Probability

Random Experiment:

An experiment where all the possible outcomes are known, but it is not possible to predict which of the outcomes will actually occur when it is performed.
Ex: The possible outcomes of a coin toss are head and tail, but we can't predict head/tail when the coin is actually tossed.

Deterministic Experiment:

An experiment where there is only one outcome.
Ex: Experiment to check the law of gravity.

Trial:

Trial is the act of performing an experiment.
Ex: Tossing a coin, throwing a dice.

Sample Space:

The set 'S' of all possible outcomes of a random experiment.
Ex:

1. Tossing a coin, the sample space, S is given by {H,T}.
2. Throwing a dice, the sample space, S is given by {1,2,3,4,5,6}, all the possible outcomes of getting each of the 6 faces.
3. If a coin is tossed twice (or two coins tossed simultaneously):
   S = {HH,HT,TT,TH}
4. If a bag contains 3 red and 4 black balls and if one ball is drawn from the bag, then S = $\{R_1,R_2,R_3, B_1,B_2,B_3,B_4,B_5\}$.

Sample Point:

Each outcome of an experiment is visualized as sample point in the sample space. i.e., if a die is thrown twice, then getting (1,2), (4,1),…. (6,6) are all sample points.

Event:

The set of one or more possible outcomes of an experiment constitutes an event. An event is a subset of the sample space.

Ex:

1. When a die is thrown, the event of getting a number of less than 5 is the set, {1,2,3,4}, the subset comprising all the possible outcomes less than 5 within the set {1,2,3,4,5,6}.

Mutually Exclusive Cases:

Cases are said to be mutually exclusive if the happening of any one of them prevents the happening of all the others in a single experiment.
Ex: In a coin tossing experiment, head and tail are mutually exclusive.

Classical Probability:

Classical probability of an event is given by:

P(A) = $\frac{Number\ of\ favoruable\ outcomes}{Total\ number\ of\ outcomes}$

Note: Probability always lies between 0 and 1.

Q1. Find the probability of getting 53 Sundays in a randomly selected non-leap year.

Q2. Out of 52 well shuffled playing cards, two cards are drawn at random. Find the probability of getting one red and one black card.

However, we have to note that the classical definition of probability fails in case where the outcomes are not equally likely or the total number of outcomes is indefinitely large.

Other Approaches to Probability Theory:

In situations where the outcomes are not equally likely or not finite, we need other approaches to calculate probability.

Consider a situation where a person is administered a sleeping pill and we are interested in finding the probability that the pill puts the person to sleep in 20 minutes. Here, we cannot say that the pill will be equally effective for all persons and hence we cannot apply classical definition here.

The solution for the above case is conducting experiments and observing the data.

Empirical Probability:

This approach to probability is called the **relative frequency** approach and it defines the statistical probability. If an event A (say) happens m times in n trials of an experiment which is performed repeatedly under essentially homogeneous and identical conditions, then the (**Statistical or Empirical**) probability of happening A is defined as P(A):

$$\lim_{n \to \infty} \frac{m}{n}$$

Laws of Probability:

Addition Law - Let S be the sample space of a random experiment and events A and B are subsets of S, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Conditional Probability:

Let A be the probability of picking a red ball from a basket containing 2 red cricket balls, 3 red tennis balls, and 5 green cricket balls. Then A is given by 5/10 = 1/2.

But, if we were told that an event B of picking a cricket ball has already occurred, then the P(A) will no longer be the same, as the denominator 'total outcomes', has changed from 10 to 7 (since a cricket ball has already been picked). That is the sample space has changed from 10 to 7.

Now the probability of picking a red ball is 2/7.

This is called conditional probability and is represented by $P(A|B)$.

Multiplication Law -

$P(A \cap B) = P(A)\, P(B|A)$ where $P(A) > 0$

$= P(B)\, P(A|B)$ where $P(B) > 0$

Independent Events -

Events are said to be independent if happening or non-happening of any one event is not affected by the happening or non-happening of other events.
Ex: If a coin is tossed several times, the outcome each time is not dependent on any past outcome.

Two events A and B are independent iff $P(B|A) = P(B)$

Thus, the multiplicative law for independent events is: $P(A \cap B) = P(A)\, P(B)$

Note: Mutually exclusive events, by definition, can never by independent.

Q1. Three cards are drawn one by one without replacement from a well shuffled pack of 52 playing cards. What is the probability that first card is jack, second is queen and the third is again a jack.

Q2. Three unbiased coins are tossed simultaneously. In which of the following cases are the events A and B independent?

(i) A be the event of getting exactly one head B be the event of getting exactly one tail

(ii) A be the event that first coin shows head B be the event that third coin shows tail

(iii) A be the event that shows exactly two tails B be the event that third coin shows head

Q3. A person is known to hit the target in 4 out of 5 shots whereas another person is known to hit 2 out of 3 shots. Find the probability that the target being hit when they both try.

Contingency Table - It helps us summarize two variables.

Ex: A factory produces certain type of output by 3 machines. The respective daily production figures are- machine X : 3000 units, machine Y: 2500 units and machine Z: 4500 units. Past experience shows that 1% of the output produced by machine X is defective. The corresponding fractions of defectives for the other two machines are 1.2 and 2 percent respectively. An item is drawn from the day's production. What is the probability that it is defective?

	Machines	

Ans:

		X	Y	Z	Total
Defect	**Yes**	30	30	90	150
	No	2970	2470	4410	9850
	Total	3000	2500	4500	10000

Convert each cell into probabilities.

		Machines			
		X	**Y**	**Z**	**Total**
Defect	**Yes**	0.003	0.003	0.009	0.015
	No	0.297	0.247	0.441	0.985
	Total	0.3	0.25	0.45	1

Marginal Probabilities - P(X), P(Y), P(Z), P(Yes) and P(No) are given by the figures in the margins:

		Machines			
		X	**Y**	**Z**	**Total**
Defect	**Yes**	0.003	0.003	0.009	**0.015**
	No	0.297	0.247	0.441	**0.985**
	Total	**0.3**	**0.25**	**0.45**	1

P(Yes) = 0.015 and P(No) = 0.985

Joint Probability - Probabilities describing a combination of attributes.

P(X and Yes) = 0.003

P(Z and No) = 0.441

Union Probability - Addition Law

P(Yes or Y) = P(Yes) + P(Y) - P(Y and Yes)

= 0.015 + 0.25 - 0.003

= 0.262

Conditional Probability - What is the probability that an item produced by the machine Z is not defective?

P(No | Z) = $\frac{P(No \cap Z)}{P(Z)}$; this is nothing but *joint probability* divided by *marginal probability*

= 0.441/0.45 = 0.98

P(Z | No) = ?

Is it the same as above?

Total Probability:

The above (machine - defective item) example can be categorized as a two stage experiment. We first pick the machine and then pick the item. Another example is: If there are two bags with bag 1 containing 2 green and 3 red balls and bag 2 containing 5 green and 6 red balls. Now, we have to randomly pick one of the bags and then pick a ball from that bag. This includes two stages.

Calculating the probability of picking a particular ball in this case is called **Total probability,** given by P(A):

$$\sum_{i=1}^{n} P(E_i)P(A|E_i)$$

Where E corresponds to the first stage

For the previous problem, using the contingency table, we can obtain the probability of picking a defective item (total probability) as 0.015.

The same result can be obtained by using the above formula:

$P(E_1) = 3000/10000 = 3/10$; $P(E_2) = 2500/10000 = 1/4$; $P(E_3) = 4500/1000 = 9/20$

$P(A|E_1) = 0.01$; $P(A|E_2) = 0.012$; $P(A|E_3) = 0.02$

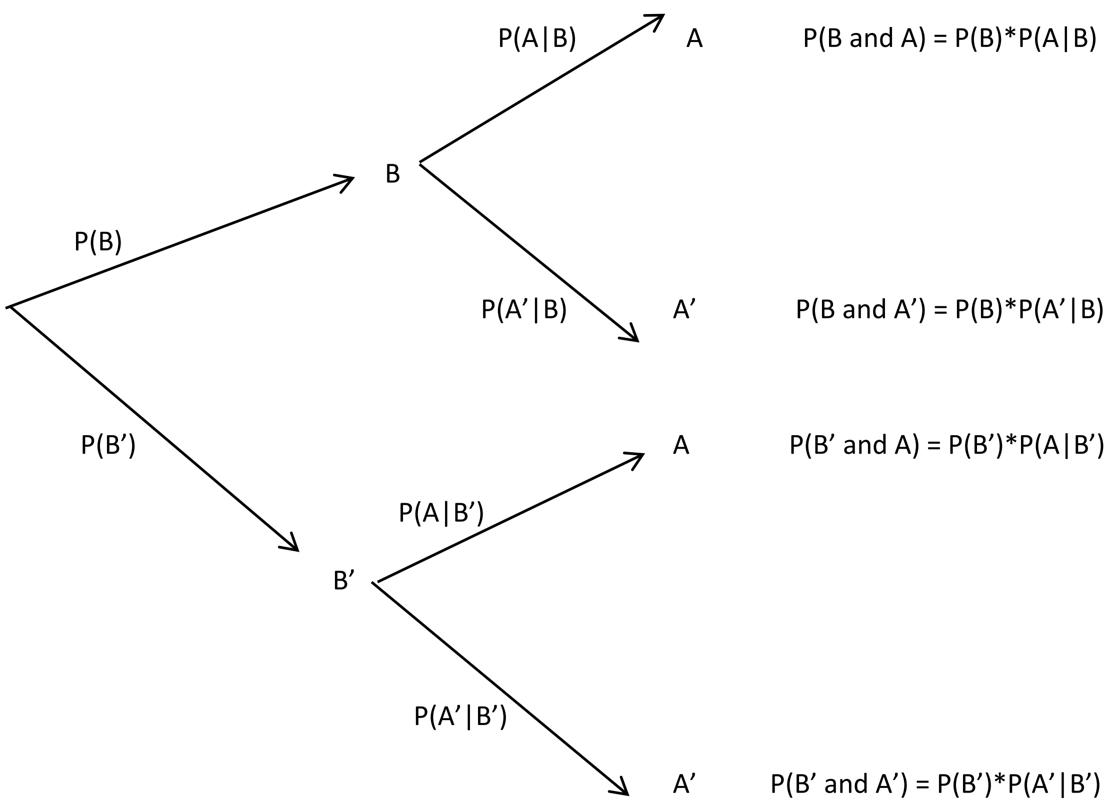
The required probability P(A) is given by:

$$P(A) = P(E_1)P(A|E_1) + P(E_2)P(A|E_2) + P(E_3)P(A|E_3)$$

$$= 3/10 * 0.01 + 1/4 * 0.012 + 9/20 * 0.02$$

$$= 0.015$$

Probability Tree:



Baye's Theorem:

In the two stage problem, if the happening of the event of second stage is given to us and on this basis we find the probability of the events of first stage, then the probability of an event of first stage is called the revised (or posterior) probability and is obtained using Bayes' theorem.

Consider our previous example - If there are two bags with bag 1 containing 2 green and 3 red balls and bag 2 containing 5 green and 6 red balls. Now, we are given the probability of picking a red ball and we are required to find the probability that the first bag was picked.

Derivation:

According to conditional probability:

$$P(E_i|A) = \frac{P(Ei \cap A)}{P(A)}$$

And now, the numerator in the above expression can be re-written as:

$P(E_i \cap A) = P(E_i)\, P(A|E_i)$, based on the multiplication law

Which gives us:

$$\boxed{P(E_i|A) = \frac{P(Ei)\ P(A\,|\,Ei)}{P(A)}}$$

Alternative form of the Baye's theorem:

$$P(E_i|A) = \frac{P(A|B) * P(B)}{P(A|B) * P(B) + P(A|not\ B) * P(not\ B)}$$

Example Problem: Consider the above machine-defect problem. If the drawn item is found defective, find the probability that it has been produced by machine Y.

$P(E_2|A) = P(E_2)\, P(A|E_2) / P(A)$

$= (1/4 * 0.012) / 0.015 = 0.003/0.015 = 1/5$

Random Variable:

A random variable is a real-valued function whose domain is a set of possible outcomes of a random experiment and range is a sub-set of the set of real numbers and has the following properties:

i) Each particular value of the random variable can be assigned some probability

ii) Uniting all the probabilities associated with all the different values of the random variable gives the value 1(unity).

Probability Distribution: The mathematical function that describes a random variable's values along with their associate probabilities is called a probability distribution.

Two types of variables - Discrete and Continuous.

Discrete Distributions	Continuous Distributions
Probability that X can take a specific value x is $P(X = x) = p(x)$.	Probability that X is between two points a and b is $P(a \leq X \leq b) = \int_a^b f(x)dx$.
It is non-negative for all real x.	It is non-negative for all real x.
The sum of $p(x)$ over all possible values of x is 1, i.e., $\sum p(x) = 1$.	$\int_{-\infty}^{\infty} f(x)dx = 1$
Probability Mass Function	**Probability Density Function**

Ex: Find the probability distribution of the number of heads when three fair coins are tossed simultaneously.

A: Let X be the number of heads in the toss of three fair coins.

As the random variable, "the number of heads" in a toss of three coins may be 0 or 1 or 2 or 3 associated with the sample space {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT},

∴ X can take the values 0, 1, 2, 3, with

P[X = 0] = P[TTT] = 1/8

P[X = 1] = P[HTT, THT, TTH] = 3/8

P[X = 2] = P[HHT, HTH, THH] = 3/8

P[X = 3] = P [HHH] = 1/8

Probability distribution of X, i.e. the number of heads when three coins:

X	0	1	2	3
P(x)	1/8	3/8	3/8	1/8

Cumulative Distribution Function:

A function F defined for all values of a random variable X by $F(x) = P[X \le x]$ is called the distribution function. It is also known as the cumulative distribution function (c.d.f.) of X since it is the cumulative probability of X up to and including the value x.

Expectation of a Random Variable:

Expected value of a r.v. X is given by E(X) =

$$\sum_{i=1}^{n} x_i p_i$$

for discrete variable and

$$\int_{-\infty}^{\infty} x f(x)\, dx$$

For continuous variable.

Ex: If it rains, a rain coat dealer can earn Rs 500 per day. If it is a dry day, he can lose Rs 100 per day. What is his expectation, if the probability of rain is 0.4?

Solution: Let X be the amount earned on a day by the dealer. Therefore, X can take the values Rs 500, and -Rs 100.

- Probability distribution of X is given by:

	Rainy Day	Dry Day
X	500	-100
P(x)	0.4	0.6

Hence, the expectation of the amount earned by him given by

$$\sum_{i=1}^{n} x_i p_i$$

is $x_i p_1 + x_2 p_2 =$ 500*0.4 + (-100)*0.6

= 200 - 60

= 140

He is expected to earn 140 on an average.

Ex1: A fair coin is tossed until a tail appears. What is the expectation of number of tosses?

Bernoulli Distribution:

A discrete random variable X is said to follow Bernoulli distribution with parameter p if its probability mass function is given by

$$P[X = x] = \begin{cases} p^x (1-p)^{1-x} & ; x = 0,1 \\ 0 & ; \text{elsewhere} \end{cases}$$

i.e. $P[X = 1] = p^1 (1-p)^{1-1} = p$ [putting $x = 1$]

and $P[X = 0] = p^0 (1-p)^{1-0} = 1-p$ [putting $x = 0$]

The Bernoulli probability distribution, in tabular form, is given as

X	0	1
$p(x)$	$1-p$	p

Mean = p
Variance = pq, where q = 1-p

Binomial Distribution:

Bernoulli distribution for n trials.

Definition:

A discrete random variable X is said to follow binomial distribution with parameters n and p if it assumes only a finite number of non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} {}^{n}C_{x} p^{x} q^{n-x}; & x = 0, 1, 2, ..., n \\ 0; & \text{elsewhere} \end{cases}$$

where, n is the number of independent trials,

x is the number of successes in n trials,

p is the probability of success in each trial, and

$q = 1 - p$ is the probability of failure in each trial.

It is represented as $X \sim B(n,p)$

Mean = np
Variance = npq

Example:

An unbiased coin is tossed six times. Find the probability of obtaining

(i) exactly 3 heads

(ii) less than 3 heads

(iii) more than 3 heads

(iv) at most 3 heads

(v) at least 3 heads

(vi) more than 6 heads