mathbb{R} = \{x : x \in (-\infty, \infty)\} = \{x : -\infty < x < \infty\}$.

(H) Set of rational numbers : $\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z},\ q \neq 0,\ g.c.d.(p, q) = 1\}$.

(I) Set of irrational numbers : $\mathbb{Q}^c = \{x : x \in \mathbb{R}\ and\ x \notin \mathbb{Q}\} = \mathbb{R} \setminus \mathbb{Q}$.

■ A set may be finite or, infinite. An infinite set may be countable or, uncountable. Let, $A = \{1, 2, \cdots, 10\}$, $B = \{1, 2, 3, \cdots\}$, $C = \{x :\ x \in [1, 2]\}$, here A is finite, B is countably infinite and C is uncountably infinite.

▶ Some terms related to set theory are given below

(A) Null Set: A set with no elements is called a null or, empty set and it is denoted by ϕ. For example, $\phi = \{x : x \in \mathbb{R}, x^2 = -1\}$, $\phi = \{x : x \in \mathbb{Z}^+, x < 0\}$ etc.

(B) Singleton Set: A set with exactly one element. For ex. $\{0\}$, $\{5\}$ etc.

(C) Finite/Infinite Set: A null set or, contains finite number of elements is called a finite set, else it is an infinite set. For ex. $A = \{x : x^2 < 10, x \in \mathbb{Z}\}$, $B = \{x : x^2 < 10, x \in \mathbb{R}\}$.

(D) Equality of Sets: Two sets are equal if they have exactly same elements (any order). For ex. $A = \{2, 4, 5, 7\}, B = \{7, 2, 4, 5\}$ are equal.

(E) Some Symbols: '$\Rightarrow$' stands for implies, $A \Rightarrow B$ means if A is true then B is also true i.e $x = 2 \Rightarrow x + 2 = 4$. '$\Leftrightarrow$' stands for implies and implied by, $A \Leftrightarrow B$ means $A \Rightarrow B$ and $B \Rightarrow A$ i.e $x = 2 \Leftrightarrow x + 2 = 4$. Note that, $x = 2 \not\Leftrightarrow x^2 = 4$. '$\wedge$' stands for 'and'. '$\vee$' stands for 'or'. '$\backsim$' stands for 'not'.

(F) Subset and Superset: If $x \in A \Rightarrow x \in B$ then A is called a subset of B and denoted by $A \subseteq B$. For ex. $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 3, 4\}$, $C = \{1, 2, 3\}$, then $A \subseteq B$, $C \subset B$. Note that, $A \subseteq B \Leftrightarrow A \supseteq B$. $A \supseteq B$ means A is a superset of B. In the example, $A \supseteq B$, $A \supset C$.

(G) Universal Set: In set theory, we assume all sets are subset of a set which is called universal set. For ex. $\mathbb{R}$ is an universal set for working with real numbers.

(H) Finite Union and Intersection: If A, B are two sets, then the set $A \cup B$ is called union of A and B and it is defined as $A \cup B = \{x : x \in A \text{ or}, \ x \in B\}$. The set $A \cap B$ is called intersection of A and B and it is defined as $A \cap B = \{x : x \in A \text{ and } x \in B\}$. For ex. $A = \{1, 2, 3, 4\}$, $B = \{2, 3, 6\}$ then $A \cup B = \{1, 2, 3, 4, 6\}$, $A \cap B = \{2, 3\}$. Finite union and intersection are denoted by $\bigcup_{i=1}^{n} A_i$, $\bigcap_{i=1}^{n} A_i$ respectively.

(I) Arbitrary Union and Intersection: The collection of subsets $\{A_\lambda : \lambda \in \Lambda\}$ of a set S is called a family of subsets of S. Here Λ is called the index set. The terms $\bigcup_{\lambda \in \Lambda} A_\lambda$, $\bigcap_{\lambda \in \Lambda} A_\lambda$ are called arbitrary union and intersection of the family S and defined by $\bigcup_{\lambda \in \Lambda} A_\lambda = \{x : x \in A_\lambda \text{ for atleast one } \lambda \in \Lambda\}$, $\bigcap_{\lambda \in \Lambda} A_\lambda = \{x : x \in A_\lambda \text{ for all } \lambda \in \Lambda\}$. For ex. if $\Lambda = \mathbb{N} = \{1, 2, 3, \cdots\}$ then $\bigcup_{\lambda \in \Lambda} A_\lambda = \bigcup_{\lambda=1}^{\infty} A_\lambda$.

(J) Disjoint Set: A, B are said to be disjoint sets if $A \cap B = \phi$. $A_1, A_2, \cdots, A_n$ are said to be disjoint if $\bigcap_{i=1}^{n} A_i = \phi$. If $A_1, \cdots, A_n$ are disjoint sets then we can write $\cup A_i = \sum A_i$.

(K) Complement: If A is a subset of S, then $A^c = \{x : x \in S \text{ and } x \notin A\}$. For ex. $A = (0, \infty)$ is a subset of $\mathbb{R}$ then $A^c = (-\infty, 0]$.

(L) Difference: If A, B two subsets of S, then $A \setminus B = \{x : x \in A \text{ and } x \notin B\} = A \cap B^c$. For ex. $A = \{1, 2, 3, 4\}$, $B = \{2, 3, 6\}$ are subsets of $\mathbb{R}$ then $A \setminus B = \{1, 4\}$.

(M) Symmetric Difference: If A, B two subsets of S, then $A\Delta B = (A \setminus B) \cup (B \setminus A)$. For ex. $A = \{1,2,3,4\}$, $B = \{2,3,6\}$ are subsets of $\mathbb{R}$ then $A\Delta B = \{1,4,6\}$.

(N) Power of a Set: If A be a non-empty set. Then the collection of all subsets of A is a family of sets and it is called the power set of A and denoted by $P(A)$. $P(A)$ has 2^n elements. For ex. $A = \{1,2,3\}$ then $P(A) = \{\phi, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, A\}$.

(O) Consistency Properties: $A \subset B \Leftrightarrow A \cup B = B \Leftrightarrow A \cap B = A$.

(P) Other Properties: Commutative: $A \cup B = B \cup A$, $A \cap B = B \cap A$. Associative: $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$. Distributive: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

(Q) De Morgan's Law: $(A \cup B)^c = A^c \cap B^c$, $(A \cap B)^c = A^c \cup B^c$.

(R) Cartesian Product: If A, B are non-empty sets. Then the Cartesian product $A \times B$ is defined as $A \times B = \{(a,b) : a \in A, b \in B\}$. For ex. $A \times A = A^2 = \{(a_1, a_2) : a_1, a_2 \in A\}$, $\mathbb{R}^n = \{(a_1, a_2, \cdots, a_n) : a_1, a_2, \cdots, a_n \in \mathbb{R}\}$. If $A = \{1,2,3\}$, $B = \{2,4\}$ then $A \times B = \{(1,2),(1,4),(2,2),(2,4),(3,2),(3,4)\}$, $B \times A = \{(2,1),(2,2),(2,3),(4,1),(4,2),(4,3)\}$.

(S) Equivalent Sets: Two sets A, B are said to be equivalent ($A \sim B$) if $\exists$ a one-one correspondence between their elements. For ex. $\mathbb{N} \sim \mathbb{Z}$ since there is a $1-1$ correspondence between their elements: $1 \leftrightarrow 0$, $2 \leftrightarrow -1$, $3 \leftrightarrow 1$,$4 \leftrightarrow -2$, $\cdots$.

(T) Enumerable/Countable Sets: A set 'A' is said to be countable/ countably infinite if ($A \sim \mathbb{N}$).

(U) Cantor Set: The Cantor ternary set (it is a three part or, ternary construction of Cantor set) $\mathcal{C}$ is created by iteratively deleting the open middle third from a set of line segments. One starts by deleting the open middle third $(1/3, 2/3)$ from the interval $[0,1]$, leaving two line segments: $[0,1/3] \cup [2/3,1]$. Next, the open middle third of each of these remaining segments is deleted, leaving four line segments: $[0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]$ and so on.

(V) The n^{th} Cantor set is $C_n = \frac{C_{n-1}}{3} \cup (\frac{2}{3} + \frac{C_{n-1}}{3})$, $n \geq 1$ and $C_0 = [0,1]$. The Cantor ternary set contains all points in the interval $[0,1]$ that are not deleted at any step in this infinite process. Cantor set has very nice properties which we will study in the course "Measure Theory".

[Do It Yourself] 1.3. *Explain the above discussed items (if applicable) through Venn diagram approach.*

[Do It Yourself] 1.4. *Is it true* $\{1\} = 1$*? Discuss.*

[Do It Yourself] 1.7. *Find finite and arbitrary union/intersection of the following sets*

(A) $A_n = [0, \frac{1}{n})$, *find* $\bigcup_{n=1}^{p} A_n, \bigcap_{n=1}^{p} A_n, \bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} A_n$.

(B) $A_n = (0, \frac{1}{n})$, *find* $\bigcup_{n=1}^{p} A_n, \bigcap_{n=1}^{p} A_n, \bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} A_n$.

(C) $A_n = (-\frac{1}{n}, 0)$, *find* $\bigcup_{n=1}^{p} A_n, \bigcap_{n=1}^{p} A_n, \bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} A_n$.

(D) $A_n = (-\frac{1}{n}, \frac{1}{n})$, *find* $\bigcup_{n=1}^{p} A_n, \bigcap_{n=1}^{p} A_n, \bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} A_n$.

1.2.6 Class

Class is a set whose elements are sets. E.g. $\mathcal{C} = \{(1,2],(2,3],\cdots\}$. A disjoint class is $\mathcal{F} = \{A_\lambda : \lambda \in \Lambda\}$, where $A_\lambda \cap A_{\lambda'} = \phi$, $\forall\ \lambda \neq \lambda'$ & $\lambda, \lambda' \in \Lambda$.

1.2.7 Field & σ-Field

■ A non-empty class $\mathcal{A}$ of the subset of Ω is called a field if

1. For all sets $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$. (Closed under complementation)
2. If $A_1, A_2, \cdots, A_n \in \mathcal{A} \Rightarrow \cup_{i=1}^n A_i \in \mathcal{A}$. (Closed under finite union)

Example 1.1. *If $\mathcal{A}$ is a field then show that $\Omega, \phi \in \mathcal{A}$.*
□ *Let $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ [Closed under complementation] $\Rightarrow A \cup A^c \in \mathcal{A}$ [Closed under finite union] $\Rightarrow \Omega \in \mathcal{A}$.*
Again $\Omega \in \mathcal{A} \Rightarrow \Omega^c \in \mathcal{A}$ [Closed under complementation] $\Rightarrow \phi \in \mathcal{A}$. Therefore, $\Omega, \phi \in \mathcal{A}$.

▶ The smallest field containing a set A is $\{\phi, \{A\}, \{A^c\}, \Omega\}$.
▶ Suppose we want to generate the smallest field containing two set A_1, A_2. Here we think about a Venn diagram and obtain $2^2 = 4$ disjoint sets of intersection e.g. $A_1 \cap A_2^c, A_1 \cap A_2, A_1^c \cap A_2, A_1^c \cap A_2^c$. Now using union we can select any combination of these 4 sets and together with the null set ϕ get $(2^4 - 1) + 1 = 16$ elements of the field.
▶ The smallest field containing two set A_1, A_2 is $\{\phi, \{A_1\}, \{A_2\}, \{A_1^c\}, \{A_2^c\}, \{A_1 \cup A_2\}, \{A_1 \cup A_2^c\}, \{A_1^c \cup A_2\}, \{A_1^c \cup A_2^c\}, \cdots, \Omega\}$. It is not easy to find 16 sets in this way.
▶ The smallest field containing two set A_1, A_2 is $\{\phi, \{P\}, \{Q\}, \{R\}, \{S\}, \{P \cup Q\}, \{P \cup R\}, \{P \cup S\}, \{Q \cup R\}, \{Q \cup S\}, \{R \cup S\}, \{P \cup Q \cup R\}, \{P \cup Q \cup S\}, \{P \cup R \cup S\}, \{Q \cup R \cup S\}, \{P \cup Q \cup R \cup S\} = \Omega\}$. Here P, Q, R, S are the 4 disjoint sets of intersection.
▶ Suppose there are n subsets $A_1, A_2, \cdots, A_n$. Then there are 2^n disjoint sets of intersections. Then the all possible unions of them are 2^{2^n} which are the smallest field. You can also extend this concept on a σ-field.

[Do It Yourself] 1.8. *For a field $\mathcal{A}$ if $A_1, A_2 \in \mathcal{A}$, then show that $A_1 \cap A_2 \in \mathcal{A}$.*

[Do It Yourself] 1.9. *Suppose $\mathcal{A}$ is a field and $A_1, A_2, \cdots, A_n \in \mathcal{A}$, then show that $\cap_{i=1}^n A_i \in \mathcal{A}$.*

■ A non-empty class $\mathcal{F}$ of the subset of Ω is called a σ-field if

1. For all sets $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$. (Closed under complementation)
2. For any sequence of sets $\{A_n, n \geq 1\}$, $A_n \in \mathcal{F}$, $\forall n \geq 1 \Rightarrow \cup_{n=1}^\infty A_n \in \mathcal{F}$. (Closed under countable union)