PROBABILITY

INTRODUCTION

Numerical study of chances of occurrence of events is dealt in probability theory.

The theory of probability is applied in many diverse fields and the flexibility of the theory provides approximate tools for so great a variety of needs.

There are two approaches to probability viz. (i) Classical approach and (ii) Axiomatic approach.

In both the approaches we use the term 'experiment', which means an operation which can produce some well-defined outcome(*s*). There are two types of experiments-

(1) **Deterministic experiment –** Those experiments which when repeated under identical conditions produce the same result or outcome are known as deterministic experiments. When experiments in science or engineering are repeated under identical conditions, we get almost the same result every time.

(2) **Random experiment –** If an experiment, when repeated under identical conditions, do not produce the same outcome every time but the outcome in a trial is one of the several possible outcomes then such an experiment is known as a probabilistic experiment or a random experiment.

In a random experiment, all the outcomes are known in advance but the exact outcome is unpredictable.

For example, in tossing of a coin, it is known that either a head or a tail will occur but one is not sure if a head or a tail will be obtained. So it is a random experiment.

DEFINITIONS OF TERMS USED IN THEORY OF PROBABILITY

(1) **Sample space –** The set of all possible outcomes of a trial (random experiment) is called its sample space. It is generally denoted by S and each outcome of the trial is said to be a sample point.

Example - (i) If a dice is thrown once, then its sample space is $S = \{1, 2, 3, 4, 5, 6\}$

(ii) If two coins are tossed together then its sample space is $S = \{HT, TH, HH, TT\}$.

(2) **Event -** An event is a subset of a sample space.

(i) **Simple event -** An event containing only a single sample point is called an elementary or simple event.

Example - In a single toss of coin, the event of getting a head is a simple event.

Here $S = \{H, T\}$ and $E = \{H\}$

(ii) **Compound events -** Events obtained by combining together two or more elementary events are known as the compound events or decomposable events.

For example, In a single throw of a pair of dice the event of getting a doublet, is a compound event because this event occurs if any one of the elementary events (1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6) occurs.

(iii) **Equally likely events -** Events are equally likely if there is no reason for an event to occur in preference to any other event.

Example - If an unbiased die is rolled, then each outcome is equally likely to happen *i.e.*, all elementary events are equally likely.

(iv) **Mutually exclusive or disjoint events -** Events are said to be mutually exclusive or disjoint or incompatible if the occurrence of any one of them prevents the occurrence of all the others.

Example - E = getting an even number, F = getting an odd number, these two events are mutually exclusive, because, if E occurs we say that the number obtained is even and so it cannot be odd *i.e.*, F does not occur.

A_1 and A_2 are mutually exclusive events if $A_1 \cap A_2 = \phi$.

(v) **Mutually non-exclusive events -** The events which are not mutually exclusive are known as compatible events or mutually non exclusive events.

(vi) **Independent events -** Events are said to be independent if the happening (or non-happening) of one event is not affected by the happening (or non-happening) of others.

Example - If two dice are thrown together, then getting an even number on first is independent to getting an odd number on the second.

(vii) **Dependent events -** Two or more events are said to be dependent if the happening of one event affects (partially or totally) other event.

Example - Suppose a bag contains 5 white and 4 black balls. Two balls are drawn one by one. Then two events that the first ball is white and second ball is black are independent if the first ball is replaced before drawing the second ball. If the first ball is not replaced then these two events will be dependent because second draw will have only 8 exhaustive cases.

(3) **Exhaustive number of cases -** The total number of possible outcomes of a random experiment in a trial is known as the exhaustive number of cases.

Example - In throwing a die the exhaustive number of cases is 6, since any one of the six faces marked with 1, 2, 3, 4, 5, 6 may come uppermost.

(4) **Favorable number of cases -** The number of cases favorable to an event in a trial is the total number of elementary events such that the occurrence of any one of them ensures the happening of the event.

Example - In drawing two cards from a pack of 52 cards, the number of cases favorable to drawing 2 queens is ${}^4C_2$.

(5) **Mutually exclusive and exhaustive system of events -** Let S be the sample space associated with a random experiment. Let $\{A_1, A_2 \dots A_n\}$ be a subsets of S such that

(i) $A_i \cap A_j = \phi$ for $i \neq j$ and (ii) $A_1 \cup A_2 \cup \dots \cup A_n = S$

Then the collection of events $A_1, A_2, \dots, A_n$ is said to form a mutually exclusive and exhaustive system of events.

If $E_1, E_2, \dots, E_n$ are elementary events associated with a random experiment, then

(i) $E_i \cap E_j = \phi$ for $i \neq j$ and (ii) $E_1 \cup E_2 \cup \dots \cup E_n = S$

So, the collection of elementary events associated with a random experiment always form a system of mutually exclusive and exhaustive system of events.

In this system, $P(A_1 \cup A_2 \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) = 1$.

CLASSICAL DEFINITION OF PROBABILITY

If a random experiment results in n mutually exclusive, equally likely and exhaustive outcomes, out of which m are favorable to the occurrence of an event A, then the probability of occurrence of A is given by

$$P(A) = \frac{m}{n} = \frac{\text{Number of outcomes favourable to } A}{\text{Number of total outcomes}}$$

It is obvious that $0 \leq m \leq n$. If an event A is certain to happen, then $m = n$, thus $P(A) = 1$.

If A is impossible to happen, then $m = 0$ and so $P(A) = 0$. Hence we conclude that

$0 \leq P(A) \leq 1$.

Further, if $\overline{A}$ denotes negative of A *i.e.* event that A doesn't happen, then for above cases m, n; we shall have

$$P(\overline{A}) = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - P(A)$$

$$\therefore \quad P(A) + P(\overline{A}) = 1.$$

PROBABILITY

Notations - For two events *A* and *B*,

(i) A' or $\overline{A}$ or A^C stands for the non-occurrence or negation of *A*.

(ii) $A \cup B$ stands for the occurrence of at least one of *A* and *B*.

(iii) $A \cap B$ stands for the simultaneous occurrence of *A* and *B*.

(iv) $A' \cap B'$ stands for the non-occurrence of both *A* and *B*.

(v) $A \subseteq B$ stands for "the occurrence of *A* implies occurrence of *B*".

ODDS IN FAVOR & ODDS AGAINST AN EVENT

As a result of an experiment if "*a*" of the outcomes are favorable to an event *E* and "*b*" of the outcomes are against it, then we say that odds are *a* to *b* in favor of *E* or odds are *b* to *a* against *E*.

Thus odds in favor of an event $E = \dfrac{\text{Number of favourable cases}}{\text{Number of unfavourable cases}} = \dfrac{a}{b} = \dfrac{a/(a+b)}{b/(a+b)} = \dfrac{P(E)}{P(\overline{E})}$.

Similarly, odds against an event $E = \dfrac{\text{Number of unfavourable cases}}{\text{Number of favourable cases}} = \dfrac{b}{a} = \dfrac{P(\overline{E})}{P(E)}$.

NOTE –

If odds in favour of an event are a - b, then the probability of the occurrence of that event is $\frac{a}{a+b}$ *and the probability of non-occurrence of that event is* $\frac{b}{a+b}$.

If odds against an event are a - b, then the probability of the occurrence of that event is $\frac{b}{a+b}$ *and the probability of non-occurrence of that event is* $\frac{a}{a+b}$.

FUNDAMENTAL THEOREM OF ADDITION

(1) **When events are not mutually exclusive -** If *A* and *B* are two events which are not mutually exclusive, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ or $P(A+B) = P(A) + P(B) - P(AB)$.

For any three events *A*, *B*, *C*

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$$

or $P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(CA) + P(ABC)$.

(2) **When events are mutually exclusive -** If *A* and *B* are mutually exclusive events, then

$n(A \cap B) = 0 \Rightarrow P(A \cap B) = 0$

$\therefore\ P(A \cup B) = P(A) + P(B)$.

For any three events *A*, *B*, *C* which are mutually exclusive,

$P(A \cap B) = P(B \cap C) = P(C \cap A) = P(A \cap B \cap C) = 0$

$\therefore\ P(A \cup B \cup C) = P(A) + P(B) + P(C)$.

The probability of happening of any one of several mutually exclusive events is equal to the sum of their probabilities, *i.e.* if $A_1, A_2 \ldots A_n$ are mutually exclusive events, then

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) \text{ i.e. } P(\sum A_i) = \sum P(A_i).$$

(3) **When events are independent -** If *A* and *B* are independent events, then $P(A \cap B) = P(A).P(B)$

$\therefore \quad P(A \cup B) = P(A) + P(B) - P(A).P(B)$.

(4) Some other theorems

(i) Let A and B be two events associated with a random experiment, then

(a) $P(\overline{A} \cap B) = P(B) - P(A \cap B)$ (b) $P(A \cap \overline{B}) = P(A) - P(A \cap B)$

If $B \subset A$, then

(a) $P(A \cap \overline{B}) = P(A) - P(B)$ (b) $P(B) \le P(A)$

Similarly if $A \subset B$, then

(a) $(\overline{A} \cap B) = P(B) - P(A)$ (b) $P(A) \le P(B)$.

Note – *Probability of occurrence of neither A nor B is* $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$.

(ii) **Generalization of the addition theorem -** If $A_1, A_2, ..., A_n$ are n events associated with a random experiment, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i,j=1 \\ i \ne j}}^{n} P(A_i \cap A_j) + \sum_{\substack{i,j,k=1 \\ i \ne j \ne k}}^{n} P(A_i \cap A_j \cap A_k) + ... + (-1)^{n-1} P(A_1 \cap A_2 \cap ... \cap A_n).$$

If all the events $A_i (i = 1, 2..., n)$ are mutually exclusive, then $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$

i.e. $P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$.

(iii) **Booley's inequality -** If $A_1, A_2 ... A_n$ are n events associated with a random experiment, then

(a) $P\left(\bigcap_{i=1}^{n} A_i\right) \ge \sum_{i=1}^{n} P(A_i) - (n-1)$ (b) $P\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} P(A_i)$

These results can be easily established by using the Principle of Mathematical Induction.

Important results to remember

Let A, B, and C are three arbitrary events. Then

Verbal description of event	*Equivalent Set Theoretic Notation*
(i) Only A occurs	*(i)* $A \cap \overline{B} \cap \overline{C}$
(ii) Both A and B, but not C occur	*(ii)* $A \cap B \cap \overline{C}$
(iii) All the three events occur	*(iii)* $A \cap B \cap C$
(iv) At least one occurs	*(iv)* $A \cup B \cup C$
(v) At least two occur	*(v)* $(A \cap B) \cup (B \cap C) \cup (A \cap C)$
(vi) One and no more occurs	*(vi)* $(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$
(vii) Exactly two of A, B and C occur	*(vii)* $(A \cap B \cap \overline{C}) \cup (\overline{A} \cap B \cap C) \cup (A \cap \overline{B} \cap C)$
(viii) None occurs	*(viii)* $\overline{A} \cap \overline{B} \cap \overline{C} = \overline{A \cup B \cup C}$
(ix) Not more than two occur	*(ix)* $(A \cap B) \cup (B \cap C) \cup (A \cap C) - (A \cap B \cap C)$
(x) Exactly one of A and B occurs	*(x)* $(A \cap \overline{B}) \cup (\overline{A} \cap B)$

PROBABILITY

CONDITIONAL PROBABILITY

Let A and B be two events associated with a random experiment. Then, the probability of occurrence of A under the condition that B has already occurred and $P(B) \neq 0$, is called the conditional probability and it is denoted by $P(A/B)$.

Thus, $P(A/B)$ = Probability of occurrence of A, given that B has already happened.

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

Similarly, $P(B/A)$ = Probability of occurrence of B, given that A has already happened.

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

FUNDAMENTAL THEOREM OF MULTIPLICATION

(i) If A and B are two events associated with a random experiment, then $P(A \cap B) = P(A)\,P(B \mid A)$, if $P(A) \neq 0$ or $P(A \cap B) = P(B)\,P(A \mid B)$, if $P(B) \neq 0$.

(ii) **Extension of multiplication theorem -** If $A_1, A_2 \ldots A_n$ are n events related to a random experiment, then

$$P(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \ldots P(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1}),$$

Where $P(A_i \mid A_1 \cap A_2 \cap \ldots \cap A_{i-1})$ represents the conditional probability of the event A_i, given that the events $A_1, A_2 \ldots A_{i-1}$ have already happened.

(iii) **Multiplication theorems for independent events -** If A and B are independent events associated with a random experiment, then $P(A \cap B) = P(A).P(B)$ *i.e.*, the probability of simultaneous occurrence of two independent events is equal to the product of their probabilities.

By multiplication theorem, we have $P(A \cap B) = P(A).P(B/A)$.

Since A and B are independent events, therefore $P(B \mid A) = P(B)$. Hence, $P(A \cap B) = P(A).P(B)$.

(iv) **Extension of multiplication theorem for independent events -** If $A_1, A_2 \ldots A_n$ are independent events associated with a random experiment, then $P(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n) = P(A_1)P(A_2) \ldots P(A_n)$.

By multiplication theorem, we have

$$P(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \ldots P(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

Since $A_1, A_2 \ldots A_{n-1}, A_n$ are independent events, therefore

$$P(A_2 \mid A_1) = P(A_2), P(A_3 \mid A_1 \cap A_2) = P(A_3) \ldots P(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1}) = P(A_n)$$

Hence, $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2) \ldots . P(A_n)$.

(2) **Probability of at least one of the n independent events -** If $p_1, p_2, p_3, \ldots\ldots, p_n$ be the probabilities of happening of n independent events $A_1, A_2, A_3, \ldots\ldots, A_n$ respectively, then

(i) Probability of happening none of them

$$= P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \ldots\ldots \cap \overline{A}_n) = P(\overline{A}_1).P(\overline{A}_2).P(\overline{A}_3) \ldots\ldots P(\overline{A}_n) = (1-p_1)(1-p_2)(1-p_3) \ldots . (1-p_n).$$

(ii) Probability of happening at least one of them

$$= P(A_1 \cup A_2 \cup A_3 \ldots . \cup A_n) = 1 - P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3) \ldots . P(\overline{A}_n) = 1 - (1-p_1)(1-p_2)(1-p_3) \ldots (1-p_n).$$

(iii) Probability of happening of first event and not happening of the remaining

$$= P(A_1)P(\overline{A}_2)P(\overline{A}_3) \ldots\ldots P(\overline{A}_n) = p_1(1-p_2)(1-p_3) \ldots\ldots . (1-p_n)$$

PROBABILITY

TOTAL & INVERSE PROBABILITY (BAYE's THEOREM)

(1) **The law of total probability -** Let S be the sample space and let $\mathrm{E}_1, \mathrm{E}_2 \ldots \mathrm{E}_\mathrm{n}$ be n mutually exclusive and exhaustive events associated with a random experiment. If A is any event which occurs with E_1 or E_2 or … or E_n, then $P(A) = P(E_1)P(A|E_1) + P(E_2)P(A|E_2) + \ldots + P(E_n)P(A|E_n)$.

(2) **Baye's rule for inverse probability -** Let S be a sample space and $\mathrm{E}_1, \mathrm{E}_2 \ldots \mathrm{E}_\mathrm{n}$ be n mutually exclusive events such that $\bigcup_{i=1}^{n} E_i = S$ and $P(E_i) > 0$ for $i = 1, 2 \ldots n$. We can think of (E_i's as the causes that lead to the outcome of an experiment. The probabilities $P(E_i)$, $i = 1, 2, \ldots, n$ are called prior probabilities. Suppose the experiment results in an outcome of event A, where $P(A) > 0$. We have to find the probability that the observed event A was due to cause E_i, that is, we seek the conditional probability $P(E_i / A)$. These probabilities are called posterior probabilities, given by Baye's rule as $P(E_i \mid A) = \dfrac{P(E_i).P(A|E_i)}{\sum_{k=1}^{n} P(E_k)P(A|E_k)}$.

BINOMIAL PROBABILITY

Binomial probability distribution - A random variable X which takes values 0, 1, 2, …, n is said to follow binomial distribution if its probability distribution function is given by $P(X = r) = {}^nC_r p^r q^{n-r}, r = 0, 1, 2, \ldots\ldots, n$ where $p, q > 0$ such that $p + q = 1$.

The notation $X \sim B(n, p)$ is generally used to denote that the random variable X follows binomial distribution with parameters n and p.

We have $P(X = 0) + P(X = 1) + \ldots + P(X = n) = {}^nC_0 p^0 q^{n-0} + {}^nC_1 p^1 q^{n-1} + \ldots + {}^nC_n p^n q^{n-n} = (q + p)^n = 1^n = 1$.

Now probability of

(a) Occurrence of the event exactly r times

$P(X = r) = {}^nC_r q^{n-r} p^r$.

(b) Occurrence of the event at least r times

$P(X \geq r) = {}^nC_r q^{n-r} p^r + \ldots + p^n = \sum_{X=r}^{n} {}^nC_X p^X q^{n-X}$.

(c) Occurrence of the event at the most r times

$P(0 \leq X \leq r) = q^n + {}^nC_1 q^{n-1} p + \ldots + {}^nC_r q^{n-r} p^r = \sum_{X=0}^{r} p^X q^{n-X}$.

(iv) If the probability of happening of an event in one trial be p, then the probability of successive happening of that event in r trials is p^r.

(i) **Mean and variance of the binomial distribution -** The binomial probability distribution is

X	0	1	2		n
$P(X)$	${}^nC_0 q^n p^0$	${}^nC_1 q^{n-1} p$	${}^nC_2 q^{n-2} p^2$	.....	${}^nC_n q^0 p^n$

The mean of this distribution is $\sum_{i=1}^{n} X_i p_i = \sum_{X=1}^{n} X.\, {}^nC_X q^{n-X} p^X = np$,

the variance of the Binomial distribution is $\sigma^2 = npq$ and the standard deviation is $\sigma = \sqrt{(npq)}$.

(ii) **Use of multinomial expansion -** If a die has m faces marked with the numbers 1, 2, 3, ….m and if such n dice are thrown, then the probability that the sum of the numbers exhibited on the upper faces equal to p is given by the coefficient of x^p in the expansion of $\frac{(x + x^2 + x^3 + ... + x^m)^n}{m^n}$.

INFINISTIC PROBABILITY

(1) **Geometrical method for probability -** When the number of points in the sample space is infinite, it becomes difficult to apply classical definition of probability. For instance, if we are interested to find the probability that a point selected at random from the interval [1, 6] lies either in the interval [1, 2] or [5, 6], we cannot apply the classical definition of probability. In this case we define the probability as follows-

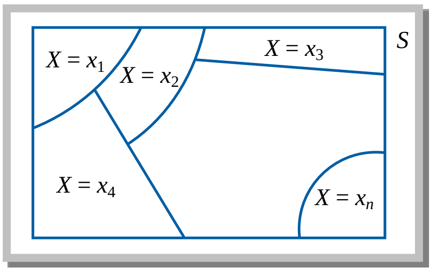
$$P\{x \in A\} = \frac{\text{Measure of region } A}{\text{Measure of the sample space } S},$$

where measure stands for length, area or volume depending upon whether S is a one-dimensional, two-dimensional or three-dimensional region.

(2) **Probability distribution -** Let S be a sample space. A random variable X is a function from the set S to R, the set of real numbers.

For example, the sample space for a throw of a pair of dice is $S = \begin{matrix} \{11, & 12, & \cdots, & 16 \\ 21, & 22, & \cdots, & 26 \\ \vdots & \vdots & \ddots & \vdots \\ 61, & 62, & \cdots, & 66\} \end{matrix}$

Let X be the sum of numbers on the dice. Then $X(12) = 3, X(43) = 7$, etc. Also, $\{X = 7\}$ is the event {61, 52, 43, 34, 25, 16}. In general, if X is a random variable defined on the sample space S and r is a real number, then $\{X = r\}$ is an event. If the random variable X takes n distinct values $x_1, x_2, ..., x_n$, then $\{X = x_1\}$, $\{X = x_2\}, ..., \{X = x_n\}$ are mutually exclusive and exhaustive events.



Now, since $(X = x_i)$ is an event, we can talk of $P(X = x_i)$. If $P(X = x_i) = P_i (1 \le i \le n)$, then the system of numbers.

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$$

is said to be the probability distribution of the random variable X. The expectation (mean) of the random variable X is defined as $E(X) = \sum_{i=1}^{n} p_i x_i$

and the variance of X is defined as $\text{var}(X) = \sum_{i=1}^{n} p_i (x_i - E(X))^2 = \sum_{i=1}^{n} p_i x_i^2 - (E(X))^2$.