Notes

Page No.
Date :

Assignment * Statistic Class-4

* In India average IQ is 100 with a standard Deviation of 15. what is the percentage would you expect to have an IQ.

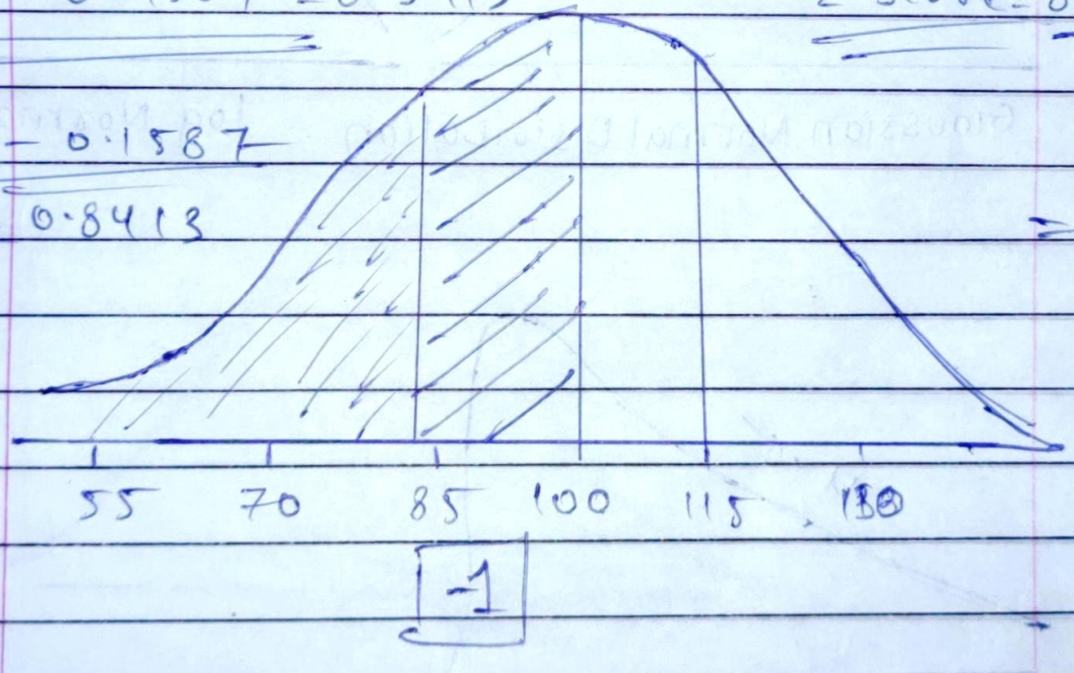
(1) Lower than 85 = 0.1587

(2) Higher than 85 = 0.8413

(3) Between 85 and 100 = 0.3413

0.5 - 0.1587 = 0.3413

1 - 0.1587
= 0.8413



-1

Z-score = 85 - 100 / 15

= -1

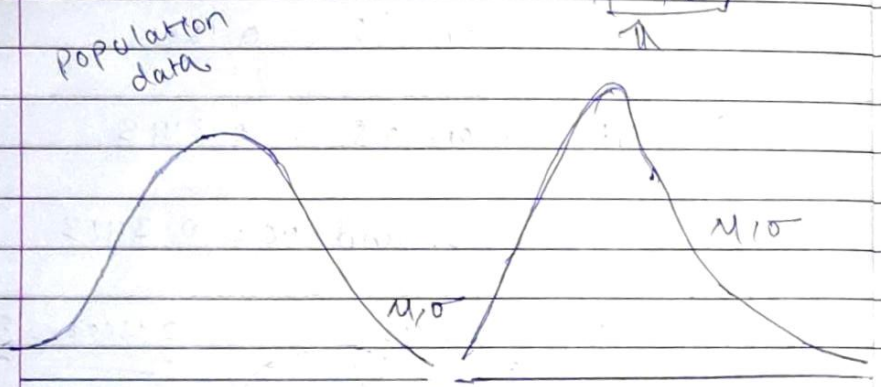
Day 4 - stats.

(1) Central limit Theorem
(2) Probability
(3) Permutation and Combination
(4) Covariance / pearson correlation / spearman Rank correlati

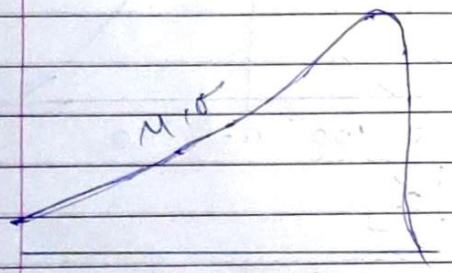
(5) Bernoulli Distribution
(6) Binomial Distribution
(7) Power Law (Pareto Distribution)

$n<=30$

* Central Limit Theorem :- $\Rightarrow n>30$



Gaussian Normal Distribution log Normal Distribution



left skewed Distribution.

central limit Theorem :- says that if Population data follows Gaussian or log Distribution it may be follows this or not follows this. sample size n is greater than equal to 30.

size of sample ↑ n The larger the value the better

No. of samples ↑ m

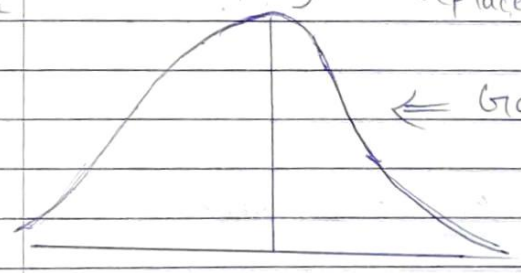
$\rightarrow S_1 \rightarrow \{x_1, x_2, x_3 \ldots\ldots x_n\} \rightarrow \bar{x}_1 = \bar{s}_1$

$\rightarrow S_2 \rightarrow \{x_3, x_4 - x_1 \ldots\ldots x_n\} \rightarrow \bar{x}_2 \quad \bar{s}_2$

$\rightarrow S_3 \rightarrow \{x_4, x_1 - \ldots\ldots x_n\} \rightarrow \bar{x}_3 \quad \bar{s}_3$

$= \quad \bar{x}_n \quad \bar{s}_n$

Sampling with replacement



⇐ Gaussian Distribution

10 different region $n>30$

↳ Assumptions

size of shark through out the world

Page No.
Date :

n<30, m↑↑↑

* Probability :- Probability is a measure of the likelihood of an event

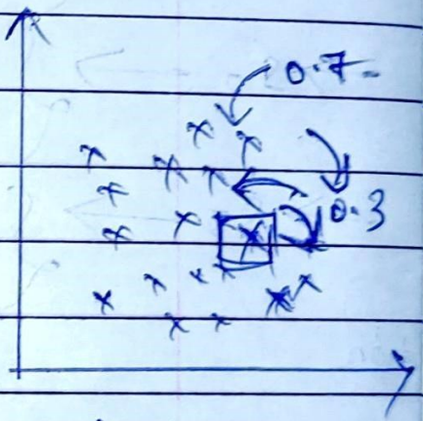
Eg :- Tossing a fair coin

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

mould Sholay → COIN P(H) = 1

unfair coin

Strong Basic Probabilities



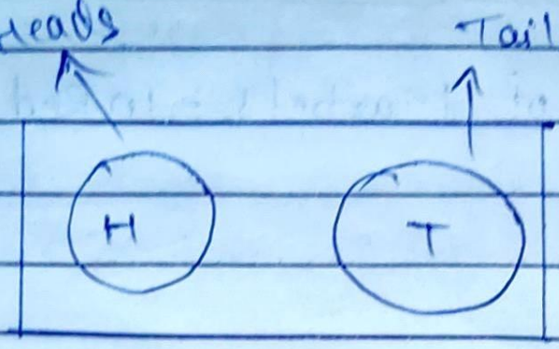
Rolling a Dice $P(1) = \frac{1}{6}$ $P(2) = \frac{1}{6}$ $P(3) = \frac{1}{6}$

① Mutual Exclusive Event

Two Event are naturally exclusive if they cannot occur at the same time

(1) Tossing a coin

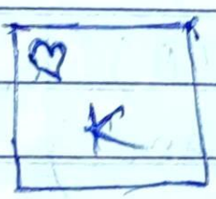
(2) Rolling a dice



(2) Non Mutual Exclusive Events

Two Events can occur at the same time

* Picking randomly a card from a deck of cards, two events "heart" and "king" can be selected



* Mutual Exclusive Event

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

Addition Rule for mutual exclusive event

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

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

Page No.
Date :

Non Mutual Exclusive Event :-

Bag of Marbels : 10 Red, 6 Green,
3 (R & Gr)

* when picking randomly from a Bag of marbles what is the probability of choosing a marbel that is red or Green ?

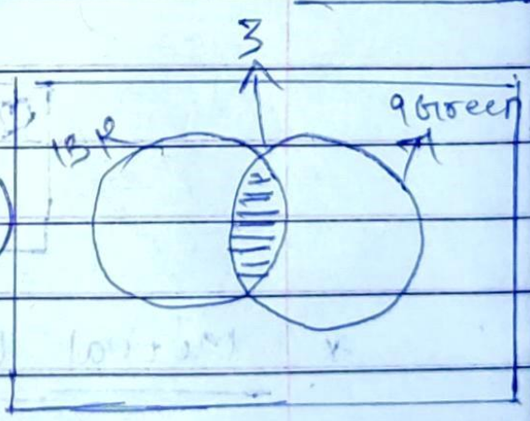
⇓

Non mutual Exclusive

Addition Rule for Non mutual Exclusive Event

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

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



deck of cards → what is the probability of choosing

♡ or Queen

Page No.
Date :

P(♡ or Queen) = P(♡) + P(Queen) −

$= \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \boxed{\frac{16}{52}}$ P(♡ and Queen)

* Multiplication Rule
* Dependents Events : Two event are dependent if they affect one another

Bag of marbels { ○ ○ ○ ✗ ○ ○ ○ }

$= P(\text{white}) = \frac{4}{7} \longrightarrow P(4) = \frac{3}{6}$

white marbels.

* what is the Probability of rolling a "5" and "3" with a normal six 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 event

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

Page No.
Date :

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

P(A or B) ⟹ → Mutual Exclusive
→ Non Mutual Exclusive

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

P(A or B) = P(A) * P(B) − P(A and B)
↓
Non Mutual Exclusive

Dependent and Independent Events

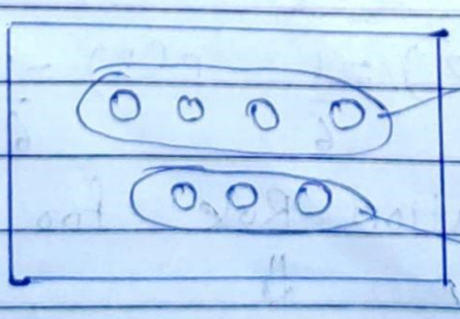
Event A, Event B

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

Tossing a coin ⟸ sample = {H & T}

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

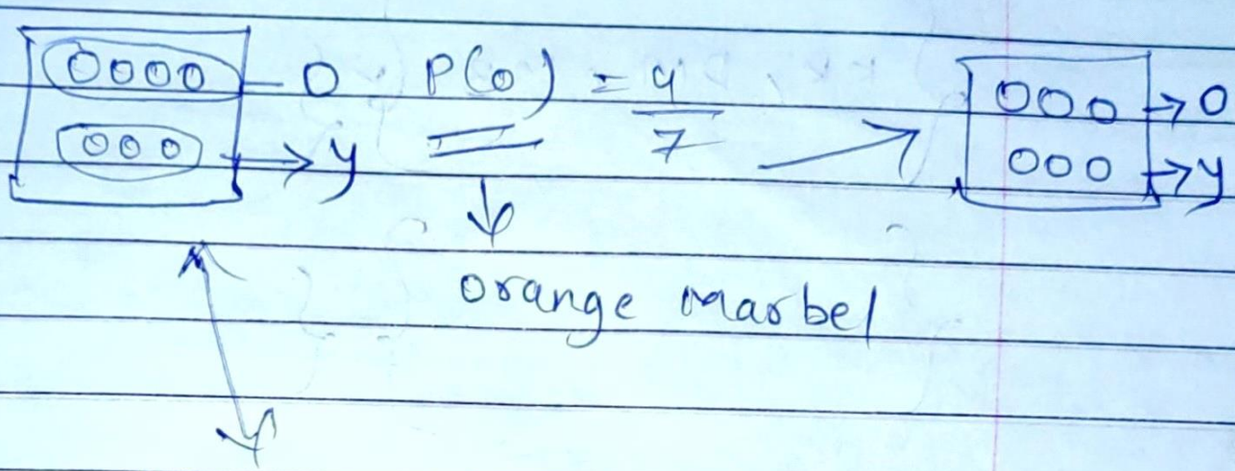
(2)



⟹ Dependent Event

Page No.
Date :

Probability of drawing a "orange" and drawing "yellow" marbel from the Bay?



P(y/o) → Conditional probability

↳ 3/6 = 1/2

P(o) and P(y) = P(o) * P(y/o)

= 4/7 * 3/6 = 12/42 = 2/7

* Permulation

school of children

5 * 4 * 3

{Diary Milk, Kit Kat, Milky Bar, Sneakers, 5star}

= 60 ⇒ peomutation

with permutation order matters

Page No.
Date :

{ Dairy Milk, KK, MB } { possible arrangement }

{ KK, DM, MB } { } { }

{ } { } { }

$n = 5$

$= {}^{n}P_{r} = \frac{n!}{(n-r)!} = \frac{5!}{(5-3)!}$ n = Total no. of objects

r = No. of selection

$= \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = 60$ ways ⇓ permutation

* Combination

{ DM KK MB }

Repetition will not occur

unique combination.

X { MB KK DM } ←

$= {}^{n}C_{r} = \frac{n!}{r!(n-r)!} = \frac{5!}{3!(2)!}$

Page No.
Date:

$$\frac{5 \times 4 \times 3 \times 2 \times 1}{3! \times 2!}$$

$$= 10$$

DREAM 11 ⟸

* Covariance { feature selection }

x (Age)	y (weight)
12	40
13	45
15	48
17	60
16	62
$\bar{x} = 15$	$\bar{y} = 51$

Quantify the relationship x & y using mathematical Questions.

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

⇓

24 ⇒ +ve

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

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

Page No.
Date :

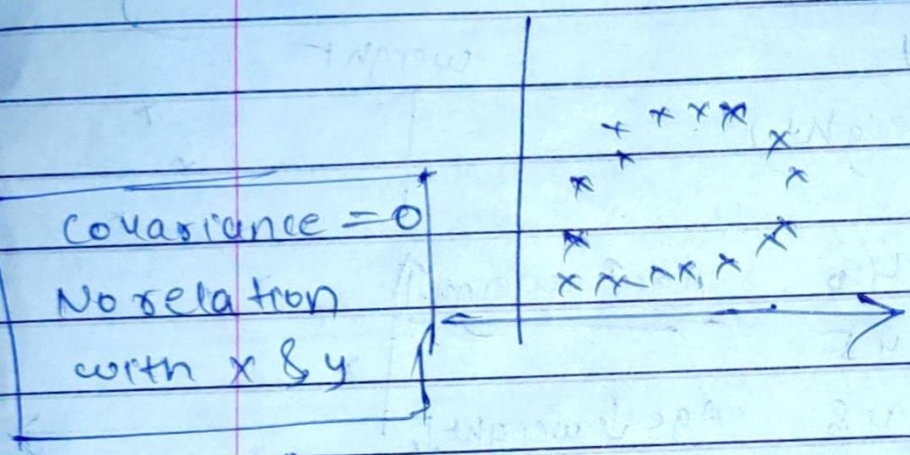
$$Cov(x,x) = var(x)$$

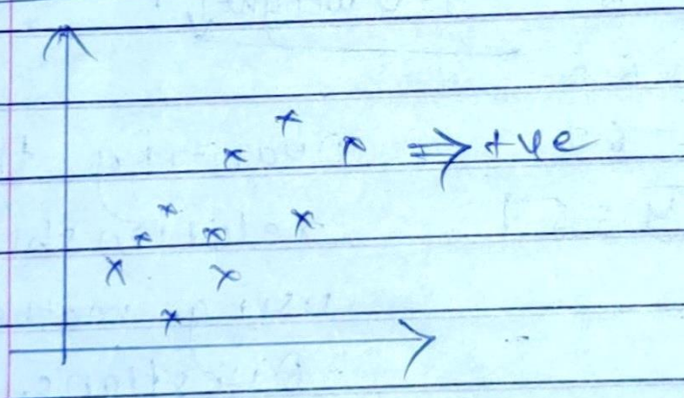
+ve covariance

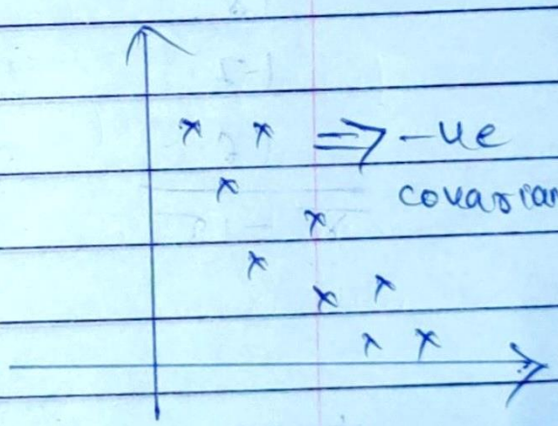
$x\uparrow$	$y\uparrow$
$x\downarrow$	$y\downarrow$

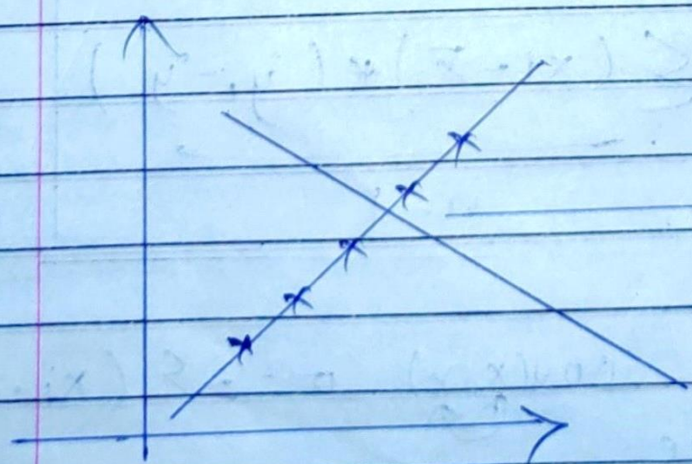
-ve covariance

$x\uparrow$	$y\downarrow$
$x\downarrow$	$y\uparrow$









Page No.
Date :

x	y
10	4
8	6
7	8
6	10
7.75	7

$cov(x,y) = -ve$

$= [(10-7.75)+(8-7.75)*(6-7)+(7-7.75)*(8-7)+(6-7.75)*(10-7)]$

$= -3.25$

$x \uparrow \; y \downarrow$

$x \downarrow \; y \uparrow$

* Pearson correlation coefficient (-1 to 1)

$\rho(x,y) = cov(x,y) / (\sigma_x \cdot \sigma_y)$

scale

covariance = +ve

-ve

+ve

-ve

x y z

x y z

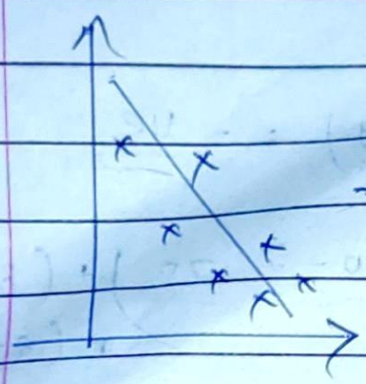
0.5

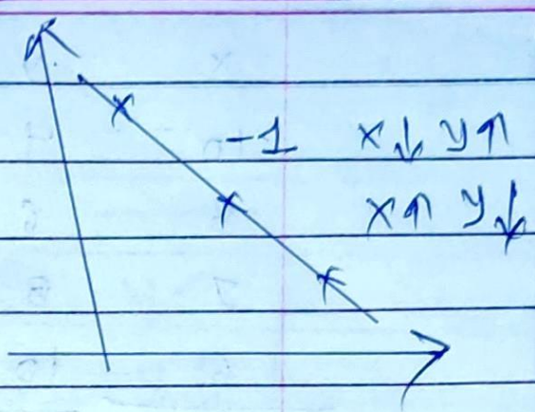
0.7

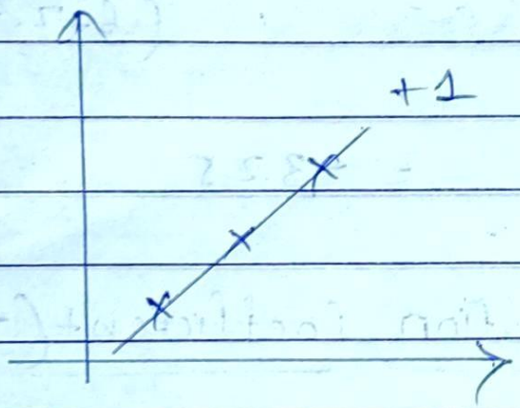
More the value towards +1 -1

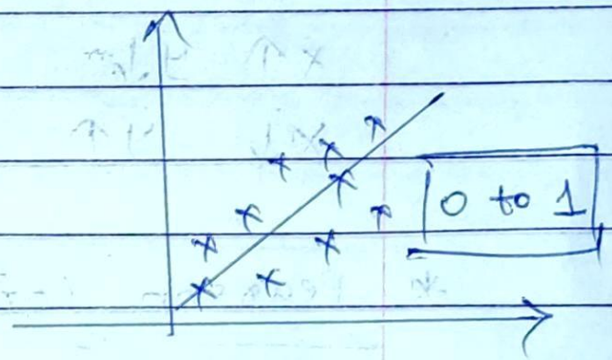
~~More the correlated it~~

More +ve corelated it is negative correlated









* Spearman Rank correlation:

$$\rho_s = \frac{\mathrm{cov}(R(x), R(y))}{\sigma(R(x)) * \sigma(R(y))}$$

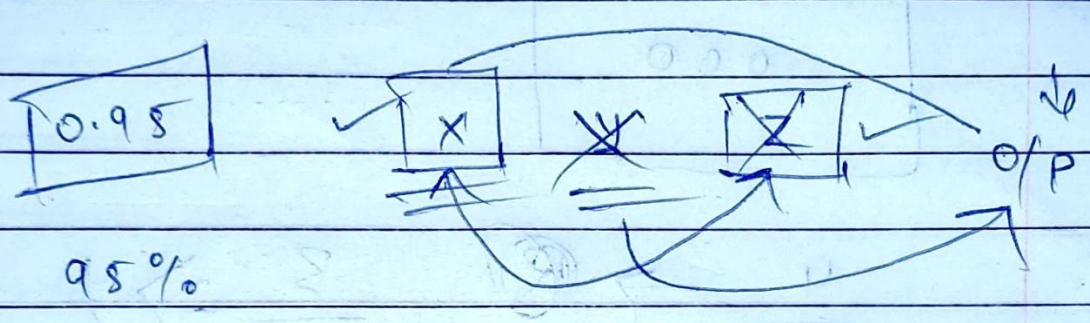
Ascending order

R(x)	R(y)
4	1
3	2
2	3
1	4

Spearman Rank correlation

Page No.
Date :

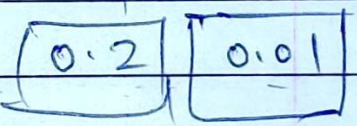
why this Correlation will be used?

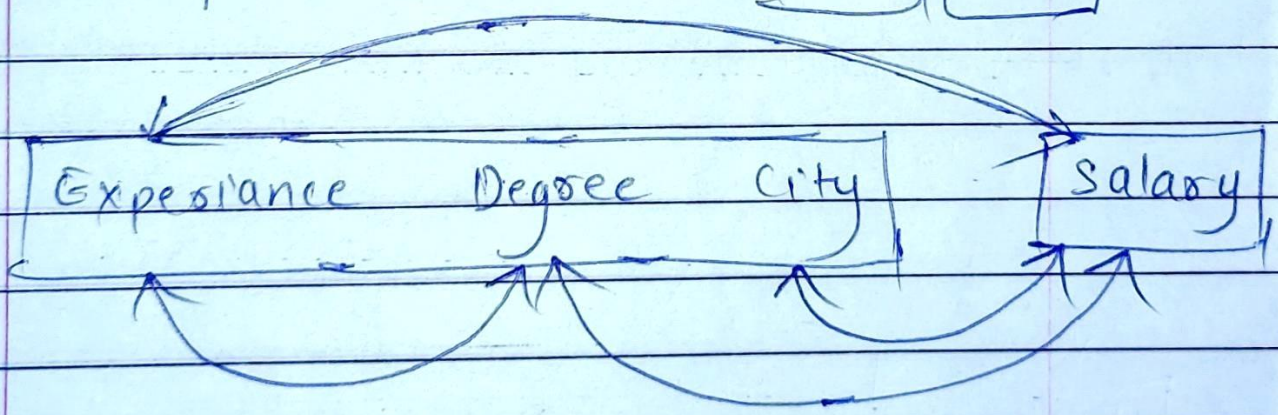


+ve
-ve
} Good

~~Exp~~ ~~Degree~~ ~~city~~

Example :-





+ve correlated

-ve correlated