0 1 n 1

Unit - V

PAGE NO :
DATE : / /
1

Curve fitting

The procedure of evaluating unknown constant with help of given data is called curve fitting

If y_i is actual value & $f(x_i)$ is estimated value $y_i - f(x_i)$ is called error.

We are minimizing of sq. of deviation the method is known as least square method.

$$S = \sum (y_i - f(x_i))^2$$

Fitting of Straight line

Let us consider straight line

$$y = a + bx$$

By principle of least sq.

$$S = \sum [y_i - (a+bx_i)]^2$$ where a and b are arbitrary constant

$\partial S/\partial a = 0 = -2 \sum y_i - (a+bx_i)$ $\partial S/\partial b = -2 \sum_{i=1}^{n} x_i (y_i - a - bx_i)$

$$\sum y_i = an + b \sum x_i \quad \text{—} \; 1$$

$$\sum x_i y_i = a \sum x_i + b \sum x_i^2 \quad \text{—} \; 2$$

n is no of observation in ~~point~~ pairs.

1 and 2 is called Normal equation.

If estimate −ve → Below straight line
+ve → Above straight line

$$\frac{\partial S}{\partial a} = 2 \sum [y_i - (a+bx_i)] = 0.$$

$$\frac{\partial S}{\partial b} = -2 \sum (y_i - (a+bx_i))$$

$$\sum y_i = \sum_i (a+bx_i) = an + b \sum x_i$$

PAGE NO. :
DATE : / /

(2)

Q) Fit a straight line to given data

x :	1	2	3	4	5	15
y :	35	68	100	138	170	511

Solⁿ Let eqⁿ of straight line be

$y = \cancel{ax+b}\ a + bx$.

Normal eqⁿs are

$\Sigma y = an + b\Sigma x$ — (2)

$\Sigma xy = a\Sigma x + b\Sigma x^2$ — (3)

x	y	xy	x^2
1	35	35	1
2	68	136	4
3	100	300	9
4	138	552	16
5	170	850	25
15	511	1873	55

$\Sigma x = 15$
$\Sigma y = 511$
$n = 5$
$\Sigma xy = 1873$
$\Sigma x^2 = 55$

$511 = 5a + 15b$
$1873 = 15a + 55b$.

$5a + 15b - 511 = 0$
$15a + 55b - 1873 = 0$

$$\frac{a}{-28095 + 28105} = \frac{b}{-7665 + 9365} = \frac{1}{275 - 225}$$

$a = 1/5$

$b = 34$

1700

$y = a + bx$

Scanned with CamScanner

③

PAGE NO.:
DATE: / /

Parabola.

$y = a + bx + cx^2$

Normal eqn

$\Sigma y = an + b\Sigma x + c\Sigma x^2$

$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$

$\Sigma x^2y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$

fit a curve $y = ax^b$ to following.

x :	1	2	3	4	5	6
y :	1200	1900	600	200	110	50

Sol) Take log

$y = ax^b$

$\log y = \log a + b\log x$.

let $Y = \log y$

$A = \log a$

$X = \log x$.

$Y = A + bX$.

$\Sigma Y = nA + b\Sigma X$.

$\Sigma XY = n\Sigma X + b\Sigma X^2$

x	y	$X = \log x$	$Y = \log y$	XY	X^2
1	1200	0	3.0991	0	0
2	1900	0.3010	2.9542	~~0.9868~~ 0.8892	0.09060
3	600	0.4771	2.7781	1.3254	0.2276
4	200	0.6021	2.3010	1.3852	0.3625
5	110	0.6989	2.0413	1.4269	0.48846
6	50	0.7782	1.6990	1.32229	0.60559

(4)

$\Sigma X = 2.8574$
$\Sigma Y =$
Σ & $\Sigma XY =$

Correlation : May be +ve/-ve
if $X \uparrow \rightarrow Y \uparrow$
if $X \downarrow \rightarrow Y \downarrow$

Karl Pearson coefficient of correlation
It is measure of degree of linear relationship b/w the variable.

given by $r(x,y) = \dfrac{cov(x,y)}{\sigma_x \sigma_y}$

$\sigma_x \rightarrow$ Standard deviation of X
$\sigma_y \rightarrow$ " " of Y

$cov(x,y) = E(XY) - E(X)E(Y)$

and
$$\boxed{r(x,y) = \dfrac{\frac{1}{n}\Sigma XY - \bar{X}\bar{Y}}{\sqrt{\frac{\Sigma X^2}{n} - \bar{X}^2}\sqrt{\frac{\Sigma Y^2}{n} - \bar{Y}^2}}}$$

if $r(X,Y) = -1$ correlation is perfect & -ve
" $= +1$ +ve
" $= 0$ X and Y are uncorelated
This indicates absence of linear relationship b/w the variable. Non-linear relationship may exist

(3)

~~$y = a+bx = 0.7+34x$~~.

Parabola

$y = a+bx+cx^2$

Form:

Q Calculate correlation coefficient for following height (in inches) of father (X) and their sons (Y)

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

Sol^n

X	Y	$U = X-68$	$V = Y-69$	U^2	V^2	UV
65	67	-3	-2	9	4	6
66	68	-2	-1	4	1	2
67	65	-1	-4	1	16	4
67	68	-1	-1	1	1	1
68	72	0	3	0	9	0
69	72	1	3	1	9	3
70	69	2	0	4	0	0
72	71	4	2	16	4	8

$$r(U,V) = \frac{\frac{1}{n}\Sigma UV - \bar{U}\bar{V}}{\sqrt{\frac{\Sigma U^2}{n} - \bar{U}^2}\sqrt{\frac{\Sigma V^2}{n} - \bar{V}^2}}$$

$\bar{U} = 0 \quad \bar{V} = 0$

$$r(U,V) = \frac{\frac{1}{8}\,24}{\sqrt{\frac{36}{8}\,\frac{44}{8}}} = \frac{2\sqrt{11}}{11}$$

$\Sigma UV = 24$

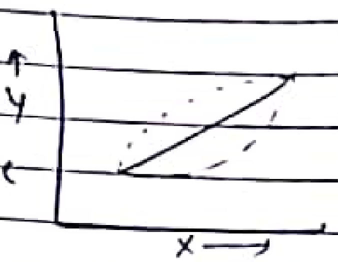
$\Sigma U^2 = 36$

$\Sigma V^2 = 44$

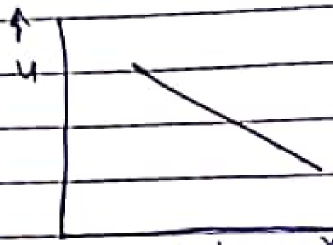
(6)

Scatter Diagram: It is graph that is used to plot data pts for 2 variables

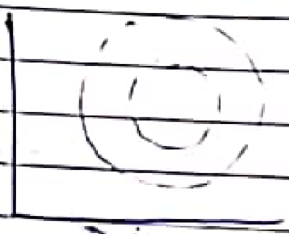
DATE: / /



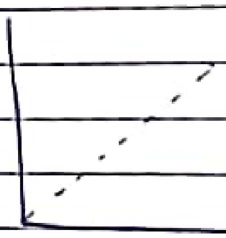
r is +ve



r is negative



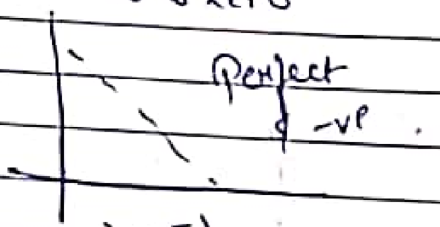
r is zero



r = ±1

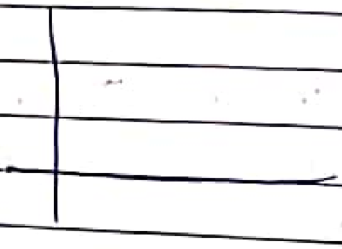
Perfect +ve positive correlation

x : 2 3 3 4 5
y : 4 6 8 8 10



Perfect -ve

r = -1



Theorem: Correlation coefficient is independent of change of origin & scale

$U = \frac{X-a}{h}$ $\quad V = \frac{Y-b}{k}$ $\quad$ h and k are +ve

Prove ~~for~~ ~~r_{xy}~~ $r_{UV} = r_{xy}$

$X = a + hU \qquad Y = b + kV.$

$E(X) = E(a+hU) \qquad E(Y) = E(b+kV)$

$= E(a) + h\,E(U) \qquad = E(b) + k\,E(V)$

$= a + h\,E(U)$

$X - E(X) = h[U - E(U)]$

$Y - E(Y) = k[V - E(V)]$

correlation

$$\sigma_x^2 = E\left[(X - E(X))^2\right]$$
$$= h^2 E\left[(U - E(U))^2\right]$$

$$\sigma_x^2 = h^2 \sigma_u^2 \qquad (h > 0)$$

$$\sigma_y^2 = k^2 \sigma_v^2 \qquad (k > 0)$$

$$\Rightarrow \boxed{\sigma_x = h\sigma_v} \qquad \boxed{\sigma_y = k\sigma_v}$$

$$r_{xy} = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \frac{E\left[(X-E(X))(Y-E(Y))\right]}{\sigma_x \sigma_y}$$

$$= \frac{hk\, Cov(U,V)}{\sigma_x \sigma_y}$$

$$= \frac{hk\, Cov(U,V)}{hk\, \sigma_u \sigma_v}$$

$$= \frac{Cov(U,V)}{\sigma_u \sigma_v} = r_{uv} = \frac{E\left[U-E(U)\right]\left[V-E(V)\right]}{\sigma_u \sigma_v}$$

Theorem:
Two independent variables are uncorrelated

Let X and Y be two independent variable.

$$r(x,y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \frac{E(XY) - E(X)E(Y)}{\sigma_x \sigma_y}$$

$$E(XY) = E(X)E(Y) \quad \because X \text{ and } Y \text{ are independent}$$

PAGE NO. :
DATE : / /

(b)

Converse of theorem is not true i.e two uncorrelated variable may not be independent

X	-3	-2	-1	1		2	3
Y	9	4	1	1		4	9
XY	-27	-8	-1	1		8	27

$$r = \frac{\frac{1}{n}\sum XY - \bar{X}\bar{Y}}{\sqrt{\frac{\sum X^2}{n} - \bar{X}^2}\sqrt{\frac{\sum Y^2}{n} - \bar{Y}^2}}$$

$$= \frac{0 - 0}{\sqrt{\,''}\sqrt{\,''}} = 0.$$

$Y = X^2$ is the relation i.e. non-linear

(9)

Q1

for[4]

Scanned with CamScanner

PAGE NO.:
DATE: / /

two
dependence

3
9
27.

(9)

PAGE NO.:
DATE: / /

Q1 The r.v. X and Y are jointly Normaly distributed and

$U = X\cos\alpha + Y\sin\alpha$

$V = Y\cos\alpha - X\sin\alpha$

Show that U and V will be uncorrelated if

$\tan 2\alpha = \frac{2r\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}$

Sol^n $\text{Cov}(U,V) = E[(U-E(U))(V-E(V))]$

$E(U) = E(X\cos\alpha + Y\sin\alpha)$
$= \cos\alpha\, E(X) + \sin\alpha\, E(Y)$

$E(V) = E(X\cos\alpha - \overset{Y}{X}\sin\alpha)$
$= \cos\alpha\, E(X) - \sin\alpha\, E(Y)$

$\text{Cov}(U,V) = E[(X\cos\alpha + Y\sin\alpha - \cos\alpha E(X) - \sin\alpha E(Y)) \times (Y\cos\alpha - X\sin\alpha - \cos\alpha E(Y) + \sin\alpha E(X))]$

$= E\{[\cos\alpha(X - E(X))] + \sin\alpha[Y - E(Y)]\}$

$= E\{\cos\alpha(X - E(X)) + \sin\alpha(Y - E(Y))\ [\cos\alpha(Y - E(Y)) - \sin(X - E(X))]\}$

$= [\cos^2\alpha\ \text{Cov}(X,Y) - \sin^2\alpha\ \text{Cov}(X,Y) - \cos\alpha\sin\alpha\ E[V(X) - V(Y)]]$ ($\sigma_X^2 - \sigma_Y^2$)

$= \cos 2\alpha\ \text{Cov}(X,Y) - \sin 2\alpha\ \frac{(\sigma_X^2 - \sigma_Y^2)}{2}$

$\because$ U and V are uncorrelated
$\text{Cov}(U,V) = 0.$

Scanned with CamScanner

Karl Pearson $r = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$.

PAGE NO
DATE / /

(10) $\cos 2\alpha \, Cov(x,y) = \frac{\sin 2\alpha}{2} (\sigma_x^2 - \sigma_y^2)$

$$\tan 2\alpha = \frac{2 Cov(X,Y)}{\sigma_x^2 - \sigma_y^2} = \frac{2 r \sigma_x \sigma_y}{\sigma_x^2 - \sigma_y^2}$$

Q2 The variable X and Y are connected by $aX + bY + c = 0$. Then show correlation (coefficient) between them is -1, if signs of a and b are alike and $+1$, if signs are different.

Sol^n Given that $aX + bY + c = 0$ — (1)

$\Rightarrow E(aX + bY + c) = 0$

$\Rightarrow aE(X) + bE(Y) + c = 0$. — (2)

(1) − (2).

$$a(X - E(X)) + b(Y - E(Y)) = 0.$$

~~$a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\,\sigma_x\sigma_y$.~~

$\Rightarrow X - E(X) = \frac{-b}{a}[Y - E(Y)]$

$\Rightarrow Cov\,XY = E[(X - E(X))(Y - E(Y))]$

$$= \frac{-b}{a}\sigma_y^2.$$

and $\sigma_x^2 = E[(X - E(X))^2] = E\left[\frac{b^2}{a^2}(Y - E(Y))^2\right]$

$$\sigma_x^2 = \frac{b^2}{a^2}\sigma_y^2$$

$$r = \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_y} = \frac{\frac{-b}{a}\sigma_y^2}{\left|\frac{b}{a}\right|\sigma_y^2} \quad \left(\because \sqrt{x^2} = |x|\right)$$

(11)

exercise

Spearman's formula

Roll
Mark
Math
Stat

(11)

$r = \begin{cases} 1, & a \text{ and } b \text{ have opp sign} \\ -1, & a \text{ and } b \text{ have same sign.} \end{cases}$

exercise
Q i.e. 10.8 and exp 10.12

Rank Correlation
[Qualitative data]

Let us suppose that a group of n individual is arrange in order of merit. Rank Correlation Coefficient between them

Spearman's formula:

$$\boxed{r = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}}$$

$n \to$ is no. of observation
$d_i \to x_i - y_i$

Q) The marks secured by students in maths & stats
~~Roll No.~~

Roll No.	1	2	3	4	5	6	7	8	9
Marks in Maths	9 10	5 15	8 12	2 17	7 13	4 16	1 24	6 14	3 22
Stats	30	42	45	46	33	34	40	35	39

Calculate Rank Correlation Coefficient

~~$r = \frac{cov(x,y)}{\sigma_x \sigma_y}$~~ =

~~$\sigma_x^2 = \frac{b^2}{a^2} \sigma_y^2$~~

17

	1	2	3	4	5	6	7	8	9
Maths	9	5	8	3	7	4	1	6	2
Stats	9	3	2	1	8	7	4	6	5
d_i	0	2	6	2	-1	-3	-3	0	-3
d_i^2	0	4	36	4	1	9	9	0	9

$r = 1 - \frac{6}{9 \times 80}\left[72\right]$

$\frac{45}{}$ $\frac{27}{72}$

$r = 1 - \frac{6}{10} = \frac{4}{10} = \frac{2}{5} = 0.4$

Q. Repeated Rank.

$$r = 1 - \frac{6\left\{\sum d_i^2 + \frac{m}{12}(m^2-1)\right\}}{n(n^2-1)}$$

$m \rightarrow$ No. of times item is repeated

eg

99	1	
92	2.5	$\frac{2+3}{2}$ $m_1 = 3$
92	2.5	$m_2 = 2$
90	5	$m_3 = 2$
90	5	$\frac{4+5+6}{3}$
90	5	
88	7	
87	8	
86	9.5	$\frac{9+10}{2}$
86	9.5	
85	11	
80	12	

PAGE NO.:
DATE: / /

Regression:

Lines of Regression: It is line which gives best estimate to value of one variable for any specific value of other variable.

* There are always two line of Regression

(a) Y on X and X on Y: They passes through common ~~poit~~ points $(\bar{x}, \bar{y})$

1) Lines of regression Y on X is

$$Y - \bar{y} = \underbrace{\frac{r\sigma_y}{\sigma_x}}_{\text{Slope}} (X - \bar{x})$$

• It is used find value of y for given value of x.

2) Lines of Regression X on Y is

$$X - \bar{x} = r\frac{\sigma_x}{\sigma_y}(Y - \bar{y})$$

It is ~~vat~~ use to find value of x for given value of y

* Regression Coefficient

$b_{yx} \cdot b_{xy} = r^2$

$\boxed{b_{yx} = r\frac{\sigma_y}{\sigma_x}}$ $\boxed{b_{xy} = r\frac{\sigma_x}{\sigma_y}}$

Sign of ~~byx and r~~ and b_{xy} ~~and r~~ are same, why

(14) because b_{yx} and $b_{xy} \to$ same sign $\to$ else complex
and r depends on b_{yx}

PAGE NO :
DATE : / /

Properties of Regression Coefficient

i) $r = \pm\sqrt{b_{yx} \cdot b_{xy}}$

r, b_{yx}, b_{xy} have same sign.

ii) If one of Regression coeffecient is greater than unity the other must be less than 1.

i.e. $b_{yx} > 1$, $b_{xy} < 1$.

$\because r^2 = b_{xy} b_{yx}$ $\quad -1 < r < 1$ so $r^2 < 1$.

iii) Modulus value of AM of regression coeff. is greater than modulus value of correlation coeff.

$$\left|\frac{1}{2}(b_{yx} + b_{xy})\right| > |r|$$

iv) Regression coeff. are ind of change of origin but not so of scale

Q Given $\sigma_x^2 = 9$

Regression eq^n are

$$8X - 10Y + 66 = 0.$$
$$40X - 18Y = 214.$$

i) Find $\bar{X}$ and $\bar{Y}$

ii) $r(X,Y)$

iii) σ_y

(1)

$b_{xy} = 16$
$b_{yx} = 1/2$.

$r^2 = 8$
$r = 2\sqrt{2}$ ✗

PAGE NO.:
DATE: / /

Sol: Let regression eqⁿ Y on X is

$$8X - 10Y + 66 = 0.$$

~~40X =~~

Regression eqⁿ X on Y is
$$40X - 18Y = 214$$

∵ both the lines of Regression passing through the Mean value $\bar{x}$ and $\bar{y}$

$$8\bar{x} - 10\bar{y} + 66 = 0$$
$$40\bar{x} - 18\bar{y} - 214 = 0$$

$$\bar{x} = 13 \qquad \bar{y} = 17.$$

(ii) ~~Mean~~

$$10Y = 8X + 66.$$
$$Y = \frac{8}{10}X + \frac{66}{10}.$$

Coefficient of X = $\boxed{b_{yx} = \frac{8}{10}}$

$$X = \frac{18}{40}Y + \frac{214}{40}.$$

$$\boxed{b_{xy} = \frac{18}{40}}$$

$$r^2 = b_{xy}\, b_{yx}$$
$$r^2 = \frac{8}{10} \cdot \frac{18}{40}$$
$$r^2 = \frac{\cancel{2}\times\cancel{3}\times\cancel{3}\times\cancel{2}\times 2\times 2}{\cancel{40}\,400} \Rightarrow r = \frac{2\times 3\times \cancel{2}}{\cancel{20}\,10} = 0.6$$

1 2 3 4 5 6 7 8 9
5 3 2 1 8 7 4 6 5

Variance = +ve. $\frac{x-a}{h}$

(1)

$\sigma_y \Rightarrow b_{yx} = \frac{r\sigma_y}{\sigma_x} \Rightarrow \frac{8}{10} = \frac{6}{10}\frac{\sigma_y}{3} \Rightarrow \boxed{\sigma_y = 4}$

Q. Regression lines Y on X and X on Y are.

$Y = aX + b$ and $X = cY + d$

Show

i) $\bar{X} = \frac{bc+d}{1-ac}$, $\bar{Y} = \frac{ad+b}{1-ac}$.

ii) $r_{xy} = \sqrt{ac}$

iii) $\frac{\sigma_x}{\sigma_y} = \sqrt{\frac{c}{a}}$

Q Find line of Regression Y on X and X on Y also obtain estimate of X for Y = 70. ~~X on Y.~~

(R)

Sol Let $U = X - a$. $V = Y - b$

$\bar{X} = \bar{U} + a$ $\bar{Y} = \bar{V} + b$

$\bar{X} = 0 + 68$ $\bar{Y} = 0 + 69$

$\sigma_x = \sigma_u = \sqrt{4.5}$

$\sigma_y = \sigma_v = \sqrt{5.5}$

$r = r(X,Y) = r(U,V) = 0.6$

Regression line. Y on X

$Y - \bar{Y} = r\frac{\sigma_y}{\sigma_x}(X - \bar{X}) = (0.6)\frac{\sqrt{5.5}}{\sqrt{}}(X - 6\ldots)$

$Y = 0.665X + 23.78.$

$X = 0.54Y + 30.74$

$X = (0.54)(70) + 30.74.$

Q Show that acute angle θ b/w line of regression is

$$\tan\theta = \frac{1-r^2}{|r|} \times \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}.$$

Discuss the case when $r=0$, $r=\pm 1$.

Sol $\tan\theta = \left|\frac{m_1 - m_2}{1+m_1 m_2}\right|$ m_1 and m_2 of Y on X and X on Y.

Regression Y on X

$$Y - \bar{y} = \frac{r\sigma_y}{\sigma_x}(x - \bar{x})$$

$$m_1 = \frac{r\sigma_y}{\sigma_x}.$$

Regression X on Y.

$$X - \bar{x} = \frac{r\sigma_x}{\sigma_y}(y - \bar{y})$$

$$m_2 = \frac{1}{r}\frac{\sigma_y}{\sigma_x} \quad (\because y = mx + c).$$

$$\tan\theta = \left|\frac{\frac{r\sigma_y}{\sigma_x} - \frac{1}{r}\frac{\sigma_y}{\sigma_x}}{1 + \frac{\sigma_y^2}{\sigma_x^2}}\right| = \left|\frac{(r^2-1)\,\sigma_y}{\sigma_x^2 + \sigma_y^2} \Big/ \frac{\sigma_y}{\sigma_x}\ \sigma_x^2\right|$$

(18)

PAGE NO. :
DATE : / /

$$\tan\theta = \left|\frac{\frac{\sigma_y}{\sigma_x}\left(r - \frac{1}{r}\right)}{\frac{\sigma_x^2 + \sigma_y^2}{\sigma_x^2}}\right|$$

$$= \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}\left|\frac{r^2 - 1}{r}\right|$$

$$= \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \frac{|r^2 - 1|}{|r|}$$

$$= \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \frac{|1 - r^2|}{|r|}$$

$\because$ acute angle $1 - r^2$ so always positive

~~θ tan~~

Case I : $r = 0$. Uncorrelated.

$$\tan\theta = \frac{\cancel{\sigma_x \sigma_y}}{\cancel{0}} \infty$$

$$\theta = \pi/2$$

Lines of Regression are $\perp$ to each other

II $r = \pm 1$

$$\tan\theta = \cancel{\sigma_x \sigma_y}\, 0$$

$$\boxed{\theta = 0, \pi}$$

~~H~~ Line of Regression are coincident if $\theta = 0$
|| if $\theta = \pi$.

* Lines can't be || because of property (a)

Ch - 5

PAGE NO.:
DATE: / /

Q1 find r.

husband	x	23	27	28	28	29	30	31	33	35 · 36
Wife	y	18	20	22	27	21	29	27	29	28 · 29

X	Y	$x = X - \bar{X}$	$y = Y - \bar{Y}$	y^2	x^2	xy
23	18	-7	-7	49	49	49
27	20	-3	-5	25	9	15
28	22	-2	-3	9	4	6
28	27	-2	2	4	4	-4
29	21	-1	-4	16	1	4
30	29	0	4	16	0	0
31	27	1	2	4	1	2
33	29	3	4	16	9	12
35	28	5	3	9	25	15
36	29	6	4	16	36	24

$\bar{X} = 30 \quad \bar{Y} = 25$

$\bar{X} = \frac{\Sigma X}{N} \quad \bar{Y} = \frac{\Sigma Y}{N}$

$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}} = \frac{123}{\sqrt{138}\sqrt{164}} = 0.82$ (approx)

Q2 Comment on r

$b_{xy} = 3.2 \quad b_{yx} = 0.6$

$r^2 = b_{xy}\, b_{yx} = 3.2 \times 0.8$

$r^2 =$

Scanned with CamScanner

Q Prove Correlation — Regression

	Correlation	Regression
①	Relationship b/w 2/more variable which ~~are~~ vary in sympathy with the other in the same or opp. dxn	Mathemical measure showing the avg relationship b/w 2 variable
②	x and y : random Variable	x is random, y is fixe.
③	If the coeff. correlation is +ve then 2 variables are +vely correlated	The regression coefficient explain that ↓ in one variable is associated with ↑ in other
④	$r = \pm 1$	Absolute fig.

Q In a contest two judges ranked seven candidate in order of their preference

	A	B	C	D	E	F	G
Judge I	2	1	4	5	3	7	6
" II	3	4	2	5	1	6	7

Calculate rank correlation coefficient.

Cand	Judge I	Judge II	$d = x - y$	d^2
	2	3	−1	1
	1	4	−3	9
	4	2	2	4
	5	5	0	0
	3	1	2	4
	7	6	1	1
	6	7	−1	1

$$R = 1 - \frac{\Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{20}{7(48)} = 0.64$$

Q. Find r.

No. of items	A	B
	12	12
Arithmetic Mean	20	15
Sum of sq. of deviat from mean	180	165

Sum of product of dev. wrt AM = 90.

Sol^n $n = 12 \quad \bar{x} = 20 \quad \bar{y} = 15$

$\Sigma(x-\bar{x})^2 = 180 \quad \Sigma(y-\bar{y}) = 165$

$$SD(x) = \sqrt{\frac{1}{n}\Sigma(x-\bar{x})^2} = \sqrt{\frac{180}{12}} = 3.87$$

$$SD(y) = \sqrt{\frac{1}{n}\Sigma(y-\bar{y})^2} = \sqrt{\frac{165}{12}} = 3.71$$

$$Cov(x,y) = \frac{1}{n}\Sigma(x-\bar{x})(y-\bar{y}) = \frac{90}{15} = 7.5$$

$$r(x,y) = \frac{Cov(x,y)}{\sigma_x \sigma_y} = \frac{7.5}{3.87 \times 3.71}$$

$$= 0.5226.$$

Ex.

Q.1 If $U = ax + by$, $V = bx - ay$
Show that U and V are uncorrelated if

$\frac{ab}{a^2 - b^2} = \frac{\rho \sigma_x \sigma_y}{\sigma_x^2 - \sigma_y^2}$, ρ is correlation coefficient b/w X and Y.

$1 - r^2$
where,