BASICS OF PROBABILITY

RANDOM VARIABLES

Random variable **X** converts the outcomes of experiments to measurable values.

(or)

A random variable is a function that maps the outcomes of a random experiment to real numbers. It assigns a numerical value to each outcome in the sample space of the experiment, allowing us to quantify and analyze uncertainty.

For example, let's say as a data analyst at a bank, you are trying to find out which of the customers will default on their loan, i.e., stop paying their loans. Based on some data, you have been able to make the following predictions:

Customer Number	Yearly Income	Amount of Loan Due	Number of Dependants	Default Prediction (Yes/No)
1	10 lakhs	75 lakhs	3	Yes
2	15 lakhs	50 lakhs	2	No
3	20 lakhs	40 lakhs	1	No

Now, instead of processing the yes/no response, it will be much easier if you define a random variable **X** to indicate whether the customer is predicted to default or not. The values will be assigned according to the following rule:

X = 1, if the customer defaults.

X = 0, if the customer does not default.

Now, the data changes to the following:

Customer Number	Yearly Income	Amount of Loan Due	Number of Dependants	Default Prediction (Yes/No)	X (Random Variable)
1	10 lakhs	75 lakhs	3	Yes	1
2	15 lakhs	50 lakhs	2	No	0
3	20 lakhs	40 lakhs	1	No	0

Now, in this form, the table is entirely quantified, i.e., converted to numbers. Now that the data is entirely in quantitative terms, it becomes possible to perform a number of different kinds of statistical analyses on it.

Random variables are a way to mathematically represent and model uncertain situations. In many real-world scenarios, outcomes are not deterministic; they are subject to chance, randomness, or variability. Random variables provide a structured way to account for this uncertainty in our models.

<u>PROBABILITY DISTRIBUTIONS</u>

A probability distribution is a form of representation that tells the probability for all the possible values of **X.**

X	P(X)
1	1/21
2	2/21
3	3/21
4	4/21
5	5/21
6	6/21

<u>Example:</u>

Let's say that a company's management is planning to invest in a certain project. Before doing this, it wants to use probability to find whether it can safely expect to make a profit. Whether the company makes a profit or not will depend on which economic cycle is going on, i.e., recession, boom, and so on.

Based on the opinions of some experts, the following table is created:

Economic Cycle	Probability
Recession	0.1
Normal	0.7
Boom	0.2

Suppose as an analyst in the investment division, you have been asked to find the answer to the question: <u>Can the company expect to make a profit or not? Should it invest in this project?</u>

However, in this form, the table is of no help at all. Hence, let's quantify it using a random variable. Since you are interested in whether the company will profit or not, let's define X as the net revenue of the project.

Now, through some calculations, a fellow analyst of the company has arrived at the net revenue for each of these scenarios. She creates a probability distribution with this data:

X (Net Revenue of Project for the scenario, in ₹ crore)	P(X)
-305	0.1
+15	0.7
+95	0.2

Now, you finally have a probability distribution for X, the net revenue of the project. Using this probability distribution, you can find the answer to our original question: Can the company expect a profit from this project? Or should it expect a loss? However, to answer this, you will have to learn the concept of **Expected Value**, which is what we will cover in the next segment.

EXPECTED VALUE

The **expected value** for a variable **X** is the value of X that we would "**expect**" to get after performing the experiment an infinite number of times. It is also called the **expectation, average or mean value**.

Mathematically speaking, for a random variable X that can take the values x1, x2, x3, ..., xn the expected value (EV) is given by:

$$EV(X)=x1*P(X=x1) + x2*P(X=x2) + x3*P(X=x3) + ... + xn*P(X=xn)$$

Answer to above question:

EV = (-305 x 0.1) + (15 x 0.7) + (95 x 0.2)

EV =-1

So, the expected return of the project is-₹1 crore. Hence, we can conclude that the project is not worth investing in.