



Probability Distribution

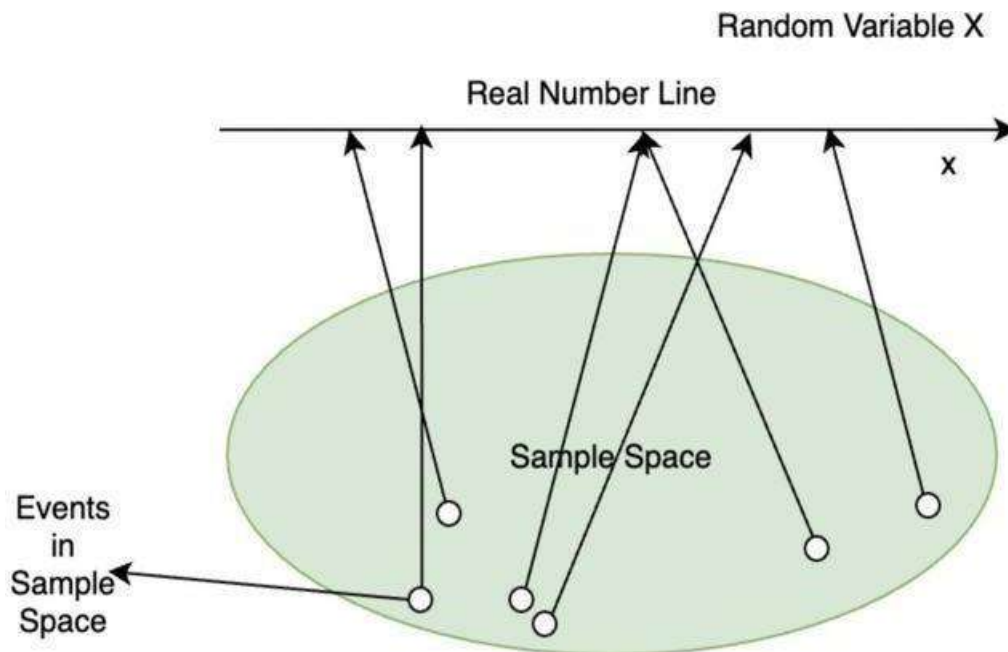
- Probability Distribution for a Random Variable shows how Probabilities are distributed over for different values of the Random Variable.
- When all values of Random Variable are aligned on a graph, the values of its probabilities generate a shape.
- The Probability distribution has several properties (example: Expected value and Variance) that can be measured.
- In Probability Distribution, A Random Variable's outcome is uncertain.
- Here, the outcome's observation is known as Realization.
- It is a Function that maps Sample Space into a Real number space, known as State Space.
- They can be Discrete or Continuous.

Random Variables:

- Random Variable is an important concept in probability and statistics.
- We need to understand it intuitively and mathematically to gain a deeper understanding of probability distributions that surround us in everyday life.
- Sometimes we are interested not only in the probabilities of the events in the experiments but also in some numbers associated with the experiment.
- Here's when we feel the need for random variables.
- Let's take an example of the coin flips. We'll start with flipping a coin and find out. We'll use H to mean 'heads' and T for 'tails'.
- So now we flip our coin 5 times, and we want to answer some questions.
- In a different scenario, suppose we are tossing two dice, and we are interested in knowing the probability of getting two numbers such that their sum is 6.
- So, in both of these cases, random variables come to our rescue. First, let's define what random variable is mathematically.

Definition

- It is an assignment of a value (number) to every possible outcome.
- In more mathematical terms, it is a function from the sample space Ω to the real numbers.
- We can choose our random variable according to our needs.



- A random variable is a real valued function whose domain is the sample space of a random experiment
- To make this more intuitive, let us consider the experiment of tossing a coin two times in succession.

The sample space of the experiment is $S = \{HH, HT, TH, TT\}$. Let's define a random variable to count events of head or tails according to our need, let X denotes the number of heads obtained. For each outcome, its values are as given below:

$$X(HH) = 2, X(HT) = 1, X(TH) = 1, X(TT) = 0.$$

More than one random variable can be defined in the same sample space. For example, let Y denote the number of heads minus the number of tails for each outcome of the above sample space S .

$$Y(HH) = 2, Y(HT) = 0, Y(TH) = 0, Y(TT) = -2$$



Thus, X and Y are two different random variables defined on the same sample.

Types of Random Variables in Probability distribution

- Discrete Random Variables
- Continuous Random Variables

Discrete Random Variables in Probability distributions:

- A discrete random variable can only take a finite number of values.

To further understand this, let's see some examples of discrete random variables:

1. $X = \{\text{sum of the outcomes when two dice are rolled}\}$. Here, X can only take values like $\{2, 3, 4, 5, 6, \dots, 10, 11, 12\}$.
2. $X = \{\text{Number of Heads in 100 coin tosses}\}$. Here, X can take only integer values from $[0, 100]$.

Continuous Random Variable in Probability distribution:

- A continuous random variable can take infinite values in a continuous domain. Let's see an example of a dart game.

Suppose, we have a dart game in which we throw a dart where the dart can fall anywhere between $[-1, 1]$ on the x-axis. So if we define our random variable as the x-coordinate of the position of the dart, X can take any value from $[-1, 1]$. There are infinitely many possible values that X can take. ($X = \{0.1, 0.001, 0.01, 1, 2, 2.112121 \dots \text{ and so on}\}$).

Probability distribution of a random variable

Now the question comes, how to describe the behavior of a random variable?

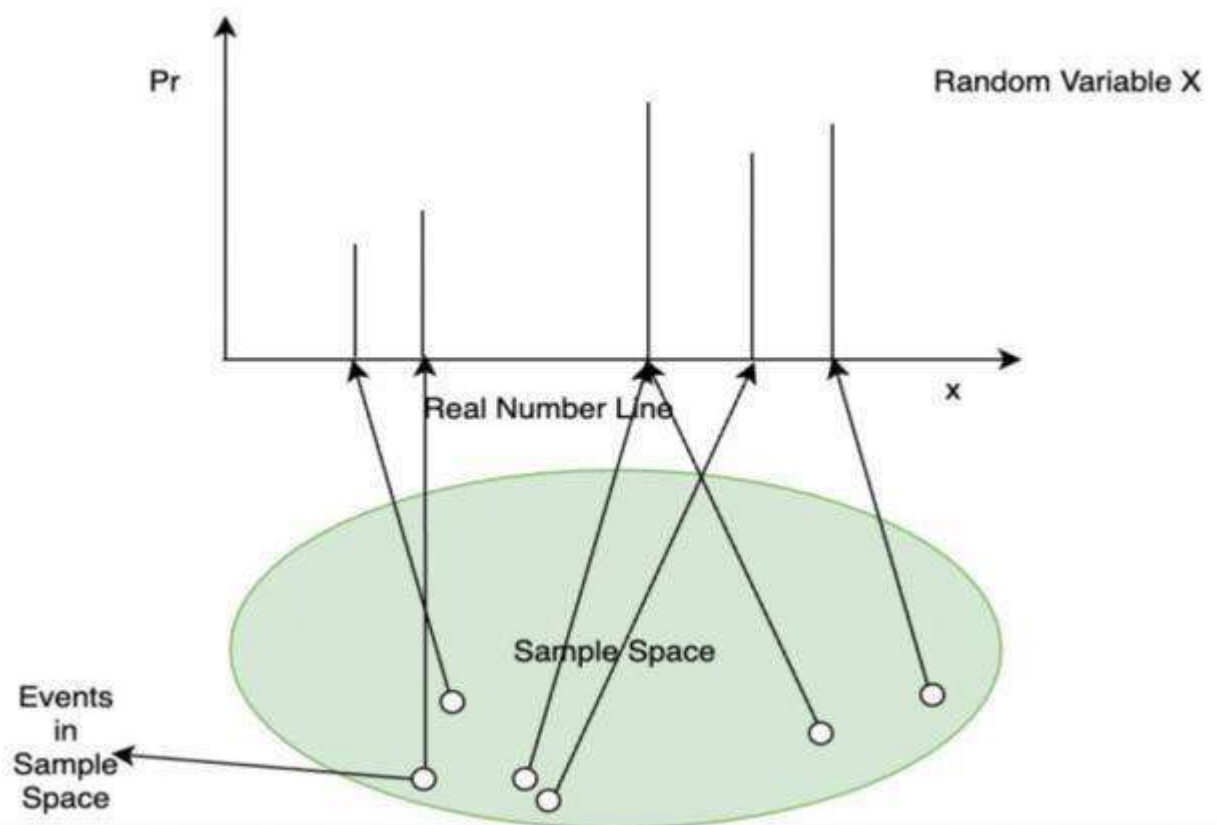
Suppose that our random variable only takes finite values, like $x_1, x_2, x_3 \dots$ and x_n . That is, the range of X is the set of n values is $\{x_1, x_2, x_3 \dots \text{ and } x_n\}$.

Thus, the behavior of X is completely described by giving probabilities for all the values of the random variable X

Event	Probability
x_1	$\Pr(X = x_1)$
x_2	$\Pr(X = x_2)$
x_3	$\Pr(X = x_3)$

The Probability Function of a discrete random variable X is the function $p(x)$ satisfying

$$p(x) = \Pr(X = x)$$



Expectation (Mean) and Variance of a Random Variable:

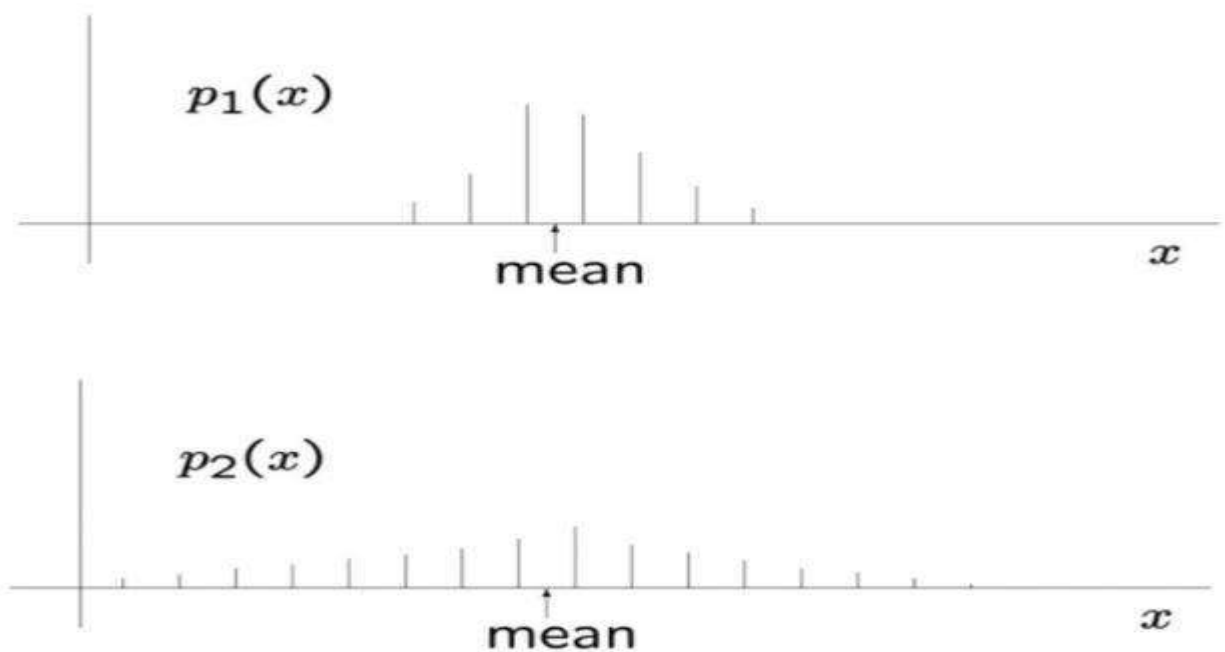
- Suppose we have a probability experiment we are performing, and we have defined some random variable(R.V.) according to our needs(like we did in some previous examples).
- Now, each time experiment is performed, our R.V. takes on a different value.

- But we want to know that if we keep on doing the experiment a thousand times or an infinite number of times, what will be the average value of the random variable?

Expectation

- The mean, expected value, or expectation of a random variable X is written as $E(X)$ or μ .
- If we observe N random values of X , then the mean of the N values will be approximately equal to $E(X)$ for large N .
- For a random variable X which takes on values $x_1, x_2, x_3 \dots x_n$ with probabilities $p_1, p_2, p_3 \dots p_n$. Expectation of X is defined as,
- i.e it is weighted average of all values which X can take, weighted by the probability of each value.

To see it more intuitively, let's take a look at this graph below,



Now in the above figure, we can see both the random variables have the almost same 'mean', but does that mean that they are equal? No. To fully describe the properties/behavior of a random variable, we need something more, right?

We need to look at the dispersion of the probability distribution, one of them is concentrated, but the other is very spread out near a single value. So we need a metric to measure the dispersion in the graph.



Variance:

- In Statistics, we have studied that the variance is a measure of the spread or scatter in data. Likewise, the variability or spread in the values of a random variable may be measured by variance.
- For a random variable X which takes on values $x_1, x_2, x_3 \dots x_n$ with probabilities $p_1, p_2, p_3 \dots p_n$ and the expectation is $E[X]$
- The variance of X or $\text{Var}(X)$ is denoted

Discrete Probability Distributions

- Discrete probability functions assume a discrete number of values.
- For example, coin tosses and counts of events are discrete functions.
- These are discrete distributions because there are no in-between values.
- We can either have heads or tails in a coin toss.
- For discrete probability distribution functions, each possible value has a non-zero probability. Moreover, probabilities of all the values of the random variables must sum to one.
- For example, the probability of rolling a specific number on a die is $1/6$.
- The total probability for all six values equals one.
- When we roll a die, we only get either one of these values.

Bernoulli trials and Binomial distributions:

- Many experiments only have either one of two outcomes.
- For example, a tossed coin shows a 'head' or 'tail', a manufactured item can be 'defective' or 'non-defective'.
- In these cases, we can call one of these outcomes "success" and another "failure".
- Let's say in the coin-tossing experiment if the occurrence of the head is considered a success, then the occurrence of the tail is a failure.
- Each time we toss a coin or roll a die or perform any other experiment, we call it a trial.
- Now we know that in our experiments coin-tossing trial, the outcome of any trial is independent of the outcome of any other trial.



- In each of such trials, the probability of success or failure remains constant. Such independent trials that have only two outcomes usually referred to as ‘success’ or ‘failure’ are called Bernoulli trials.

Binomial Distribution:

- It is a random variable that represents the number of successes in “N” successive independent trials of Bernoulli’s experiment.
- It is used in a plethora of instances like including the number of heads in “N” coin flips, and so on.
- Let P and Q denote the success and failure in the Bernoulli trial.
- Let’s Suppose we are interested in finding different ways in which we have 1 success in all six trials.

Clearly, six cases are available as listed below:

PQQQQQ, QPQQQQ, QQPQQQ, QQQPQQ, QQQQPQ, QQQQQP

Likewise, 2 success and 4 failures will show combinations. So many combinations are very difficult to list. Henceforth, calculating probabilities of 0, 1, 2, ..., n number of successes can be long and time-consuming. To avoid such lengthy calculation along with a listing of all possible cases, for probabilities of the number of successes in n-Bernoulli’s trials, a formula is made:

If Y is a Binomial random variable, we denote this $Y \sim \text{Bin}(n, p)$, where p is the probability of success in a given trial, q is the probability of failure, Let ‘n’ be the total number of trials and ‘x’ be the number of successes. A binomial random variable has the following properties:

$$P(Y) = {}^n C_x q^{n-x} p^x$$

Now the probability function P(Y) is known as the probability function of the binomial distribution.