Chapter 2

Basic Probability Theory

Basics in Probability

Experiment
The process of obtaining an observation is called an experiment. Hitting a target, checking the boiling point of a liquid, taking examination for a student, conducting interviews for some jobs, tossing of a coin, rolling of a die, hitting a ball of a batsman, sale of some products, chemical reaction of elements, are few examples of experiments.

Trial
A single performance of an experiment is called trial. If a batsman plays a single ball, if a bowler bowls, if a student solves a single question, single rolling of a die, all these are the examples of a trial.

Outcome
An outcome is the result of an experiment. Each possible distinct result of an experiment is referred as outcome. Hitting or not hitting a target, making some scores, leaving a ball or being out for a batsman, boiling of water on 100°C are the examples of outcomes relevant to the experiments discussed above.

Random Experiment
An experiment is called random experiment if its outcomes cannot be predicted in advance even if it is performed under similar conditions. Any random experiment has the following properties:

(i) It has at least two outcomes.
(ii) The number of all possible outcomes is known in advance.
(iii) It can be repeated any number of times under similar conditions.

Among above examples of experiments, hitting a target, taking examination for a student, tossing of a coin, rolling of a die, hitting a ball of a batsman, sale of some product are examples of Random Experiments while checking the boiling point of a liquid and chemical reaction of elements will be nonrandom experiments.

Sample Space

The set of all possible outcomes of a random experiment is called a sample space. It is usually denoted by Ω (omega).
In the random experiment in which a student takes a examination, suppose the result of examination can be in form of grades, 'A', 'B', 'C', 'D' and 'F' then Ω = {'A', 'B', 'C', 'D','F'}.
In rolling a die, $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Event

Any subset of sample space is called event. It is a collection of outcomes of an experiment. Events may be either simple or composite. Formally, if an event consists of single sample point of a sample space, it is called elementary or simple event and in case of two or more sample points, it is referred as composite event.
Simple Events: A student passes an examination, a batsman makes shot for six, a die show a number 2, etc.
Composite Event: A motor accident for rash driving and failure of brakes, a ball results in 1 run and a run-out during an over in a cricket match, etc.

Event Space

A set of all events relevant to a sample space is called event space. It is usually denoted by $\mathcal{A}$.

Equally Likely Events

Two events A_1 and A_2 relevant to a sample space are said to be equally likely if their probabilities of occurrences are equal. In other words, if event A_1 is as likely to occur as A_2 does then A_1 and A_2 are said to be equally likely events.

Mutually Exclusive or Disjoint Events

Two events A_1 and A_2 relevant to a sample space are said to be mutually exclusive or disjoint if they cannot occur simultaneously. Thus tossing of a coin is associated with mutually exclusive events. If the event 'head' occurs, the event 'tail' cannot occur at the same time. In above example of grading

of a student in an examination, if all the five grades refer to five different events then they can not occur simultaneously as a student can take, obviously any one of the grade, so these events are also disjoint.
Mathematically, if A_1 and A_2 are two mutually exclusive events then

$$P(A_1 \cap A_2) = 0.$$

Collectively Exhaustive Events

The events $A_1 A_2, \ldots, A_k$ are said to be collectively exhaustive if all A_i's ($i = 1, 2, \ldots, k$) are Disjoint and $A_1 \cup A_2 \cup \ldots A_k = \Omega$.
In above example of grades, all the five events are collectively exhaustive.

Complementary Events

Complementary event for an event A is the event that A does not occur. For event A it is denoted by A' or A^c. For example, passing of a student is complementary event to the failing of that student.

Independent Events

Two events A_1 and A_2 relevant to a sample space are said to be statistically independent if the occurrence of A_1 does not affect the probability of occurrence or non-occurrence of A_2.
Symbolically,

$$P(A_1 \cap A_2) = P(A_1) \,.\, P(A_2).$$

The passing (or failing) of one student is statistically independent to the passing or failing of other student(s) in the same examination, score on a current ball is statistically independent to the result of previous ball in a cricket match, are the examples of statistically independent events.

Probability

Probability may be defined as the likelihood of the occurrence of an event. A probability provides a quantitative description of the likely occurrence of a particular event. In other words, it is a numerical measure of uncertainty. Probability is conventionally expressed on a scale from 0 to 1; a rare event has a probability close to 0, a very common event has a probability close to 1.

Subjective Probability

A subjective probability describes an individual's personal judgment about how likely a particular event is to occur. It is not based on any precise computations but is often a reasonable assessment by a knowledgeable person. A person's subjective probability of an event describes his/her degree of belief in the event. For example, a cricket expert says that there are more than 80% chances that the team A will win the tournament. A planning minister guesses that at least 3/4 villages of the country will be supplied electricity by the end of next year.

Objective Probability

A probability that can be established theoretically or from historical data. The objective probability has following main approaches to define Probability:

(i) *Classical (Priori) Definition of Probability*
(ii) *Relative –Frequency Definition of Probability*
(iii) *Axiomatic Definition of Probability*

Classical (Priori) Definition of Probability

If an experiment can produce n different mutually exclusive results all of which are equally likely, and if m of these results are considered favorable (or result in event A), then the probability of a favorable result (or the probability of event A) is m/n.

In tossing of coin, the probability of occurring 'head' or 'tail' and the probability of selection a female student among a group of 20 male and 30 female students, follow classical definition of probability.

Relative–Frequency Definition of Probability

This definition of probability states that if in n trials, an event A occurs n_A times, its probability, $P(A)$ is approximately n_A/n:

$$P(A) = n_A/n,$$

provided that n is sufficiently large and the ratio n_A/n is nearly constant as n increases.

The probability of having number 6 on an irregular die, the probability of survival of a 40 years old Hepatitis C patient, the probability for wining of a tournament by a team, are the examples of relative frequency definition of probability.

Axiomatic Definition of Probability

Let Ω be the Sample Space, the probability of an event A is, by definition, a number $P(A)$ assigned to A. This number satisfies the following three axioms:

(i) $P(A)$ is a non-negative number; $P(A) \geq 0$.
(ii) Probability of the event Ω (sure or certain event) is equal to 1;
$P(\Omega) = 1$.

(iii) If two events A_1 and A_2 have no common element (i.e. Mutually Exclusive) then

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

Conditional Probability

In many situations, once more information becomes available; we are able to revise our estimates for the probability of further outcomes or events happening. For example, suppose you go out for lunch at the same place and time every Friday and you are served lunch within 15 minutes with probability 0.9. However, given that you notice that the restaurant is exceptionally busy, the probability of being served lunch within 15 minutes may reduce to 0.7. This is the conditional probability of being served lunch within 15 minutes given that the restaurant is exceptionally busy. The effect of such information is to reduce the sample space by excluding some outcomes as being impossible which before receiving the information were believed possible.

Formally, the conditional probability of A given B is the probability of event A occurring, given that event B has already occurred. It can be found by dividing the probability of events A and B both occurring by the probability of event B as shown below

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

It is obvious to note that if A and B are statistically independent then

$$P(A \mid B) = P(A).$$

Some Probability Rules

Addition Rule of Probability

If A_1 and A_2 be any two events relevant to a sample space Ω, then

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

In words, the probability that either event A_1 or event A_2 or both events occur equals the probability that event A_1 occurs plus the probability that event A_2 occurs minus the probability that both occur.

Generally, for n events:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n+1} P(\bigcap_{i=1}^{n} A_i)$$

Multiplication Rule of Probability

For any two events A_1 and A_2 relevant to a sample space Ω, the multiplication rule is a result used to determine the probability that two events, A_1 and A_2, both occur. This rule follows the definition of conditional probability as

$$P(A_1 \cap A_2) = P(A_1).P(A_2 \mid A_1).$$

Generally, for n events:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1).P(A_2 \mid A_1).P(A_3 \mid A_1 \cap A_2) \ldots \cap P(A_n \mid A_1 \cap A_2 \cap \ldots \cap$$

Law of Total Probability

Suppose, A_1 and A_2 be mutually exclusive and exhaustive events with non-zero probabilities then for any event B (with non-zero probability)

$$P(B) = P(A_1).P(B \mid A_1) + P(A_2).P(B \mid A_2).$$

Generally, for n events:

$$P(B) = \sum_{i=1}^{n} P(A_i).P(B \mid A_i).$$

Baye's Theorem

Suppose, A_1 and A_2 be mutually exclusive and exhaustive events with non-zero probabilities then for any event B (with non-zero probability)

$$P(A_1 \mid B) = \frac{P(A_1).P(B \mid A_1)}{P(A_1).P(B \mid A_1) + P(A_2).P(B \mid A_2)}$$

Generally, for n events:

$$P(A_i \mid B) = \frac{P(A_i).P(B \mid A_i)}{\sum_{i=1}^{n} P(A_i).P(B \mid A_i)}, \quad (i = 1, 2, ..., n)$$

Counting Rules

Rule of Multiplication

If there are K procedures and ith procedure may be performed in n_i ways (i = 1, 2, ..., k) then all the k procedures may be performed in $n_1 \times n_2 \times ... \times n_k$ ways. For example, a person has 3 different pairs of shoes and 4 different pairs of socks then he may use all of these pairs in 3 x 4 = 12 different ways.

Rule of Addition

If there are K procedures and ith procedure may be performed in n_i ways (i = 1, 2, ..., k) then the number of ways in which one can perform procedure 1 or procedure 2 ... or procedure k given by $n_1 + n_2 + ... + n_k$ (assuming that one procedure can be performed one time or no two procedures can be performed together).

Suppose, a group of students is planning a trip and thinking about either bus or train to use for that. If there are 3 different routes available when using bus for the trip and 2 different routes for train then there are 3 + 2 = 5 routes available for that trip.

Permutations

A permutation of n different objects taking r at a time ($0 \le r \le n$)

$^nP_r = \frac{n!}{(n-r)!}$. In permutations, order of objects is important or meaningful.

For example, if one wants to calculate the ways in which 6 persons may be seated on a bench having a capacity of 4 seats then in this case, $n = 6$, $r = 4$ and answer is $^6P_4 = \frac{6!}{(6-4)!} = 6 \times 5 \times 4 \times 3 = 360$.

Combinations

A combination of n different objects taking r at a time ($0 \le r \le n$)

$^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$. In combinations, order of objects is not meaningful.

To differentiate between the case of permutation and combination, we consider the example that out of 7 Statisticians a committee is to be formed of 4 then in $^7C_4 = \binom{7}{4} = \frac{7!}{4!(7-4)!} = 35$ different ways the committee can be formed. Now, if we specify the positions of committee members like, one president, one secretary, one treasurer and one speaker then this will be a case of permutation because here the positions (order) are meaningful and the number of ways in which such committee can be constituted is $^7P_4 = \frac{7!}{(7-4)!} = 840$.