Probability

What is Probability?

Probability is the measure of how likely an event will occur.

- Probability is the ratio of desired outcomes to total outcomes:

(desired outcomes) / (total outcomes)

- Probabilities of all outcomes always sums to 1

Example:

- On rolling a dice, you get 6 possible outcomes
- Each possibility only has one outcome, so each has a probability of 1/6
- For example, the probability of getting a number '2' on the dice is 1/6

Terms used in Probability

Random Experiment

An experiment or a process for which the outcome cannot be predicted with certainty

Sample Space

The entire possible set of outcomes of a random experiment is the sample space (S) of that experiment

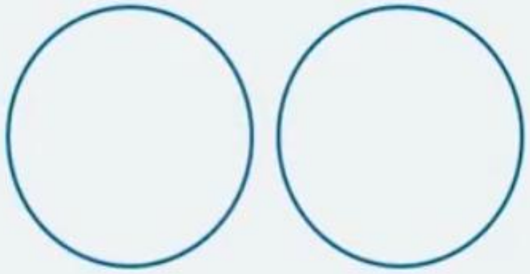
Event

One or more outcomes of an experiment. It is a subset of sample space(S)

Types of events

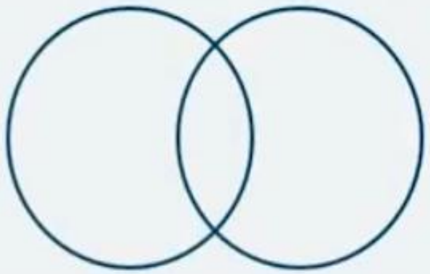
Disjoint Events *do not have any common outcomes.*

- The outcome of a ball delivered cannot be a sixer and a wicket
- A single card drawn from a deck cannot be a king and a queen
- A man cannot be dead and alive

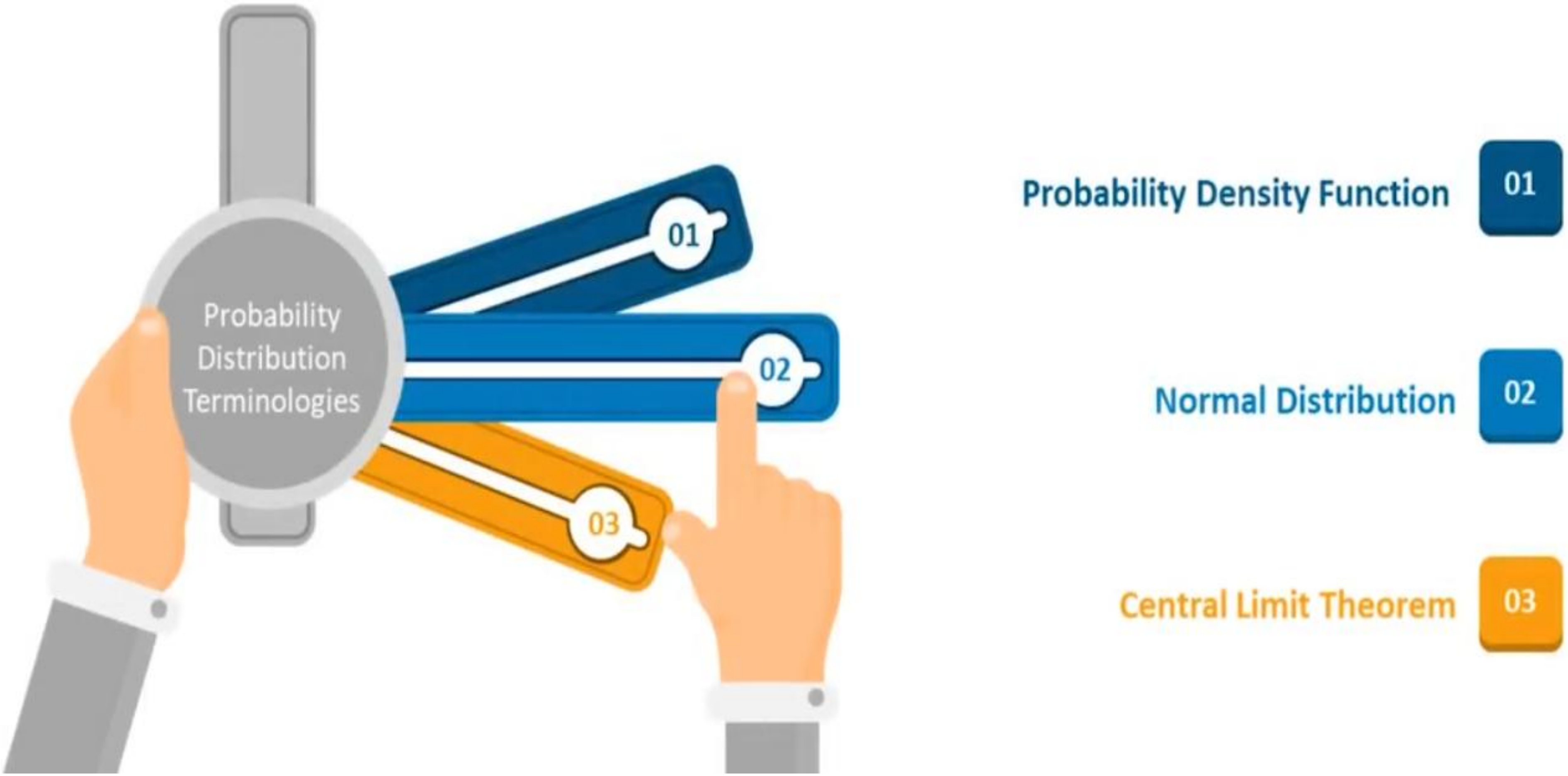


Non-Disjoint Events *can have common outcomes*

- A student can get 100 marks in statistics and 100 marks in probability
- The outcome of a ball delivered can be a no ball and a six

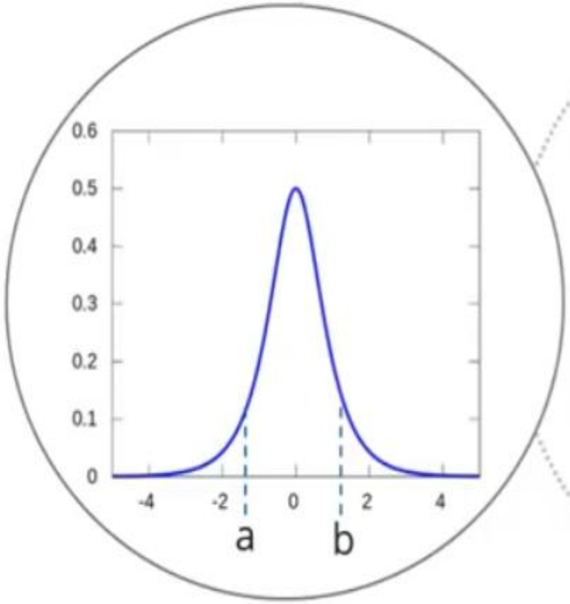


Probability Distribution



Probability Density Function

The equation describing a continuous probability distribution is called a Probability Density Function



Property 01

Graph of a PDF will be continuous over a range

Property 02

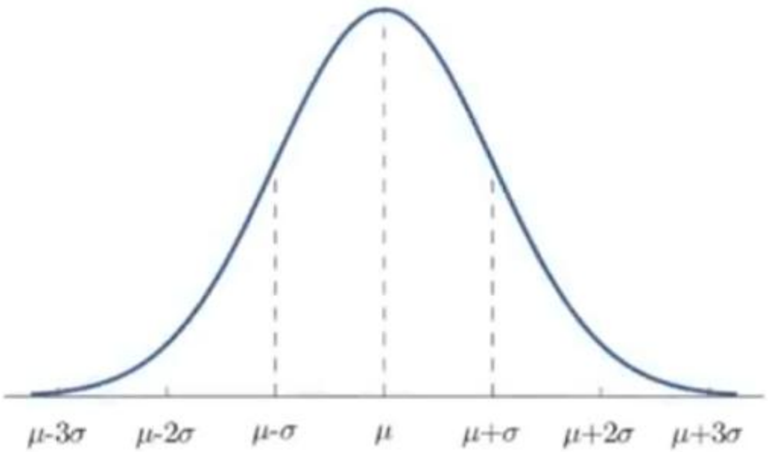
Area bounded by the curve of density function and the x-axis is equal to 1

Property 03

Probability that a random variable assumes a value between a & b is equal to the area under the PDF bounded by a & b

Normal Distribution

The Normal Distribution is a probability distribution that associates the normal random variable X with a cumulative probability



$$Y = [\ 1/\sigma * \mathrm{sqrt}(2\pi)\ ] * e^{-(x - \mu)2/2\sigma 2}$$

Where,

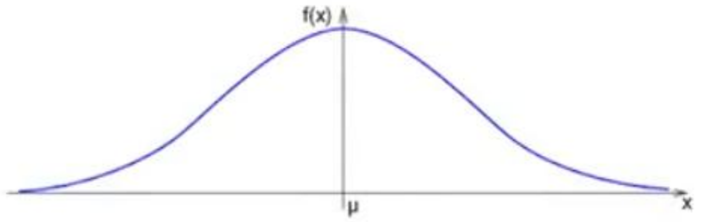
- *X* is a normal random variable
- μ is the mean and
- σ is the standard deviation

Note: Normal Random variable is variable with mean at 0 and variance equal to 1

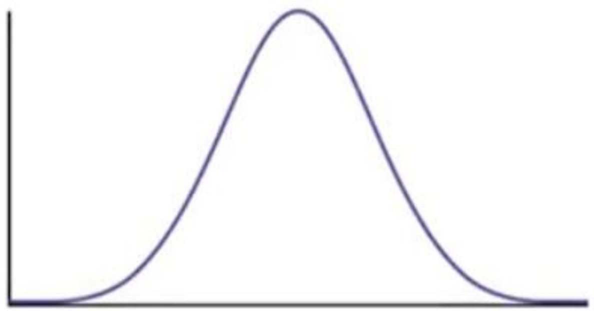
Standard Deviation and Curve

The graph of the Normal Distribution depends on two factors: the *Mean* and the *Standard Deviation*

- **Mean:** *Determines the location of center of the graph*
- **Standard Deviation:** *Determines the height of the graph*



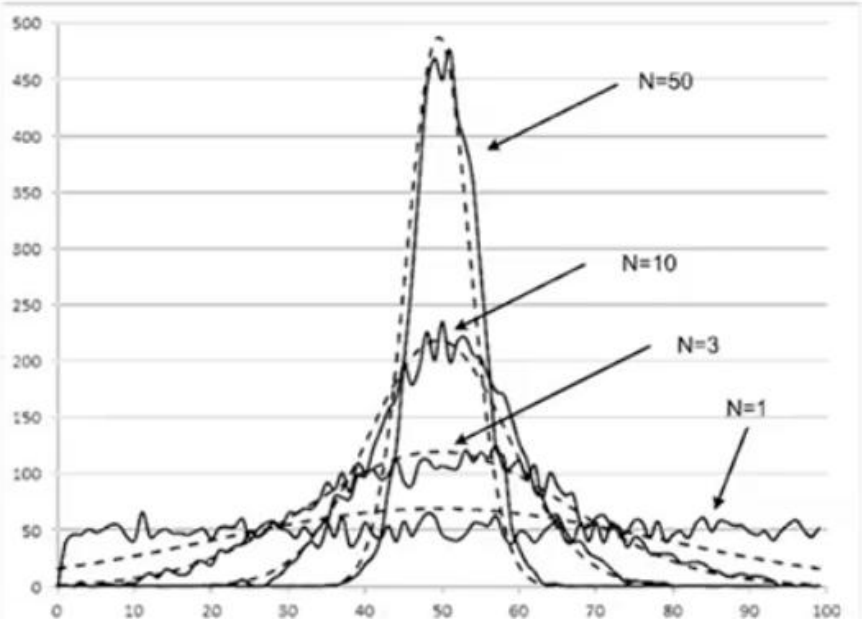
If the standard deviation is large, the curve is short and wide.



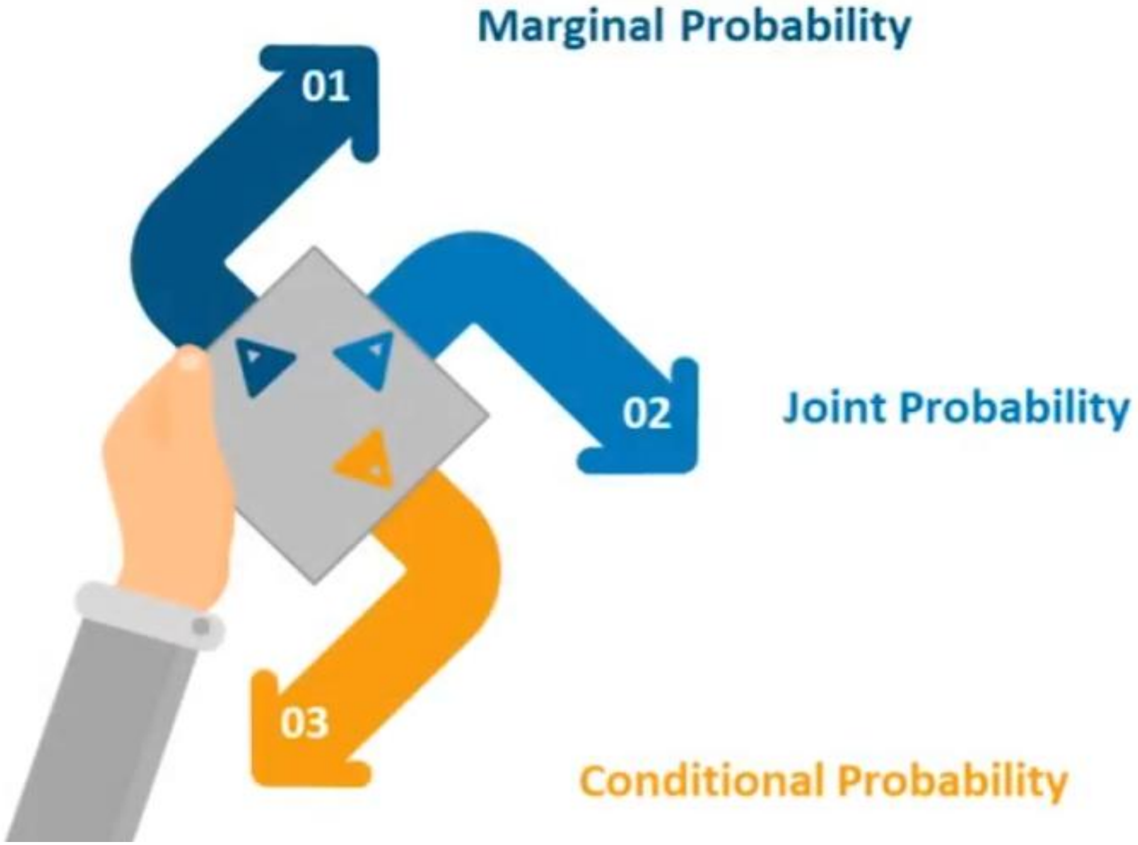
If the standard deviation is small, the curve is tall and narrow.

Central Limit Theorem

The ***Central Limit Theorem*** states that the sampling distribution of the mean of any independent, random variable will be normal or nearly normal, if the sample size is large enough

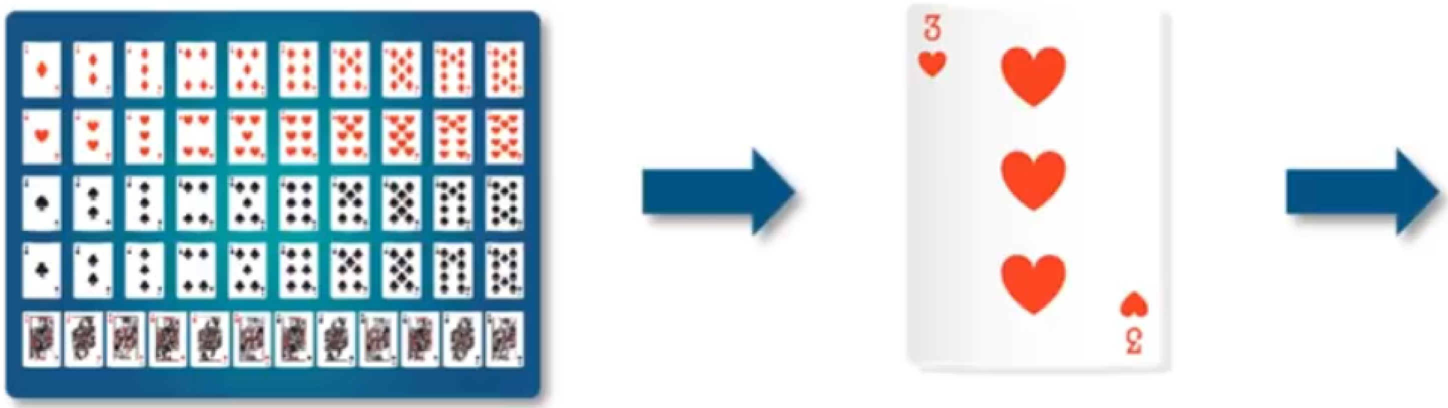


Types of Probability



Marginal Probability

Marginal Probability is the probability of occurrence of a single event.



$$\text{Marginal Probability} = \frac{13}{52}$$

It can be expressed as: P (A) = $\sum_{i=1}^{k} P(x_i)$

Joint Probability

Joint Probability is a measure of two events happening at the same time

Example: The probability that a card is an Ace of hearts = P (Ace of hearts)
(There are 13 heart cards in a deck of 52 and out of them one in the Ace of hearts)

Conditional Probability

- *Probability of an event or outcome based on the occurrence of a previous event or outcome*
- *Conditional Probability of an event B is the probability that the event will occur given that an event A has already occurred*

If A and B are dependent events then the expression for conditional probability is given by:
P (B|A) = P (A and B) / P (A)

If A and B are independent events then the expression for conditional probability is given by:
P(B|A) = P (B)

Example

EDUREKA'S TRAINING STUDY
TRAINING AND SALARY PACKAGE OF CANDIDATES Study examines salary package and training undergone by candidates
Sample: 60 candidates without training and 45 candidates with edureka's training
Task to do: Assessment of training with salary package

Results		Training		
		Without Edureka Training	With Edureka Training	Total
Salary Package obtained by participant	Very Poor Package	5	0	5
	Poor Package	10	0	10
	Average Package	40	10	50
	Good Package	5	30	35
	Excellent Package	0	5	5
	Total	60	45	105

Marginal Probability

Finding the probability that a candidate has undergone Edureka's training

Results		Training		
		Without Edureka Training	With Edureka Training	Total
Salary Package obtained by participant	Very Poor Package	5	0	5
	Poor Package	10	0	10
	Average Package	40	10	50
	Good Package	5	30	35
	Excellent Package	0	5	5
	Total	60	45	105

The probability that a candidate has undergone Edureka's training

P(Edu.Training) = 45 /105 ≈ 0.42

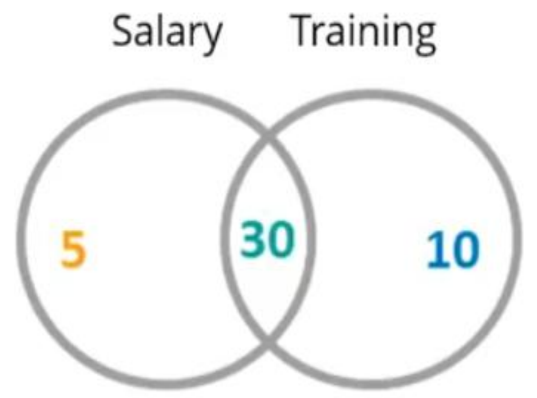
Results		Training		
		Without Edureka Training	With Edureka Training	Total
Salary Package obtained by participant	Very Poor Package	5	0	5
	Poor Package	10	0	10
	Average Package	40	10	50
	Good Package	5	30	35
	Excellent Package	0	5	5
	Total	60	45	105

Joint Probability

Finding the probability that a candidate has attended Edureka's training and also has good package.

Results		Training		
		Without Edureka Training	With Edureka Training	Total
Salary Package obtained by participant	Very Poor Package	5	0	5
	Poor Package	10	0	10
	Average Package	40	10	50
	Good Package	5	30	35
	Excellent Package	0	5	5
	Total	60	45	105

Results		Training		
		Without Edureka Training	With Edureka Training	Total
Salary Package obtained by participant	Very Poor Package	5	0	5
	Poor Package	10	0	10
	Average Package	40	10	50
	Good Package	5	30	35
	Excellent Package	0	5	5
	Total	60	45	105



P (Good Package & Edu.Training) =
30 /105 ≈ 0.28

Conditional Probability

Finding the probability that a candidate has a good package given that he has not undergone training

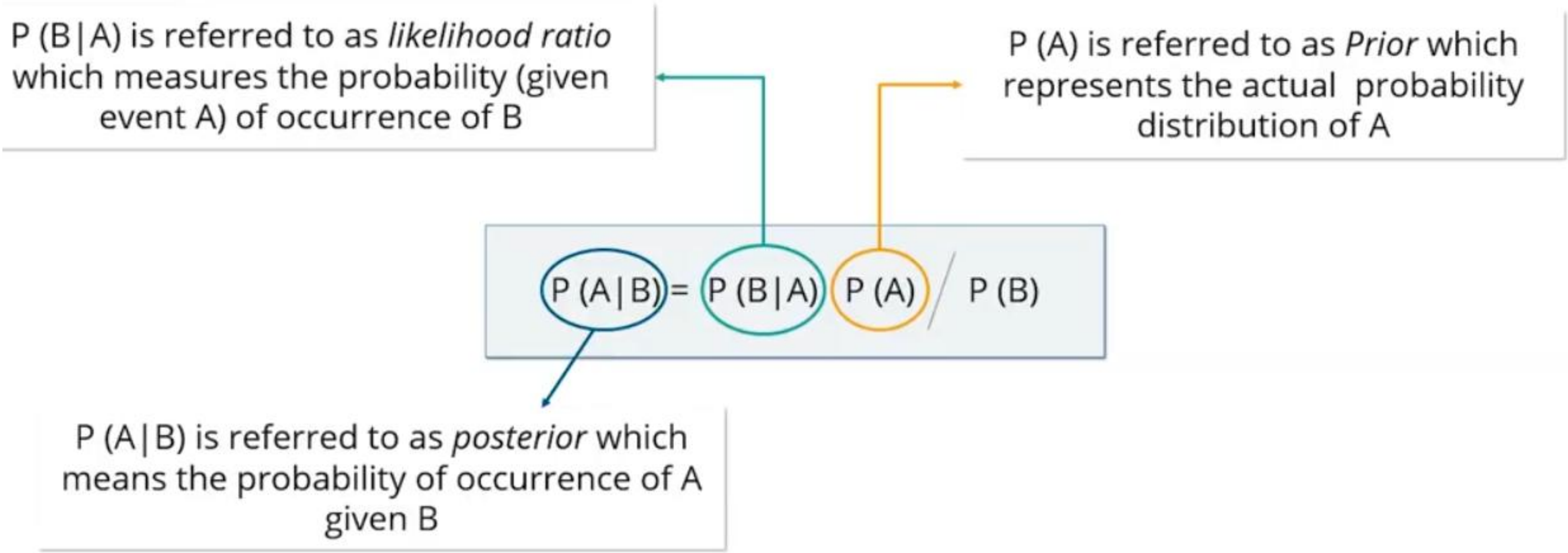
Results		Training		
		Without Edureka Training	With Edureka Training	Total
Salary Package obtained by participant	Very Poor Package	5	0	5
	Poor Package	10	0	10
	Average Package	40	10	50
	Good Package	5	30	35
	Excellent Package	0	5	5
	Total	60	45	105

Results		Training		
		Without Edureka Training	With Edureka Training	Total
Salary Package obtained by participant	Very Poor Package	5	0	5
	Poor Package	10	0	10
	Average Package	40	10	50
	Good Package	5	30	35
	Excellent Package	0	5	5
	Total	60	45	105

P (Good Package | Without Edureka) = 5 / 60 ≈ 0.08

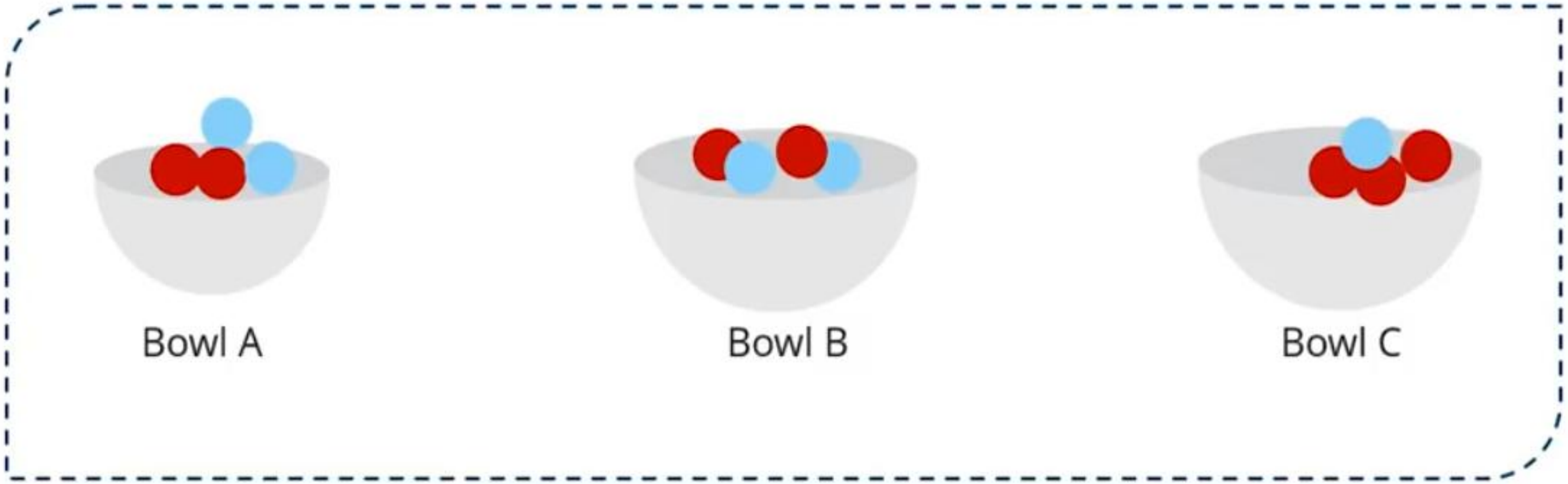
Bayes Theorem

Shows the relation between one conditional probability and its inverse



Bayes Theorem Example

" Consider 3 bowls. Bowl A contains 2 blue balls and 4 red balls; Bowl B contains 8 blue balls and 4 red balls, Bowl C contains 1 blue ball and 3 red balls. We draw 1 ball from each bowl. What is the probability to draw a blue ball from Bowl A if we know that we drew exactly a total of 2 blue balls? "



- Let *A* be the *event of picking a blue ball* from bag A, and let *X* be the *event of picking exactly two blue balls*
- We want Probability(A|X), i.e. probability of occurrence of event A given X

By the definition of *Conditional Probability*,

$$\Pr(A|X) = \frac{Pr(A \cap X)}{\Pr(X)}$$

- We need to find the two probabilities on the right-side of equal to symbolop

Steps to execute the problem

Step 1:

First find Pr(X). This can happen in three ways:

(i) white from A, white from B, red from C

(ii) white from A, red from B, white from C

(iii) red from A, white from B, white from C

Step 2:

Next we find Pr(A∩X).

This is the sum of terms (i) and (ii) above