ENGINEERING MATHEMATICS

UNIT-3 : *Statistical Techniques – I*

Lec-9

Today's Target

- **Correlation and Co-efficient of Correlation**
- PYQ
- DPP

Correlation

- ***Correlation is a statistical tool which deal with two or more variables(or two or more series).***
- **The degree of relationship between the variables under consideration is measured through the corelation analysis**

x	y
1	10
2	11
3	12
4	13
5	12

x	y
5	15
4	14
3	13
2	12
1	11

x	y
1	10
2	9
4	8
5	7
6	6

Types of Correlation

(i) *Positive corelation* :

If two variables move in same direction then corelation is said to be positive corelation.

(i) $x \uparrow$ $y \uparrow$ (ii) $x \downarrow$ $y \downarrow$

(ii) *Negative corelation* :

If two variables move in opposite direction then corelation is said to be negative corelation.

(i) $x \uparrow$ $y \downarrow$ (ii) $x \downarrow$ $y \uparrow$

Method to measure Correlation

Co – efficient of corelation (γ_{xy})

OR

Karl pearson's Co – efficient of corelation

OR

Product moment Co – efficient of corelation

$$\gamma_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

$$cov(x,y) = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{n}$$

$$\sigma_x = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n}}$$

$$\sigma_y = \sqrt{\frac{\Sigma (y-\bar{y})^2}{n}}$$

$$\gamma_{xy} = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma (x-\bar{x})^2}\sqrt{\Sigma (y-\bar{y})^2}}$$

Direct Method

Where

$\bar{x}$ = Mean of x

$\bar{y}$ = Mean of y

$(x - \bar{x})$ = Deviation from mean

$(y - \bar{y})$ = Deviation from mean

Another Form

$$r_{xy} = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

Direct Method

Shortcut Method

$$d_x = x - A$$

$$d_y = y - A$$

$$r_{xy} = \frac{n\Sigma d_x d_y - \Sigma d_x \Sigma d_y}{\sqrt{n\Sigma d_x^2 - (\Sigma d_x)^2}\sqrt{n\Sigma d_y^2 - (\Sigma d_y)^2}}$$

Step deviation Method

$$u = \frac{x - A}{h}$$

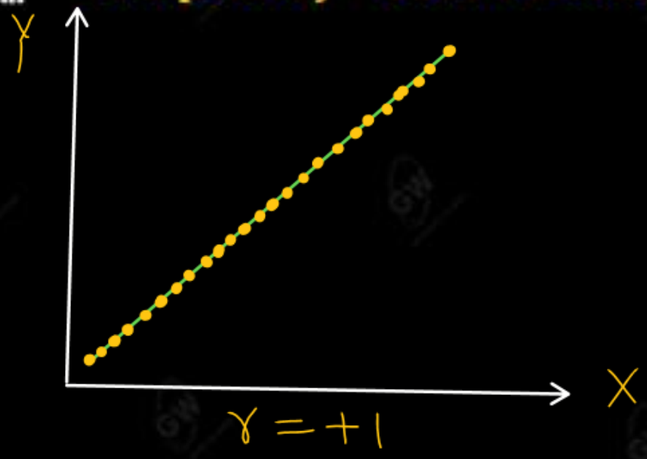
$$v = \frac{y - A}{h}$$

$$r_{xy} = \frac{n\Sigma uv - \Sigma u \Sigma v}{\sqrt{n\Sigma u^2 - (\Sigma u)^2}\sqrt{n\Sigma v^2 - (\Sigma v)^2}}$$

Important points

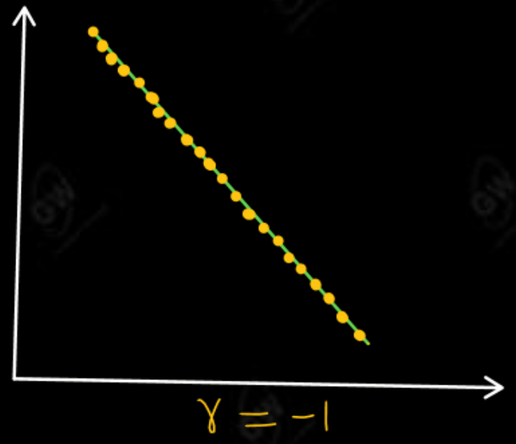
1. $-1 \leq r \leq 1$
2. Positive correlation
   $0 < r < 1$
3. Negative correlation
   $-1 < r < 0$
4. Perfect +ve correction
   $r = +1$
5. Perfect Negative correlation
   $r = -1$
6. No correlation ($r = 0$)

Graphs of correlation / scattered or dotted diagrams of correlation



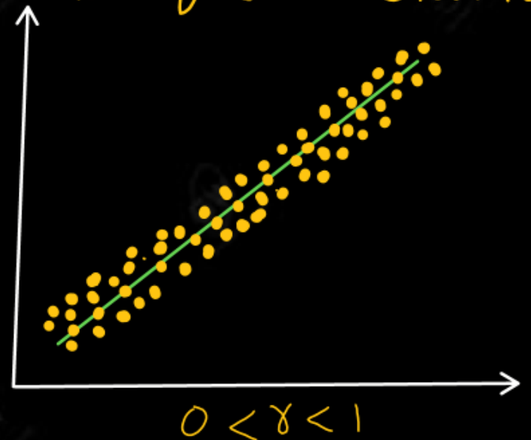
$r = +1$

Perfect +ve correlation



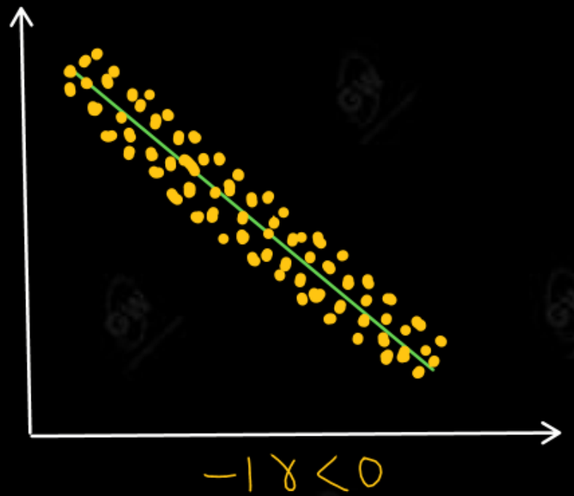
$r = -1$

Perfect -ve correlation



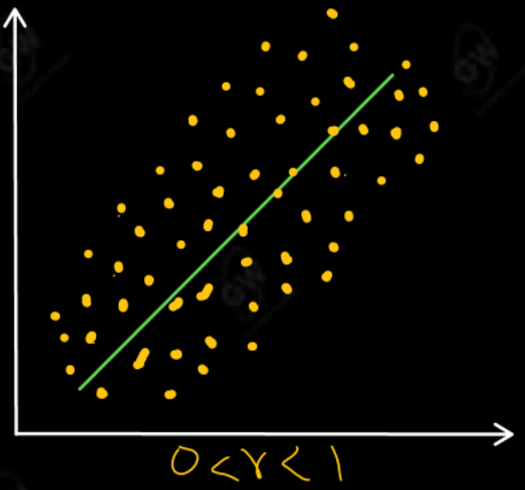
$0 < r < 1$

Strong +ve correlation



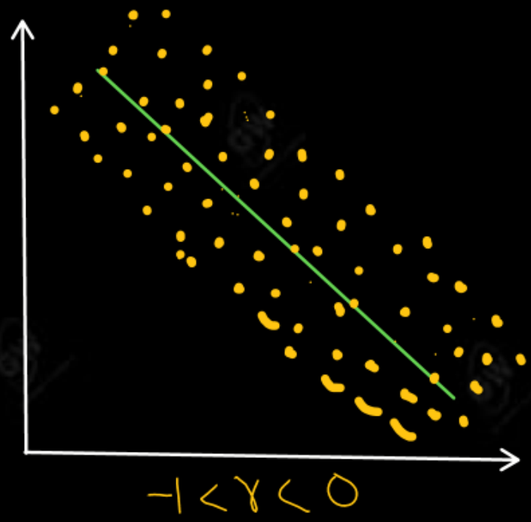
$-1 < r < 0$

Strong -ve correlation



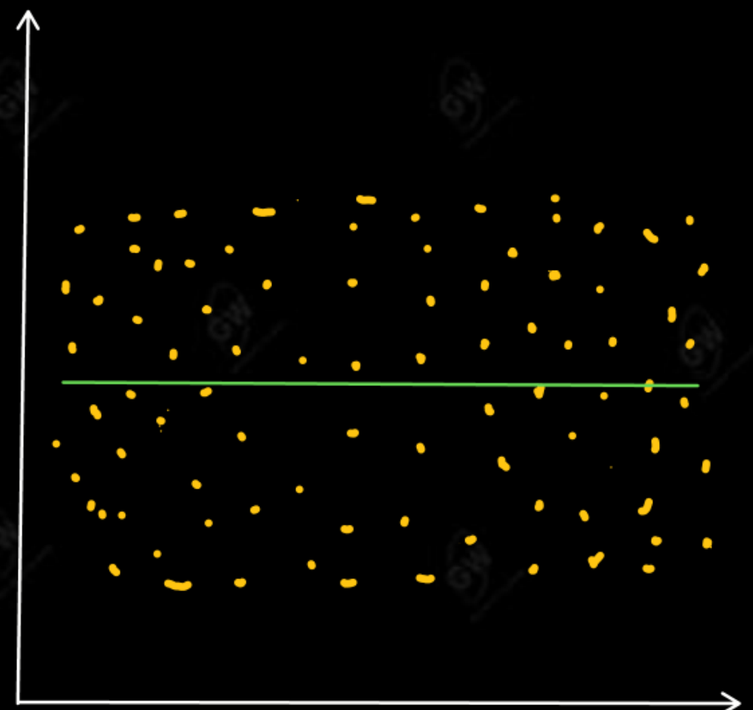
$0 < r < 1$

Weak +ve correlation



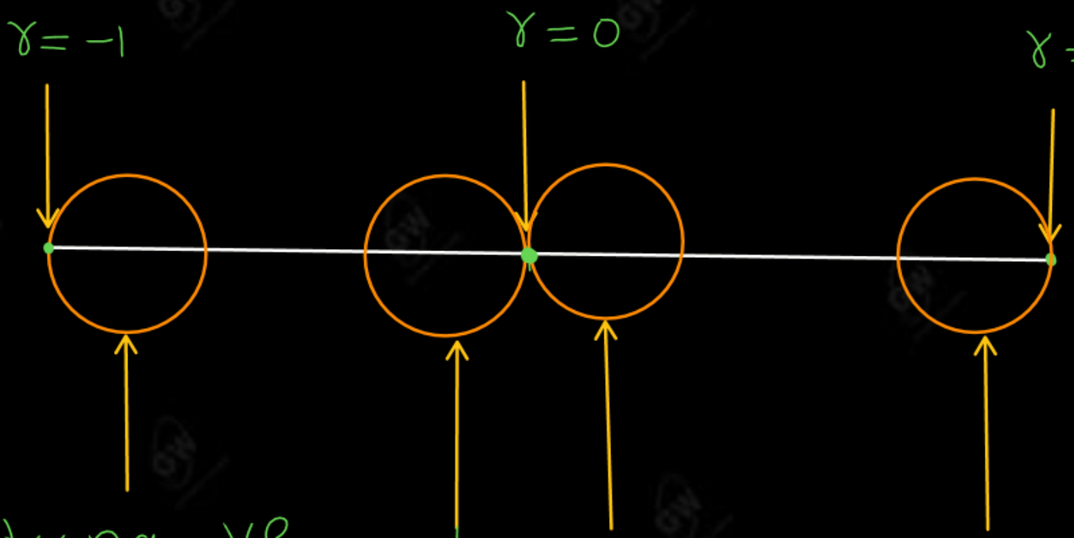
$-1 < r < 0$

Weak -ve correlation



$\gamma = 0$

No correlation



Strong −ve correlation

Weak −ve correlation

Weak +ve correlation

Strong +ve correlation

Q. 1:– Form the data given below, find the number of items n:

$r_{xy} = 0.5,\ \sum XY = 120,\ \sum X^2 = 90, \sigma_y = 8$ *where X and Y are deviations from the arithmetic mean.*

(AKTU-2016)

Given

$r_{xy} = 0.5$

$\Sigma XY = 120$

$\Sigma X^2 = 90$

$\sigma_y = 8$

$X = x - \bar{x}$

$Y = y - \bar{y}$

$\Sigma (x-\bar{x})(y-\bar{y}) = 120$

$\Sigma (x-\bar{x})^2 = 90$

$\sigma_y = 8$

$\sqrt{\frac{\Sigma(y-\bar{y})^2}{n}} = 8$

$\frac{\Sigma (y-\bar{y})^2}{n} = 64$

$\Sigma (y-\bar{y})^2 = 64n$

$$r_{xy} = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma (x-\bar{x})^2}\sqrt{\Sigma (y-\bar{y})^2}}$$

$$0.5 = \frac{120}{\sqrt{90}\ \sqrt{64n}}$$

$$\frac{0.5}{\frac{1}{2}} = \frac{120}{\sqrt{90} \times 8\sqrt{n}}$$

$$\sqrt{n} = \frac{120 \times 2}{8 \times \sqrt{90}}$$

$$\sqrt{n} = \frac{30}{\sqrt{90}}$$

$$n = \frac{900}{90}$$

$$n = 10$$

Q. 2:- *Ten students got the following percentage of marks in Principles of Economics and Statistics:*

Roll Nos.:	1	2	3	4	5	6	7	8	9	10
(x) Marks in Econmics	78	36	98	25	75	82	90	62	65	39
(y) Marks in Statistics	84	51	91	60	68	62	86	58	53	47

Calculate the co – efficient of correlation

Mean

$$\bar{x} = \frac{\Sigma x}{n}$$

$$= \frac{650}{10}$$

$$\boxed{\bar{x} = 65}$$

$$\bar{y} = \frac{\Sigma y}{n}$$

$$= \frac{660}{10}$$

$$\boxed{\bar{y} = 66}$$

$$r_{xy} = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma (x-\bar{x})^2}\sqrt{\Sigma (y-\bar{y})^2}}$$

$$r_{xy} = \frac{2734}{\sqrt{5398}\sqrt{2224}}$$

$$\boxed{r_{xy} = 0.789}$$

$n = 10$ $\quad \bar{x} = 65$ $\quad \bar{y} = 66$

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
78	84	13	18	169	324	234
36	51	-29	-15	841	225	435
98	91	33	25	1089	625	825
25	60	-40	-6	1600	36	240
75	68	10	2	100	4	20
82	62	17	-4	289	16	-68
90	86	25	20	625	400	500
62	58	-3	-8	9	64	24
65	53	0	-13	0	169	0
39	47	-26	-19	676	361	494
$\Sigma x =$ 650	$\Sigma y =$ 660			$\Sigma(x-\bar{x})^2 =$ 5398	$\Sigma(y-\bar{y})^2 =$ 2224	$\Sigma(x-\bar{x})(y-\bar{y}) =$ 2734

Q. 3:- Find the coefficient of correlation between the values of x and y

x	1	3	5	7	8	10
y	8	12	15	17	18	20

$n = 6$

x	y	x^2	y^2	xy
1	8	1	64	8
3	12	9	144	36
5	15	25	225	75
7	17	49	289	119
8	18	64	324	144
10	20	100	400	200
$\Sigma x = 34$	$\Sigma y = 90$	$\Sigma x^2 = 248$	$\Sigma y^2 = 1446$	$\Sigma xy = 582$

$$r_{xy} = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

$$r_{xy} = \frac{6 \times 582 - 34 \times 90}{\sqrt{6 \times 248 - (34)^2}\sqrt{6 \times 1446 - (90)^2}}$$

$$r_{xy} = 0.9879$$

Q. 4:- *A computer while calculating correlation co – efficient between two varibales X and Y from 25 pairs of observations obtained the following results :*

$n = 25, \quad \sum X = 125, \quad \sum X^2 = 650$

$\sum Y = 100, \quad \sum Y^2 = 460, \quad \sum XY = 508.$

It was, however, later discoveredе at the time of checking that he had copied down two pairs as

X	Y
6	14
8	6

while the correct values were

X	Y
8	12
6	8

Obtain the correct value of correlation co – efficient

$n = 25$

Corrected values

$\Sigma X = 125 - 6 - 8 + 8 + 6 = 125$

$\Sigma Y = 100 - 14 - 6 + 12 + 8 = 100$

$\Sigma x^2 = 650 - 36 - 64 + 64 + 36$

$\Sigma x^2 = 650$

$\Sigma y^2 = 460 - 196 - 36 + 144 + 64$

$\Sigma y^2 = 436$

$$\Sigma XY = 508 - 6\times14 - 8\times6 + 8\times12 + 6\times8$$

$$\Sigma XY = 520$$

$$r_{xy} = \frac{n\Sigma XY - \Sigma X \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$r_{xy} = \frac{25\times520 - 125\times100}{\sqrt{25\times650 - (125)^2}\sqrt{25\times436 - (100)^2}}$$

$$\boxed{r_{xy} = 0.67}$$

Q. 5:- Find the coefficient of correlation for the following table: **(UPTU – 2011)**

x:	10	14	18	22	26	30
y:	18	12	24	6	30	36

x	y	$d_x = x-22$	$d_y = y-24$	$u = \frac{x-22}{4}$	$v = \frac{y-24}{6}$	u^2	v^2	uv
10	18	−12	−6	−3	−1	9	1	3
14	12	−8	−12	−2	−2	4	4	4
18	24	−4	0	−1	0	1	0	0
22	6	0	−18	0	−3	0	9	0
26	30	4	6	1	1	1	1	1
30	36	8	12	2	2	4	4	4
$\Sigma x =$ 120	$\Sigma y =$ 126			$\Sigma u = -3$	$\Sigma v = -3$	$\Sigma u^2 = 19$	$\Sigma v^2 =$ 19	$\Sigma uv =$ 12

$n = 6$

$$r_{xy} = \frac{n\Sigma uv - \Sigma u \Sigma v}{\sqrt{n\Sigma u^2 - (\Sigma u)^2}\sqrt{n\Sigma v^2 - (\Sigma v)^2}}$$

$$r_{xy} = \frac{6 \times 12 - (-3)(-3)}{\sqrt{6 \times 19 - (-3)^2}\sqrt{6 \times 19 - (-3)^2}}$$

$$r_{xy} = \frac{72 - 9}{\sqrt{105} \times \sqrt{105}} = \frac{63}{105}$$

$$r_{xy} = 0.6$$

Topic : Correlation and Coefficient of Correlation

Q. 1 *Find the correlation coefficient between x and y for the following data:*

x	60	34	40	50	45	41	22	43
y	75	32	34	40	45	33	12	30

Q. 2 *Calculate the coefficient of correlation between the following ages of husband* (x) *and wife* (y) *by taking* 30 *and* 28 *as assumed mean in case of x and y respectively:*

x:	24	27	28	28	29	30	32	33	35	35	40
y:	18	20	22	25	22	28	28	30	27	30	32

GATEWAY CLASSES

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

GATEWAY CLASSES