PROBABILITY

What is Probability ?

Probability is a branch of Mathematics that deals with calculating the likelihood of a given event to occur.

Simple Examples:

1. Roll a Dice
2. Toss a coin
3. Bag containing different coloured balls

Loss function depends on probability a lot.

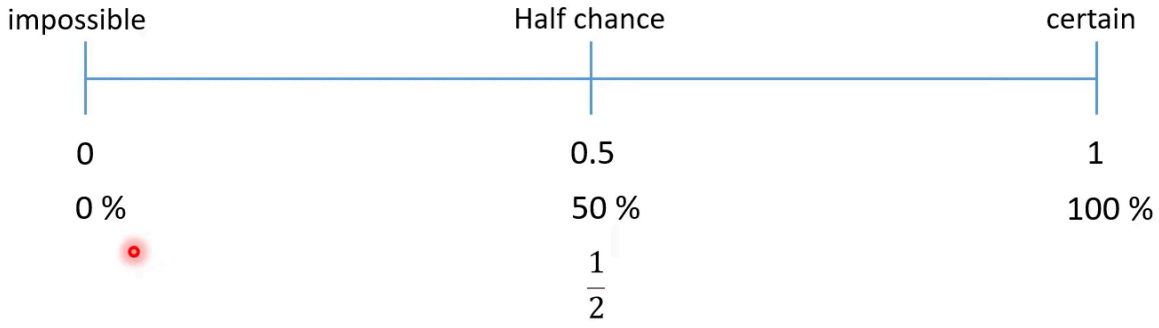
overall syllabus:

Topics covered in this module:

1. Basics of Probability
2. Random Variables
3. Probability Distributions
4. Maximum Likelihood
5. Bayes Theorem
6. Information Theory
7. Cross Entropy
8. Information Gain

The value of prob. lies between 0 and 1.

The Probability value lies between 0 and 1.



$$\text{Probability of an event to occur} = \frac{\text{Number of ways an event can occur}}{\text{Total number of outcomes}}$$

(H, T)

Possible Outcomes

$P(H) = \frac{1}{2}$

$P(T) = \frac{1}{2}$

(1, 2, 3, 4, 5, 6)

Possible Outcomes

$P(5) = \frac{1}{6}$

$P(5) = 0.16$

$P(even) = \frac{3}{6}$

$P(even) = 0.5$

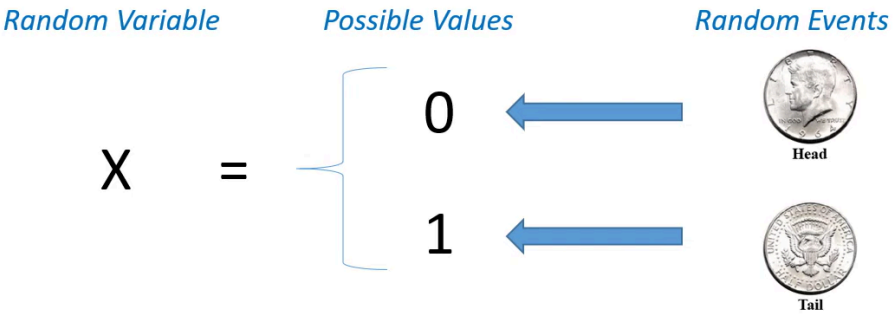
RANDOM VARIABLES

Random Variables

A Random Variable is a numerical description of the outcomes of Random events.

In other words, a random variable maps the outcomes of random events to numerical values.

Consider Tossing a Coin



Few Examples of Random Variables:

X = 0, if Heads; 1, if Tail

Y = Weight of a random person in a class

P (Weight of a random person in a class is less than 60 kg)

P (Y < 60)

Applications:

- Turnover of a company in a given time period.
- Price change of an asset over a given time period

Two types of random variables:

Random Variables

Discrete

A discrete random variable takes only discrete or distinct values.

Examples: Coin toss, Colour of the ball.

Continuous

A continuous random variable can take any value in a given range.

Examples: weight of a random person in a class.

PROBABILITY DISTRIBUTION FOR RANDOM VARIABLES

Probability Distributions

The **probability distribution** for a random variable describes how the probabilities are distributed over the values of the random variable.

Tossing 3 Coins

X = Sum of number of Heads when 3 coins are tossed

HHH = 3
TTT = 0
HHT = 2
HTH = 2
THH = 2
TTH = 1
HTT = 1
THT = 1

TTT , HHH , etc are random variables.

HHH	= 3	THH	= 2
TTT	= 0	TTH	= 1
HHT	= 2	HTT	= 1
HTH	= 2	THT	= 1

X (No. of Heads)	P(X = x)	P(X = x)
0	1/8	0.125
1	3/8	0.375
2	3/8	0.375
3	1/8	0.125

Normal Distribution and Skewness

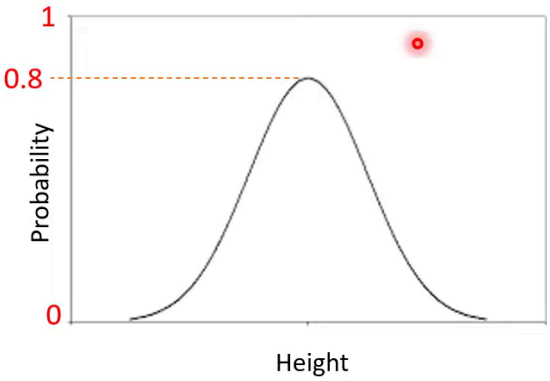
Data set = entire data

Data point = individual data

Normal Distribution

A **normal distribution** is an arrangement of a data set in which most of the data points lie in the middle of the range and the rest taper off symmetrically toward either extreme.

Normal Distribution is also known as **Gaussian Distribution**.

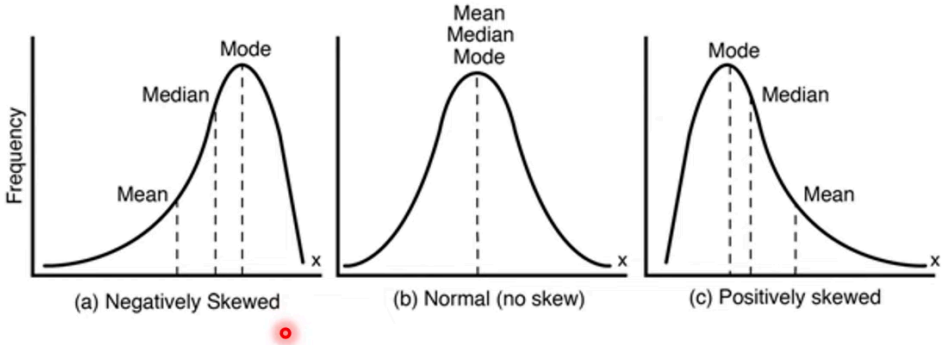


Bell Shaped Curve

The curve is symmetrical about the central vertical axis.

Skewness

A data is considered **skewed** when the distribution curve appears distorted or skewed either to the left or to the right, in a statistical distribution.



Example: Average income of people in different cities

POISSON DISTRIBUTION

Poisson Distribution

Poisson Distribution is a probability distribution that measures how many times an event is likely to occur within a specified period of time.

Poisson distribution is used to understand independent events that occur at a constant rate within a given interval of time.

Sic

Examples of Poisson Distribution

- Number of accidents occurring in a city from 6 pm to 10 pm
- Number of Patients arriving in an Emergency Room between 10 pm to 12 pm
- How many views does your blog gets in a day

Poisson Distribution

$$p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

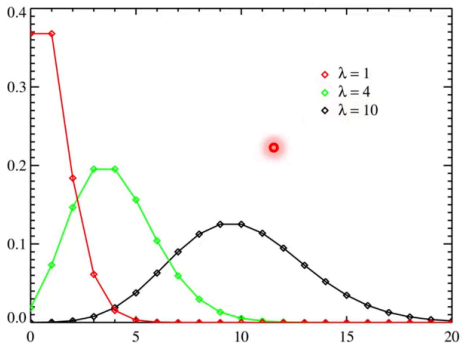
x --> Number of times the event occurs

p(x) --> Probability

λ --> Mean number of events

x! --> Factorial of x

e --> Euler's Number (2.71828)



Siddhardhan

X = number of times the event is happening

Y = the probability that the event will occur