Probability, Set, Distributions and Random Variable

AccioJob/ DataScience-Fundamentals

Content

- Probability
- Sets and Venn Diagrams
- Bayes Theorem and Conditional Probability
- Mutually Exclusive Events and Independent Events
- Random variables
- Probability Distributions - Bernoulli, Poisson, Binomial, Normal
- Permutations and Combinations

Probability

Probability Fundamentals

What is the Probability of an Event?

Probability is a measure of '**uncertainty**' or '**likelihood**' or '**chance**' that a particular event or outcome will occur. It represents the **ratio** of the '**favorable outcomes**' to the '**total possible outcomes**' in a given situation.

With the concept of probability, we can **quantify uncertainty** and **randomness** that we experience in our daily lives. **With any random phenomenon, the probability of a particular outcome is the proportion of times that the outcome would occur in a long run of observations.**

$$P(A) = \frac{n(A)}{n(S)}$$

Probability of Event A

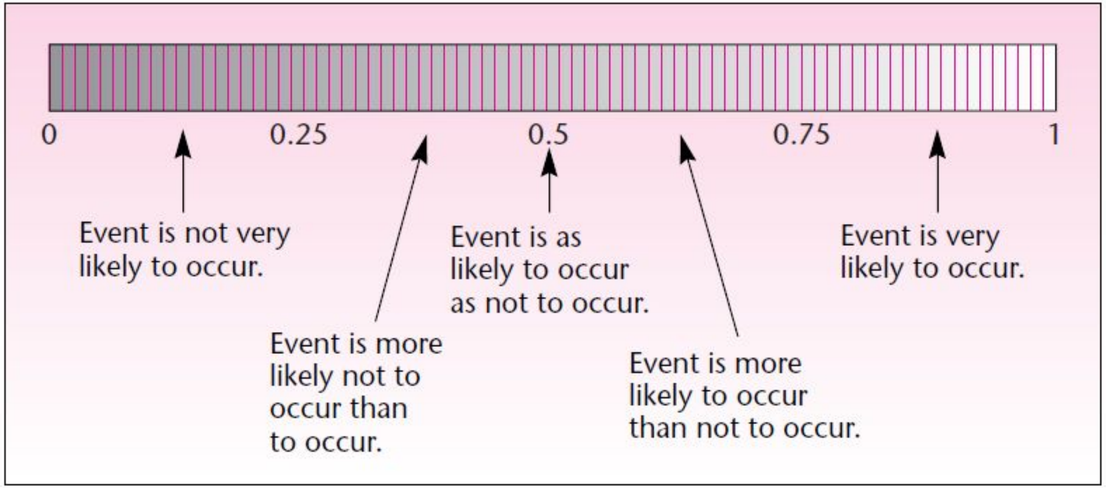
number of elements** in the set of the **Event A

total no. of elements** in the **Sample space S

Probability - Range of values

For any event A, the probability $P(A)$ satisfies

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



- Probability lies between 0 and 1
- Greater the Probability higher the confidence of occurence of event and vice versa
- Probability of 0.75 means 75% chance of occurence of an event

Probability - The Coin experiment - Single flip

What happens when we flip a fair Coin? What is the probability of Heads and Tails?

Probability of Heads

P(H)

$$\frac{\text{Favorable outcome (count of outcomes with Heads)}}{\text{Total possible outcomes}}$$

$$\frac{n(\text{Heads})}{n(\text{Total})}$$

1/2

Probability of Tails

P(T)

$$\frac{\text{Favorable outcome (count of outcomes with Tails)}}{\text{Total possible outcomes}}$$

$$\frac{n(\text{Tails})}{n(\text{Total})}$$

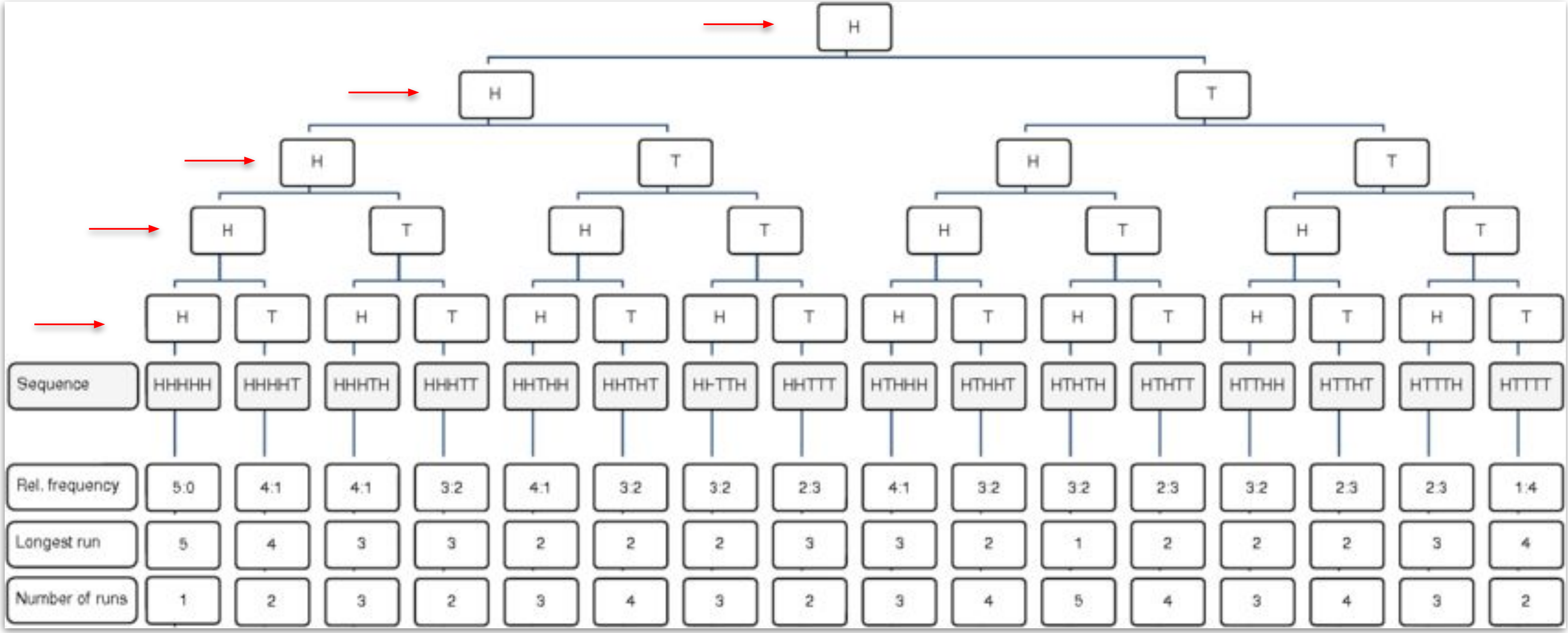
1/2

Outcome Matrix

All outcomes	Heads	Tails
Symbol of outcomes	H	T
Count of outcomes	1	1

Probability Tree of Coin experiment - Multiple flips

The probability tree illustrates the possible outcomes of a sequence of 5 coin tosses. For the sake of brevity only the 16 sequences starting with heads and their characteristics-relative frequency of heads and tails, longest run within the sequence, number of runs within the sequence, and alternation rate-are displayed.



Probability - The Dice Experiment

What happens when we roll 2 dice? How does probability work for different outcomes? What is the probability of rolling a 4 or a 7?

Total Outcomes (The Sample Space)

[1][1], [1][2], [1][3], [1][4], [1][5], [1][6],
[2][1], [2][2], [2][3], [2][4], [2][5], [2][6],
[3][1], [3][2], [3][3], [3][4], [3][5], [3][6],
[4][1], [4][2], [4][3], [4][4], [4][5], [4][6],
[5][1], [5][2], [5][3], [5][4], [5][5], [5][6],
[6][1], [6][2], [6][3], [6][4], [6][5], [6][6].

n(Total)
36

Favourable Outcomes (4 or 7)

[1][1], [1][2], **[1][3]**, [1][4], [1][5], **[1][6]**,
[2][1], **[2][2]**, [2][3], [2][4], **[2][5]**, [2][6],
[3][1], [3][2], [3][3], **[3][4]**, [3][5], [3][6],
[4][1], [4][2], **[4][3]**, [4][4], [4][5], [4][6],
[5][1], **[5][2]**, [5][3], [5][4], [5][5], [5][6],
[6][1], [6][2], [6][3], [6][4], [6][5], [6][6].

n(4 or 7)
9

$$\frac{n(4 \text{ or } 7)}{n(\text{Total})} = 9/36 = 0.25$$

The outcome Matrix for roll of 2 dice

	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2	3	4	5	6	7
⚁	3	4	5	6	7	8
⚂	4	5	6	7	8	9
⚃	5	6	7	8	9	10
⚄	6	7	8	9	10	11
⚅	7	8	9	10	11	12

Two (6-sided) dice roll outcome table

Outcome of Dice	Favourable Outcomes	Total Outcomes	Probability
2		36	
3		36	
4		36	
5		36	
6		36	
7		36	
8		36	
9		36	
10		36	
11		36	
12		36	

- **What is the probability of getting a 5 and 8 in a roll of 2 dice?**
- **What is the probability of getting a value > 10 in a roll of 2 dice?**
- **What is the probability of getting a value < 5 in a roll of 2 dice?**

Probability - The Cards Experiment

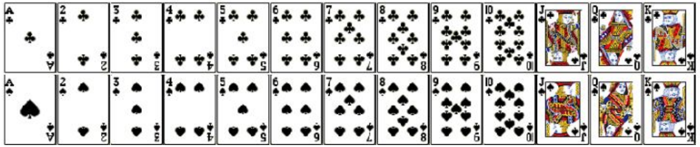
What happens when we draw a card(s) from a pack of playing cards ? How does probability work for different outcomes?

Understanding the deck of cards

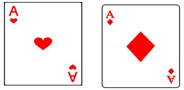
- The number of **Spades (Black)**, **Hearts (red)**, **Diamonds (red)**, and **Clubs (back)** is same in every pack of 52 cards. These are called **Suits.** Each suit has **13 cards.**
- There are 13 cards of each suit, consisting of 1 Ace, 3 face cards, and 9 number cards.
- There are 4 Aces, 12 face cards, and 36 number cards in a 52 card deck
- Probability of drawing any card will always lie between 0 and 1.

Calculating the probabilities from the deck of cards

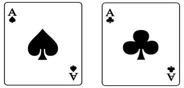
P(Black Card) = 26/52 = 1/2



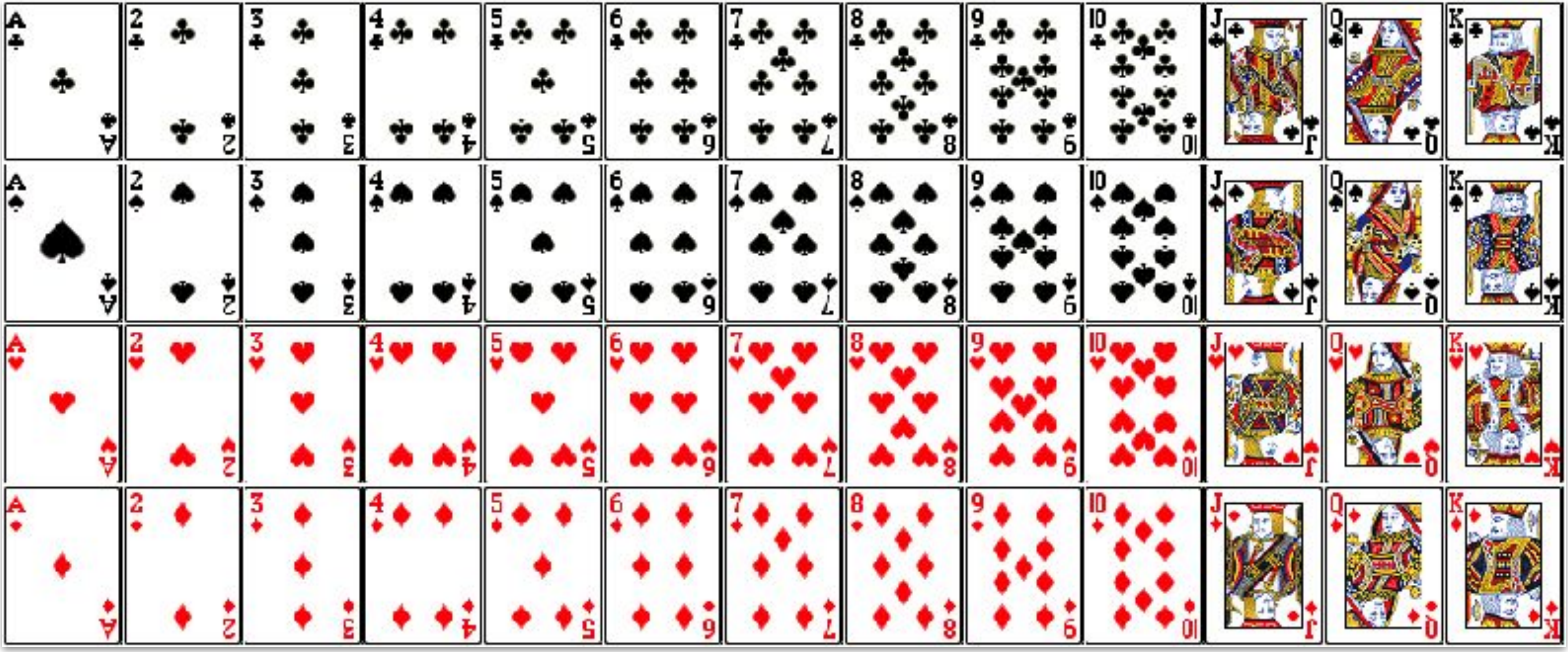
P(Ace) = 4/52



P(Black Ace) = 2/52



The Cards Experiment - All the cards



Simple and Compound Event

SIMPLE EVENT

A simple event is one that can only happen in one way - in other words, it has a single outcome.

Example of tossing a coin: we get one outcome that is a Head or a Tail.

Calculating simple event probability starts with two key points.

1) Know the desired outcome.

2) Know the total possible outcomes.

Example - Find the probabIlity of picking the months with the name starting from - 'J' from the bag having then names of all months.

Desired Outcomes - Jan, June and July - 3
Total Outcomes - 12

Probability = 3/12 = 0.25

Simple and Compound Event

COMPOUND EVENT

A compound event is more complex than a simple event, as it involves the probability of more than one outcome. It is as a combination of two or more simple events.

Example: The probability of finding an even number less than 5.

We have a combination of two simple events: finding an even number, and finding a number that is less than 5.

Example - Find the probability of Flipping Heads and Rolling number Greater Than 4.

Possible outcomes of coin **{H,T}**
Possible outcomes of dice **{1,2,3,4,5,6}**

Total list of outcomes of the experiment
{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6} = 12

Success outcomes
{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6} = 2

Thus the probability of the successful outcomes = **2/12**

Sets and Venn Diagrams

Basic Definitions: Events, Sample Space, and Probabilities

What are Sets?

- A set is defined as a **well-defined collection of objects.** It is represented by a capital letter.
- A set can have any group of items, be it a collection of numbers, days of a week, types of vehicles, and so on.
- Every item in set is called an **'Element'** of the set.
- **Curly brackets** are used while writing a set.
- An element that is contained in a set is represented by the symbol '∈'. In the example, 2 ∈ A. If an element is not a member of a set, then it is denoted using the symbol '∉'. For example, 3 ∉ A.

$$\mathbf{A} = \{ 2,\ 6,\ 10,\ 100,\ 12000 \}$$

Set Name = A

Elements of the Set = 2,6,10,100 and 12,000

2 ∈ A, 6 ∈ A, 10 ∈ A, 100 ∈ A, 12,000 ∈ A

Symbols used in Sets

Symbol	Name	Meaning	Example
{ }	Set	A collection of elements	A = {1, 7, 9, 13, 15, 23}, B = {9, 13, 21, 23}
A ∪ B	Union	Elements that belong to set A or set B	A ∪ B = {1, 7, 9, 13, 15, 21, 23}
A ∩ B	Intersection	Elements that belong to both the sets, A and B	A ∩ B = {9, 13, 23 }
A ⊆ B	Subset	subset has few or all elements equal to the set	{7, 15} ⊆ {7, 13, 15, 21} {7, 15} ⊆ {7, 15}
A ⊄ B	Not subset	left set is not a subset of right set	{1, 20} ⊄ B
A ⊇ B	Superset	Set A has more elements or equal to the set B	{1, 7, 9, 13, 15, 23} ⊇ {7, 13, 15, 23}
x ∉ A	Not element of	No set membership	A = {1, 7, 9, 13, 15, 23}, 5 ∉ A
Ø	Empty set	Ø = { }	C = {Ø}

Symbols used in Sets

Symbol	Name	Meaning	Example
a ∈ B	Element of	Set membership	B = {7, 13, 15, 21}, 13 ∈ B
x **∉ A**	Not element of	no set membership	A = {1, 7, 9, 13, 15, 23}, 5 ∉ A
A = B	Equality	Both sets have the same members. Order need not be same.	{7, 13,15} = {13, 7, 15}
Ā or **A’** or $\mathbf{A^c}$	Complement	All the objects that do not belong to set A	If U = {1, 2, 7, 9, 13, 15, 21, 23, 28, 30}; A = {1, 7, 9, 13, 15, 23} $\mathbf{A^c}$ = {2, 21, 28, 30}

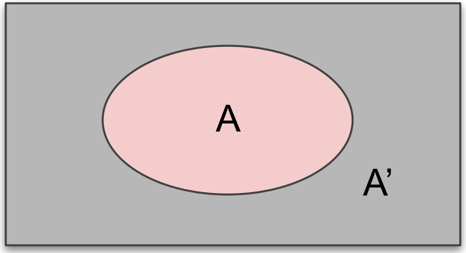
Types of Sets

Set Type	Definition	Example
Singleton Sets	A set that has only one element is called a singleton set or also called a unit set.	Set A = { k | k is an integer between 7 and 9} which is A = {8}.
Finite Sets	A set with a finite or countable number of elements is called a finite set.	Set B = {k | k is a prime number less than 15}, which is B = {2,3,5,7,11,13}
Infinite Sets	A set with an infinite number of elements is called an infinite set.	Set C = {Multiples of 10}
Empty or Null Sets	A set that does not has any element in it is called an Empty set or a Null set. An empty set is denoted using the symbol **'∅'**.	Set X = { }.
Disjoint Sets	Two Sets are said to be disjoint when they do not have any common element.	A = {1,2,3,4} B = {5,6,7,8}. Here, set A and set B are disjoint sets.

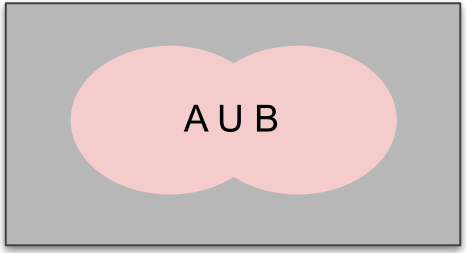
Venn Diagram

Venn Diagram

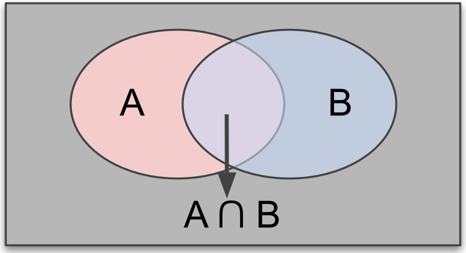
- Schematic drawing of sets that demonstrates the relationships between different sets
- Sets are shown as circles, or other closed figures, within a rectangle corresponding to the universal set, U.
- The elements are either shown inside the circles (if they belong to set) or outside if they di not belong to set.



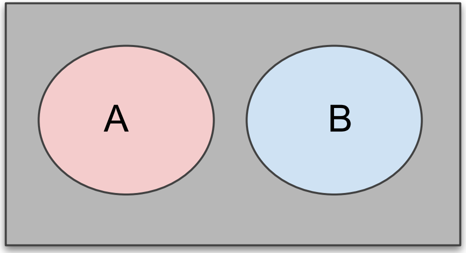
Set A and its Complement.



AUB is A Union B.



A∩B - A Intersection B.



A and B are Disjoint sets.

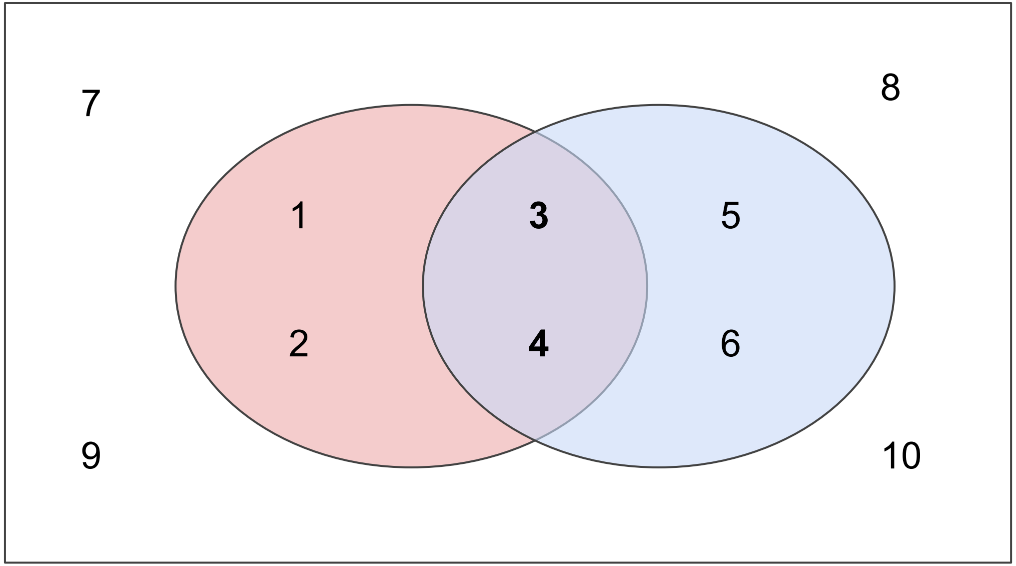
Union, Intersection and Complement - An Example

Example -

U = {1,2,3,4,5,6,7,8,9,10}

Set A = {1,2,3,4}

Set B = {3,4,5,6}



Union, Intersection and Complement

- **A U B** = {1,2,3,4,5,6}
- **A ∩ B** = {3,4}
- **A'** = {7,8,9,10}

Formula for Union/ Intersection for number of elements

Union

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

Intersection

$$n(A \cap B) = n(A) + n(B) - n(A \cup B)$$

Formula for Union/ Intersection for number of elements

For two Disjoint sets

Union

$$n(A \cup B) = n(A) + n(B)$$

Intersection

$$n(A \cap B) = \phi$$

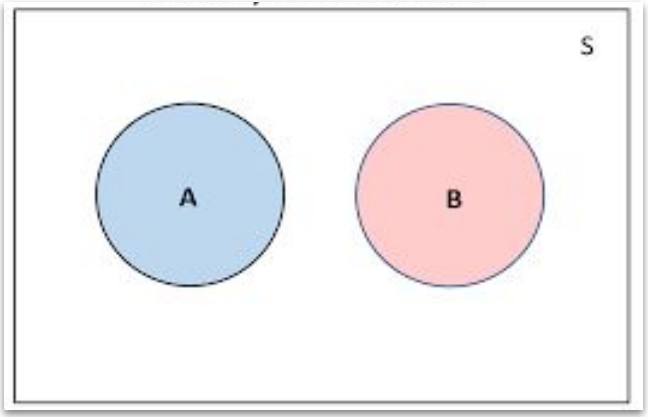
Difference

$$n(A - B) = n(A)$$

Mutually Exclusive and Disjoint Events

Mutually Exclusive Events

- Mutually exclusive events are the **events that cannot occur or happen at the same time.**
- Thus, the probability of the events happening at the same time is zero.
- They are also called **Disjoint Events.**
- The sum of the probability of Mutually Exclusive events can never be greater than 1. It can be less than 1 (in case the 2 events are not exhaustive) or equal to 1 (if the events are exhaustive).
- In this case, the sum of their probability is exactly 1.



The formula for the probability of an event A occurring or the probability of event B occurring is given as P(A) + P(B). Thus -

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

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

Example of Mutually Exclusive Events

Tossing a coin has 2 outcomes - Heads and Tails. They both cannot occur together thus the event of 'Head' outcome and event of 'Tail' outcome are Mutually Exclusive. Hence the probability of H and T occurring together is 0.

In a six-sided die, the events '1', '2', '3', '4', '5' and '6' are mutually exclusive events. We cannot get any of these events together at the same time when we threw one die.

In a deck of 52 cards, drawing a Red card and Jack are NOT Mutually Exclusive Events - because we can draw 2 cards which can be RED and JACK.

Mutually Exclusive and Exhaustive

Mutually Exclusive AND Exhaustive

The two events are said to be Mutually Exclusive and Exhaustive when the sum of their probabilities = 1.

Eg - The event of Head or Tail when you toss a coin are Mutually Exclusive and Exhaustive. Thus -

P (H) + P (T) = 0.5 + 0.5 = 1

Mutually Exclusive AND NOT Exhaustive

The two events are said to be Mutually Exclusive and Exhaustive when the sum of their probabilities < `1.

Eg - In a six-sided die, the events '2', and '6' are M.E. events but NOT EXHAUSTIVE as there are other outcome also that can come up.

P (2 and 6) = 0.33 < 1

Bayes Theorem

Bayes Theorem

- It was named after the 18th-century British mathematician Thomas Bayes
- It is the Theorem that describes the **Conditional Probability.**
- Conditional probability means - **probability of occurrence of an event related to any condition.**
- It is defined as the **likelihood that an event will occur, based on the occurrence of a previous outcome.**

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Where -

P(A) = probability of happening of Event A

P(B) = probability of happening of Event B

P(A|B) = probability of event A given B

P(B|A) = probability of event B given A

Example for Bayes Theorem

What are the chances of Rain during the day knowing it is cloudy?

P(R) = Probability of Rain = 10%

P(C) = Probability of Cloud = 40%

P(C|R) = Probability of Cloud knowing it rained = 70%.

Find P(R|C) = Probability of Rain knowing it is cloudy?

$$P(\text{🌧}|\text{🌥}) = P(\text{🌥}|\text{🌧}) \times \frac{P(\text{🌧})}{P(\text{🌥})}$$

=((0.7)*(0.1))/0.4 =0.175

Confusion Matrix for Bayes Theorem

What is the probability that a person selected is a Girl knowing the person plays cricket?

P(A) = probability of selecting a Girl =(12+40)/(12+40+35+10) = **0.536**

P(B) = probability of playing cricket =(35+12)/(35+10+12+40) = **0.485**

P(B|A) = probability of cricket knowing girl **= 12/52 = 0.231**

P(A|B) = probability of girl knowing the person plays cricket **=** (0.231*0.536)/0.485 or **=** 12/(35+12)

	Cricket	Other Games
Boy	35	10
Girls	12	40

Independent Events

Independent Events

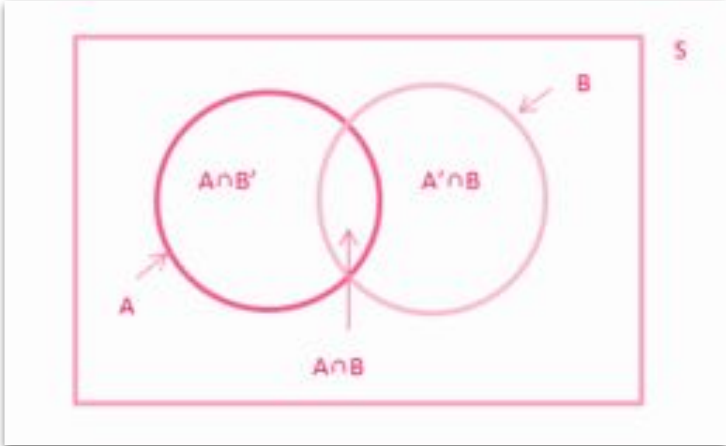
When the occurence of one event does not control the happening of the other event then it is termed as an independent event.

There is **no influence of an occurrence with another** and they are independent of each other.

If the probability of occurrence of an event A is not affected by the occurrence of another event B, then A and B are said to be independent events.

For an Independent Event

$$P(A \cap B) = P(A).\ P(B)$$



Independent Events

When rolling a dice -

Event A - Rolling an Odd number **{1,3,5}** and **Event B -** Rolling multiple of 3 **{3,6}.** Thus -

P(A)= 3/6 = 1/2 and **P(B)** = 2/6 = ⅓ and **P(A ∩ B)** =1/6 i.e. **{3}**

P (A|B) is the probability of Event A knowing B has already occurred.

P(A|B) = P(A ∩ B)/ P(B) = (1/6)/(1/3) = 0.5

Since P(A) and P(A|B) are equal to 0.5 - means that the occurrence of Event B does not impacted the probability of occurrence of event A.

That means both the events are independent.

Thus - If A and B are independent events, then P(A|B) = P(A)

Independent Events

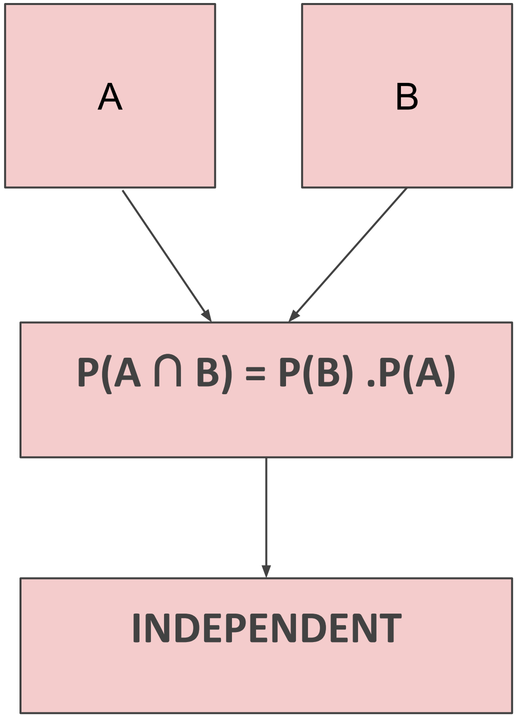
Using multiplication Rule -

$$P(A \cap B) = P(B) \,.P(A \mid B)$$

$$P(A \cap B) = P(B) \,.P(A)$$

Thus - A and B are two events associated with the same random experiment, then A and B are known as independent events if -

$$P(A \cap B) = P(B) \,.P(A)$$



Independent Events

Difference between Mutually exclusive and independent events	
Mutually exclusive events	Independent events
When the occurrence is not simultaneous for two events then they are termed as Mutually exclusive events.	When the occurence of one event does not control the happening of the other event then it is termed as an independent event.
The non-occurrence of an event will end up in the occurrence of an event.	There is no influence of an occurrence with another and they are independent of each other.
The mathematical formula for mutually exclusive events can be represented as P(X and Y) = 0	The mathematical formula for independent events can be defined as P(X and Y) = P(X) P(Y)
The sets will not overlap in the case of mutually exclusive events.	The sets will overlap in the case of independent events.

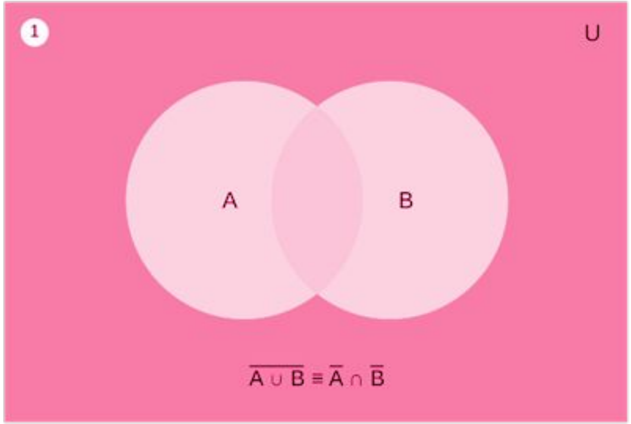
De Morgans’s law

First law states that - The complement of the union of two sets is the intersection of their complements

LHS Complement of Union of two sets

RHS Intersection of their Complements

$$(A \cup B)' = (A)' \cap (B)'$$



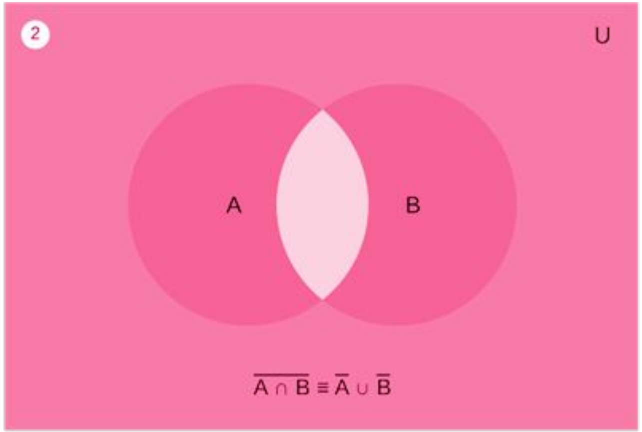
De Morgans's law

Second law states that - the complement of the intersection of two sets is the union of their complements

LHS Complement of intersection of two sets

RHS Union of their Complements

$$(A \cap B)' = (A)' \cup (B)'$$



Example of De Morgans's law

If U = {1, 2, 3, 4, 5, 6, 7, 8,9,10}, A ={3, 4, 5,6}, B = {4, 5, 6,7}.

Then show that

a) (AUB)’ = A’∩B’
b) (A∩B)’ = A’U B’.

(AUB)’ = A’ ∩ B’

LHS

(A U B) - {3,4,5,6,7}
(A U B)' - {1,2,,8,9,10}

RHS

(A)' - {1,2,7,8,9,10}
(B)' - {1,2,3,8,9,10}
(A)' ∩ (B)' - {1,2,8,9,10}

(A∩B)’ = A’ U B’

LHS

(A ∩ B) - {4,5,6}
(A ∩ B)' - {1,2,3,7,8,9,10}

RHS

(A)' - {1,2,7,8,9,10}
(B)' - {1,2,3,8,9,10}
(A)' U (B)' - {1,2,3,78,9,10}

Terminology related to Probability

Term	Definition
Experiment	An activity whose outcomes are not known is an experiment. Every experiment has a few favorable outcomes and a few unfavorable outcomes.
Random Experiment	A random experiment is an experiment for which the set of possible outcomes is known, but which specific outcome will occur on a particular execution of the experiment cannot be said prior to performing the experiment. Eg - Tossing a coin, rolling a die, and drawing a card from a deck are all examples of random experiments.
Trial	The numerous attempts in the process of an experiment are called trials. In other words, any particular performance of a random experiment is called a trial. For example, tossing a fair coin 2 times are 2 trials.
Event	An event is an outcome or a set of outcomes that we are interested in. It can be a single outcome or a combination of outcomes.
Random Event	An event that cannot be easily predicted is a random event. For such events, the probability value is very less.The formation of a rainbow during the rain is a random event.

Terminology related to Probability

Term	Definition
Sample Space	The sample space represents the set of all possible outcomes in an experiment or a random process. It is denoted by the symbol "Ω" or sometimes "S".
Outcome	An outcome is a specific result that can occur in an experiment or a random process. It is a member of the sample space.
Possible Outcome	The list of all the outcomes in an experiment can be referred to as possible outcomes. In tossing a coin, the possible outcomes are heads or tails.
Equally likely Outcomes	An experiment in which each of the outcomes has an equal probability, such outcomes are referred to as equally likely outcomes. In the process of rolling a six-faced dice, the probability of getting any number is equal to 1/6.
Probability	Probability is a numerical measure that quantifies the likelihood of an event occurring. It is a value between 0 and 1, where 0 indicates impossibility and 1 indicates certainty.

Terminology related to Probability

Term	Definition
Probability Distribution	A probability distribution describes the probabilities of all possible values that a random variable can take. It can be represented by a table, a graph, or a mathematical function.
Independent Events	Independent events are events where the occurrence of one event does not affect the probability of the other event occurring. The probability of two independent events happening together is the product of their individual probabilities.
Dependent Events	Dependent events are events where the occurrence of one event affects the probability of the other event occurring. The probability of two dependent events happening together is the product of the probability of the first event and the conditional probability of the second event given the first event.
Complementary Events	The complementary event of an event A is the event that A does not occur. The probability of the complementary event is equal to 1 minus the probability of the original event. For an event with probability $P(A)$, its compliment is $P(\bar{A})$ i.e. $P(A) + P(\bar{A}) = 1$.

Terminology related to Probability

Term	Definition
Mutually Exclusive Events	Two events such that the happening of one event prevents the happening of another event are referred to as mutually exclusive events. In other words, two events are said to be mutually exclusive events, if they cannot occur at the same time. For example, tossing a coin can result in either heads or tails. Both cannot be seen at the same time.
Conditional Probability	Conditional probability is the probability of an event occurring given that another event has already occurred. It is denoted as P(A|B), where A and B are events. Thus P(A|B) means Probability of A knowing B has already occurred. This is alao known as Bayes Theorem.
Expected Value	The expected value of a random variable is the long-term average value that would be obtained if the random experiment or process were repeated many times. It is also known as the mean or the average.

Random Variable and Distributions

Random Variable

Concept

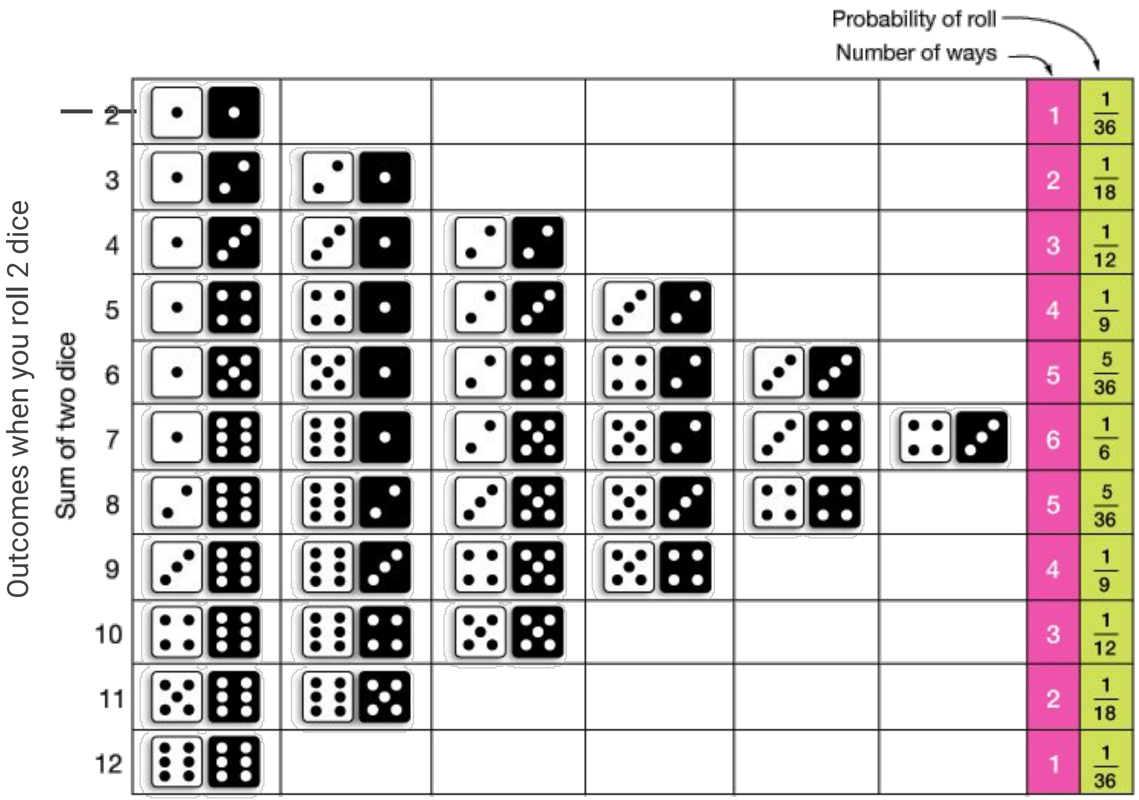
Random Variable is

- Whose value **cannot be determined till the last** moment (or till the time it occurs) thus it can take up may values
- Whose value is **probabilistic in nature** i.e. its value is determined in probability
- Function that maps the **outcomes of a random experiment** to numeric quantities, typically real numbers thus with random variables you quantify outcomes

Random Variable has a probability law - a rule that assigns probabilities to different values of random variable. This probability law is called **Probability Distribution** of the Random Variable. We usually donate Random Variable with capital letter let's say X. Then Probability Distribution of X is donate by P(X).

Random Variable - Example of rolling of 2 dice

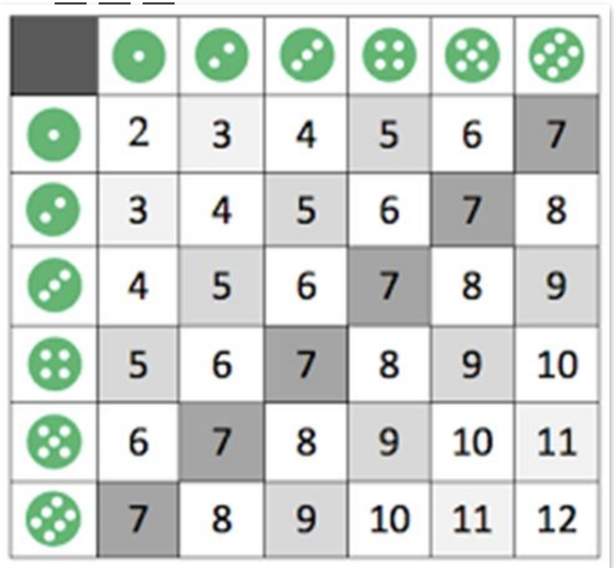
Example



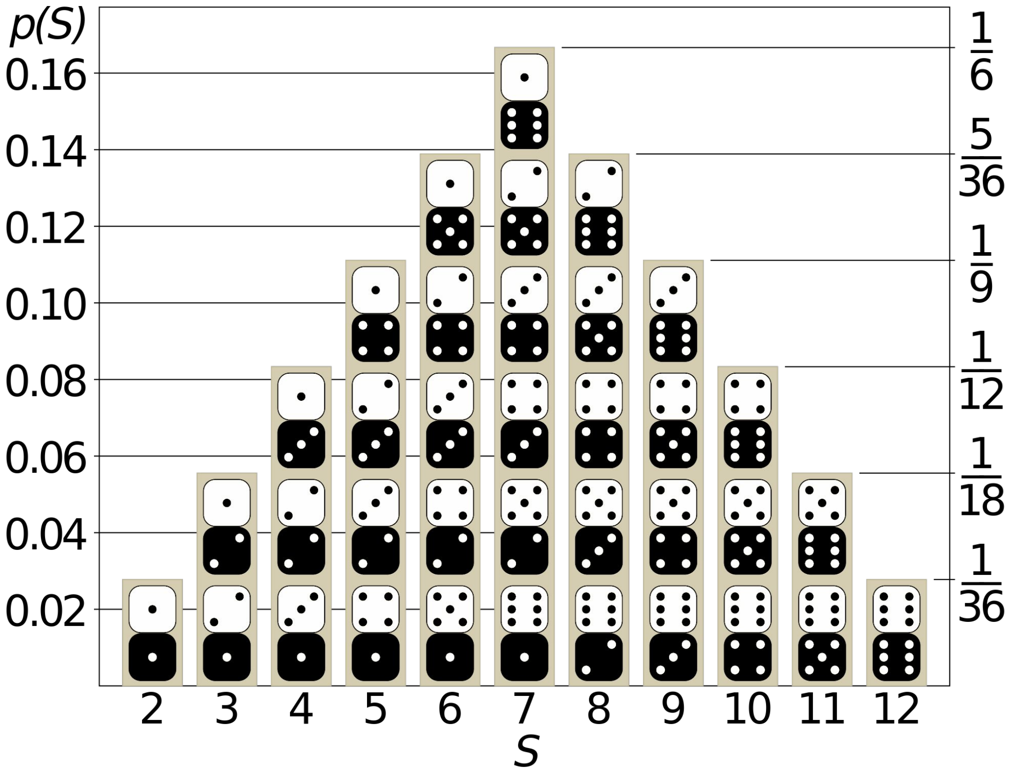
	2	3	4	5	6	7
	3	4	5	6	7	8
	4	5	6	7	8	9
	5	6	7	8	9	10
	6	7	8	9	10	11
	7	8	9	10	11	12

Histogram of rolling dice

Concept



	2	3	4	5	6	7
	3	4	5	6	7	8
	4	5	6	7	8	9
	5	6	7	8	9	10
	6	7	8	9	10	11
	7	8	9	10	11	12



Random Variable - Example of flipping 4 coins

Example

Probability Distribution of Heads and Tails when flipping a coin **4 times**.

S =

H,H,H,H	H,T,H,H	T,H,H,H	T,T,H,H
H,H,H,T	H,T,H,T	T,H,H,T	T,T,H,T
H,H,T,H	H,T,T,H	T,H,T,H	T,T,T,H
H,H,T,T	H,T,T,T	T,H,T,T	T,T,T,T

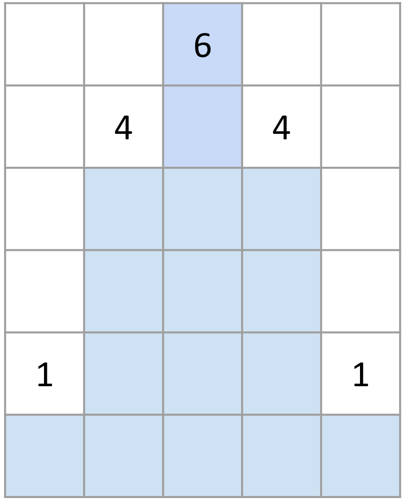
Sample Space of flipping a coin 4 times

Let **X** be the number of Heads in the event (also called **Number of Successes**).

X (Outcome)	**Options**	**Probability [P(X)]**
0	TTTT	**1/16 = 0.0625**
1	TTTH, TTHT, THTT, HTTT	**4/16 = 0.2500**
2	HHTT, TTHH, HTHT, THTH, HTTH, THHT	**6/16 = 0.3750**
3	HHHT, HHTH, HTHH, THHH	**4/16 = 0.2500**
4	HHHH	**1/16 = 0.0625**

Sum = 1

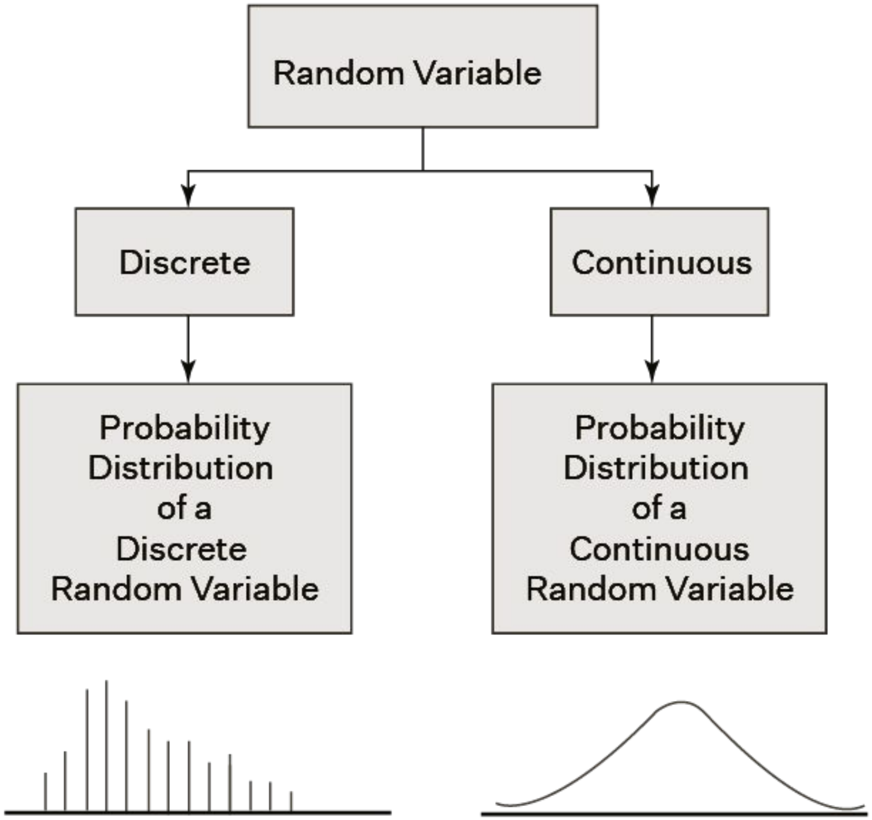
Histogram showing the Heads



Number of Heads

Random Variable - Types

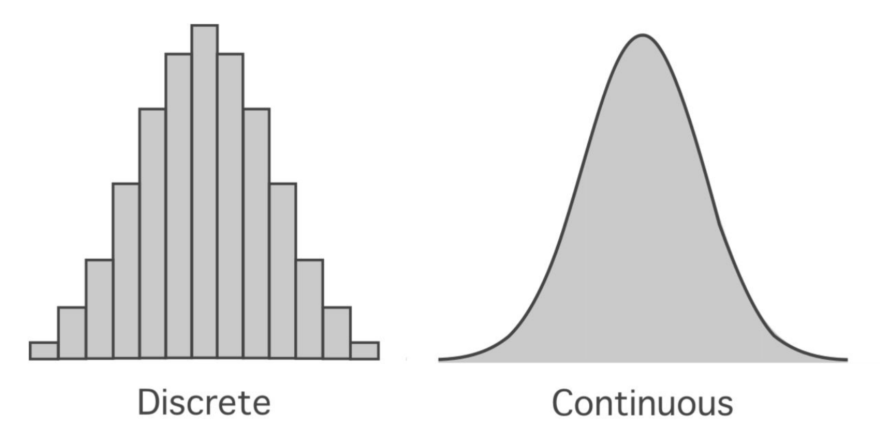
Concept



Random Variable - Probability Distribution

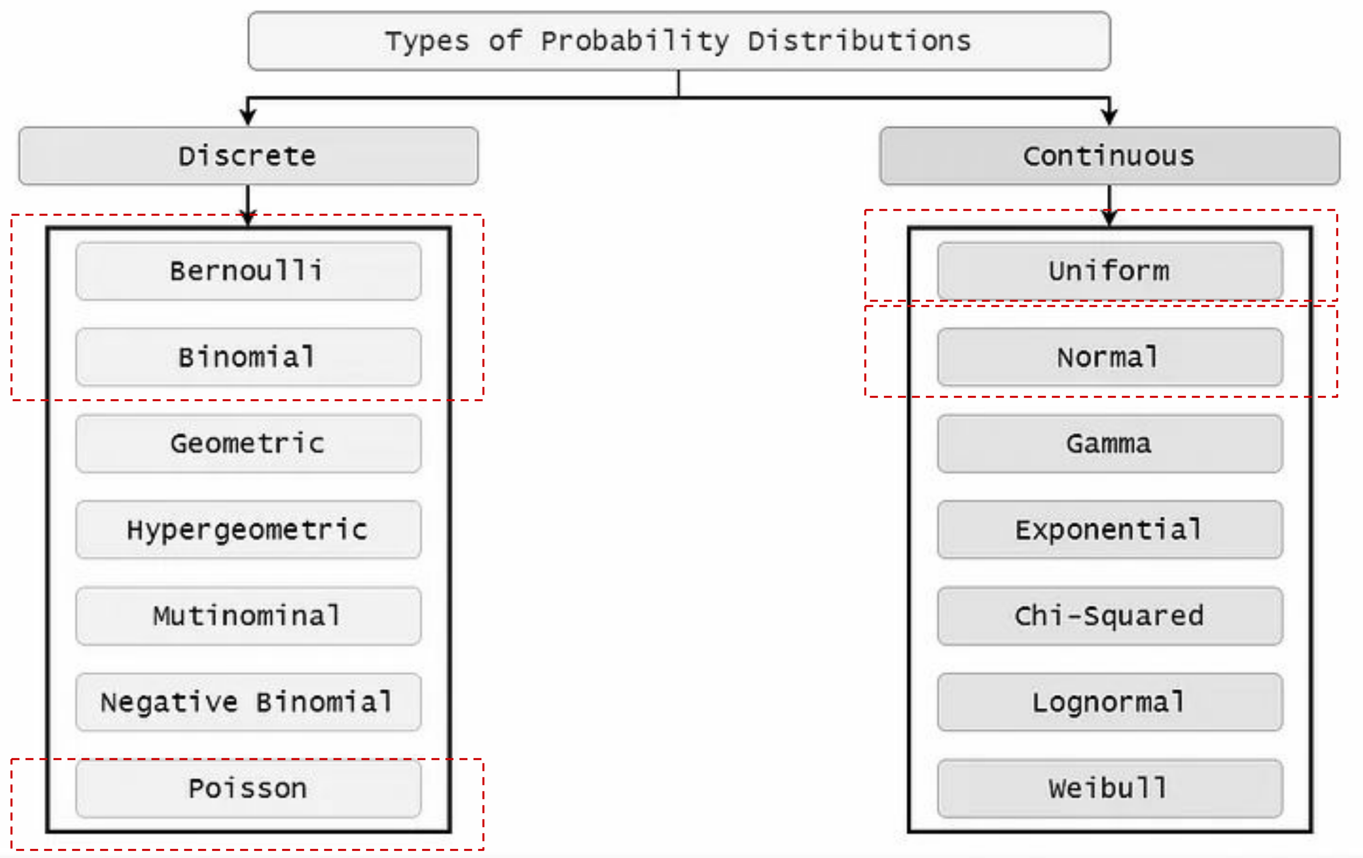
Concept

A probability distribution is a **statistical function** that describes the **likelihood** of obtaining all possible values that a random variable can take. In other words, the values of the Random Variable vary based on the underlying probability distribution. We usually write Random Variable with capital letter like X. Then Probability Distribution of X is donate by P(X).



Probability Distribution - Types

Concept

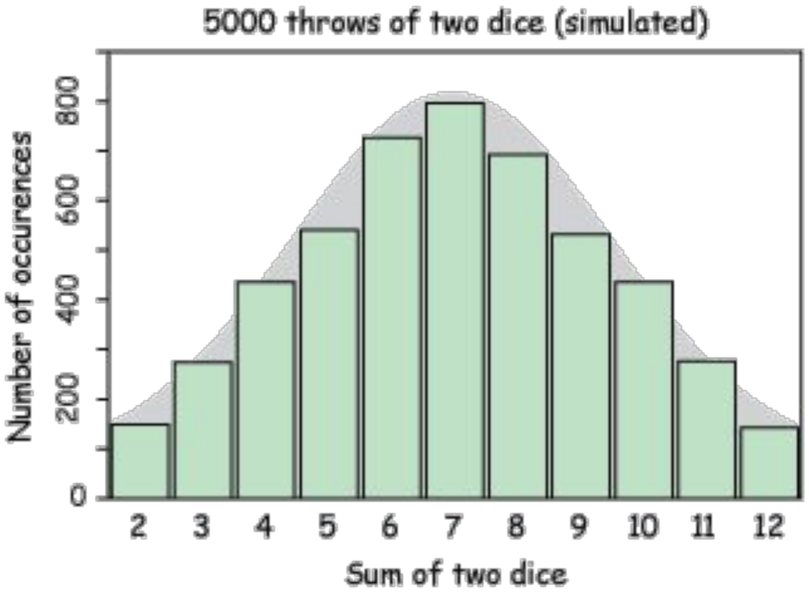


Discrete Probability Distribution

Concept

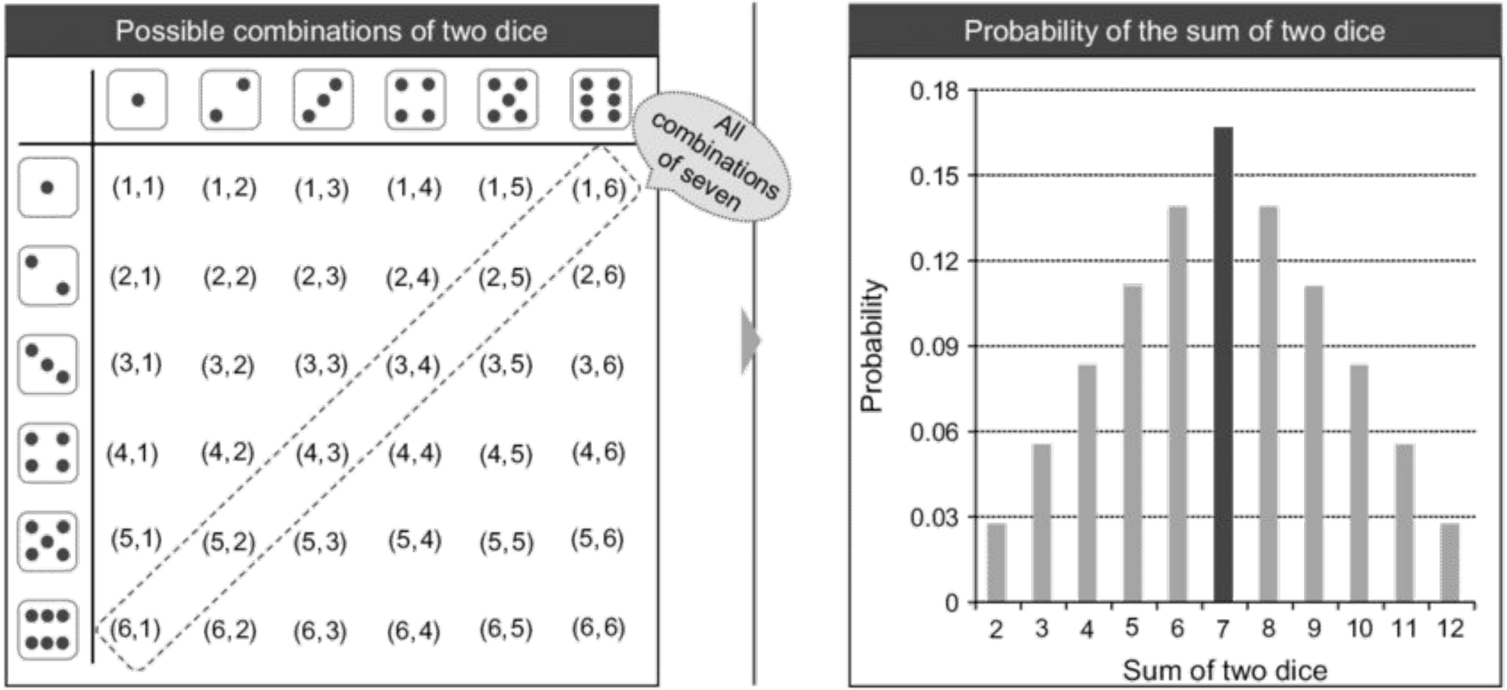
A **discrete random** variable is a variable that can take on a finite number of distinct values. For example, coin tosses and counts of events are discrete functions. These are discrete distributions because there are no in-between values.

For **discrete probability distribution** functions, each possible value has a **non-zero likelihood.** Furthermore, the **probabilities for all possible values must sum to one.** Because the total probability is 1, one of the values must occur for each opportunity.



Discrete Probability Distribution - Example

Concept



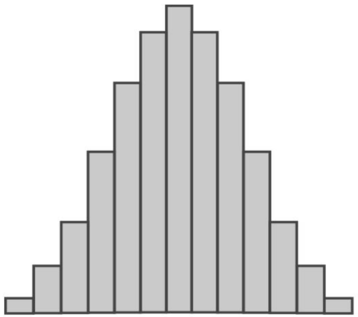
Continuous Probability Distribution

Concept

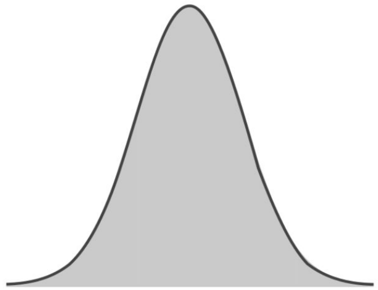
Continuous probability functions are also known as **probability density functions.** You know that you have a continuous distribution if the **variable can assume an infinite number** of values between any two values. Continuous variables are often measurements on a scale, such as height, weight, and temperature.

Unlike discrete probability distributions where each particular value has a non-zero likelihood, **specific values in continuous probability distribution functions have a zero probability.** For example, the likelihood of measuring a temperature that is exactly 32 degrees is zero.

The values of the Continuous PD **can be measured** and not counted. It will take infinite amount of time to count all the outcomes of a continuous variable.



Discrete



Continuous

Continuous Probability Distribution - Example

Concept

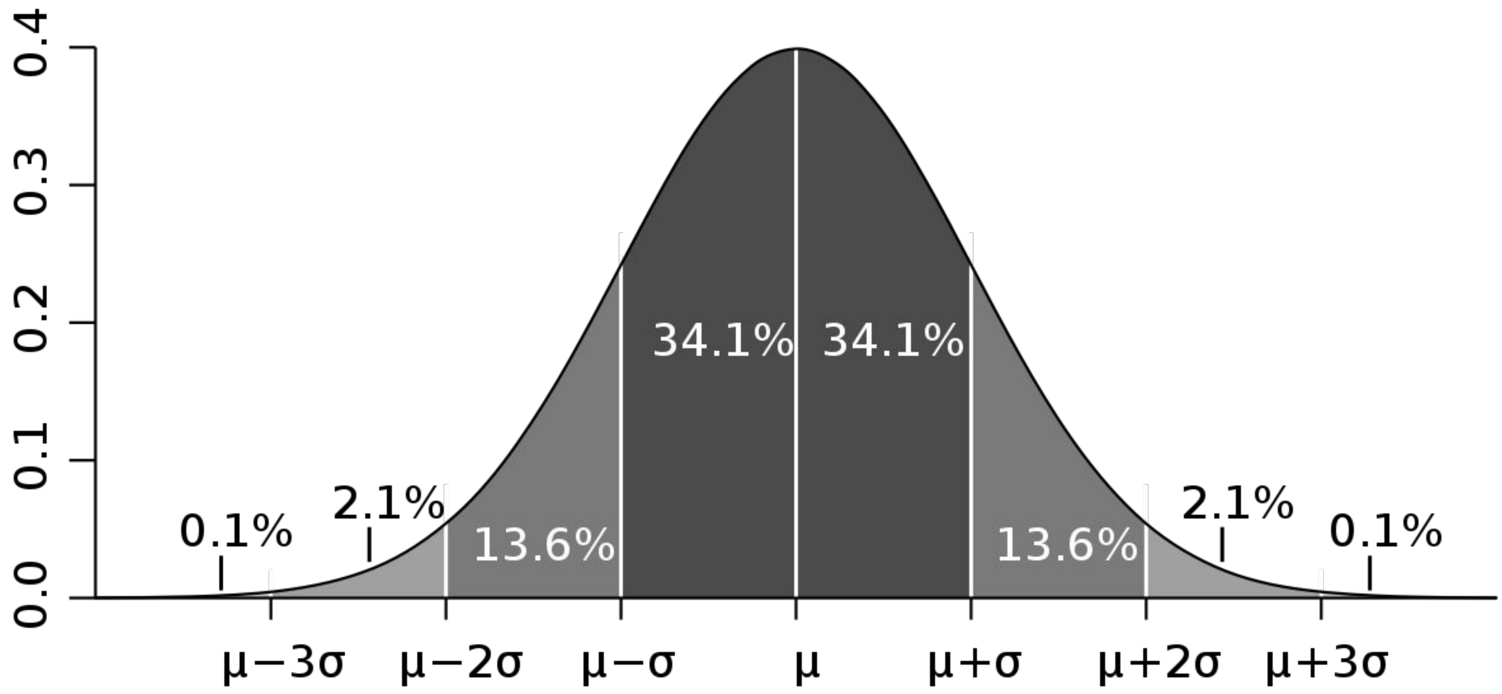
Distance travelled by my car from city A to B
50 kms
50 kms and 450 meters
50 kms 450 meters and 100 cms
50 kms 450 meters 100 cms and 25 mm

For continuous variables, **we can become more precise (or sharp) but** we **cannot count the exact** distance, age or time just the way we count discrete variables. That is why it is a continuous variable.

In other words, we can always find another data point between 2 data points of an infinite variable eg - between 50 kms and 51 kms there is 50.5 kms. Further between 50 kms and 50.5 kms we can have 50.25 kms and so on.

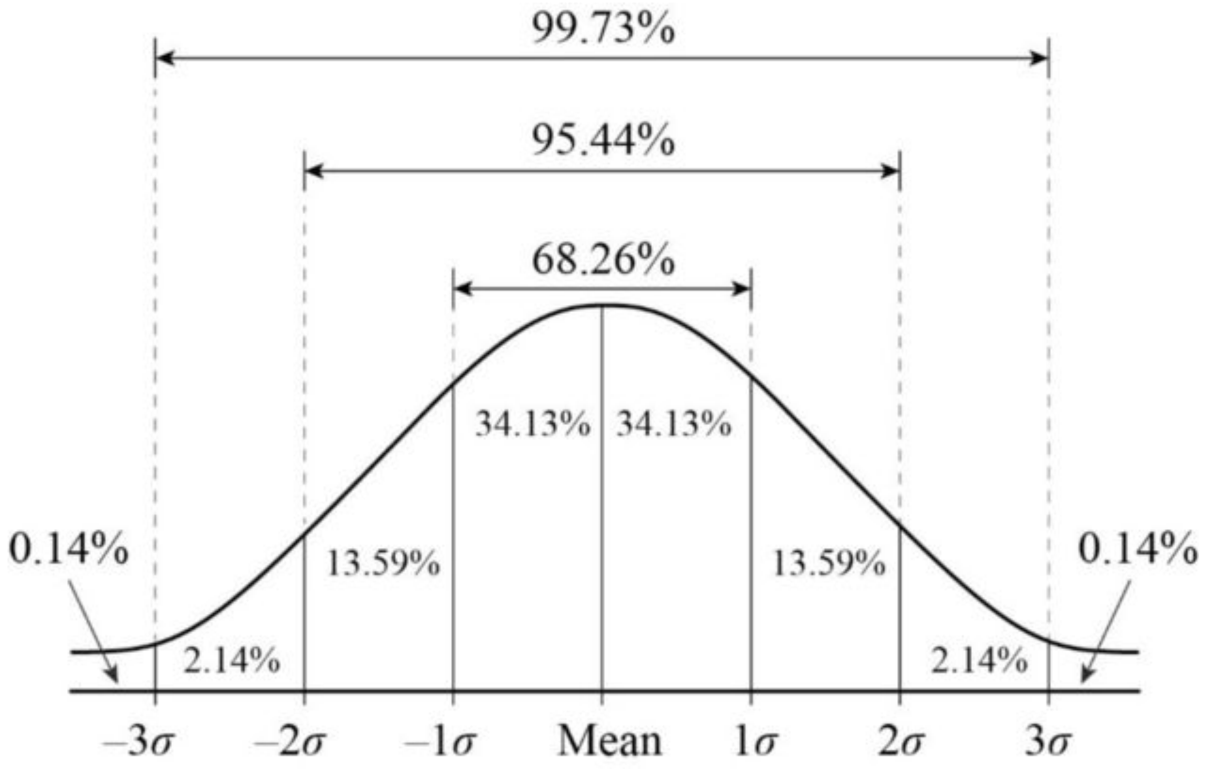
Continuous Probability Distribution - Representation

Concept



Continuous Probability Distribution - Example

Concept



Random Variable

Concept

Cumulative Distribution Function (CDF) for a Discrete Random Variable

The **cumulative distribution function,** $F(x)$, of a discrete random variable X is

$$F(x) = P(X \leq x) = \sum_{\text{all } i \leq x} P(i)$$

Random Variable - Expected Value

Concept

Expected Values of Discrete Random Variables

- The Mean of a Probability Distribution of a Random Variable is a measure of Centrality
- It is a measure that considers both -
  - The values of Random Variable
  - The probabilities of the possible outcomes of Random Variable
- It is the Weighted Average of the possible outcomes of the Random Variable
- The **Mean of the Probability Distribution of a Random Variable is called the Expected Value** of the Random Variable (sometimes called the Expectation)
- It is denoted by μ or E(X)

The **expected value** of a discrete random variable X is equal to the sum of all values of the random variable, each value multiplied by its probability.

$$\mu = E(X) = \sum_{\text{all } x} xP(x)$$

Expected Value - Example

Example

For flipping coin example following is the Cumulative distribution function (Discrete Variable)

X (Outcome)	Options	Probability [P(X)]	CDF	μ = E(X)= $\sum$x*[P(X)]
0	TTTT	1/16 = 0.0625	0.0625	0*(0.0625) = 0.00
1	TTTH, TTHT, THTT, HTTT	4/16 = 0.2500	0.3125	1*(0.2500) = 0.25
2	HHTT, TTHH, HTHT, THTH, HTTH, THHT	6/16 = 0.3750	0.6875	2*(0.3750) = 0.75
3	HHHT, HHTH, HTHH, THHH	4/16 = 0.2500	0.9375	3*(0.2500) = 0.75
4	HHHH	1/16 = 0.0625	1.0000	4*(0.0625) = 0.25

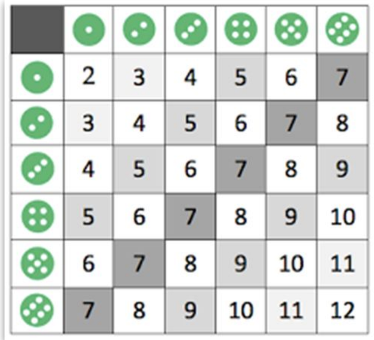
Expected Value = E(X) =μ = 2.00

Expected Value - Example for the roll of 2 dice

Example

Calculate the following for the roll of 2 dice:

1) **Expected Value and**
2) **Variance and S.D.**



1) Expected Value

X	2	3	4	5	6	7	8	9	10	11	12
P(X=x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
Expected Value $\mu = E(X) = \sum x*[P(X)]$	= 2*(1/36)+3*(2/36)+4*(3/36)+5*(4/36)+6*(5/36)+7*(6/36)+8*(5/36)+9*(4/36)+10*(3/36)+11*(2/36)+12*(1/36)										
	= 252/36 = 7										

Random Variable - Variance and Standard Deviation

Concept

Variance and Standard Deviation (SD) for a Random Variable

- Variance of Random Variable is the **Expected Squared Deviation** of the random variable from its **Mean (Expected Value)**.
- This is similar to that of the Variance of a data set or a population
- Probabilities of the values of the random variable are used as weights in the computation of the expected squared deviation from the mean of a discrete random variable
- The definition of the variance follows. As with a population, we denote the variance of a random variable by $\boldsymbol{\sigma}^2$.

The **variance** of a discrete random variable X is given by

$$\sigma^2 = V(X) = E[(X - \mu)^2] = \sum_{\text{all } x} (x - \mu)^2 P(x)$$

Computational formula for the variance of a random variable:

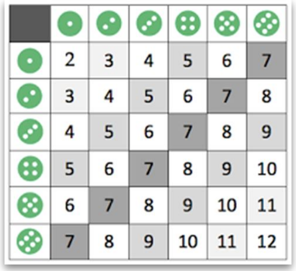
$$\sigma^2 = V(X) = E(X^2) - [E(X)]^2$$

Variance - Example for the roll of 2 dice

Example

Calculate the following for the roll of 2 dice:

1) **Expected Value and**
2) **Variance and S.D.**



2) Variance =

X	2	3	4	5	6	7	8	9	10	11	12
P(X=x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
EV= μ = E(X)= $\sum x*[P(X)]$	= 2*(1/36)+3*(2/36)+4*(3/36)+5*(4/36)+6*(5/36)+7*(6/36)+8*(5/36)+9*(4/36)+10*(3/36)+11*(2/36)+12*(1/36)										
$E(X^2)= \sum x^2_i*[P(X)]$	= 2^2 x(1/36)+3^2 x(2/36) + 4^2 x (3/36) +5^2 x (4/36) +6^2 x (5/36) +7^2 x (6/36) +8^2 x (5/36) +9^2 x (4/36) +10^2 x (3/36) +11^2 x (2/36)+12^2 x (1/36) =										
$V(X) = E(X^2) - [E(X)]^2$	= 54.83 (-) 49 = 5.83										

Bernoulli Distribution

Concept

Bernoulli Random Variable

- It is the most basic discrete Random Variable Distribution
- Named in the honour of mathematician - Jakob Bernoulli (1654–1705)
- The basic form of Bernoulli Distribution for Random Variable x where x can only take 2 values (0 and 1) -
  - x = 1 with probability p
  - x = 0 with probability (1 - p) or q
  - Event x = 1 is called **'Success'** and Event x = 0 is called **'Failure'**
- Outcome of Bernoulli trial can only be either a 'Success' or a 'Failure'
- Bernoulli trial is a single experiment
- Some examples of Bernoulli are - Tossing a coin, Yes/ No, Male/ Female etc.

Bernoulli Distribution

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

For any Bernoulli Distribution -

$$
\begin{aligned}
E(X) &= 1 * p + 0 * (1 - p) = p \\
E(X^2) &= 1^2 * p + 0^2 * (1 - p) = p \\
V(X) &= E(X^2) - [E(X)]^2 = p - p^2 = p(1 - p)
\end{aligned}
$$

Where -

E(X) *- Mean of the outcomes*

V(X) *- Variance of the outcomes*

Bernoulli Distribution - Example

Concept

Q - A striker in a game of soccer can shoot with the probability of 0.7. Find the Mean and Variance of the striker in the game?

Here $p = 0.7$ and $q = 0.3$; **Mean =** $E(X) = 1*(0.7) + 0*(0.3) = 0.7$; Variance = $(0.7)*(0.3) = 0.21$

Q - An AI bot has a 0.9 probability of correctly predicting the outcome of an event. What is the Mean and Variance of the AI bot in prediction?

Here $p = 0.9$ and $q = 0.1$; **Mean =** $E(X) = 1*(0.9) + 0*(0.1) = 0.9$; Variance = $(0.9)*(0.1) = 0.09$

Binomial Distribution

Concept

Binomial Random Variable

The Binomial distribution represents the probability for **'r'** successes of an experiment in **'n'** trials, given a success probability **'p'** for each trial at the experiment.

An X (Random Variable) that counts the number of **'Successes'** in multiple Independent, identical Bernoulli trials is called a Binomial Random Variable. Thus in Binomial Distribution the experiment is repeated multiple times.

In a single experiment when n = 1, the Binomial distribution is called a Bernoulli distribution.

Properties of Binomial Distribution -

- Experiment consist of n repetition
- Each trial has only 2 possible outcomes
- Probability of Success, denoted p, is the same for each trial.
- Each trial is independent.

Binomial Probability Distribution

$$P(r) = {}_nC_r\,(p)^r(1-p)^{n-r}$$

Mean $\mu = np$

Standard Deviation $\sigma = \sqrt{np(1-p)}$

p- probability of success
r- number of successes
n- number of trials

Binomial Distribution - Example

Concept

Q - If a coin is tossed 10 times, use binomial find the probability of: (a) Exactly 4 heads (b) At least 7 heads.

Solution - No of trials = n = 10; Probability of head: p = 1/2 and probability of tail, q = 1/2

Probability of exactly 4 heads:

X = 4

P(x =4) = $\mathbf{{}^{10}C_4\ p^4\ q^{10-4}}$ = (10!/ (4!x 6!)) × $(½)^4$× $(½)^6$

P(x =4) = 210*0.000976563 = **0.205**

Probability of exactly 7 or more heads:

P(X >= 7) = P(x=7) + P(x=8) + P(x=9) + P(x=10)

Use COMBIN in excel to get the value of)

= $(\mathbf{{}^{10}C_7+{}^{10}C_8+{}^{10}C_9+{}^{10}C_{10}})*(2)^{\mathbf{10}}$

= 0.1171875 + 0.043945313 + 0.009765625 + 0.000976563= **0.171875**

Poisson Distribution

Concept

Poisson distribution is a discrete distribution function that is useful for **characterizing events with very low probabilities of occurrence within some definite time or space.** It is applicable in events that have a large number of rare and independent possible events. It is used to estimate how many times an event is likely to occur within the given period of time. Some examples of Poisson Distribution are -

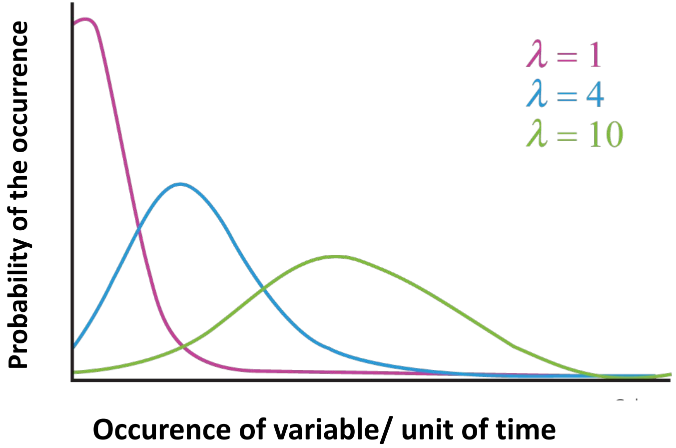
a) On a toll, a car arrives at a mean rate of 10 per min. What is the probability of arrival of 50 cars in 1 minute.

b) 2% of the phones manufactured by Samsung are defective, what is the probability that 10 phones turn defective in a sample of 200.

c) Mc'donalds restaurant has 200 customers coming every evening. What is the probability that on a day 400 customers turn up

It is used under the condition when - the **no. of trails 'n' tends to infinity** the probability of **Success 'p' tends to zero.**

Thus np = 1 i.e. a finite number.



Poisson Distribution

Concept

Properties of Poisson Distribution -

- Events are independent and occur at random
- The number of trails is infinitely large
- The variable of interest has a discrete outcome
- Mean = Variance = λ
- np = λ is finite, where λ is constant

Applications of Poisson Distribution -

- Count the number of defects of a finished product
- Count the number of deaths in a country by any disease or natural calamity
- Count the number of infected plants in the field
- Count the number of bacteria in the organisms or the radioactive decay in atoms
- Calculate the waiting time between the events.

Poisson Probability Distribution

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

λ- mean number of successes over a given interval

$Var(X) = \lambda$

Where -

λ - is the average number of times an event occurs

e - is Euler's constant (approximately 2.718)

Poisson Distribution - Example

Concept

1) On a toll, a car arrives at a mean rate of 10 per min. What is the probability of arrival of 15 cars in 1 minute.

Mean = λ = 10. Use =POISSON.DIST(15,10,FALSE) in excel. Answer = 0.034718

2) 2% of the phones manufactured by Samsung are defective, what is the probability that 8 phones turn defective in a sample of 200.

Mean = λ = 2%. Use =POISSON.DIST(4,2,FALSE) in excel. Answer = 0.090224

3) Mc'donalds restaurant has 200 customers coming every evening. What is the probability that on a day 220 customers turn up?

Mean = λ = 200. Use =POISSON.DIST(220,200,FALSE) in excel. Answer = 0.01021

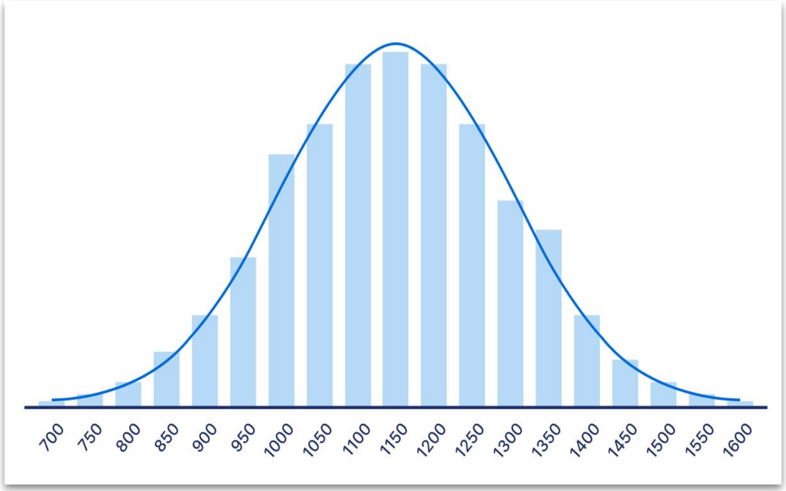
The Normal Distribution

Concept

The **Normal Distribution**, also known as the Gaussian distribution, is the most important probability distribution in statistics for independent, random variables. Most people recognize its familiar **bell-shaped curve** in statistical reports. The data is symmetrically distributed with no skewness. **When plotted on a graph, the data follows a bell shape, with most values clustering around a central region and tapering off as they go further away from the center.**

Why Normal Distribution?

- All kinds of variables in **natural and social sciences are usually or approximately normally distributed**. Height, birth weight, reading ability, job satisfaction, or SAT scores are just a few examples of such variables.
- Because normally distributed variables are so common, many **statistical tests for inferences are designed for normally distributed populations.**
- It is imperative to understand the properties of normal distributions to **use inferential statistics to compare different groups** and make estimates about populations using samples.

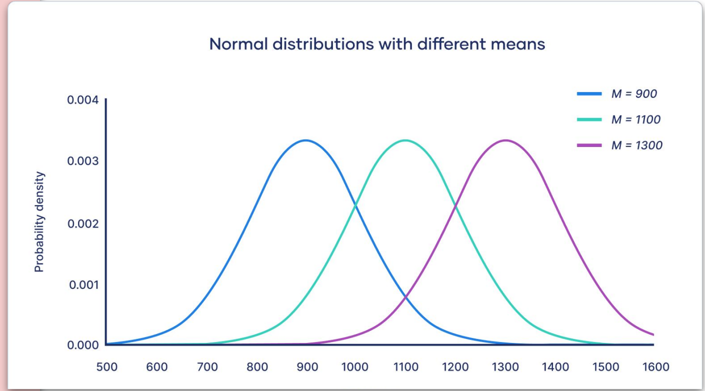


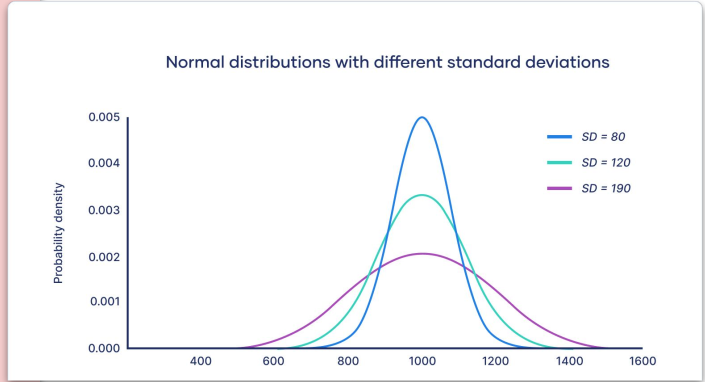
Properties of Normal Distribution (1/2)

Concept

Normal Distribution is easy to identify

- Mean = Median = Mode
- Distribution is symmetric about the mean - half the values lie below the mean and half above the mean
- The normal curve is asymptotic i.e. the curve approaches x-axis on both the ends but does not meet the x-axis
- Distribution can be defined by 2 parameters - Mean and S.D.
- Mean is the location parameter. It determines where the peak of the curve is centered. As the mean shifts right the Normal curve shifts right and as the mean shifts left the curve shift left.
- S.D. is the scale parameter. It decides how much the curve is spread or is squeezed. Smaller the S.D. narrow the curve and larger S.D. widens the curve



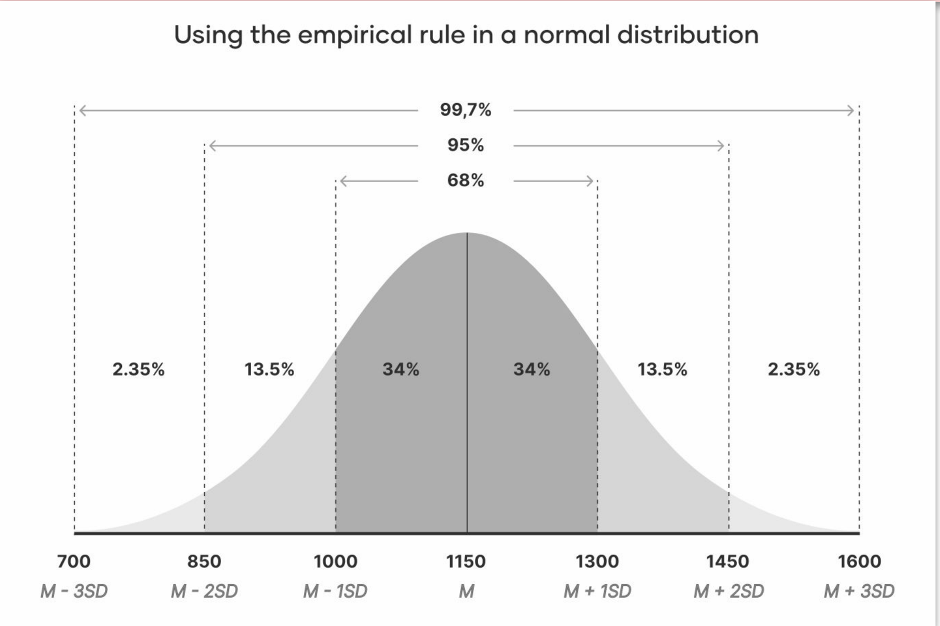


Properties of Normal Distribution (2/2)

Concept

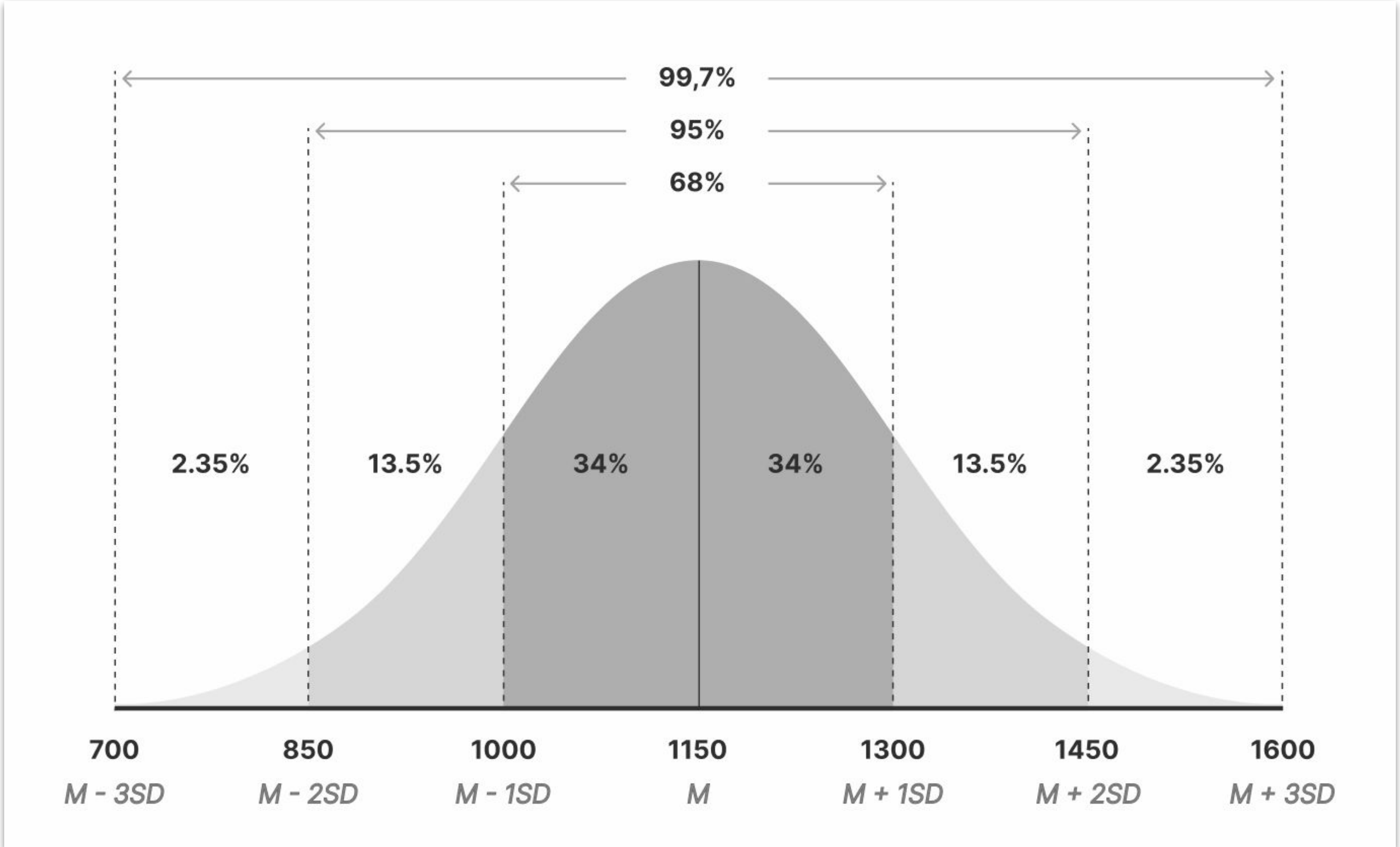
Empirical Rule or 3 Sigma Rule or the **68-95-99.7 rule** applies to Normal curve. It states that -

- Around **68%** of values are within **1 S.D.** from the mean.
- Around **95%** of values are within **2 S.D.** from the mean.
- Around **99.7%** of values are within **3 S.D.** from the mean.



Properties of Normal Distribution (2/2)

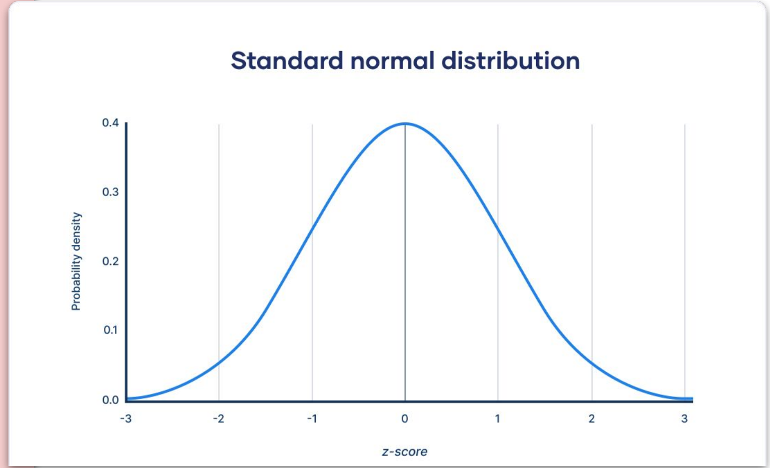
Concept



Standard Normal Distribution - SND (z-distribution)

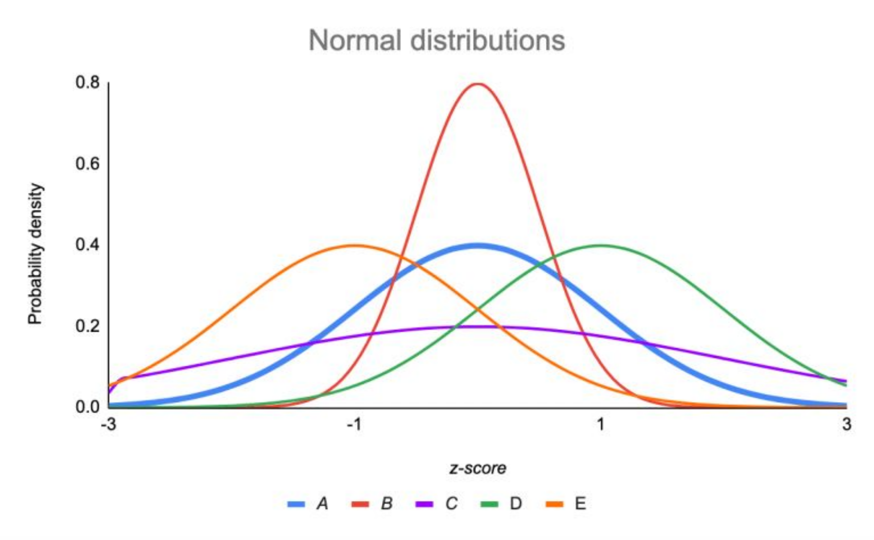
Concept

- The standard normal distribution, also called the **z-distribution**, is a **special normal distribution** where the **Mean is 0** and the **S.D. is 1. Area under curve =1 or 100%.**
- Every normal distribution can be created as a version of the SND that's stretched or squeezed and moved horizontally right or left. In other words every Normal Distribution can be converted into an SND by adjusting the scale and the central position
- In normal distribution the observations are referred as x. In the z-distribution they are referred to as z. The normal distribution is converted into a z-distribution by turning x values into z scores on the x-axis.
- The z-score tells you how many standard deviation away from the mean each value lies.



Converting Normal Distribution into SND (z-distribution)

Concept



Curve	Position or Shape (in relation to SND)
A (*M* = 0, *SD* = 1)	**Standard Normal Distribution**
B (M = 0, SD = 0.5)	*Squeezed, because SD < 1*
C (M = 0, SD = 2)	*Stretched, because SD > 1*
D (M = 1, SD = 1)	*Shifted right, because M > 0*
E (M = −1, SD = 1)	*Shifted left, because M < 0*

- Data points are referred to as **x** in a normal distribution, they are called **z** or **z-scores** in the z-distribution. A z-score is a standard score that tells you how many standard deviations away from the mean an individual value (x) lies`

Permutations and Combinations

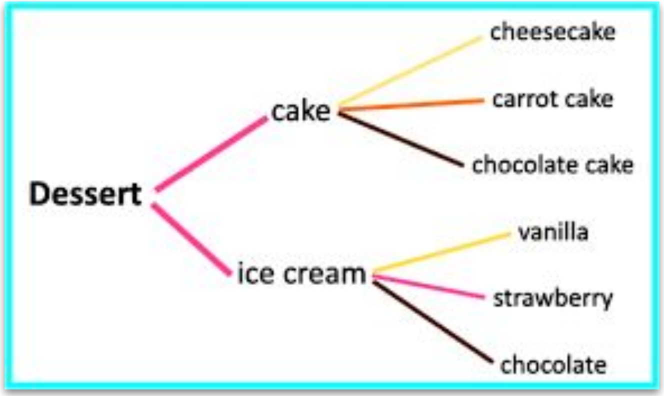
Fundamentals Principle of Counting

What is the Fundamental Principle of Counting?

The principle states that - if **an event can occur** in **m different ways**, and another event can occur in **n different ways**, then the **total number of occurrences** of the events is **m × n**.

Example - I have 3 books to read and 4 different places I can read these books - Home, Office, Park and Car. Total number of ways I can read books are 3 X 4 = 12 (as shown below)

Books/ Places	**Home**	**Office**	**Park**	**Car**
Book-1	Book1/ Home	Book1/ Office	Book1/ Park	Book1/ Car
Book-2	Book2/ Home	Book2/ Office	Book2/ Park	Book2/ Car
Book-3	Book3/ Home	Book3/ Office	Book3/ Park	Book3/ Car



Understanding Factorial

What are Factorials?

- Factorials was first discovered by Daniel Bernoulli.
- It is a simple concept which only involves multiplication.
- **Factorial is a multiplication operation of natural numbers with all the natural numbers that are less than it.**
- Used in - Probability, Permutation and Combination.
- Indicated by **'!'.**
- **n Factorial is written as**
- **$n! = n \times (n-1) \times (n-2) \times (n-3) \times \ldots . \times 3 \times 2 \times 1$**
- **$n! = n \times (n-1)!$**

Factorials from 0 till 8

$0! = 1$

$1! = 1 \text{ X } 1$

$2! = 2 \times 1 = 2$

$3! = 3 \times 2 \times 1 = 6$

$4! = 4 \times 3 \times 2 \times 1 = 24$

$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$

$7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5040$

$8! = 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 40320$

Permutations

- Permutation relates to **the act of ‘arranging’ all the members of a set into some sequence or order.**
- In other words, if the set is already ordered, then the rearranging of its elements is called the **process of permuting**.
- They often arise when different orderings on certain finite sets are considered.
- Thus Permutations are - **ARRANGEMENTS.**
- A permutation is the **choice of 'r' things** from a **set of 'n' things without replacement** and where the order matters.

$$P(n,r) = \frac{n!}{(n-r)!}$$

Combination

- Combination is a **way of selecting items** from a collection, such that (unlike permutations) **the order of selection does not matter.**
- **Combination refers to the combination of 'n' things taken 'k' at a time without repetition.**
- To refer to combinations in which repetition is allowed, the terms k-selection or k-combination with repetition are often used.

$${}_nC_r = \binom{n}{r} = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!}$$

More examples of Permutations and Combination

Pick 2 letters from - a,b,c,d.

1) In how many different ways can you pick the 2 letters.

ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc.

This can be calculated by **P(4,2)** i.e. **4!/ (4-2)! = (4x3x2x1)/(2x1) = 12**

2) What are the different combination of selecting 2 different letters from a,b,c,d.

ab, ac, ad, bc, bd, cd.

This can be calculated by **C(4,2)** i.e. **4!/ 2! x (4-2)! = (4x3x2x1)/2x(2x1) = 6**

In Nutshell -

a) **In Permutation - ‘ab’ and ‘ba’ are not same as the ORDER is important**
b) **In Combination - ‘ab’ and ‘ba’ are same. i.e. order is NOT important**

Use of Permutation and Combination

PERMUTATION is used for the list of data (where the order of the data matters)
COMBINATION is used for a group of data (where the order of data doesn't matter)

Simulator for Permutation and Combination

Click

Permutation Examples	Combination Examples
Arranging people, digits, numbers, alphabets, letters and colours	Selection of menu, food, clothes, subjects and teams
Picking a team captain, batsmen, bowler from a group in that order	Picking three team members from a group.
Picking 2 favourite colors, in order, from a color card	Picking 2 colors from the color card.
Picking first,second and third place	Picking top 3 winners.

Examples of Permutations and Combination

For the word - CODE - write all the Permutations possible.

Total Permutations $= {}^{n}P_{r}$

$n = 4$ and $r = 4$

$= 4!/\,(4-4)!$

C,D,E,O	D,C,E,O	E,C,D,O	O,C,D,E
C,D,O,E	D,C,O,E	E,C,O,D	O,C,E,D
C,E,D,O	D,E,C,O	E,D,C,O	O,D,C,E
C,E,O,D	D,E,O,C	E,D,O,C	O,D,E,C
C,O,D,E	D,O,C,E	E,O,C,D	O,E,C,D
C,O,E,D	D,O,E,C	E,O,D,C	O,E,D,C

Examples of Permutations and Combination

For the word - AGAIN - calculate the number of words which can be made using all the letters of the word AGAIN?

Words that can be created using 5 letters

Total Permutations

= ${}^{n}P_{r}$ where **n = 5** and **r = 5**

Now since there are 2 A in the word we will divide by 2! Because the position of the 2 A's does not make a difference in the arrangement. Hence the final answer = **5!/(5-5)!*2!** = **60**

A,A,G,I,N	**A,I,A,G,N**	**G,A,A,I,N**	**I,A,A,G,N**	**N,A,A,G,I**
A,A,G,N,I	**A,I,A,N,G**	**G,A,A,N,I**	**I,A,A,N,G**	**N,A,A,I,G**
A,A,I,G,N	**A,I,G,A,N**	**G,A,I,A,N**	**I,A,G,A,N**	**N,A,G,A,I**
A,A,I,N,G	**A,I,G,N,A**	**G,A,I,N,A**	**I,A,G,N,A**	**N,A,G,I,A**
A,A,N,G,I	**A,I,N,A,G**	**G,A,N,A,I**	**I,A,N,A,G**	**N,A,I,A,G**
A,A,N,I,G	**A,I,N,G,A**	**G,A,N,I,A**	**I,A,N,G,A**	**N,A,I,G,A**
A,G,A,I,N	**A,N,A,G,I**	**G,I,A,A,N**	**I,G,A,A,N**	**N,G,A,A,I**
A,G,A,N,I	**A,N,A,I,G**	**G,I,A,N,A**	**I,G,A,N,A**	**N,G,A,I,A**
A,G,I,A,N	**A,N,G,A,I**	**G,I,N,A,A**	**I,G,N,A,A**	**N,G,I,A,A**
A,G,I,N,A	**A,N,G,I,A**	**G,N,A,A,I**	**I,N,A,A,G**	**N,I,A,A,G**
A,G,N,A,I	**A,N,I,A,G**	**G,N,A,I,A**	**I,N,A,G,A**	**N,I,A,G,A**
A,G,N,I,A	**A,N,I,G,A**	**G,N,I,A,A**	**I,N,G,A,A**	**N,I,G,A,A**

Examples of Permutations and Combination

In how many ways 4 letter words (no meaning needed) can be drawn from the word - 'NUMBERS'?

Words that can be created using 4 letters **Total Permutations**

= $^{n}P_{r}$ where **n = 7** and **r = 4**

= **P(7,4)**