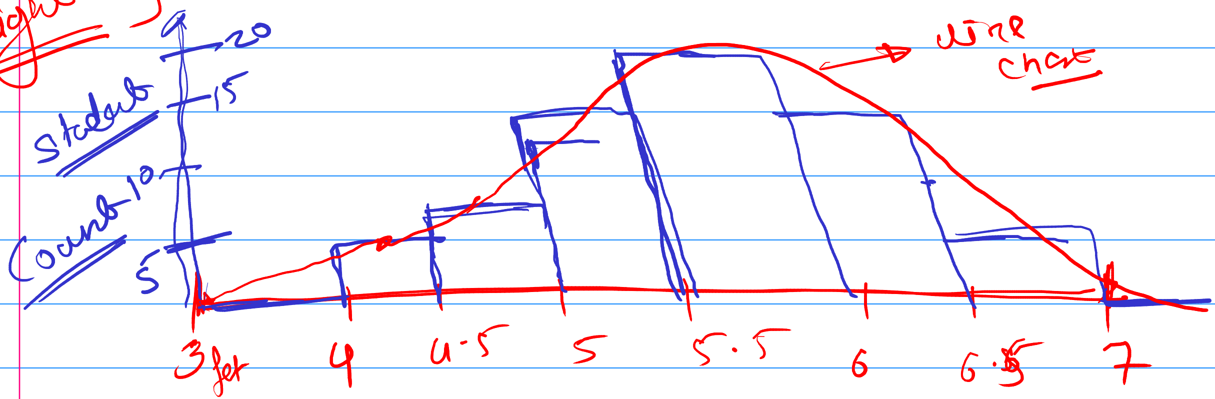
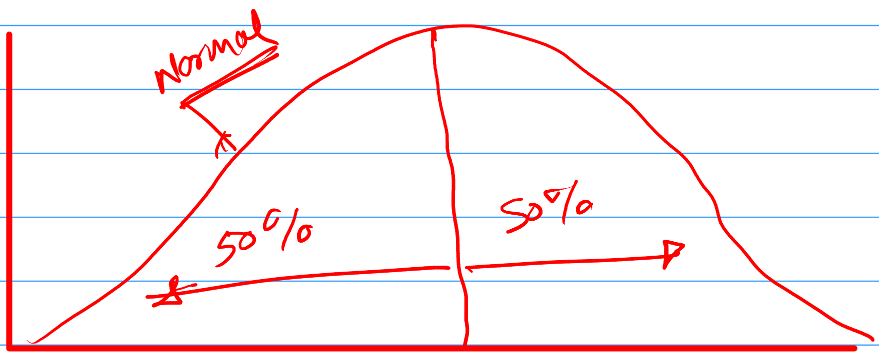
Height } BCA = 6 feet — 4 feet — 5 feet
Weight } map 2

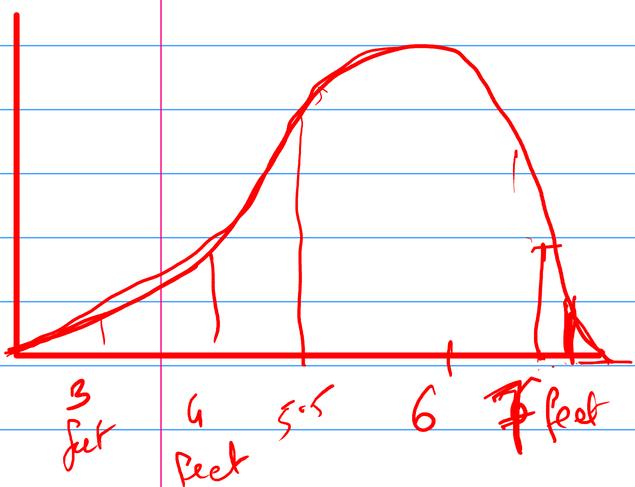


KDE - kernel density Estimator.

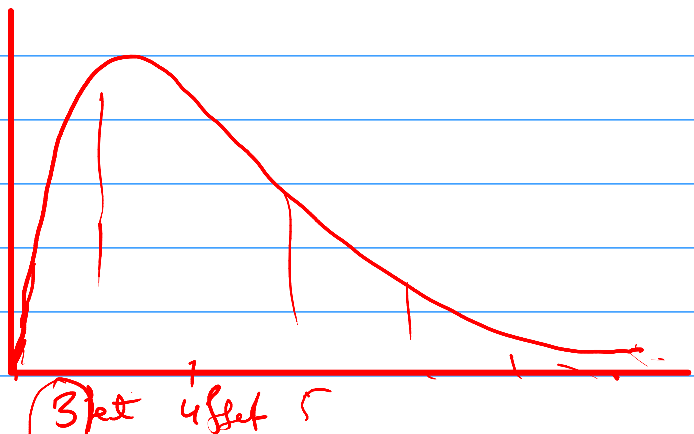
KDE = TRUE



mean
median
mode



Left skew



Outlier

Bernoulli Distribution: -

whole no

- Bernoulli distribution is a discrete probability distribution
- it's concerned with discrete random variables {PMF}
- Bernoulli distribution applies to events that have **one trial** and **two possible outcomes**. These are known as Bernoulli trials.

E.g.: -

- Tossing a coin {H,T}

  P_r(H)=0.5 = p

  P_r(T)=0.5 = 1-p=q

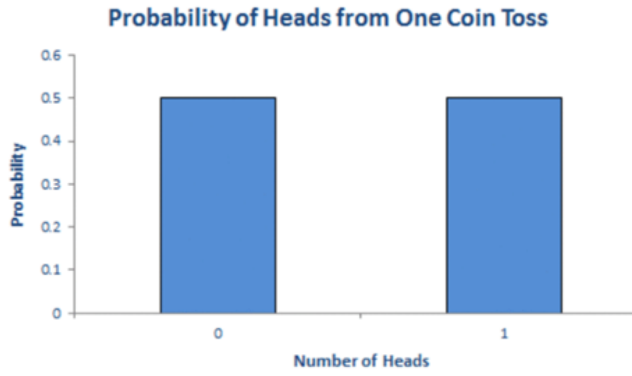
$P(H)=P=0.5$

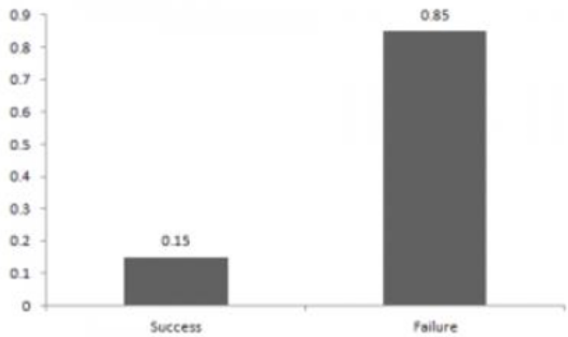
$P(T)=1-P$
$=1-0.5$
$=0.5$

- Whether the person will Pass/Fail

  P_r(Pass)=0.85 = p

  P_r(Fail)= 1-p = 0.15 = q



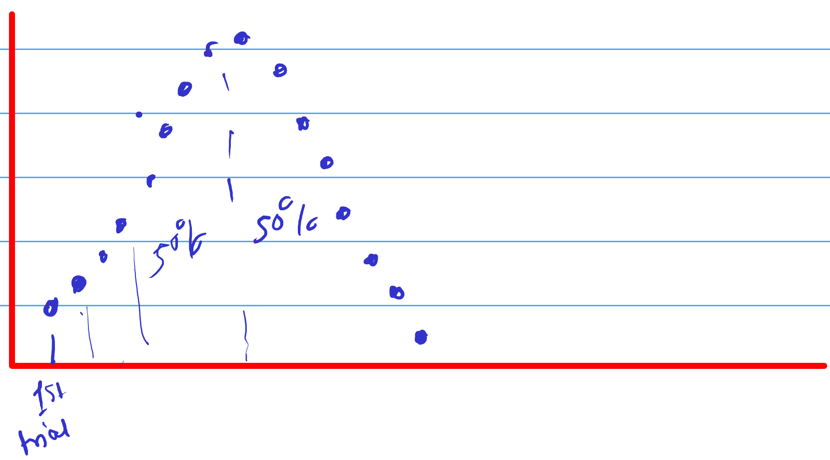


Binomial Distribution: -

- it's concerned with discrete random variables {PMF}
- There are two possible outcomes: true or false, success or failure, yes or no.
- These Experiments is Performs for n trials
- Every trial is an independent trial, which means the outcome of one trial does not affect the outcome of another trial.

E.g.: -

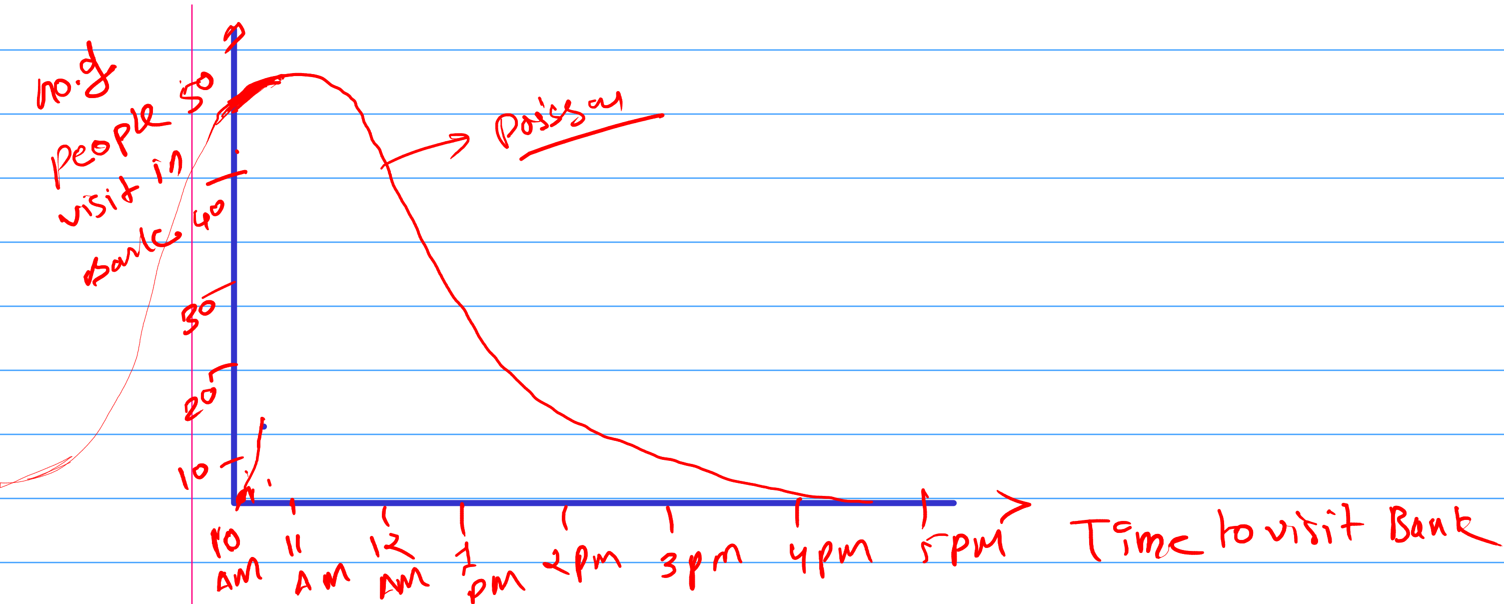
Tossing a Coin 10 times



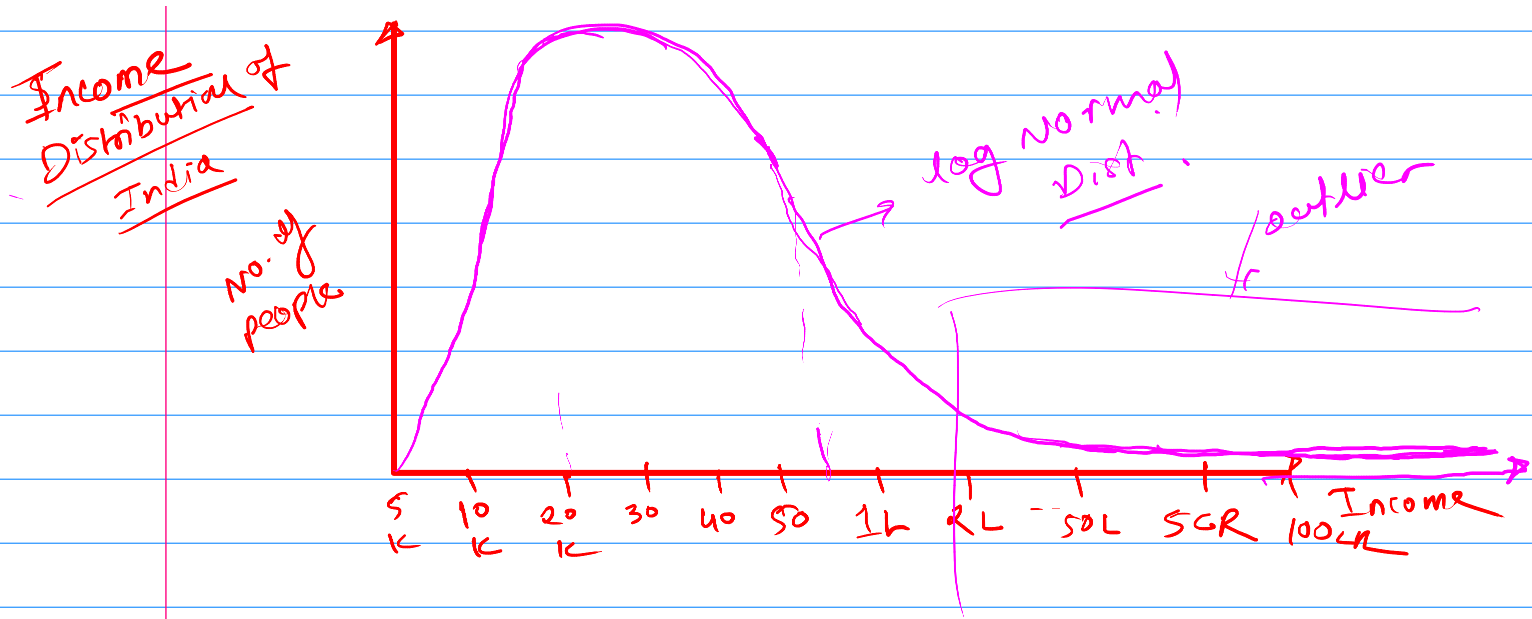
Poisson Distribution: -

- it's concerned with <u>discrete random variables {PMF}</u>
- Describe the number of events occurring in a fixed time interval

E.g.: - No. of people visiting hospital every hour
No. of people visiting bank at 11am



Log-Normal Distribution: - A log-normal distribution is a continuous distribution of random variable y whose natural logarithm is normally distributed. For example, if random variable $y = \exp\{y\}$ has log-normal distribution then $x = \log(y)$ has normal distribution.



Uniform Distribution

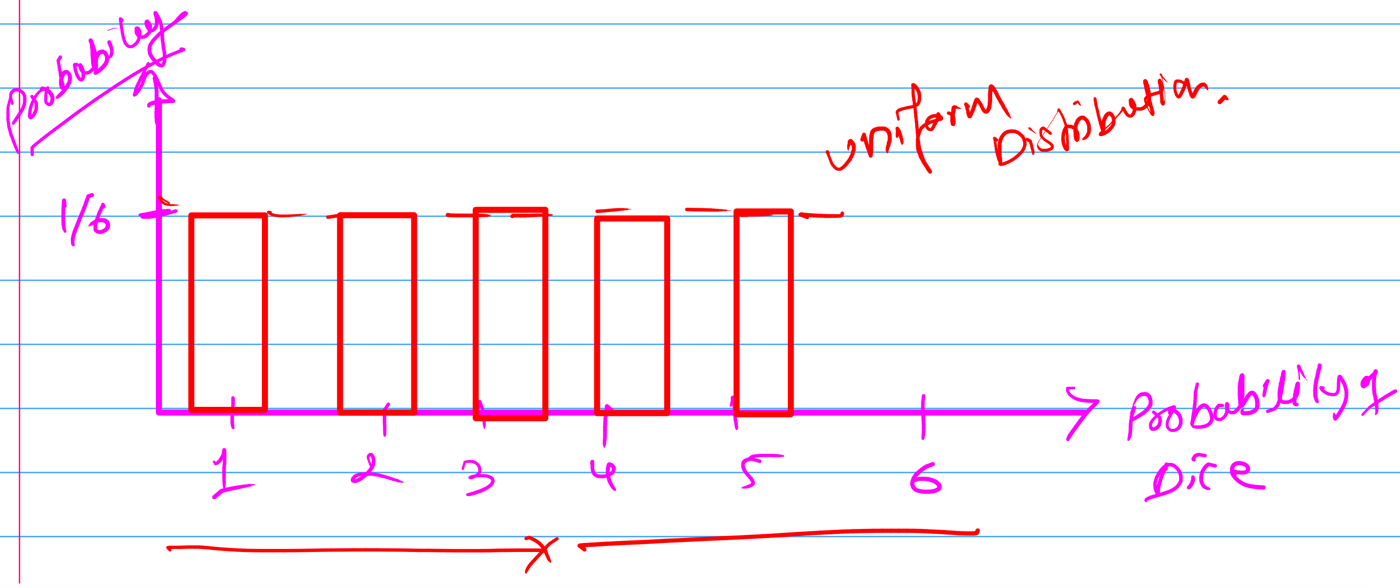
eg - Rolling a Dice = {1, 2, 3, 4, 5, 6}

Total outcome = 6 outcome

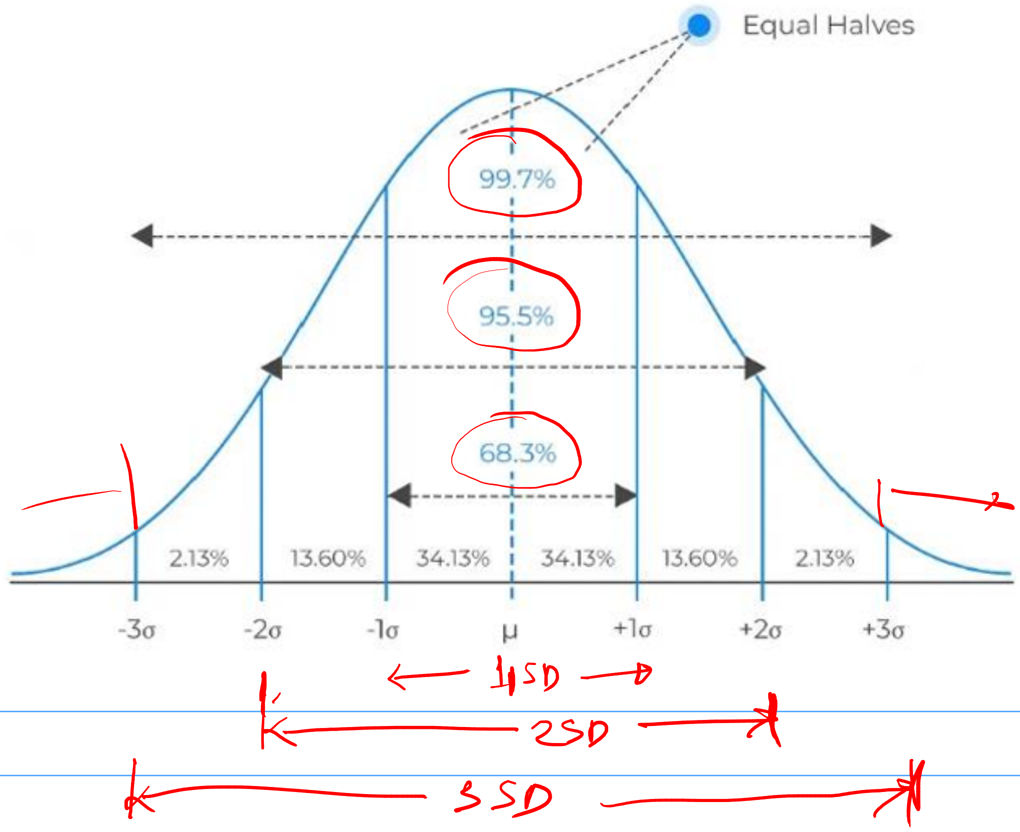
Possibility of 1 = 1/6 $\quad$ P(4) = 1/6

P(2) = 1/6 $\quad$ P(5) = 1/6

P(3) = 1/6 $\quad$ P(6) = 1/6

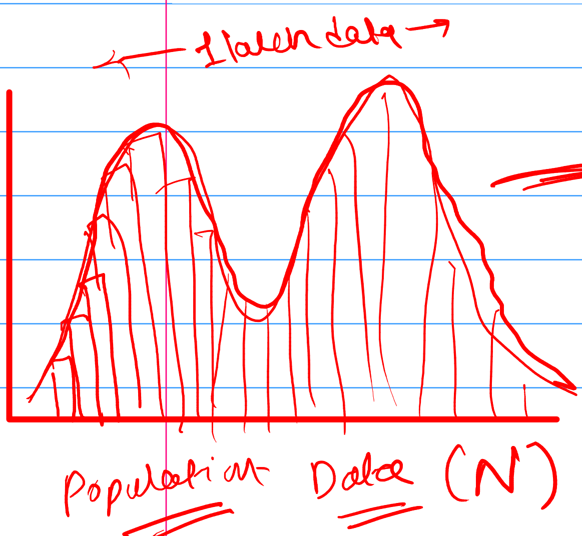


Empirical Rule of Normal Distribution: - The empirical rule in statistics, also known as the 68 95 99 rule, states that for normal distributions, 68% of observed data points will lie inside one standard deviation of the mean, 95% will fall within two standard deviations, and 99.7% will occur within three standard deviations.



Central limit Theorem: - For large sample sizes, the sampling distribution of means will approximate to normal distribution even if the population distribution is not normal.

1. The sample size is **sufficiently large**. This condition is usually met if the size of the sample is $n \geq 30$.
2. The samples are **independent and identically distributed, i.e., random variables**. The sampling should be random.
3. The population's distribution has a **finite variance**. The central limit theorem doesn't apply to distributions with infinite variance.



data $x_1 = 30$ sample $= \bar{x}_1$

$x_2 = 30 - = \bar{x}_2$

$x_3 = 30$ sample $= \bar{x}_3$

$x_4 = \quad = \bar{x}_4$

x_5

$x_{20} \quad \bar{x}_{20}$