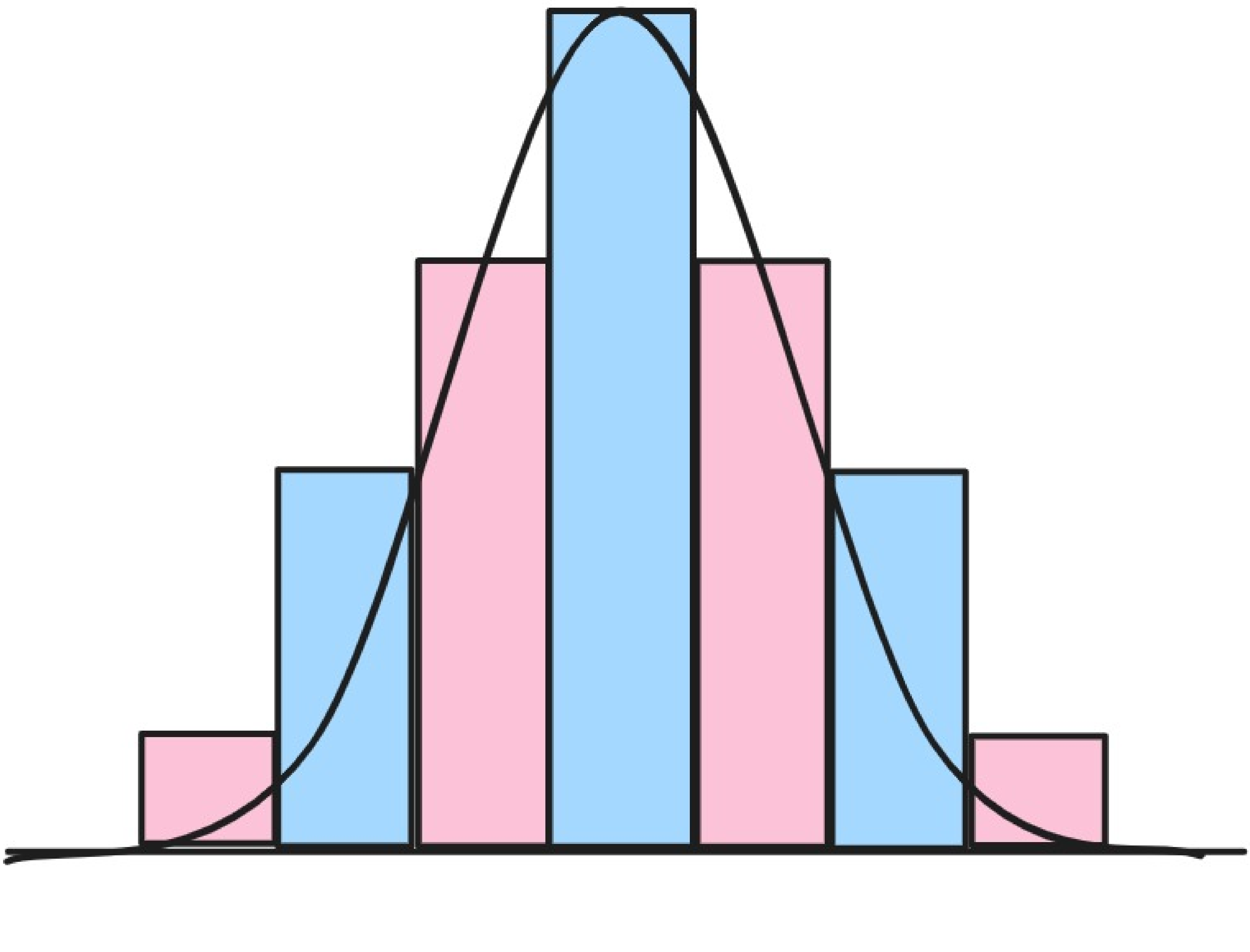
Simplifying AI

Probability in Machine Learning
Part-1



Contents

Terminology in Probability

1. Random Experiment

An Experiment is called random if it satisfies 2 conditions:

i) It has more than one possible outcome.
ii) It is not possible to predict the outcome in advance.

2. Trial

Trial is single execution of random experiment.
Each Trial produces an outcome.

3. Outcome

Outcome is single possible result of trial.

4. Sample space

Sample space is all set of possible outcome that can occur.
One random experiment will have one sample space.

5. Event

Event is a specific set of outcomes from a random experiment.
It's a subset of the sample space.
Event can include a single outcome or it can include multiple outcomes.
Onc random experiments can have multiple events.

Terminology Example

Example: Let's consider a 6-side die that can show number from number 1-6.

Random Experiment	Rolling a Die
Trial	Rolling the die once
Outcome	Any possible result of rolling a die Eg: {5}
Sample Space	All Possible outcome. Eg: {1, 2, 3, 4, 5, 6}
Event	Getting a 3. Eg: {3} Getting a number greater then 4. Eg: {5, 6} Getting an even number. Eg: {2, 4, 6}

Types of Events

1. Simple Event

A simple event is when there's only one outcome.

Eg: When rolling a die, the event of "rolling a 2" is a simple event.

2. Compound Event

A compound event consists of two or more simple events.

Eg: When rolling a die, the event of "rolling an even number" is a compound event. It consists of three simple events, namely rolling any of {2, 4, 6}.

3. Independent Event

Independent events are two events where the occurrence of one event doesn't affect the probability of the other event.

Eg: When flipping a coin and rolling a die, the event of "tossing a tail" and "rolling an even number" are independent events. The outcome of the coin flip doesn't affect the outcome of rolling the die.

4. Dependent Event

Dependent events are two events where the occurrence of one event affects the probability of the other event.

Eg: When drawing a card without replacement, the event of "drawing the first card as black and the second card as red" represents dependent events. The outcome of the first card affects the outcome of the second card because there will be fewer cards remaining compared to the first draw.

Types of Events

5. Mutually Exclusive event

Mutually exclusive events are events that cannot both occur simultaneously.

Eg: When rolling a die, the event of "rolling a 1" and "rolling a 2" are mutually exclusive as single roll can't have both outcome.

6. Exhaustive event

Exhaustive events are events where at least one event from a set must occur when an experiment is performed.

Eg: When rolling a die, the events "rolling an even number" or "rolling an odd number" are exhaustive because one of these events is certain to happen, covering the complete sample space of possible outcomes.

7. Impossible event

An impossible event is an outcome that cannot happen.

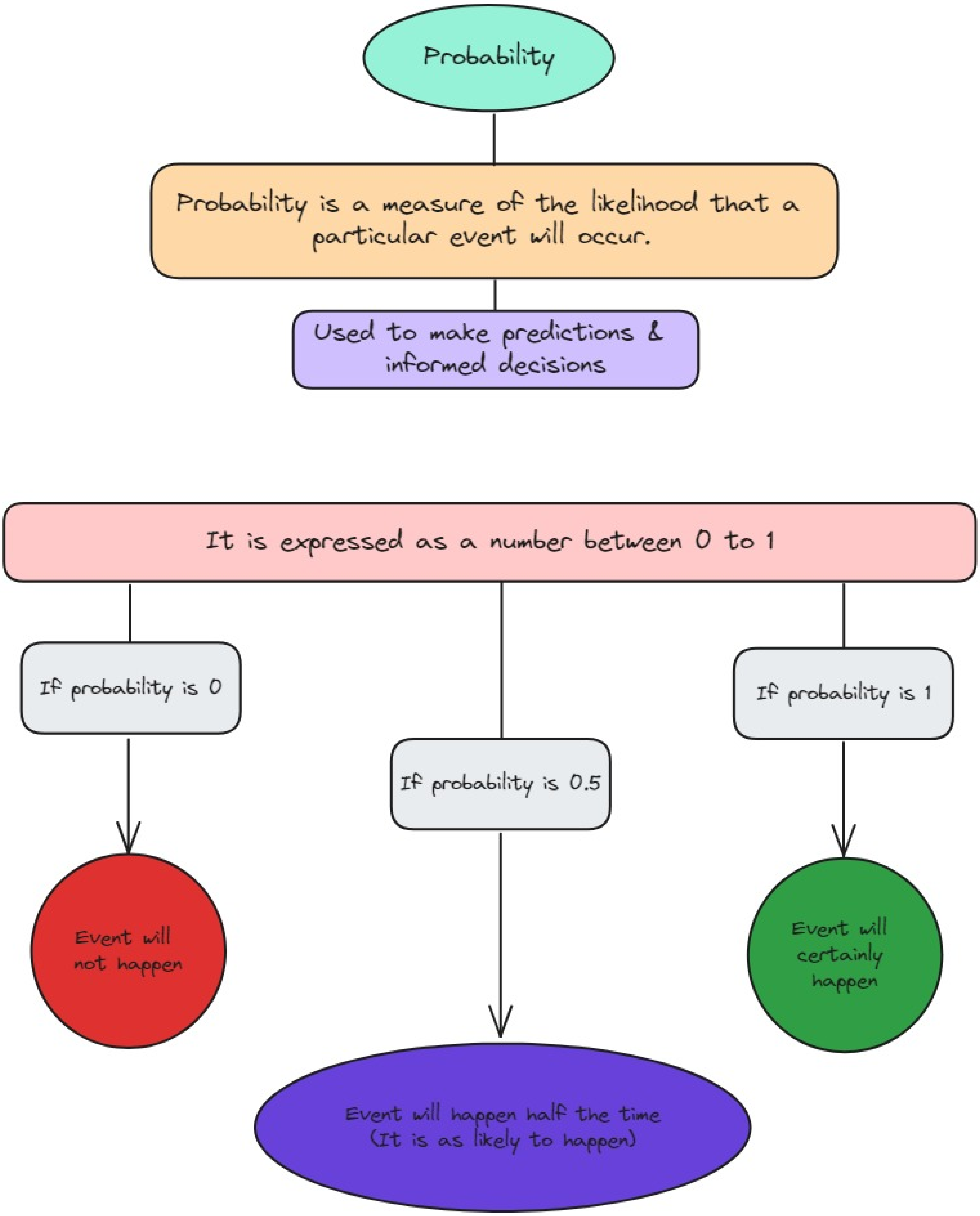
Eg: When rolling a die, the event of "rolling a number greater than 6" is an impossible event because the outcome cannot be greater than 6.

8. Certain event

certain event is an outcome that is guaranteed to happen.

Eg: When rolling a die, the event of getting a number between 1 and 6 is certain because the outcome will always be within that range.

Probability



Empirical vs Theoretical Proabability

Empirical Probability

A probability measure that is based on observed data, rather than theoretical assumptions.
It's calculated as the ratio of the number of times a particular event occurs to the total number of trials.
It's also known as Experimental Probability.

Eg. : If we toss a coin 100 times and get heads 60 times and tails 40 times, the empirical probability of getting heads is 60/100, which is 0.6, and the empirical probability of getting tails is 40/100, which is 0.4.

vs

Theroretical Probability

A probability that is used when each outcome in a sample space is equally likely to occur.
It's calculated for any event as the ratio of the number of favorable outcomes to the total number of outcomes in the sample space.
It is also known as Classical Probability.

Eg: If we consider flipping a coin 100 times, the theoretical probability of getting heads and tails is equal. The theoretical probability of getting heads is 50/100, which is 0.5, and the theoretical probability of getting tails is also 50/100, which is 0.5.

As the number of trials increases, empirical probability approaches theoretical probability.

In an infinite number of trials,

Empirical Probability ≈ Theoretical Probability

Examples of Random Variable

Let's consider rolling two dice at once, which will result in a sample space of 36 possible outcomes.

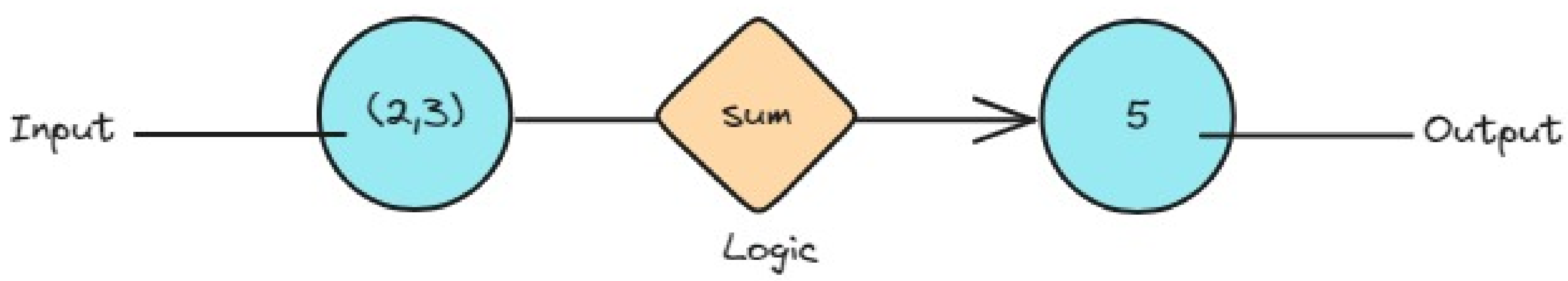
(1,1), (1,2), (1,3), (1.4). (1,5). (1,6),
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)

Eg:1

Event : Getting a sum of 7

To represent this event, we will capture all the outcomes of our sample space and add the numbers obtained on both dice.

Here, the logic will be to "add the numbers".

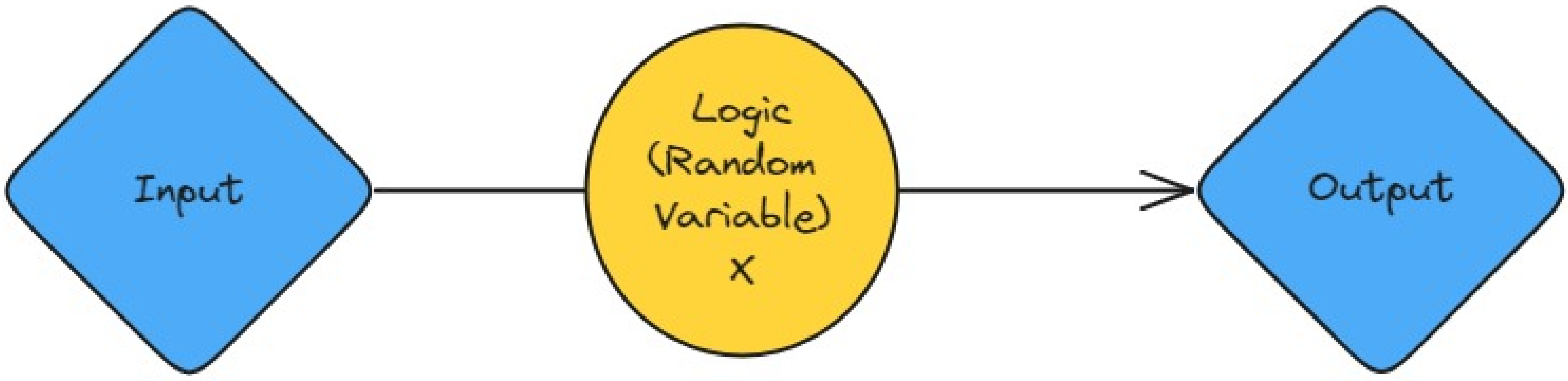


In this case, the random variable will consist of the following outcomes:
X = {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

When determining the probability of an event, the random variable will help identify the pattern.

Random Variable

Random variable is a function that maps the outcome of a random process (from the sample space) to a set of real numbers.
It is denoted by "X".



Input: The Input of the function is an outcome from the sample space of a random process

Ouput: The output of the function is a real number that we assign to each possible outcome

The transformation from input to output in the function of a random variable is determined by how we choose to define the random variable. The choice of definition depends on the specific aspect of the random process or event we want to study.

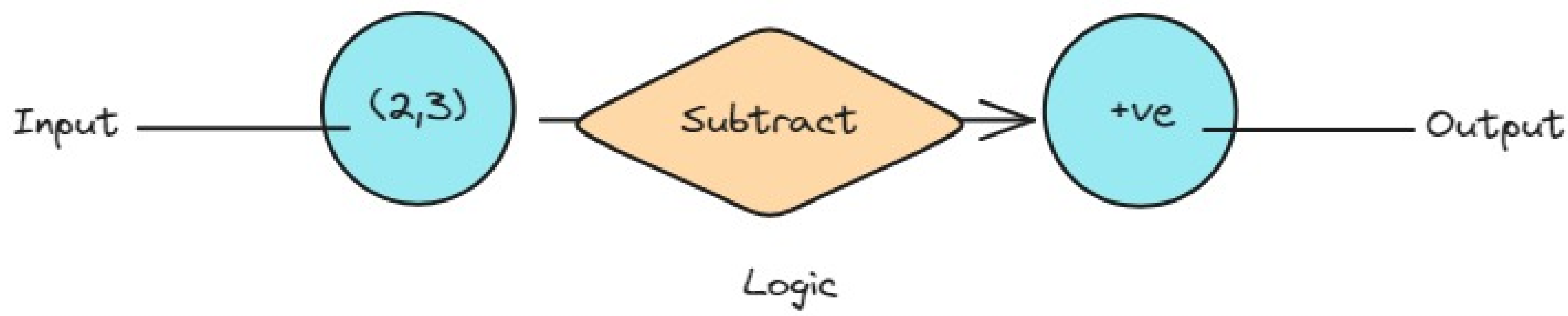
Examples of Random Variable

Eg:2

Event: Getting a smaller number on the first die than on the second die.

To represent this event, we will capture all the outcomes of our sample space and compare the numbers by subtracting the number on the first die from the number on the second die.

Here, the logic will be to "subtract the numbers"



In this case, the random variable will consist of the following outcomes:
X = {Positive integer if the number on the second die is less than the first die}
X = 0 { if the numbers on both dice are the same}
X = {Negative integer if the number on the second die is greater than the first die}
In this case, outcomes of (X) with negative integers will be the favorable outcomes given the condition of the event.

Types of Random Variable

There are 2 types of Random Variable

Discrete Random Variable

When the output of the random variable is a discrete number, it is known as a Discrete Random Variable.

Consider rolling two dice and aiming for the sum to be less than 5. In this case, the variable can take discrete values from 2 to 12.

Eg: 2, 4, 5, 11, 12

Continuous Random Variable

When the output of the random variable is a continuous number, it is known as a Continuous Random Variable.

Consider the CGPA of a student, with a minimum CGPA of 7. In this case, the variable can take continuous values from 0 to 10.

Eg: 7.12, 8.21, 9.81, 8.72, 6.52, 8.90

Probability Distribution of Random Variable

Probability Distribution is a list of all possible outcomes of a random variable along with their corresponding probability values.

Eg:1

Let's consider rolling a single die.

Rolling a single die will result in the following random variable output:
$X = \{1, 2, 3, 4, 5, 6\}$
A Probability Distribution Table will be formed with the occurrence probability for each output.

Random Variable (X)	1	2	3	4	5	6
Probability (P(X))	1/6	1/6	1/6	1/6	1/6	1/6

Eg:2

Let's consider rolling two dice at once.

Rolling two dice will result in the sum random variable output. The summation will occur by adding the numbers obtained on the two dice, and a probability distribution will be formed with the probability of occurrence of each possible sum.

Die 1

Die 2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Random Variable (sum) output of 2 dice:
$X = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

X	2	3	4	5	6	7	8	9	10	11	12
P(X)	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

Probability Distribution of Random Variable (Rolling 2 Dice)

When a function is derived for the probability against the output of the random variable, it is called a Probability Distribution Function (PDF).

When a graph is plotted against a PDF, it is called a Probability Density Function (PDF) if the random variable is continuous, or a Probability Mass Function (PMF) if the random variable is discrete.

Mean of Random Variable

Mean of Random variable (Expected Value) is essentially the average outcome of a random process that is repeated many times

It's a weighted average of the possible outcomes of the random variable, where each outcome is weighted by it's probability of occurence.

Let's consider the event of rolling a single die, where the probability of each number coming up is equal, i.e., 1/6 for each number from 1 to 6

Mean of the following random variable will be equal to:

1/6 (1) + 1/6 (2) + 1/6 (3) + 1/6 (4) + 1/6 (5) + 1/6 (6) = 3.5

↓

This signifies that when this event is performed many times, the average value (mean) of the random variable (X) will be 3.5.

The mean of a random variable E(X) is calculated by adding the product of the probability and outcome together.

Mean of Random Variable: $E(X) = \sum_{i=1}^{n} xi * P(xi)$

Variance of Random Variable

Variance of Random Variable how much individual observations in a group differ from the mean (expected value).

Let's consider the event of rolling a single die, where the probability of each number coming up is equal, i.e., 1/6 for each number from 1 to 6, where expected is equal to 3.5

Variance of the following random variable will be equal to:

((1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3/5)^2)/6 = 2.92

This signifies that when this event is performed many times, the variance of the random variable (X) will be 2.92.

The variance of a random variable Var(X) is calculated by taking the expected value (mean) of the squared difference between the random variable and its expected value.

Variance of Random Variable: Var(X) = E((X - E(X))^2)

Variance of Random Variable is also calculated as:

Var(X) = E(X^2) - (E(X))^2

Expected value & variance formula is applicable for both discrete & continuous random variable

Follow Ayush Gandhi for more such content !!