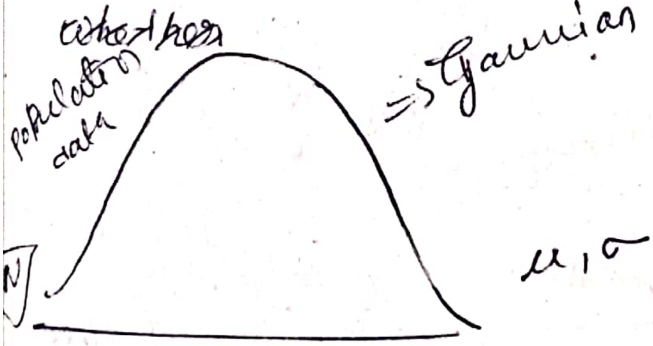
Agenda:

* Central limit Theorem
* Probability
* Permutation and combination
* Covariance
* Pearson correlation
* Spearman Rank correlation

Central limit Theorem

The Sampling distribution of the mean will always be normal / Gaussian distributed as long as the sample size is large enough.



$\rightarrow S_1 \rightarrow \{x_1, x_2, x_3 \ldots x_n\} \rightarrow \bar{x}_1 = \bar{S}_1$

$\rightarrow S_2 \rightarrow \{x_3, x_4, x_1 \ldots x_n\} \rightarrow \bar{x}_2 = \bar{S}_2$

$\rightarrow S_3 \rightarrow \{x_4, x_1, x_2 \ldots x_n\} \rightarrow \bar{x}_3 = \bar{S}_3$

$\vdots$

$\bar{x}_m \quad \bar{S}_m$

$\boxed{n \geq 30}$

$\boxed{n} \rightarrow$ Size of Sample

$\boxed{m} \rightarrow$ No. of samples (The larger the value the better

Importance:

Based on this concept lot of assumptions can be made.

→ n should be greater than or equal to 30. Then only the Gaussian distribution is formed.

* If we consider $n < 30$ then the gaussian distribution is not formed.

Probability: Probability is a measure of ~~prop~~ the likelihood of an event.

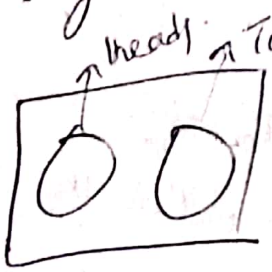
Eg: 1) Tossing a fair coin $P(H) = 0.5$ $P(T) = 0.5$

2) Rolling a dice $P(1) = \frac{1}{6}$ $P(2) = \frac{1}{6}$ $P(3) = \frac{1}{6}$ ...

1). Mutual Exclusive Event

Two Events are mutually exclusive if they cannot occur at the same time.

① Tossing a coin ② Rolling a dice



The head and tail cannot occur at a time so it is mutual exclusive event,

2) Non mutual Exclusive Events

Two events can occur at the same time

Eg: picking randomly a card from a deck of cards, two events "heart" and "king" can be selected.

Addition Rule:

mutual exclusive Event

① what is the probability of coin landing on heads or tails.

⇓

Addition rule for ~~φ~~ mutual exclusive events

$$P(A \text{ or } B) = P(A) + P(B)$$

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

② what is the probability of getting 1 or 6 or 3 while rolling a dice?

$$P(1 \text{ or } 6 \text{ or } 3) = P(1) + P(6) + P(3)$$

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

Addition rule for non mutual exclusive Event

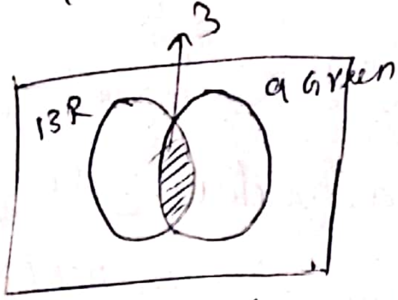
Bag of marbles:

10 Red, 6 Green, 3(R & G)

(1) when picking randomly from a bag of marbles what is the probability of choosing a marble that is red or green ?

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

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



(2) Deck of cards → what is the probability of choosing heart or queen.

$$P(\heartsuit \text{ or Queen}) = P(\heartsuit) + P(\text{Queen}) - P(\heartsuit \text{ and Queen})$$

$$= \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52}$$

* Multiplication Rule

* Dependents Events ! Two events are dependent if they a affect one another.

Bag of marble $\left\{ \begin{matrix} w & w & w & w \\ Y & Y & Y & Y \end{matrix} \right\}$ (w) → white, (Y) → yellow

$\Rightarrow P(w) = \frac{4}{7} \longrightarrow P(Y) = \frac{3}{6}$

↑ white
7 marble

→ after picking one (w) then the probability of picking (Y) is $\frac{3}{6}$

* w Independent event

→ what is the probability of rolling a "5" and then a "3" with a normal 6 sided dice!

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

Here the events are Independent, not dependending on previous events.

multiplication rule for Independent events

$$P(A \text{ and } B) = P(A) * P(B)$$

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

$P(A \text{ or } B) \Rightarrow$
- → mutual exclusive
- → Non mutual exclusive

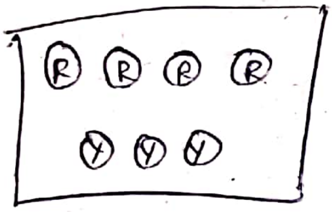
$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$ → Non mutual exclusive

$P(A \text{ or } B) = P(A) + P(B)$ (mutual exclusive)

Dependent and Independent Events

$$P(A \text{ and } B) = P(A) * P(B)$$

(2)



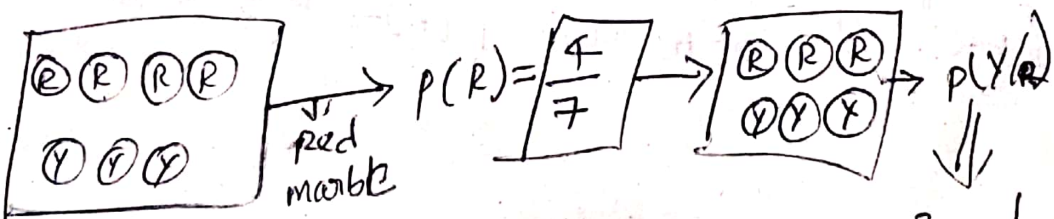
⇒ Dependent Events

R → Red marbles

Y → yellow marbles

probability of drawing a "Red" and then drawing a "yellow" marble from the bag?

Hw:



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

$$P(R \text{ and } Y) = P(R) * (P\,Y/R)$$

$$= \frac{4}{7} * \frac{3}{6} = \frac{4}{14} = \frac{2}{7}$$

* Permutation

~~Arranging~~ of objects in a definite order.

The elements or members of sets are arranged here in a sequence or linear order.

Ex: set $A = \{1, 6\}$

we can arrange them in two ways.

$\{1, 6\}$ $\{6, 1\}$

Ex2: School of children should pick a chocolate in a set of chocolate.

~~{Dairy~~ {Dairy milk, kit kat, milky Bar, Snickers, 5 star}

first person has a chance to pick 5 different chocolate

second has a chance of 4 different chocolate

has a chance of 3 different chocolate

$\underline{5} \times \underline{4} \times \underline{3}$

= 60 ways => permutation

Formula:

$n = 5$

n = Total no. of objects

r = no. of selection

$$^{n}P_{r} = \frac{n!}{(n-r)!} = \frac{5!}{(5-3)!}$$

$$= \frac{5 \times 4 \times 3 \times \cancel{2!}}{\cancel{2!}} = \boxed{60}$$

combination:

~~Rep~~ Here repetition will not occur

* unique combination.

Formula:

$$ {}^{n}C_{r} = \frac{n!}{r!\,(n-r)!} = \frac{5!}{3!(2)!} = \frac{5\times4\times3\times2\times1}{3!\times2} $$

$$ = 10 $$

* Covariance:

X Age	Y weight
12	40
13	45
15	48
17	60
18	62
$\bar{x}=15$	$\bar{y}=51$

$\boxed{\begin{matrix}\text{Age} \uparrow \ \text{weight} \uparrow \\ \text{Age} \downarrow \ \text{weight} \downarrow\end{matrix}}$

$\Uparrow$

Quantify the relationship x & y using mathematical equation

$$ \boxed{cov(x,y) = \frac{\sum(x_i-\bar{x}) * (y_i-\bar{y})}{n-1}} $$

$\Downarrow$

24 => +ve covariance

$$ \sigma^2 = \frac{\sum(x_i-\bar{x})^2}{n-1} $$

$$ s^2 = \frac{\sum(x_i-\bar{x})(x_i-\bar{x})}{n-1} $$

$\Downarrow$

$cov(x,x)$

$$\boxed{cov(x,x) = Var(x)}$$

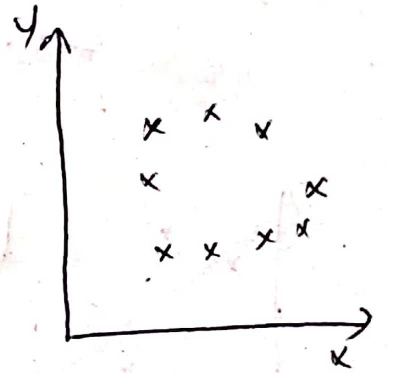
+ve covariance $\boxed{\begin{matrix} x\uparrow & y\uparrow \\ x\downarrow & y\downarrow \end{matrix}}$

-ve covariance $\boxed{\begin{matrix} x\uparrow & y\downarrow \\ x\downarrow & y\uparrow \end{matrix}}$

If covariance = 0 then there is no relation with x & y.

+ve covariance

-ve covariance



No relation

~~Ex for -ve cova~~

X	Y
10	4
8	6
7	8
6	10

$\bar{x} = 7.75 \quad \bar{y} = 7$

$cov(x,y) = -ve$

$$= (10-7.75)(4-7) + (8-7.75)(6-7) + (7-7.75)(8-7) + (6-7.75)(10-7)$$

$$= \underline{\underline{-3.25}}$$

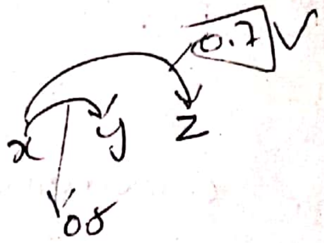
$\boxed{\begin{matrix} x\uparrow & y\downarrow \\ x\downarrow & y\uparrow \end{matrix}}$

Scanned with CamScanner

* Pearson Correlation coefficient (-1 to 1)

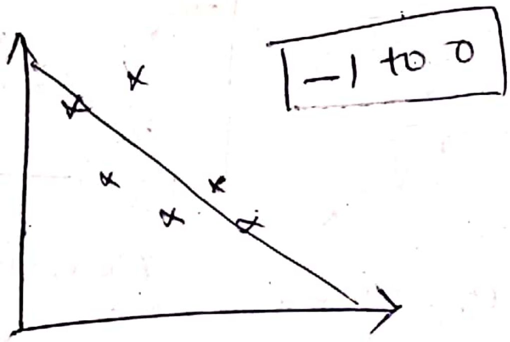
+ve covariance
-ve covariance

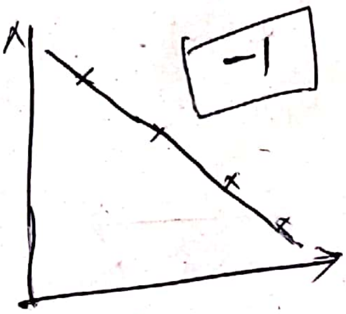
$$\rho(x,y) = \frac{\operatorname{cov}(x,y)}{\sigma_x \cdot \sigma_y}$$

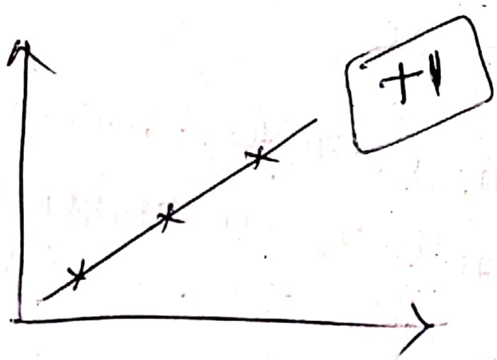


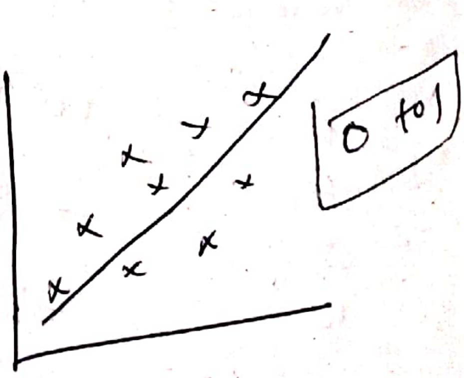
more the value towards +1 then more +ve correlated

If more the value towards -1 then more -ve correlated.









* Spearman Rank Correlation

$$r_s = \frac{cov(R(X), R(Y))}{\sigma(R(X)) * \sigma(R(Y))}$$

X	Y
10	4
8	6
7	8
6	10

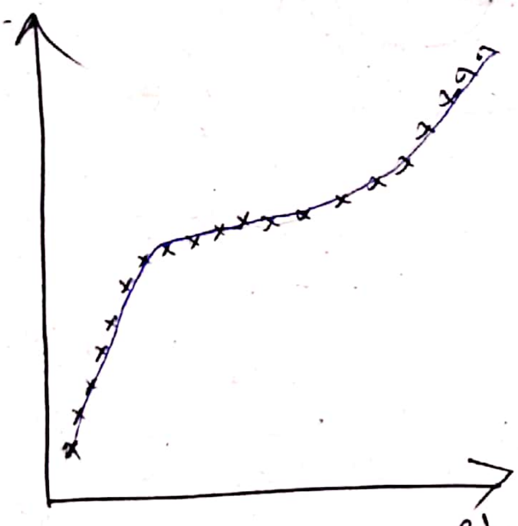
$\Rightarrow$

R(X)	R(Y)
4	1
3	2
2	3
1	4

Spearman Rank Correlation

$\Downarrow$

giving the ranks in Ascending order



$\rightarrow$ spearman correlation = 1.
pearson correlation = 0.88

Here it is not linear but it is a +ve ~~correlation~~ correlation

why this correlation is used?

It measures the strength and direction of association between two ranked variables. It basically gives the measures of monotonicity of the relation between two variables i.e., how well the relationship between two variables could be represented using a monotonic function.