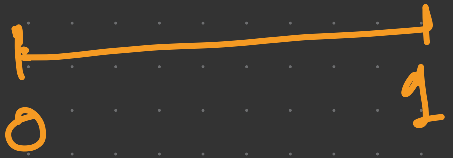
Probability Distributions

function

pmf cdf pdf

Rolling a dice

Sample = {1, 2, 3, 4, 5, 6}



1	$P(1) = 1/6$
2	$P(2) = 1/6$
3	$P(3) = 1/6$
4	$P(4) = 1/6$
5	$P(5) = 1/6$
6	$P(6) = 1/6$

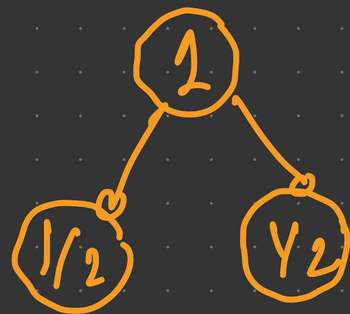
$= 6/6 = 1$

total probability = 1

$P(H) = 1/2$

$P(T) = 1/2$

$2/2 = 1$



distribution → function

- pmf
- pdf

```
def get_probability(face):
    return 1/6 [formula]

prob_1 = get_probability(1) = 1/6
```

Introduction

formula = function

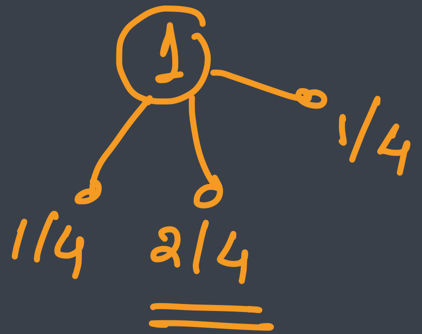
- Theoretical listing of outcomes and probabilities which can be obtained from a mathematical model representing some phenomenon of interest → event
- An empirical listing of outcomes and their observed relative frequencies
- A subjective listing of outcomes associated with their subjective probabilities representing the degree of conviction of decision maker as to the likelihood of the possible outcomes
- It is possible to deduce mathematically what frequency distributions pf certain populations should be
- Such distributions are expected on the basis of previous experience or theoretical considerations known as "theoretical distributions" or "probability distributions"
- A probability distribution for a discrete random variable is a mutually exclusive listing of all possible numerical outcomes for that random variable such that a particular probability of occurrences is associated with each observation

Example

- Probability distribution of results of rolling one fair dice

Probability function

- Probability function P(x) is used to calculate probability distributions by using different values of X
- When a random experiment is performed, the totality of outcomes of the experiement forms a set which is called as sample space (S) of the experiment
- E.g.
  - Let the random experiment be tossing a coin 2 times
  - S = [(T, T), (T, H), (H, T), (H, H)] event
  - If we replace T by 0 and H by 1 then number of heads obtained in the both the trials will be
    - (T, T) – 0
    - (T, H) – 1
    - (H, T) – 1
    - (H, H) – 2
  - The sample space can be written as [0, 1, 2]
  - P(X = 0) = P(T, T) = 1 / 4
  - P(X = 1) = P[(T, H), (H, T)] = 2 / 4 = 1 / 2
  - P(X = 2) = P(H, H) = 1 / 4



Types of distribution

- Probability Mass Function (PMF) 4, 5.5, 100
  - If the variable X is of type discrete, the probability function is called as PMF
  - The distribution is called as discrete probability distribution
  - E.g.
    - Binomial Distribution
    - Poisson Distribution
    - Negative Binomial Distribution
    - Geometric Distribution

$$\frac{\text{add all probabilites}}{\text{total probabilites}} = \underline{1}$$

- Probability Density Function (PDF) 5~15
  - If the variable X is of type continuous, the probability function is called as PDF
  - The distribution is called as continuous probability distribution
  - E.g.
    - Hypergeometric Distribution
    - Normal Distribution ****
    - Uniform Distribution
    - Exponential Distribution

area under curv is 1

Binomial Distribution

1. Each trial gives two outcomes < success / failure
2. find probality of r success in n trials
3. formula $= {}^nC_r \, q^{(n-r)} p^r$

Binomial Distribution

two outcomes

yes, 1, true, success

No, 0, false, failure

- It is also known as Bernoulli Distribution which is associated with name of a Swiss Mathematician Jacob Bernoulli
- This distribution is probability distribution expressing the probability of one set of dichotomous alternatives i.e. Success or Failure
- The binomial distribution is defined as a probability distribution related to a binomial experiment where the binomial random variable specifies how many successes or failures occurred within that sample space
- It's important for data scientists and professionals in other fields to understand this concept as binomials are used often in business applications

success failure

Bernoulli Trial

success
failure

- In the binomial experiment, the outcome of each trial in an experiment could take one of the two values which are either success or failure
- Each trial in the binomial experiment can also be termed as a Bernoulli trial
- For a single trial, binomial distribution can also be termed as Bernoulli distribution
- Examples of Bernoulli trial
  - In tossing a coin, the outcome could be either success (HEADS) or failure (TAILS)
  - In finding defective items, the outcome could be either success (item is defective) or failure (item is non-defective)
  - In rolling a die, the outcome could be either success (one of the numbers out of 1-6 (say, six-6)) or failure (any of the numbers except) otherwise
- The outcome of interest in a trial of an experiment is often termed as a **success**
- The binomial random variable could be the number of successes in an experiment consisting of N trials

Assumptions

- An experiment is performed under same conditions for fixed number of trials (n)
- In each trial there are only two possible outcomes: success (1) and failure (0)
  - Sample space = [Success, Failure]
- The probability of a success denoted by p remains constant from trial to trial
- The probability of a failure denoted by q is equal to 1 – p
- The trials are statistically independent

① toss a coin = [H, T]

$p(H) = p(T) = 0.5$

② Rolling a dice = {1, .. 6}

$p(1) = p(2) .. = p(6) = 1/6$

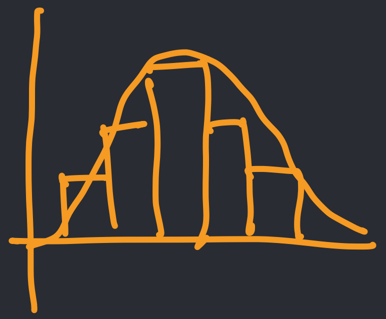
Binomial Distribution

```python
from scipy.stats import binom
import matplotlib.pyplot as plt
import numpy as np

X = np.arange(0,21)
p = 0.6
n = 20

binomial_distribution = binom.pmf(X, n, p)

plt.plot(X, binomial_distribution, 'o', label='binomial pmf')
plt.ylabel("Probability", fontsize="18")
plt.xlabel("X - No. of Successes", fontsize="18")
plt.title("Binomial Distribution - No. of Successes Vs Probability", fontsize="18")
plt.vlines(X, 0, binomial_distribution, colors='b', lw=5, alpha=0.5)
plt.show()
```



Binomial Distribution Function

$$P(r) = {}^{n}_{r}C\, q^{n-r} p^{r}$$

- Where
  - p = probability of success in a single trial
  - q = probability of failure in a single trial (1 – p)
  - n = number of trials
  - r = number of success in n trials

Example 1

$$^6C_4 = \frac{n!}{r!(n-r)!} = \frac{6!}{4! \times 2!} = \frac{\overset{3}{\cancel{6}} \times 5 \times \cancel{4!}}{\cancel{4!} \times \cancel{2}!} = 15$$

- A coin is tossed six times. What is the probability of obtaining four or more heads ?

$n = 6$, success = $\underbrace{\text{getting 4 heads}}_{1}$, $\underbrace{\text{5 heads}}_{2}$, $\underbrace{\text{6 heads}}_{3}$

$p(H) = 0.5$, $q = 1 - p = 0.5$

$$p(4\text{ heads}) = {}^nC_r\, q^{(n-r)} p^r = {}^6C_4\, q^{(6-4)} p^4 = 15 \times (0.5)^2 \times (0.5)^4$$

$$p(4\text{ heads}) = 0.234$$

$$p(5\text{ heads}) = {}^6C_5\, q^{6-5} p^5 = 6 \times 0.5^1 \times 0.5^5 = \underline{0.09}$$

$$p(6\text{ heads}) = {}^6C_6\, q^0 p^6 = 1 \times 1 \times 0.5^6 = \underline{0.01}$$

$$p(4 \text{ or } 5 \text{ or } 6 \text{ heads}) = p(4H) + p(5H) + p(6H) = \underline{\underline{0.343}}$$

Example 2

- The incidence of a certain disease is such that on the average 20% of workers suffer from it. If 10 workers are selected at random, find the probability that
  - Exactly 2 workers suffer from the disease
  - Not more than 2 workers suffer from the disease

$n=10$, success = worker is diseased

$p(\text{success}) = 0.2$, $q = 1-p = 0.8$

① $p(2 \text{ diseased workers}) = {}^{n}C_{r}\, q^{(n-r)}\, p^{r} = {}^{10}C_{2}\, 0.8^{8} \times 0.2^{2} = 45 \times 0.167 \times 0.04$

$= \underline{\underline{0.304}}$

$${}^{10}C_{2} = \frac{10!}{2!\,8!} = \frac{\cancel{10}^{5} \times 9}{\cancel{2}} = 45$$

② $p(0 \text{ or } 1 \text{ or two diseased workers}) = p(0) + p(1) + p(2)$

$= \underline{\underline{0.67}}$

Poisson Distribution

Introduction

- It is a discrete probability distribution and is widely used in statistical calculations
- It was developed by ~~Fresh~~ French mathematician Simeon Denis Poisson
- It can be expected in cases where the chance of any individual event being a success is small
- The distribution is used to describe the behavior of rare events such as number of printing mistakes in a book, number of accidents on road etc.
- It is also called as Law of improbable events

Poisson Distribution

$$P(r) = \frac{e^{-m} m^{r}}{r!}$$

- Where
  - r = 0, 1, 2, 3, 4 .....
  - e = 2.7183
  - m – mean of Poisson distribution

Role of Poisson Distribution

- Used in quality control statistics to count the number of defects of an item
- In biology to count the number of bacteria
- In physics to count the number of particles emitted from radioactive substance
- In insurance problems to count number of causalities
- In wait time problem to count number of incoming telephone calls on incoming customers
- In general Poisson distribution explains the behavior of those discrete variables where the probability of occurrence of the event is small and number of possible cases is sufficiently large

Example

- Ten per cent of the tools produced in a certain manufacturing process turn out to be defective. Find the probability that in a sample of 10 tools chosen, exactly two will be defective. (use m = np for calculating mean)

$n = 10$, success = item is defective $p = 0.1$, $q = 0.9$, $m = np = 1$

Binomial

$$P(2 \text{ defective}) = {}^{n}C_{r}\, q^{(n-r)} p^{r}$$

$$= {}^{10}C_{2}\, q^{8} p^{2} = 45 \times 0.9^{8} \times 0.1^{2}$$

$$= \underline{\underline{0.193}}$$

Poisson

$$P(2 \text{ defective}) = \frac{e^{-m} m^{r}}{r!} = \cancel{\frac{2.71^{-1} \times 1^{2}}{2!}}$$

$$= \frac{1}{2.71 \times 2}$$

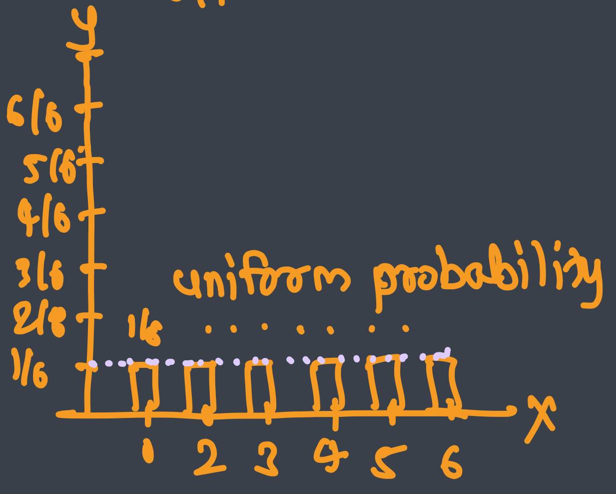
$$= \frac{0.271 \times 1}{2 \times 1} = \underline{\underline{0.18}}$$

Geometric distribution

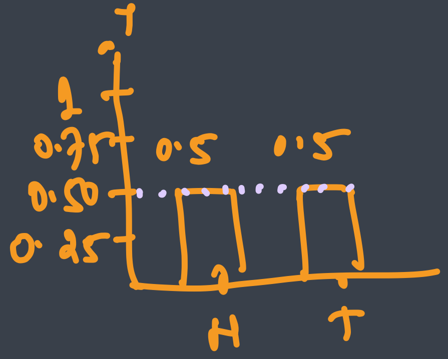
- Is a distribution of the number of trials needed to get the first success in the repeated Bernoulli trials
- Every trial results in one of two possible outcomes
  - Success
  - Failure
- P(Success) = p (stays constant)
- P(Failure) = 1 – p
- X represents the number of trails needed to get the first success
- For the first success to occur on the xth trial:
  - The first x – 1 trials must be failures
  - The xth trial must be a success
- Probability mass function

$$P(x) = (1 - p)^{x-1} * p$$

Roll a dice = $\{1, 2, 3, 4, 5, 6\}$



toss a coin = $\{H, T\}$



probability

Uniform Distribution

pmf

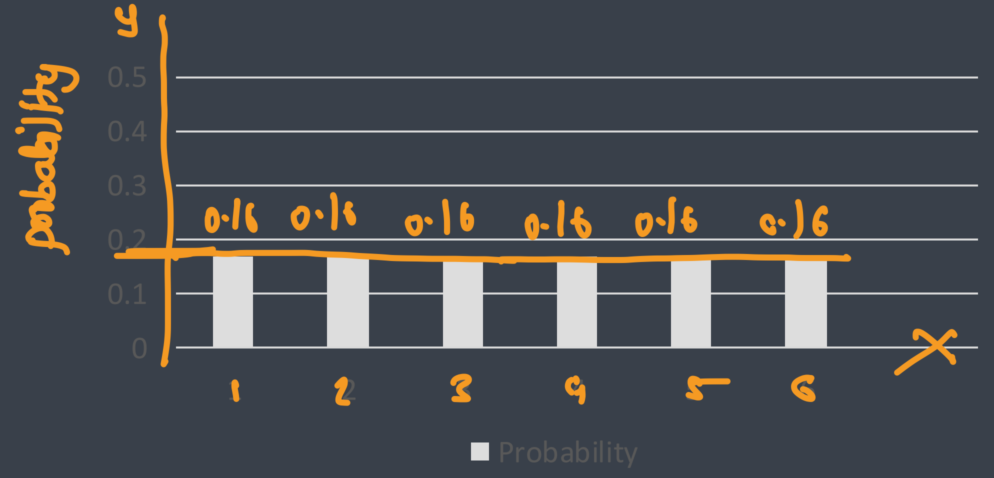
(discrete)

pdf

(continuous)

Uniform distribution

- In statistics, uniform distribution refers to a type of probability distribution in which all outcomes are equally likely → every outcome will get same probability
- A deck of cards has within it uniform distributions because the likelihood of drawing a heart, a club, a diamond, or a spade is equally likely
- A coin also has a uniform distribution because the probability of getting either heads or tails in a coin toss is the same
- The uniform distribution can be visualized as a straight horizontal line, so for a coin flip returning a head or tail, both have a probability p = 0.50 and would be depicted by a line from the y-axis at 0.50



Key Takeaways

every outcome must have same probability

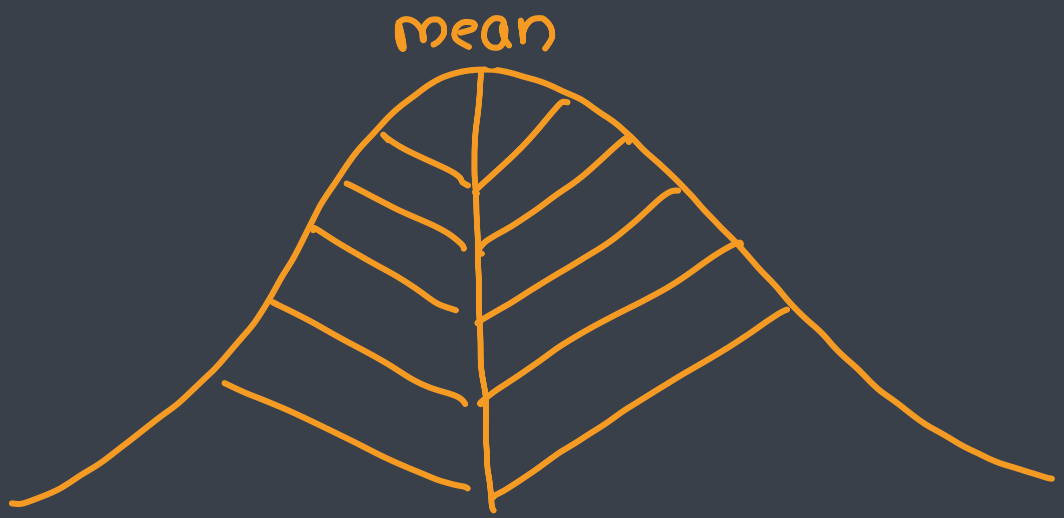
- Uniform distributions are probability distributions with equally likely outcomes
- In a discrete uniform distribution, outcomes are discrete and have the same probability
- In a continuous uniform distribution, outcomes are continuous and infinite 5-15
- In a normal distribution, data around the mean occur more frequently
- The frequency of occurrence decreases the farther you are from the mean in a normal distribution

Types

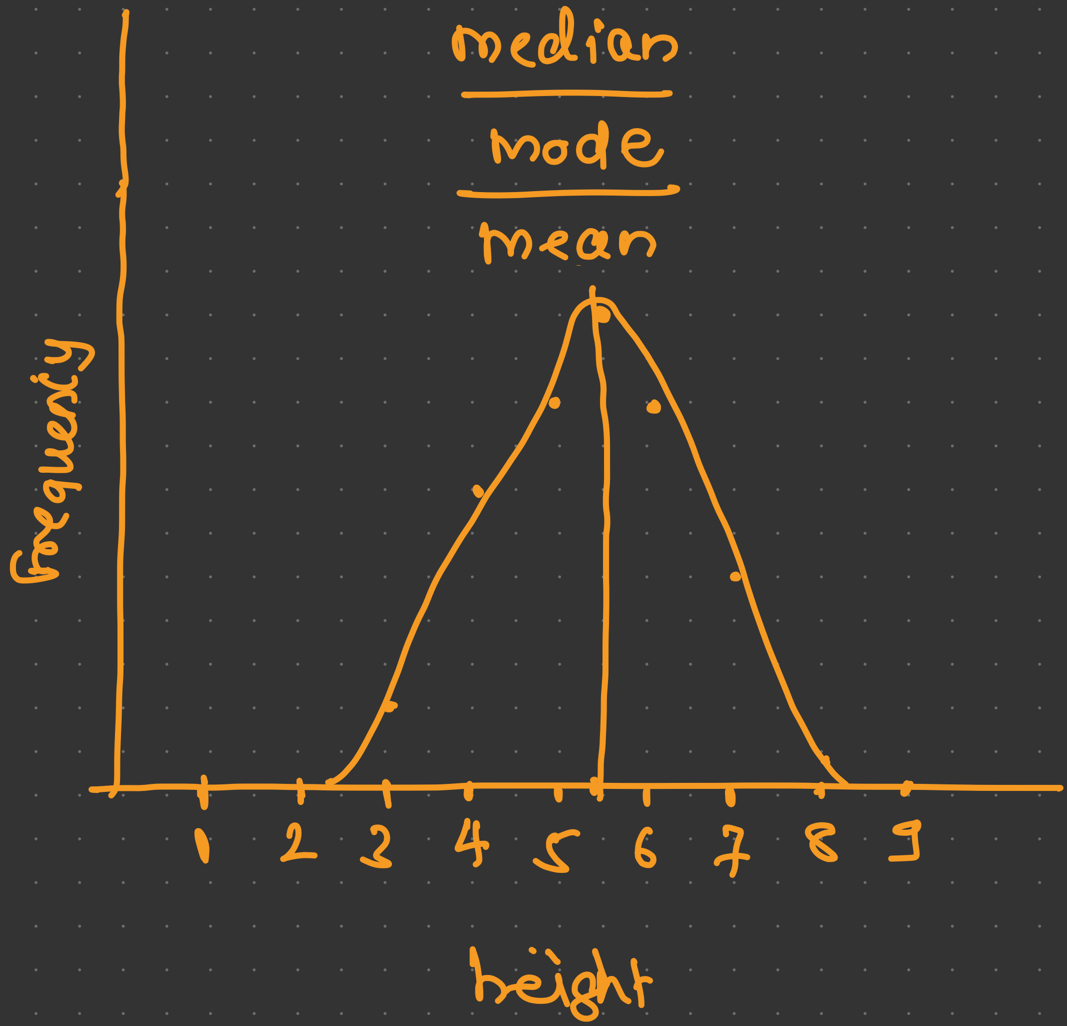
- There are two types of uniform distributions: **discrete** and **continuous**
- **Discrete Uniform Distribution**
  - The possible results of rolling a die provide an example of a discrete uniform distribution
  - it is possible to roll a 1, 2, 3, 4, 5, or 6, but it is not possible to roll a 2.3, 4.7, or 5.5
  - Therefore, the roll of a die generates a discrete distribution with $p = 1/6$ for each outcome
  - There are only 6 possible values to return and nothing in between
- **Continuous Uniform Distribution**
  - Some uniform distributions are continuous rather than discrete
  - An idealized random number generator would be considered a continuous uniform distribution
  - With this type of distribution, every point in the continuous range between 0.0 and 1.0 has an equal opportunity of appearing, yet there is an infinite number of points between 0.0 and 1.0

Examples

- **Guessing a birthday**
  - If you randomly approach a person and try to guess his/her birthday, the probability of his/her birthday falling exactly on the date you have guessed follows a uniform distribution
  - This is because every day of the year has equal chances of being his/her birthday or every day of the year is equally likely to be his/her birthday
- **Deck of cards**
  - The total number of cards present in the deck of playing cards is equal to 52
  - If you select a card randomly from a fair deck of playing cards, the probability that the drawn card would be either a diamond, spade, heart, or club follows a uniform distribution because the probability of choosing a spade is equal to 0.25, which is same as the probability of choosing a diamond, heart, or a club card
- **Lucky Draw Contest**
  - The probability of a person winning a lucky draw contest is equal for every other person participating in the contest
  - Hence, such a distribution is known as the uniform probability distribution because the winning chances of every person are equal



Normal Distribution



$$\frac{1000}{5.5}$$

Introduction

class

bell-shaped

- It is also called as Normal probability distribution
- It happens to be the most useful theoretical distribution for continuous variable
- Many statistical data concerning business and economic problems are displayed in the form of normal distribution. In fact, it is the cornerstone of modern statistics
- It was first described by English mathematician Abraham De Moivre as limiting form of Binomial
- It is an approximation to Binomial distribution

Normal Distribution Curve

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{1}{3}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- Where
  - μ = mean of the normal distribution
  - σ = standard deviation
  - $\sqrt{2\pi}$ = 2.5066
  - e = 2.7183

Importance of Normal Distribution

- It has the remarkable property stated in Central Limit Theorem, according to which as sample size n increases, the distribution of mean of random sample taken from particularly any population approaches a normal distribution
- As n becomes large, the normal distribution serves as a good approximation of many discrete distribution like Binomial or Poisson
- In theoretical statistics, many problems can be solved only under the assumption of a normal distribution
- It has many mathematical properties which makes it popular and comparatively easy to understand
- It is extensively used in statistical quality control industry in setting up control limits

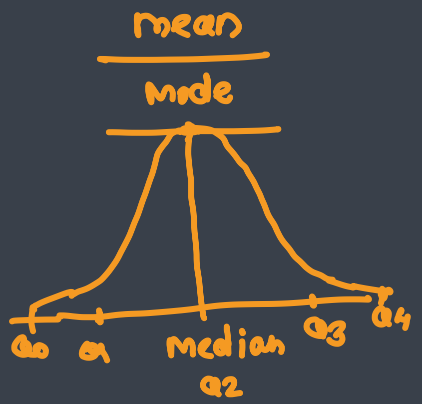
Properties of Normal Distribution

- Normal curve is bell-shaped and symmetrical in it appearance
- The height of normal curve is at its maximum at the mean, hence the mean and mode of normal distribution coincide

mean = median = mode

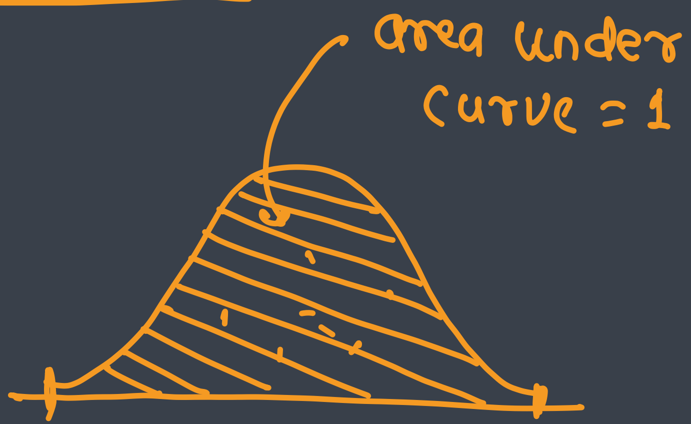
peak

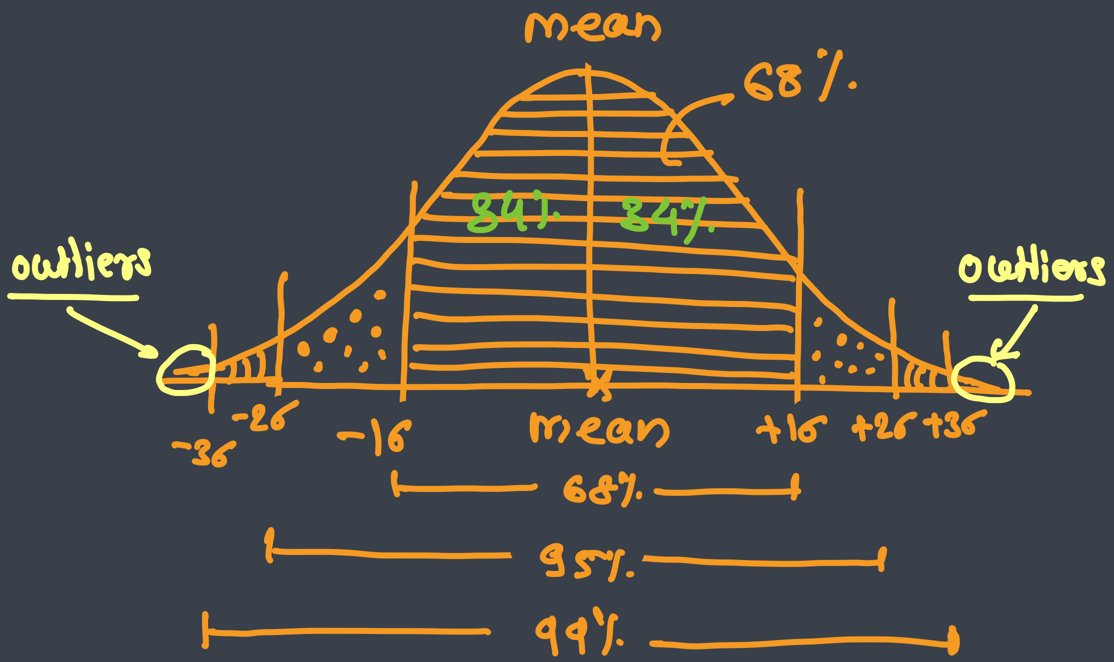
- There is one maximum point of the normal curve which occurs at the mean
- The height of the curve declines as we go in eight direction from the mean
- Since there is only one maximum point, the normal curve is unimodal
- As distinguished from Binomial and Poisson, the normal curve is continuous
- The first and third quartiles are equidistant from mean → $Q2 - Q1 = Q3 - Q2$
- The mean deviation is $4/5^{th}$ of standard deviation



Area under normal curve

- Equation of normal curve gives the ordinate of the curve corresponding any of x
- We are usually interested in area under the normal curve instead of its ordinate
- The are under the curve gives us the proportion of the case falling between two numbers or the probability of getting a value between two numbers
- The total area under the curve is always summed to 1
- The area under normal curve is distributed as follows
  - Mean $\pm 1\sigma$ covers 68.27% area (34.135% on either side)
  - Mean $\pm 2\sigma$ covers 95.45% area
  - Mean $\pm 3\sigma$ covers 99.73% area



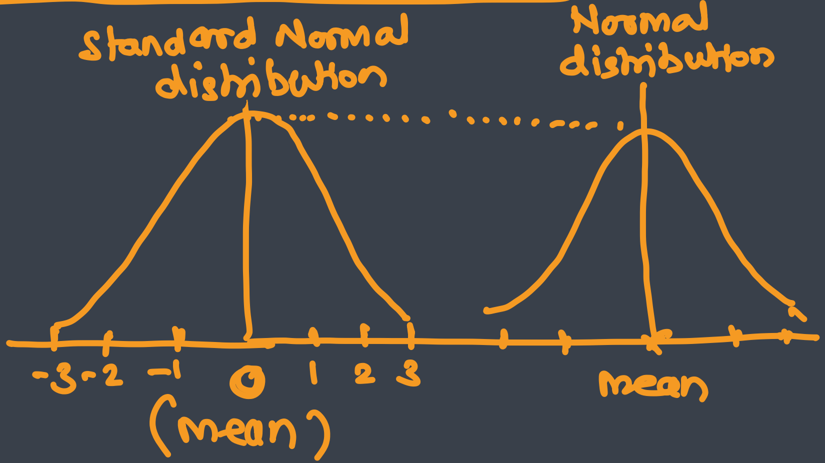


Standard normal distribution

- The equation of normal curve depends on mean and standard deviation and for different values of mean and standard deviation we will obtain different curves
- This would necessitate separate tables of normal curve areas for each pair of population mean and sample mean and an infinite number of tables would be required
- Fortunately this problem is solved by standardizing the data and only one table will be needed
- We will be able to determine normal curve areas regardless of population mean and sample mean by tabulating only the area under normal curve having mean = 0 and standard deviation = 1
- Such a curve with zero mean and unit standard deviation is known as standard normal curve
- A normal curve with mean $\bar{X}$ and standard deviation σ can be converted to standard normal distribution by performing change of the scale

[Standard Scaler]

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



Central Limit Theorem

- The mean values from a group of samples will be normally distributed about the population mean, even if the population itself is not normally distributed
- That is, 95% of all sample means should fall within 2σ of the population mean