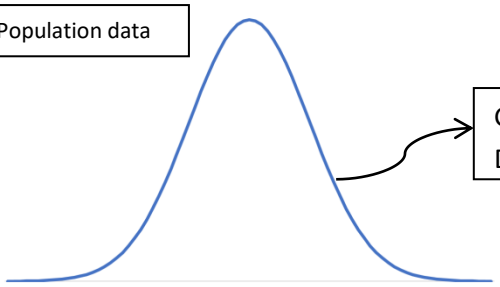
Central Limit Theorem

- Central Limit Theorem states that whether you have your population data (with mean μ and standard deviation σ) is Gaussian/Normally Distributed or Log normally distributed or Left skewed distributed, if we will take a sample data of size n >= 30 for m number of times then the plotted Histogram will be normally distributed.

- For all the samples sample size should be same.

Population data



Gaussian/Normal Distribution

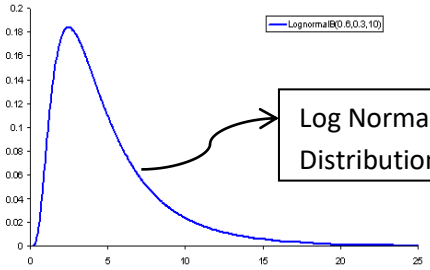
Size of Sample

n

No. of Samples

m

The larger the size the better the result



Log Normal Distribution

s_1-> $\{x_1, x_2, x_3, ...x_n\}$ -> $\overline{s_1}$

s_2-> $\{x_2, x_3, x_4, ...x_n\}$ -> $\overline{s_2}$

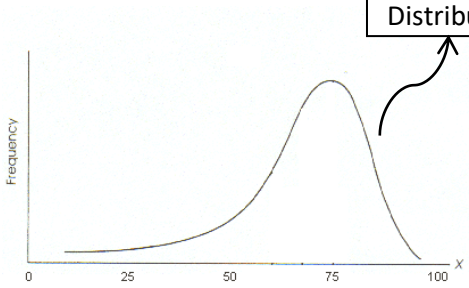
s_3-> $\{x_3, x_4, x_5, ...x_n\}$ -> $\overline{s_3}$

:

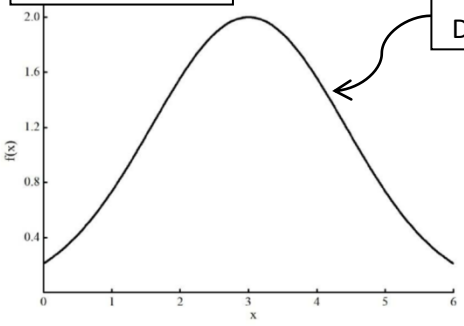
$\overline{s_n}$

Sampling with replacement

Left Skewed Distribution



Gaussian Distribution



- If n < 30 it won't follow the theorem.

Probability :

- Probability is the measure of likelihood of an event.

Eg; Tossing a fair coin. P(H) = 0.5 P(T) = 0.5

Sholay ⟶ COIN ⟶ Unfair coin P(H) = 1

Eg : Rolling a dice

P(1) = $\frac{1}{6}$, P(2) = $\frac{1}{6}$, P(3) = $\frac{1}{6}$

1. Mutual Exclusive Events :

➡ Two events are mutually exclusive if they can't occur at the same time.

Eg :

1. Tossing a coin
2. Rolling a dice

Q. What is the probability of coin landing on heads or tails?

⬇

Addition rule to mutual exclusive events

P(A or B) = P(A) + P(B)

$= \frac{1}{2} + \frac{1}{2} = 1$

Q. What is the probability of getting 1 or 6 or 3 while rolling a dice?

P(1 or 6 or 3) = P(1) + P(6) +P(3)

$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$

2. Non-mutual Exclusive events :

➡ Two events are non-mutually exclusive if they can occur at the same time.

Eg :

Picking randomly a card from a deck of cards, two events "heart" and "king" can be selected at the same time.

Q. Bag of Marbles : 10 Red, 6 Green, 3 (R&G)

When picking randomly from a bag of marbles what is the probability of choosing a marbles that is red or green?

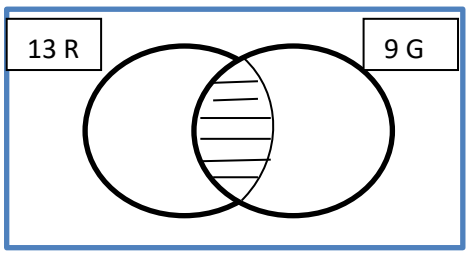
⬇

Non mutual Exclusive

➡ Addition Rule for Non mutual Exclusive events :

P(A or B) = P(A) + P(B) – P(A and B)

$= \frac{13}{19} + \frac{9}{19} - \frac{3}{19} = \frac{19}{19} = 1$



Q. Deck of cards => What is the probability of choosing ♥ or Queen?

P(♥ or Queen) = P(♥) + P(Queen) – P(♥ and Queen)

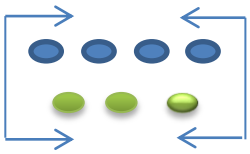
$$= \frac{1}{13} + \frac{4}{13} - \frac{1}{52} = \frac{19}{52}$$

Multiplication Rule for Non mutual Exclusive events :

Dependent events : Two events are dependent if they affect one another.

Eg :

Bag of marble :



P(B) = $\frac{4}{7}$ → P(G) = $\frac{3}{6}$

1 white marble taken

Q. What is the probability of rolling a '5' and then a '3' with a normal 6 sided dice?

P(1) = $\frac{1}{6}$, P(2) = $\frac{1}{6}$, P(3) = $\frac{1}{6}$, P(4) = $\frac{1}{6}$

Multiplication Rule for Independent events

P(A and B) = P(A) * P(B)

$$= \frac{1}{6} * \frac{1}{6} = \frac{1}{36}$$

P(A or B) => Mutual Exclusive

Non Mutual Exclusive

P(A or B) = P(A) + P(B) – P(A and B) → Non Mutual Exclusive; Common part got minus

P(A or B) = P(A) + P(B) → Mutual Exclusive

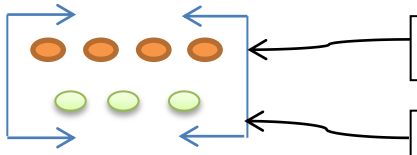
- Addition is done for Mutual Exclusive events and Non Mutual Exclusive events. (or)
- Multiplication is done for Dependent and Independent events. (and)

Tossing a COIN Sample{H & T}

P(H) = 0.5 P(T) = 0.5

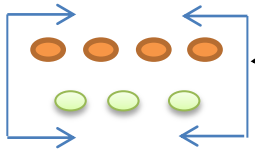
Dependent and Independent Events

P(A and B) = P(A) * P(B)



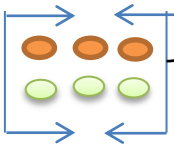
Dependent events

Q. Probability of drawing an “orange” and then drawing a “Yellow” marble from the bag?



P(O) = $\frac{4}{7}$

Orange



P(Y/O)

Conditional Probability

$\frac{3}{6} = \frac{1}{2}$

Algorithm

Naïve Baye’s

P(O and Y) = P(O) * P(Y/O)

$= \frac{4}{7} * \frac{3}{6} = \frac{2}{7}$

Permutation :

Permutation is arrangement of objects where order matters.

5 x 4 x 3 = 60 Ways => Permutation

School of children

Dairy Milk, KitKat, MIlkyBar, Snickers, 5Star

With permutation order matters

{DM KK MB} {.. .. ..} {.. .. ..}

{KK DM MB} {.. .. ..} {.. .. ..}

Possible arguments

P(n,r) = $\frac{n!}{(n-r)!} = \frac{5!}{(5-3)!}$ = 60 ways

n = Total no. of objects

r = number of selection

Combination :

Combination is selection of items from a set where repetition is not allowed. (It should be of unique members where order of selection doesn't matter)

{DM KK MB}

{KK DM MB} ⟵ X

$$C(n,r) = \frac{n!}{r!(n-r)!} = \frac{5!}{3!\,(5-3)!} = 10 \text{ ways}$$

Covariance : {Feature Selection}

Q.

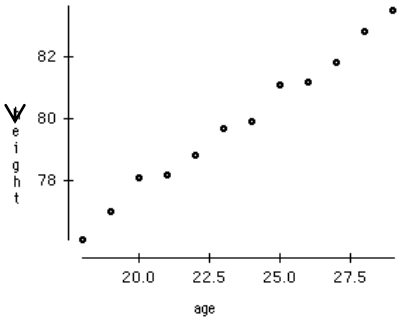
Age	Weight
12	40
13	45
15	48
17	60
18	62

Age ↑ Weight ↑

Age ↓ Weight ↓

Quantify the relationship between X & Y using mathematical equation.

$$Cov(X,Y) = \frac{\sum(Xi-\bar{X})(Yi-\bar{Y})}{n-1} = \frac{96}{4} = 24$$



$$\text{Cov}(X,Y) = \frac{\sum(Xi-\bar{X})(Yi-\bar{Y})}{n-1}$$

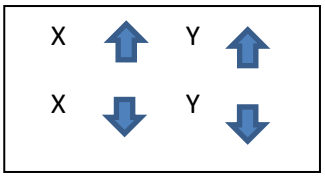
$$S^2 = \frac{\sum(Xi-\bar{X})^2}{n-1}$$

Cov(X,Y) ⟵ $S^2 = \frac{\sum(Xi-\bar{X})(Xi-\bar{X})}{n-1}$

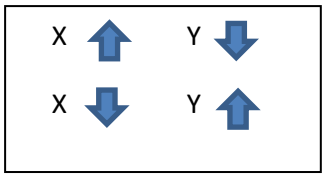
Q. What is the relationship between Cov(X,X) and var(X)?

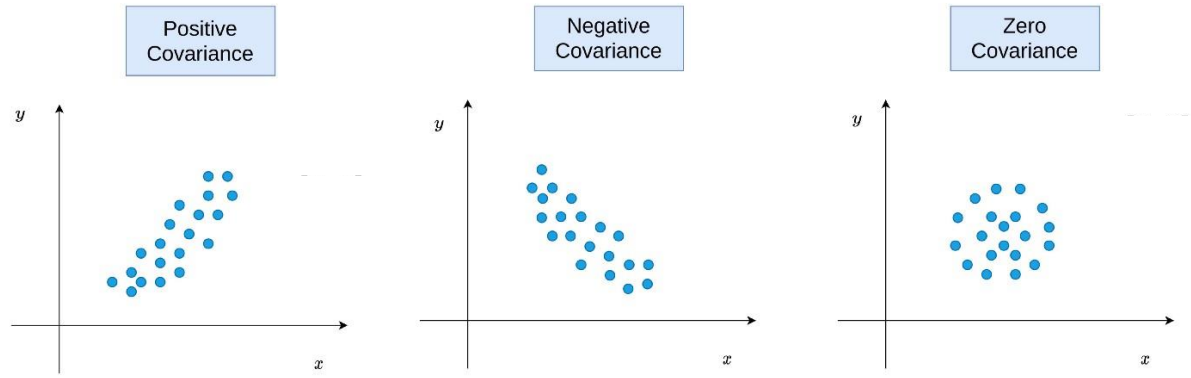
Cov(X,X) = Var(X)

+ve Covariance



+ve Covariance





X	Y
10	4
8	6
7	8
6	10

Cov(X,Y) = -ve

$$= \frac{\sum(Xi-\bar{X})(Yi-\bar{Y})}{n-1}$$

= -3.25