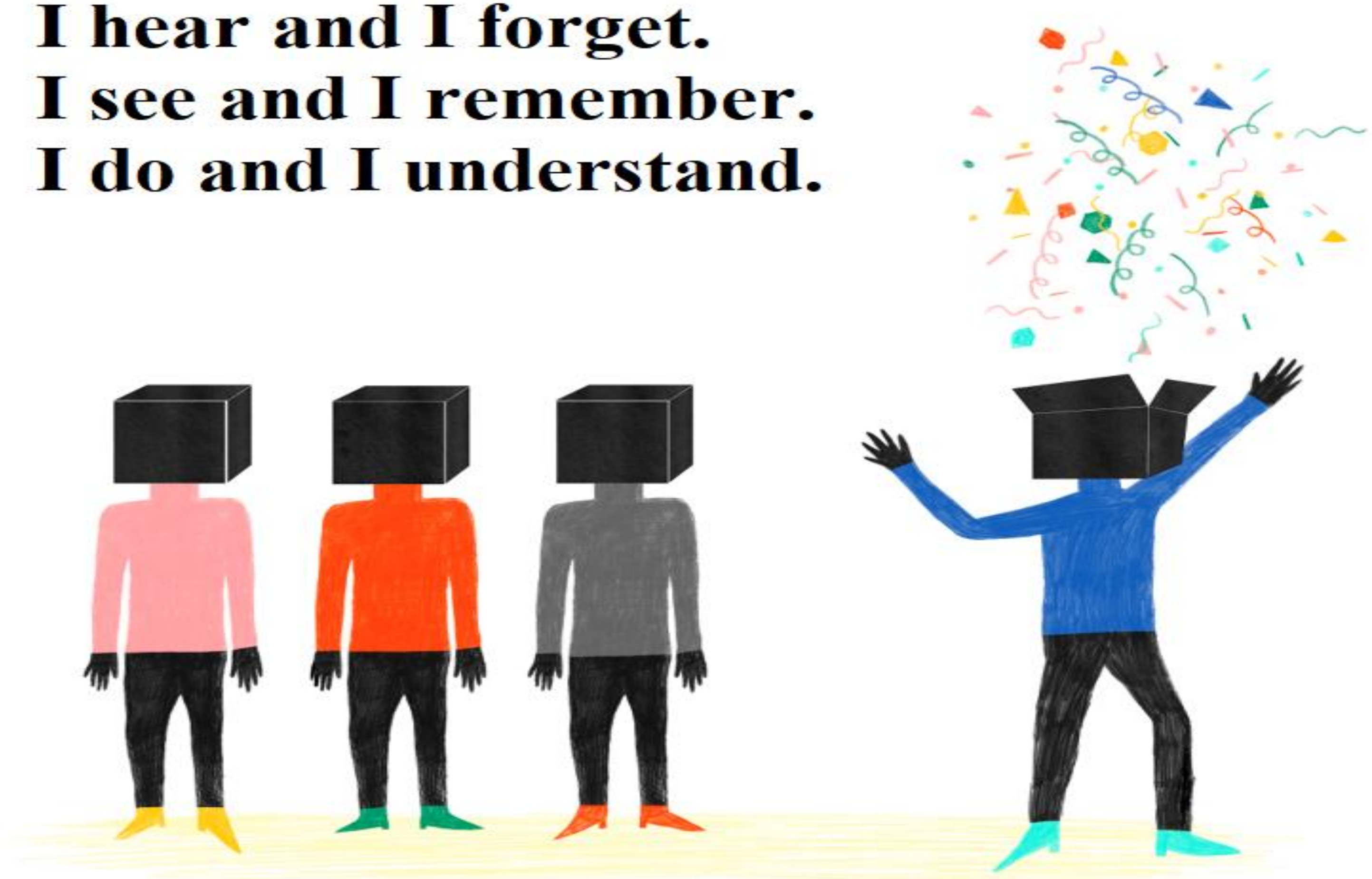
I hear and I forget.
I see and I remember.
I do and I understand.



Chandan Verma

Corporate Trainer(Machine Learning,AI,Cloud Computing,IOT)

www.facebook.com/verma.chandan.070

Random Variables

Random Variables

Random Variable is a real-valued function of the outcome of the Experiment.

A random variable is a number assigned to every outcome of an Experiment.

In probability and statistics, random variables are used to quantify outcomes of a random occurrence.

Random Variables

x=Number of Random Variable

$$X = \begin{cases} \mathbf{1 \text{ if Head}} \\ \mathbf{0 \text{ if tail}} \end{cases}$$

Suppose that a coin is tossed twice.
X is Random variable ,that represent number of head come up.

Sample Point	HH	HT	TH	TT
X	2	1	1	0

Types of Random Variables

Discrete Random Variables

A random variable is called discrete if it has either a finite or a countable number of possible values.

Continuous Random Variables

A continuous random variable is one which takes an infinite number of possible values.

Continuous random variables are usually measurements.

DATA

Discrete Data, as the name suggests, can take only specified values. For example, when you roll a die, the possible outcomes are 1, 2, 3, 4, 5 or 6 and not 1.5 or 2.45.

Continuous Data can take any value within a given range. The range may be finite or infinite. For example, A girl's weight or height, the length of the road. The weight of a girl can be any value from 54 kgs, or 54.5 kgs, or 54.5436kgs.

Random Variables

- Speed of train
- Number of students getting A grade
- Height of men in Alaska
- Error in measurement
- The duration of the next outgoing telephone call from a business office.
- The number of arrivals at an emergency room between midnight and 6:00 a.m.
- The number of applicants for a job.

Probability Density Functions

The most important way to characterize a random variable is through the probabilities of value that it can take.

The probability distribution of a random variable tells what the possible values of X are and how probabilities are assigned to those values

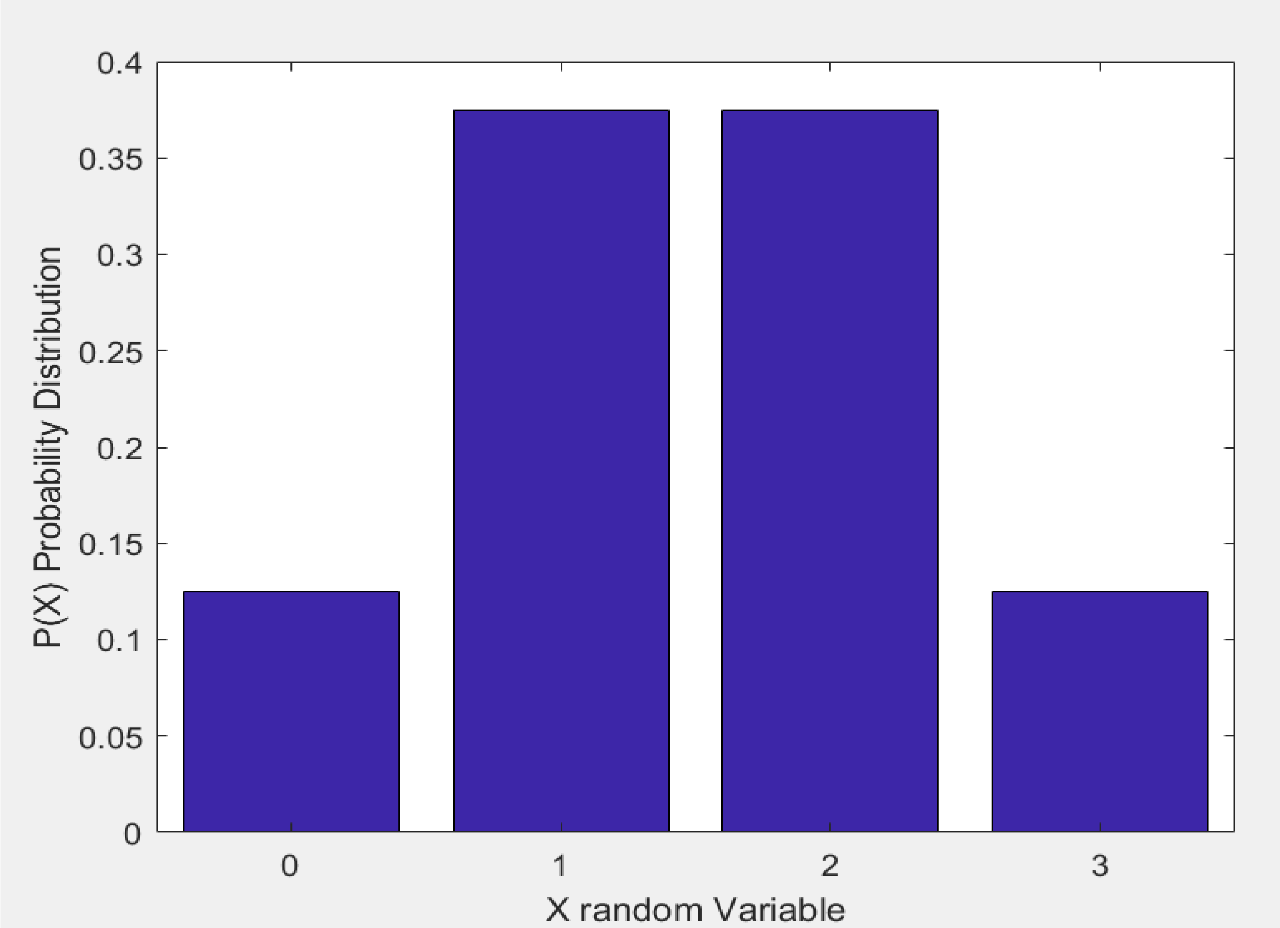
Discrete Probability Density Functions

A probability distribution is a table or an equation that links each outcome of a statistical experiment with its probability of occurrence.

Flip 3 coins at same time.
Let random variable X be the heads Showing.

HHH	3 HEAD
HHT	2 HEAD
HTH	2 HEAD
HTT	1 HEAD
THH	2 HEAD
THT	1 HEAD
TTH	1 HEAD
TTT	0 HEAD

$$\sum P(x) = 1$$



X	0	1	2	3
P(X)	1/8	3/8	3/8	1/8

Mean of a discrete random variable

The mean of a discrete random variable x is the average value that we would expect to get if the experiment is repeated a large number of times.

The mean is obtained using the formula

$$\boldsymbol{\mu = \Sigma x P(x)}$$

Another name for the mean of a discrete random variable is expected value.

$$\mathbf{E(x) = \Sigma x P(x)}$$

The mean measures where the distribution is centered.

X	1	2	3	4
P(X)	.1	.3	.2	.4

$$\mathbf{E(x)=}\ 1(0.1)+2(0.3)+3(0.2)+4(0.4)$$
$$= 0.1+0.6+0.6+1.6$$
$$\mathbf{=2.9}$$

X	0	1	2	3	4
P(X)	.30	.20	.25	.15	.10

$$\mathbf{E(x)=}\ 0 + 0.20 + 0.50 + 0.45 + 0.40$$
$$= 1.55$$

Variance of a discrete random variable

The variance of a random variable is a measure of Dispersion, or how "spread out" the data is, defined by the sum of the squared distance from the data to the mean.

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

X	2	3	4	5	6	7	8
P(X)	1/16	2/16	3/16	4/16	3/16	2/16	1/16

$E(x) = \Sigma xP(x)$

$=2/16+6/16+12/16+20/16+18/16+14/16+8/16$

$=80/16$

$=5$

$E(x^2) = \Sigma x^2P(x)$

$=4/16+18/16+48/16+100/16+108/16+98/16+64/16$

$=440/16$

$Var=440/16 -25$

$=2.5$

Discrete Distribution

A discrete probability distribution is made up of discrete variables. Specifically, if a random variable is discrete, then it will have a discrete probability distribution.

Discrete random variables give rise to discrete probability distributions. For example, the probability of obtaining a certain number x when you toss a fair die is given by the probability distribution table below.

Discrete Distribution

- Bernoulli Distribution
- Binomial distribution.
- Geometric Distribution
- Hypergeometric distribution.
- Multinomial Distribution.
- Negative binomial distribution.
- Poisson distribution.

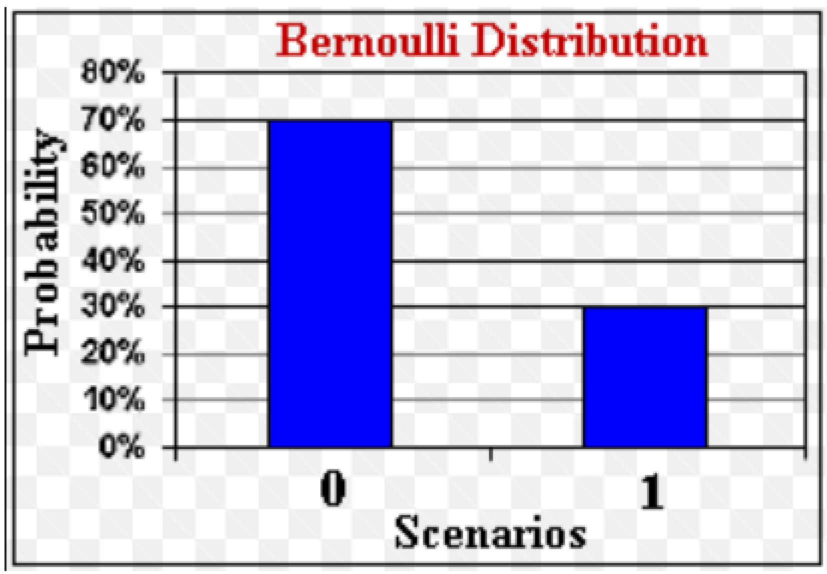
Bernoulli Distribution

The Bernoulli distribution is a model for an experiment that has only two possible outcomes. When a random variable must assume one of the two values, 0 or 1, such a variable is called a Bernoulli random variable.

Suppose a random experiment has two outcomes, namely Success and Failure. Let probability of success be p and probability of failure be $\mathbf{q = 1 - p}$. Such an experiment is called Bernoulli experiment or Bernoulli trial.

Trials of random experiment are called Bernoulli trials, if they satisfy the following conditions

1. There is finite number of trials
2. The trials should be independent.
3. Each trial has exactly two outcomes: success or failure.
4. The probability of success remains the same in each trial.



Mean,Variance

The expected value for a random variable, X, from a Bernoulli distribution is:

$$E[X] = p.$$

For example, if p = .04, then E[X] = 0.4.

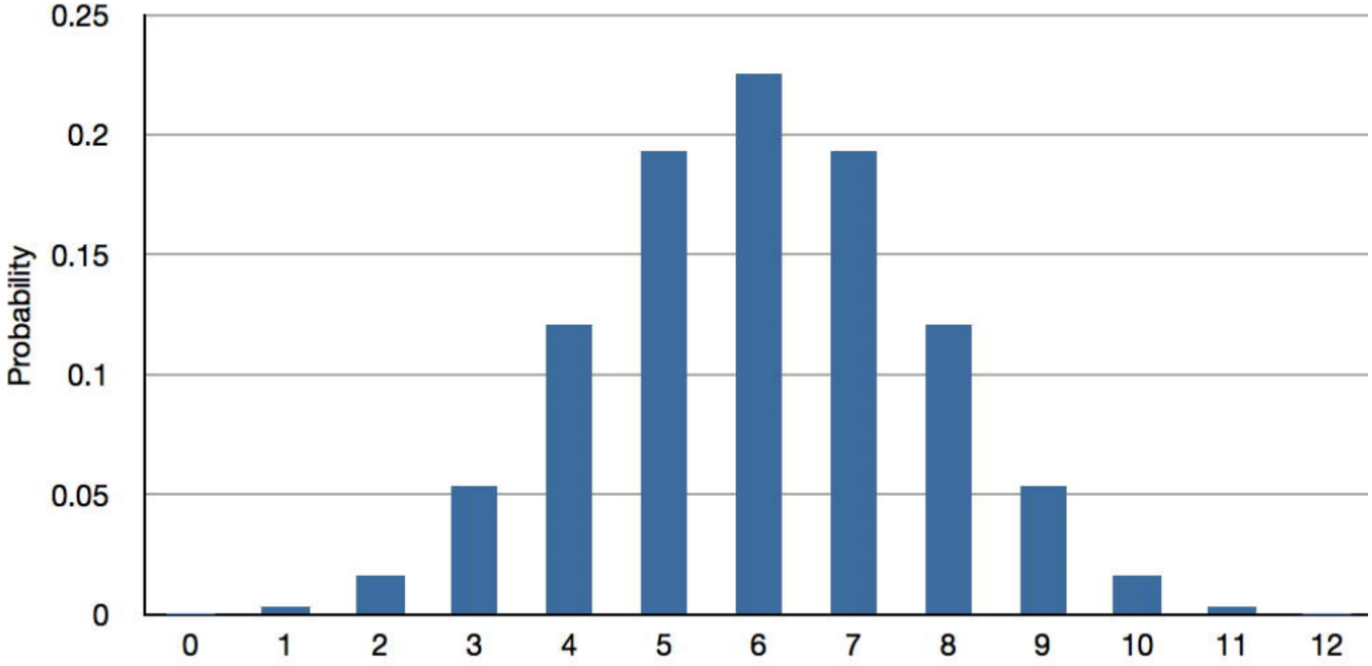
The variance of a Bernoulli random variable is:

$$Var[X] = p(1 – p).$$

Binomial distribution

A binomial distribution can be thought of as simply the probability of a SUCCESS or FAILURE outcome in an experiment or survey that is repeated multiple times. The binomial is a type of distribution that has two possible outcomes (the prefix “bi” means two, or twice).

1: The number of observations n is fixed.
2: Each observation is independent.
3: Each observation represents one of two outcomes ("success" or "failure").
4: The probability of "success" p is the same for each outcome.



a binomial random variable with parameter n=1is equivalent to a Bernoulli random variable, i.e. there is only one trial.

Binomial Distribution - Formula

$$P(B=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

n is the number of trials (sample size)
k is the number of successes;
p is s the probability of success for a single trial or the (hypothesized) population proportion.

$$\mu_X = np$$

$$\sigma_X^2 = np(1\text{-}p)$$

Poisson distribution.

Poisson distribution, in statistics, a distribution function useful for characterizing events with very low probabilities of occurrence within some definite time or space.

A Poisson distribution is a tool that helps to predict the probability of certain events from happening when you know how often the event has occurred. The Poisson distribution gives us the probability of a given number of events happening in a fixed interval of time.

1) The outcomes of the experiment can be easily classified as either success or failure.

2) The average of the number of successes within a region that is specified is known.

3) The probability of occurrence of a success is always proportional to the size of the specified region.

4) The probability of occurrence of success in a very small region is zero virtually.

It is to be noted that the region that is specified can take different forms like area, length, time period etc

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

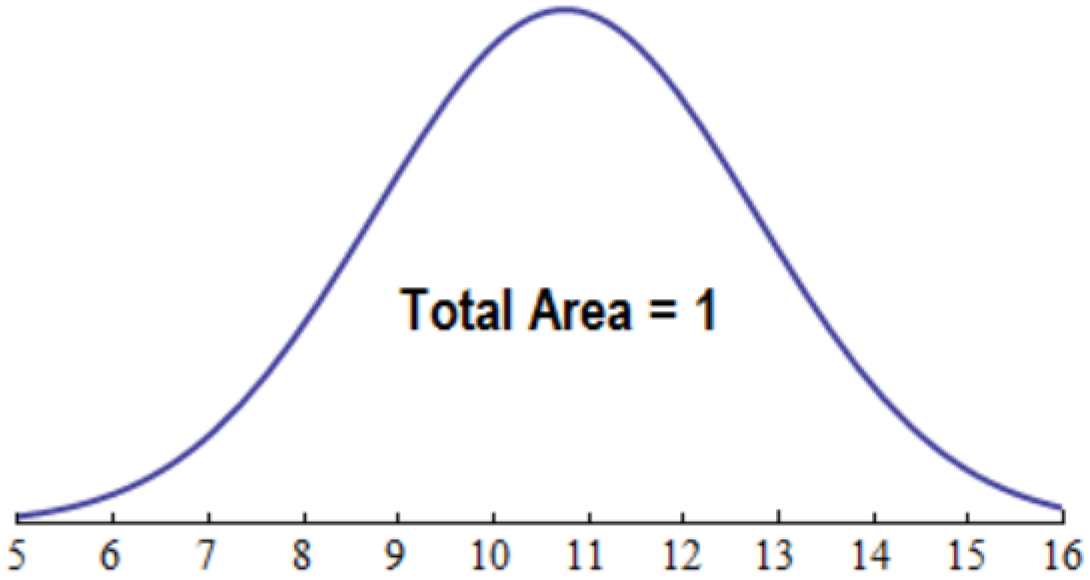
Λ is the average number of events per interval
e is the number 2.71828.
x takes values 0, 1, 2, …

Continuous Probability Distribution

If a random variable is a continuous variable , its probability distribution is called a continuous probability distribution .

$$P[a \le x \le b] = \int_a^b f(x)dx$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$



Normal Distribution

Normal distribution represents the behavior of most of the situations in the universe (That is why it’s called a “normal” distribution).

Any distribution is known as Normal distribution if it has the following characteristics:

- The mean, median and mode of the distribution coincide.
- The curve of the distribution is bell-shaped and symmetrical about the line $x=\mu$.
- The total area under the curve is 1.
- Exactly half of the values are to the left of the center and the other half to the right.

$$p(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \; e^{\frac{-(X-\mu)^2}{2\sigma^2}}$$

u is the mean or expectation of the distribution
σ is the standard deviation, and
σ 2 is the variance.

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

when $\mu = 0$
$\sigma = 1$