

Basic Concepts of Probability

The term probability as its origin relates with the games of chances in the seventh century. Tossing of a coin, throwing of a die, drawing cards from a packet etc. in which the outcome of a trial is uncertain. But the total numbers of possible outcomes are known before the trial. Although the outcomes of a particular trial are **uncertain**, there are **predictable** long term outcome. This type of uncertainty and long term regularity often occurs in the experimental sciences also. As for example,

- a) Difficult to say whether tomorrow will be rain or not, but after a long term observations, one can easily predict the probability of this event.
- b) In the science of genetics, it is uncertain whether a newly born child will be male or female. But the male and female birth rate of community can be known satisfactorily.
- c) In demography it is difficult to say that Mr. X will die at age 30, but life-table of a community can be constructed at every age.

Roughly a numerical measure of **uncertainty** of an event of an experiment is probability.

Experiment: An experiment is an act that can be repeated under certain given identical conditions. Examples of experiment are tossing a coin, throwing a die, drawing a card from a well shuffled packet etc.

Trial: When an experiment is executed single time then it is known as trial. Such as tossing a coin single time, throwing a die single time etc.

Outcome: An outcome is the result obtained from an experiment. Such as, if you throw a fair coin single time then outcome will be either H or T or if you throw a dice single time then the outcomes are any numbers from 1 to 6.

Random Experiment: An experiment is called random experiment if the occurrence of an outcome has specific probability and cannot be predicted with certainty. Let us consider the sex of two newly born baby in a locality then the outcomes are $S=\{BB, BG, GB, GG\}$ each having probability $\frac{1}{4}$.

Others examples are tossing a coin, throwing a die, drawing of card from a pack etc.

Sample Space: The collection or totality of the all possible outcomes of a random experiment is called sample space.

Sample space is usually denoted by Ω or S . Universal set can be considered as sample space. Sample space is sometime called sure event.

Example: If a coin is tossed twice then the sample space $\Omega=\{HH, HT, TH, TT\}$. If you consider the sex of two newborn babies in a locality then the sample space $\Omega=\{GG, GB, BG, BB\}$ etc.

Event: An event is a subset of the sample space and is usually denoted by capital letter A, B, X, Y etc. Let us consider a dice throwing experiment then $\Omega=\{1, 2, 3, 4, 5, 6\}$. Let A be an event only odd number will occur. Then $A=\{1, 3, 5\}$. We have two types of events: (i) simple event (ii) compound event.

Simple event: An event is called simple event if it contains only one sample point. Such as $A=\{HH\}$, $B=\{HT\}$.

Compound event: An event is called compound event if it contains more than one sample point or it is the union of simple event. Such as $A=\{HH, HT\}$, $B=\{2, 4, 6\}$.

Event space: The class of all events associated with a given experiment is defined to be the event space. It is denoted by script Latin letter such as $\mathcal{A}, \mathcal{B}, \mathcal{H}$ etc.

Example: Suppose a fair coin is tossed twice or two coins are tossed simultaneously, then the sample space of the experiment will be $\Omega = \{HH, HT, TH, TT\} = \{w_1, w_2, w_3, w_4\}$. Where H and T denotes the head and tail of the coin. Here, the event space will contain $2^4 = 16$ events. The event space will be:

$$A = \left\{ (w_1), (w_2), (w_3), (w_4), (w_1, w_2), (w_1, w_3), (w_1, w_4), (w_2, w_3), (w_2, w_4), (w_3, w_4), (w_1, w_2, w_3), (w_1, w_2, w_4), (w_1, w_3, w_4), (w_2, w_3, w_4), \Omega, \emptyset \right\}$$

Mutually exclusive event: If A and B are two events in \mathcal{A} , then they are said to be mutually exclusive if $A \cap B = \emptyset$. That is two events are said to be mutually exclusive if they have **no common point**.

Complementary event: Let A be any event defined on a sample space Ω then the complementary of A , denoted by \bar{A} or A^c is the event consisting of all the sample points in Ω but not in A .

Let us consider an experiment of tossing a coin twice, then $A = \{\text{exactly two head will occur}\}$, $A = \{HH\}$, \bar{A} or $A^c = \{HT, TH, TT\}$.

Probability space: In probability measure **three** things must be specified, viz, the sample space Ω , the event space \mathcal{A} , (an algebra of events) and the probability $P[.]$ with domain \mathcal{A} .

Definition: The triplet $(\Omega, \mathcal{A}, P[.])$ is called probability space, where Ω is the sample space, \mathcal{A} is collection of events and $P[.]$ is a probability function with domain \mathcal{A} .

Definition of Probability:

There are mainly three definitions of probability. They are:

1. Mathematical or Classical or a priori definition of probability

2. Statistical or Empirical or posteriori definition of probability
3. Axiomatic definition of probability

Mathematical or Classical or a priori definition of probability: If a trial results in n exhaustive, mutually exclusive and equally likely outcomes and m of these outcomes are favorable to a particular event A , then the probability P of happening the event A is given by:

$$P = P(A) = \frac{\text{favorable number of outcomes or cases}}{\text{total number of possible outcomes or cases}} = \frac{m}{n}$$

Example 1: A card is drawn from a pack of 52 cards. Find the probability that the selected card is:

- (a) a red card (b) a spade (c) not a spade (d) a king or a queen

Solution: When a card is drawn from a pack of 52 cards, then the total numbers of equally likely, mutually exclusive and exhaustive outcomes are 52. That is here $n=52$.

- (a) Let A be the event of drawing a red card. There are 26 black and 26 red cards in a pack and any one of the red card can be drawn in 26 ways. That is $m=26$. Therefore,

$$P(A) = \frac{26}{52} = \frac{1}{2}$$

- (b) Let B be the event of drawing a spade. Here the favorable numbers of cases of the event B are $m=13$ and total numbers of cases are $n =52$. Therefore,

$$P(B) = \frac{13}{52} = \frac{1}{4}$$

- (c) Let C be the event of drawing a card of not spade. Here the favorable numbers of cases of the event C are $m=39$ and total numbers of cases are $n =52$. Therefore,

$$P(c) = \frac{39}{52} = \frac{3}{4} \quad \text{or} \quad P(C) = 1 - P(B) = 1 - \frac{1}{4} = \frac{3}{4}$$

(d) Let D be the event of that the drawn card is king or queen. Here in the 52 cards there are 4 kings and 4 queens, so the favorable numbers of cases of the event D are $m=8$ and total numbers of cases are $n=52$. Therefore,

$$P(D) = \frac{8}{52} = \frac{2}{13}$$

Example 2: Three contractors A, B and C are bidding for the construction of a new Auditorium. Some experts in construction industry believes that A has exactly half of the chance that B has, B in turn, is $\frac{4}{5}$ th as likely as C to win the contract. What is the probability for each to win the contract if the experts estimates are accurate?

Solution: Let us suppose that the probability of winning the contract of C's is x . Then, $P(C) = x$, $P(B) = \frac{4}{5}x$ and $P(A) = \frac{4}{5} \cdot \frac{1}{2}x$. Since the total probability is 1. So we can write, $P(A) + P(B) + P(C) = 1$ or, $x + \frac{4}{5}x + \frac{2}{5} = 1$

$$\therefore x = \frac{5}{11}$$

Hence, $P(C) = \frac{5}{11}$, $P(B) = \frac{4}{11}$ and $P(A) = \frac{2}{11}$

Example: Tickets are numbered from 1 to 100. They are well shuffled and a ticket is drawn at random. What is the probability that the drawn ticket has:

- 1) An odd number
- 2) A number 4 or multiple of 4
- 3) A number which is greater than 70 and

4) A number which is square?

Solution: Since there are 100 tickets, the total number of exclusive mutually exclusive and equally likely case is 100.

1) Let A denote the event that the ticket drawn an odd number. Since there are 50 odd number tickets so the number of cases favorable to the event A is 50.

$$\therefore P(A) = \frac{50}{100} = 0.5$$

2) Let B denote the event that the ticket drawn has a number 4 or multiple of 4. The numbers favorable to event B are 4, 8, 12, 16, 20, ..., 92, 96, 100. The total number of cases will be

$$\frac{100}{4} = 25$$

$$\therefore P(B) = \frac{25}{100} = 0.25$$

3) Let C denote the event that the drawn ticket has a number greater than 70. Since the number greater than 70 are 71, 72, 73, ..., 100. Therefore, 30 cases are favorable to the event C .

$$\therefore P(C) = \frac{30}{100} = 0.3$$

4) Let D denote the event that the drawn ticket has a number which is a square. Since the squares between 1 and 100 are 1, 4, 9, 16, 25, 36, 49, 64, 81 and 100. So the cases favorable to event D are 10 in number. Hence,

$$\therefore P(D) = \frac{10}{100} = 0.1$$

Limitations of Classical Probability: Following are the limitations of classical probability:

a. The classical probability fails to define probability when the total numbers of possible outcomes are infinite.

- b. The classical definition of probability breakdown if the outcomes of the trial are not equally likely. Such as if a coin is biased in favor of head (it is bent so that head is more likely to appear than tail). Here two outcomes are not equally likely.
- c. It is not always possible to enumerate all the equally likely outcomes.
- d. It is not possible to say the probability that a newly born baby will be a boy.
- e. It is not possible to say the probability that a person will survive up to the age of 85 years.

Statistical or Empirical or posteriori definition of probability: 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 infinitely large is called the statistical or empirical probability. It assumed that the limit is finite and unique.

Symbolically, if in n trials an event A happens m times, then the probability P of the happening A is given by:

$$P = P(A) = \lim_{n \rightarrow \infty} \frac{m}{n}$$

Example: A fair coin was tossed 100, 500, 1000 times and the outcomes were recorded in the table 1 and a fair die was tossed 1200 times and the outcomes were recorded in the table 2. The important thing is that the relative frequencies of heads and tails are very close to $\frac{1}{2}$ (as trial increases it is more closer to 0.5). Similarly the relative frequencies of faces showing points 1, 2, 3, 4, 5 and 6 are very close to $\frac{1}{6}$.

Table 1: Showing outcomes of tossing a fair coin 100, 500 and 1000 times

Outcomes	Relative (observed frequency			Expected relative frequency
	100	500	1000	
Head	0.46 (46)	0.49 (245)	0.494	0.50
Tail	0.54 (54)	0.51 (510)	0.506	0.50

Table 2: Showing outcomes of throwing a fair dice 1200 times

Outcomes	Observed	Relative frequency	Expected relative frequency
1	204	0.1700	0.1667
2	212	0.1767	0.1667
3	192	0.1600	0.1667
4	204	0.1700	0.1667
5	196	0.1633	0.1667
6	192	0.1600	0.1667

From both the experiment it is reasonable to assume that if we toss a coin or thrown a fair dice, then there exists a number say P will be the probability of head or a face point of the die. If the coin is fair, this number will be equal to $\frac{1}{2}$ and in case of dice it is $\frac{1}{6}$.

Limitations: Following are the limitations of Statistical or Empirical or posteriori definition of probability:

- Practically it is not possible to repeat the experiment an infinite number of times under the identical conditions to get the probability.
- Even if it is possible to repeat an experiment an infinite number of times, it is conceivable that a different infinite sequence of performance of the same experiment could produce a different value of P .
- It is not clear how large n should be before we are certain that the probability P is close to the limiting value of $\frac{m}{n}$ as $n \rightarrow \infty$.

Axiomatic definition of probability: The axiomatic probability was proposed by A.N. Kolmogorov in 1933. Axiomatic definition of probability is based on statement that can be deduced either from axioms or from previously proved theorem. The axioms are:

- i. $P(A) \geq 0$ for every event $A \in \mathcal{A}$,
- ii. $P(\Omega) = 1$
- iii. Let A_1 and A_2 be mutually exclusive events in \mathcal{A} , then

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

- iv. Let $A_1, A_2, \dots, A_n, \dots$ be a sequence of mutually exclusive events in \mathcal{A} , and if $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

The above axioms are termed as the axioms of positiveness, certainty and union respectively.

Subjective or Judgment Probability: The probability that a person assigns to an event which is the possible outcomes of some processes on the basis of his own judgment, belief and information about the process is known as subjective probability. As this probability depends on individual judgment, belief and perception, so it varies from individual to individual. As for example, one fine morning Mr. may well prepared for rain but his friend Mr. Jami may not.

Geometric Probability (Definition): If Q is sample region with well defined measure, the probability, that a point chosen at random lies in a sub region A of Q is called the geometric probability and is defined by the ratio equal to P as:

$$P = P(A) = \frac{\text{measure of specific part } A \text{ of } Q}{\text{measure of the whole region } Q}$$

where measure refers to the length, area or volume of the region if we are dealing with one, two or three dimensional space respectively.

Relative Frequency: Relative frequency is the ratio of the individual cell frequency and the total number of observations. As the limits of relative frequency varies from 0 to 1, but we can't treat this as probability. It is the approximation of probability. But when the numbers of observation are sufficiently large then relative frequency can be treated as probability.

Probability of Sample space: Let Ω be sample space associated with a random experiment. Then the probability of the sample space is always 1 that is $P(\Omega) = 1$.

Example: Two coins are tossed simultaneously then the sample space is $\Omega = \{HH, HT, TH, TT\}$. Now $P(\Omega) = 1$.

Probability of an Event: Probability of an event is always greater than zero. Probability equals 0 if the event is null or empty event.

Example: Consider three items from a lot of a production process. If the defective item is denoted by D and non-defective as N then the sample space is: $\Omega = \{NNN, NND, NDN, DNN, DDN, DND, NDD, DDD\}$. Let us define an event A that two items are non-defective then $A = \{NND, NDN, DNN\}$ and $P(A) = \frac{3}{8}$. Similarly the probability of every events defined from the above sample space will be greater than 0.

Tree diagram: Very often the sample points of an experiment consist of more than one element. This happens when the experiment is conducted with more than one-steps. The sample space of such experiment can be represented by a diagram which we call tree diagram. Sample points of such experiment can be displayed by a tree diagram.

Example: Let us assume that two coins are tossed simultaneously, then show the sample points with a tree diagram and construct the sample space of the experiment with their probabilities.

Solution: The experiment can be considered as two steps. There are two possible outcomes for the first coin and two possible outcomes for the second coin. The total number of outcomes will be 2×2 and the tree diagram of the sample space will be as follows:

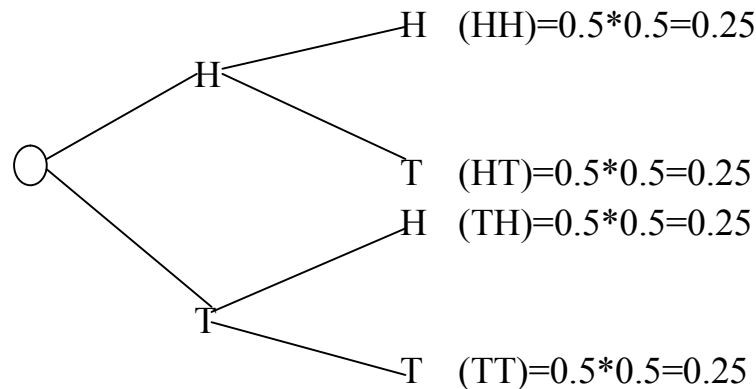
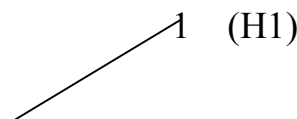


Figure: Tree diagram showing outcomes of tossing two coins simultaneously.

The sample space of the experiment is $S = \{HH, HT, TH, TT\}$. There are four sample points of this random experiment with probability $\frac{1}{4}$ and each outcome consists with two elements one from each coin.

Example: Toss a coin and a dice simultaneously. Show the sample points with a tree diagram and construct the sample space.

Solution: [The experiment can be considered as two steps. There are two possible outcomes for the coin and six possible outcomes for the die. The total number of outcomes will be $2 \times 6 = 12$. The sample space of the experiment is: $S = \{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6\}$]



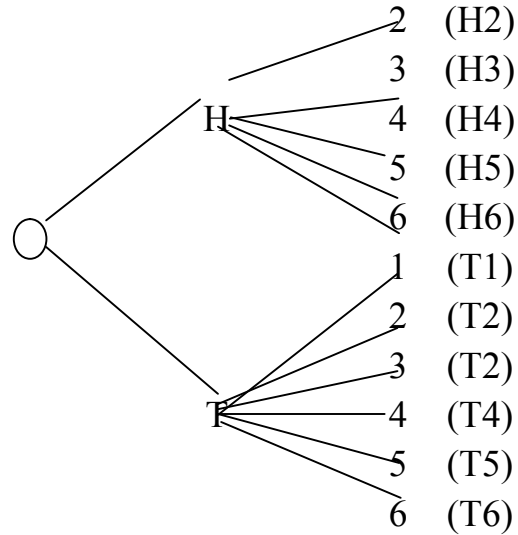


Figure: Tree diagram showing outcomes of a coin and a die.

Example: Suppose that three items are selected at random from a manufacturing process. Each item is inspected and classified as defective, D, or non defective N. List the elements of the sample space with a tree diagram and construct the sample space of the experiment. Also compute the probabilities of the outcomes from tree diagram.

Solution: The experiment can be considered at three steps. There two possible outcomes in the 1st step i.e. the selected 1st item may be D or N. Similarly the selected 2nd and 3rd item may be D or N. The total number of outcomes will be $2 \times 2 \times 2 = 8$. The total number the experiment is:

$$S = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}$$

The tree diagram of the outcomes and the probabilities are shown as follows:

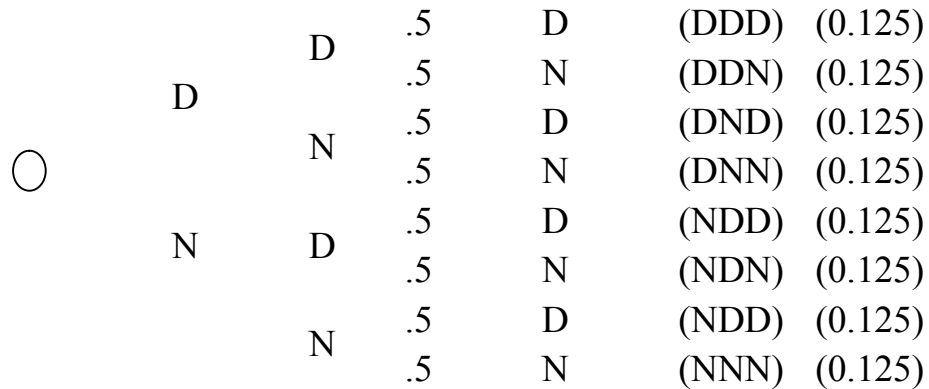


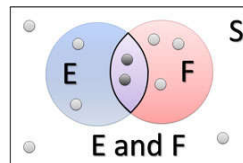
Figure: Tree diagram showing possible outcomes of three items.

Laws of Probability: There are **two** important rules or laws of probability. They are:

- i) Additive laws of probability
- ii) Multiplication laws of probability

Additive Laws (General Rules of Addition): If A and B are two events, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof: Let A and B be two events, \bar{A} and \bar{B} be their complementary events respectively. From the Venn-diagram it is seen that:



$$A = AB \cup A\bar{B}, \text{ then } P(A) = P(AB \cup A\bar{B}) = P(AB) + P(A\bar{B}) \dots (1)$$

Since AB and $A\bar{B}$ are mutually exclusive.

Similarly, $B = AB \cup \bar{A}B$, then $P(B) = P(AB \cup \bar{A}B) = P(AB) + P(\bar{A}B) \dots (2)$

It also seen that $A \cup B = A\bar{B} \cup AB \cup \bar{A}B$

Then $P(A \cup B) = P(A\bar{B}) + P(AB) + P(\bar{A}B) \dots (3)$

Since $A\bar{B}$, AB and $\bar{A}B$ are mutually exclusive. By adding equation (1) and (2) we get then we use the general rule for addition. The rule is

$$\begin{aligned} P(A) + P(B) &= P(A \cap B) + P(A \cap \bar{B}) + P(B) + P(\bar{A} \cap B) \\ &= P(A \cap B) + P(A \cap \bar{B}) + P(B) + P(\bar{A} \cap B) \\ &= P(A \cap B) + P(A \cup B) \end{aligned}$$

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

Example: The probability that a contractor will get a plumbing contract is $2/3$ and the probability that he will get an electric contract is $4/9$. The probability of getting both the contract is $14/45$, what is the probability that he will get at least one contract?

Solution: Let A and B be two events that a contractor will get a plumbing and electric contract respectively. Then $P(A) = 2/3$, $P(B) = 4/9$ and $P(A \cap B) = 14/45$. Now the contractor will get at least one contract is:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 2/3 + 4/9 - 14/45 = \frac{30 + 20 - 14}{45} = \frac{36}{45} = \frac{4}{5} = 0.75$$

Example: Mr. X feels that the probability that he will pass Mathematics is $\frac{2}{3}$ and Statistics is $\frac{5}{6}$. If the probability that he will pass both the course is $\frac{3}{5}$. What is the probability that he will pass at least one of the courses?

Solution: Let M and S be the events that he will pass the courses Mathematics and Statistics respectively. The event $M \cup S$ means that at least one of M or S occurs. Therefore

$$\begin{aligned} P(M \cup S) &= P(M \text{ or } S) \\ &= P(\text{he pass at least one of the course}) \end{aligned}$$

$$\begin{aligned}
 &= P(M) + P(S) - P(M \text{ or } S) \\
 &= \frac{2}{3} + \frac{5}{6} - \frac{3}{5} = \frac{9}{10}.
 \end{aligned}$$

Example: Mr. Y feels that the probability that he will get A in Calculus is $\frac{3}{4}$, A in Statistics is $\frac{4}{5}$ and A in both the courses is $\frac{3}{5}$. What is the probability that Mr. Y will get: (a) at least one A (b) No A 's.

Solution: Let C be the event that Mr. Y will get A in Calculus and S be the event that he will get A in Statistics. We have, $P(C) = \frac{3}{4}$, $P(S) = \frac{4}{5}$ and $P(C \text{ and } S) = P(C \cap S) = \frac{3}{5}$

$$\begin{aligned}
 \text{a) } P(\text{at least one } A) &= P(C \cup S) = P(C) + P(S) - P(C \cap S) = \frac{3}{4} + \frac{4}{5} - \frac{3}{5} \\
 &= \frac{19}{20}
 \end{aligned}$$

$$\text{b) } P(\text{no } A's) = \overline{P(C \cup S)} = 1 - P(C \cup S) = 1 - \frac{19}{20} = \frac{1}{20}$$

Additive Laws (Special rule): If two events A and B are mutually exclusive, the special rule of addition states that the probability of one or the other events occurring equals the sum of their probabilities i., e.,

$$\begin{aligned}
 P(\text{at least one event will Occur}) &= P(A \text{ or } B) = P(A \cup B) \\
 &= P(A) + P(B)
 \end{aligned}$$

For three mutually exclusive events designated A , B and C the rule is written as

$$P(A \text{ or } B \text{ or } C) = P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

Example: If we toss a coin then what is the probability of head or tail?

Solution: Here there are two events, namely event $A=H$ and event $B=T$.
So that

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) = \frac{1}{2} + \frac{1}{2} = 1$$

Multiplication Law (General Rule): The general rule of multiplication states that for two events A and B , the joint probability that both events will happen is found by multiplying the probability of event A will happen by the conditional probability of event B occurring given that event A has occurred. Symbolically, the joint probability is:

$$P(A \text{ and } B) = P(A \cap B) = P(A) * P(B|A) = P(B) * P(A|B)$$

Proof: From the definition of conditional probability, we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ provided } P(B) > 0$$

It follows that $P(A \cap B) = P(B) * P(A|B)$

$$\text{Similarly } P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{ provided } P(A) > 0$$

It follows that $P(A \cap B) = P(A) * P(B|A)$

$$\therefore P(A \cap B) = P(A) * P(B|A) = P(B) * P(A|B)$$

Example: There are 10 rolls of film in a box, 3 of which are defective. Two rolls are to be selected one after another. What is the probability of selecting a defective roll followed by another defective roll?

Solution: The first roll of film selected from the box being found defective is event D_1 and the second roll selected being found defective is event D_2 .

$$P(D_1) = \frac{3}{10} \text{ and } P(D_2|D_1) = \frac{2}{9}$$

Since, after the first selection was found to be defective, only 2 defective rolls of film remained in the box containing 9 rolls.

So the probability of two defectives is:

$$P(D_1 \text{ and } D_2) = P(D_1) * P(D_2|D_1) = \frac{3}{10} \cdot \frac{2}{9} = \frac{6}{90} = 0.07$$

Multiplication Law (Special Rule): Let the two events A and B are independent (two events are independent if the occurrence of one event does not alter the probability of the occurrence of the other event).

For two independent events A and B , the probability that A and B will both occur is found by multiplying the two probabilities i., e.,

$$P(A \text{ and } B) = P(A) * P(B).$$

For three events A , B and C the special rule of multiplication used to determine the probability that all the events will occur is:

$$P(A \text{ and } B \text{ and } C) = P(A) * P(B) * P(C)$$

Example: A company has two large computers. The probability that the newer one will breakdown on any particular month is 0.05, the probability that the older one will breakdown on any particular month is 0.1. What is the probability that both the computer will breakdown in a particular month?

Solution: Let, Event A is the newer one will breakdown and Event B is the older one will breakdown. So that $P(A) = 0.05$ and $P(B) = 0.1$.

$$\therefore P(A \text{ and } B) = P(A) * P(B) = 0.05 * 0.1 = 0.005.$$

Compound event (Definition): An event is called compound event if it contains more than one sample point or it is the union of sample events.

Example: Suppose a fair coin is tossed twice. Let H and T denotes the head and tail of the coin respectively. The sample space of the experiment is

$$\Omega = \{HH, HT, TH, TT\} = \{w_1, w_2, w_3, w_4\}$$

In this example, there are four sample vents which are:

$$w_1 = \{HH\}, w_2 = \{HT\}, w_3 = \{TH\}, w_4 = \{TT\}$$

Let A be the event of head of the first coin, then A will contain the sample points $A = \{HH, HT\} = \{w_1, w_2\}$. Here A is compound event, since it contains two sample points.