41. What are the different types of Probability Distribution used in Data Science?

Ans.

There are two types of probability distribution

i. Discrete probability distribution

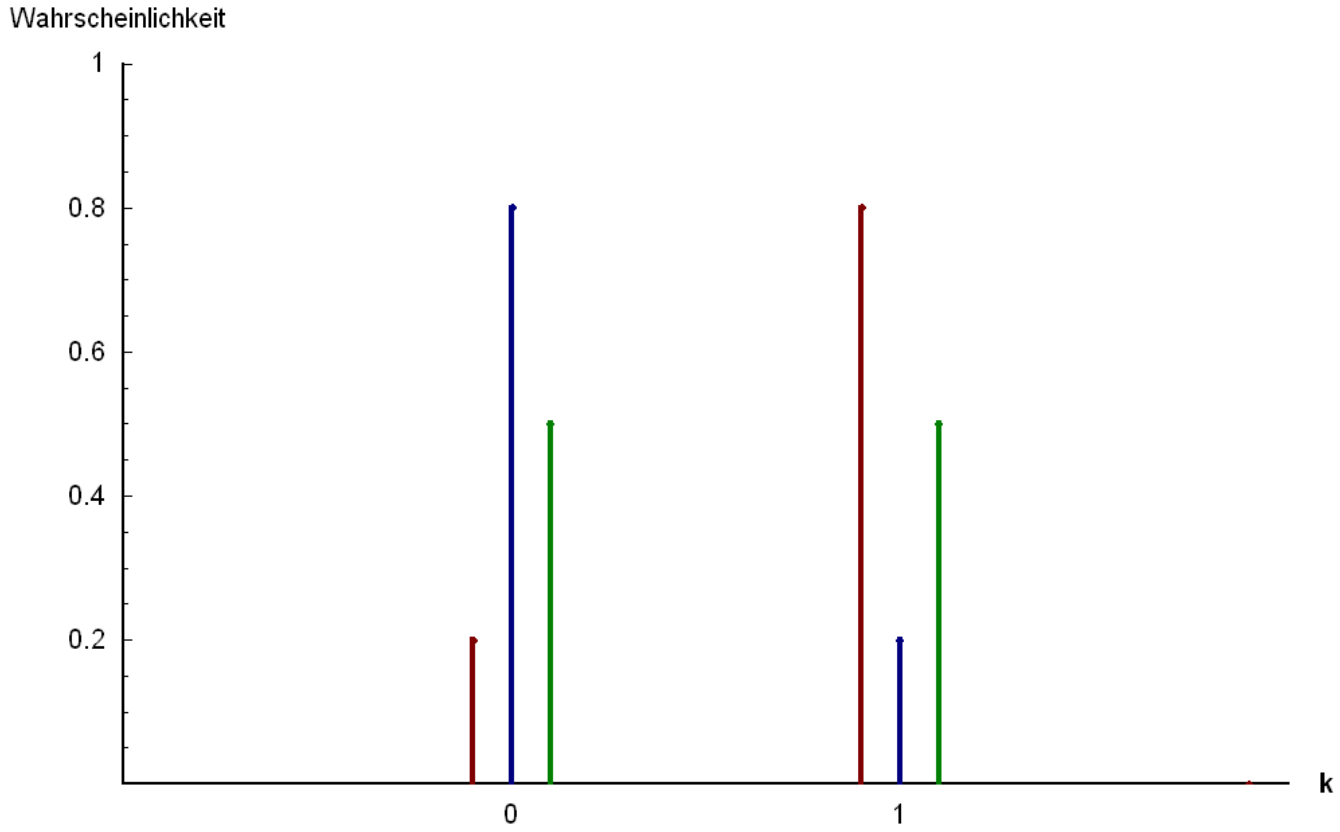
ii. Continuous probability distribution

i. Discrete probability distribution

A discrete probability distribution describes the probability of occurrence of each value of discrete random variables.

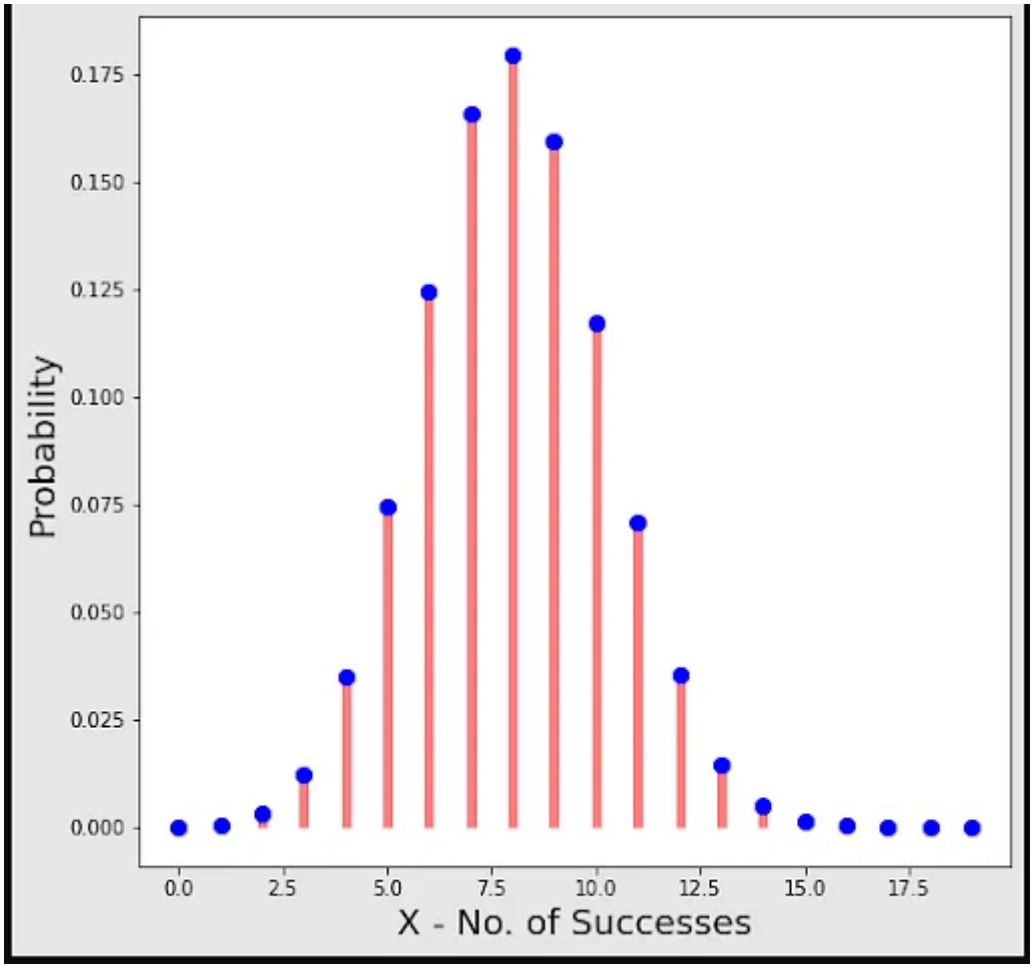
1. Bernoulli Distribution

Bernoulli distribution has two outcomes. It is also known as probability of success where only one experiment is conducted based on single observation.



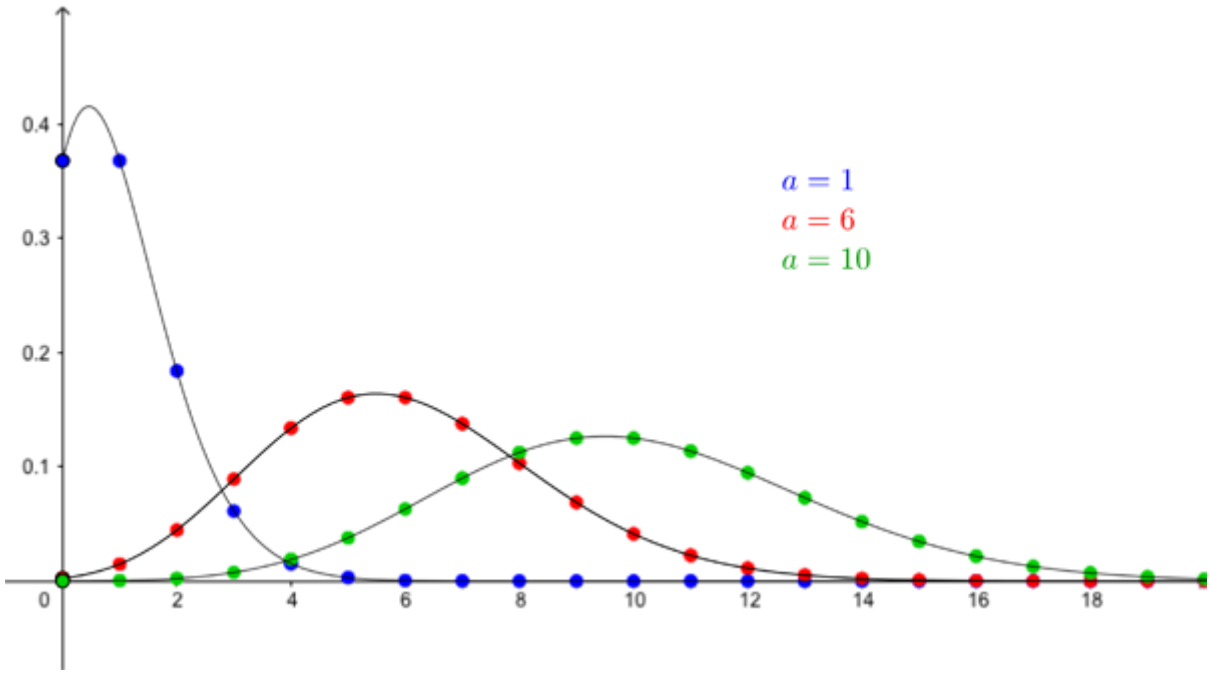
2. Binomial Distribution

Binomial distribution is the combination of finite number of observation taken from Bernoulli distribution. Where the outcome will be true and false.



3. Poisson Distribution

Poisson distribution is a probability distribution used to show how many times an event is likely to occur over a given period of time. It also known as 80-20 rule, which states that 80% of the consequences occur because of 20% of the event.

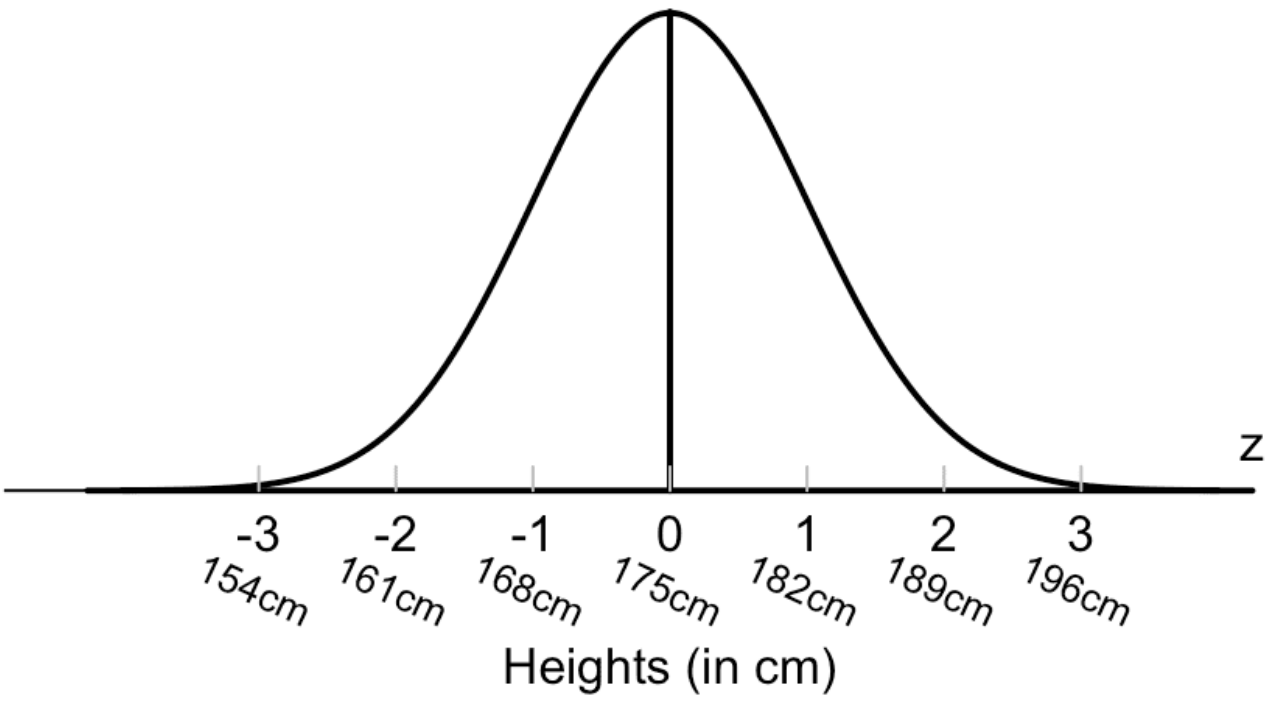


ii. Continuous Probability Distribution

Continuous probability distribution has infinite number of values. So we represent them in a range. The are under the curve used to calculate it's probability.

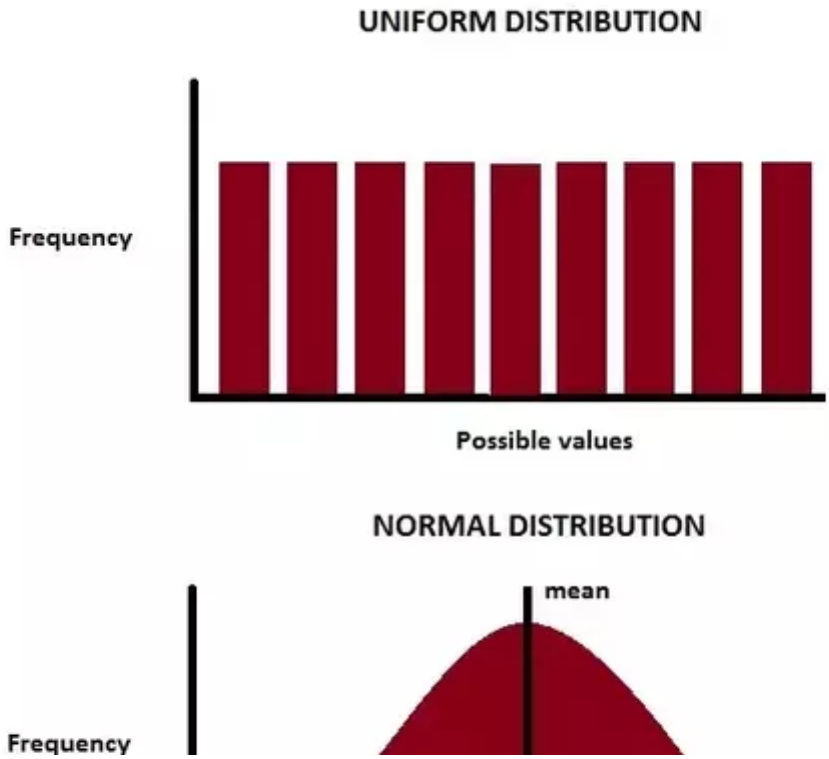
1. Normal Distribution

It is a basic gaussian distribution which is defined with mean and standard deviation. Here the distribution around the mean is symmetric. It also states data close to mean occur more frequent than away from the mean with mean value zero and standard deviation value is a finite value.



2. Continuous uniform distribution

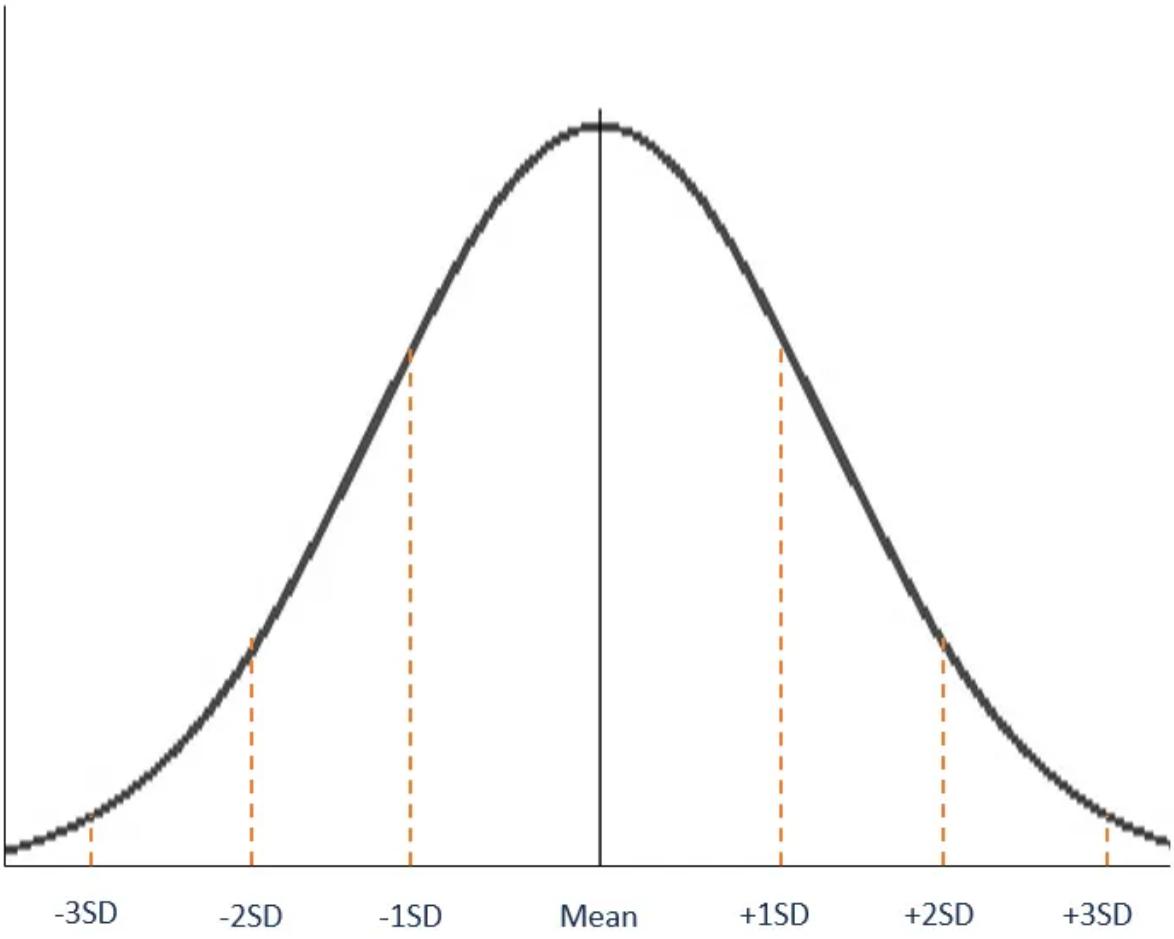
In continuous uniform distribution all outcomes are equally possible. Each variable has same chance of being hit.



42. What do you understand by the term Normal Distribution or What is a bell-curve distribution?

Ans.

Normal distribution is the probability distribution of range of values which is defined with mean and standard deviation. It states that data point are symmetric to the mean and data points nearer to the mean are more frequent than away from the mean.



43. Can you state the formula for normal distribution?

Ans

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

44. What type of data does not have a normal distribution or a Gaussian distribution?

Ans.

Normal distribution or a Gaussian distribution doesn't have discrete data.

45. What is the relationship between mean and median in a normal distribution?

Ans.

In normal distribution the mean, median, mode are equal.

46. What are some of the properties of a normal distribution?

Ans.

i. In normal distribution mean, median, mode are equal.

ii. Data points are more symmetric to the mean.

iii. The distribution describes with two values : mean and standard deviation.

47. What is the assumption of normality?

Ans.

Assumption of normality states that the distribution of sample across mean is normal.

Researcher also plot Q-Q plot to check the normality, if the plotted values vary from the straight line then it is not normally distributed.

48. How to convert normal distribution to standard normal distribution?

Ans.

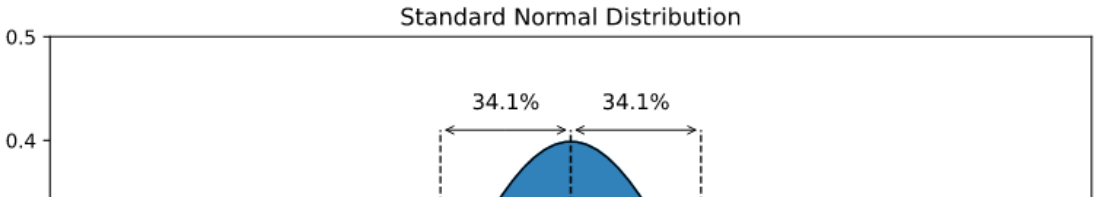
We can convert the normal distribution to standard normal distribution using z-score.

$$Z = \frac{x - \mu}{\sigma}$$

49. Can you tell me the range of the values in standard normal distribution?

Ans.

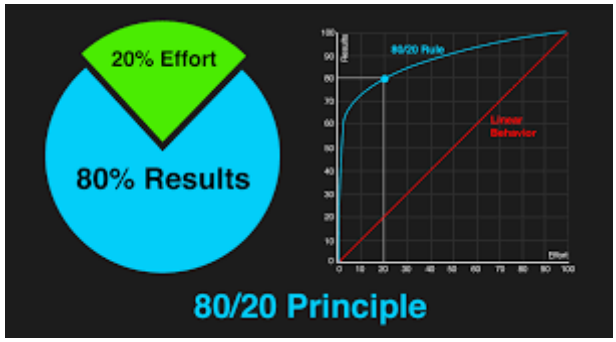
In standard normal distribution most of the data points range in between -3 to +3. Here the mean is 0 and standard deviation is 1.



50. What is the Pareto principle?

Ans.

Pereto principle is also known as 80-20 rule. Which states that 80% of the consequences are for 20% of the causes.



In [ ]:

```
1
```

Thanks

github: https://github.com/saisubhasish (https://github.com/saisubhasish)

In [ ]:

```
1
```