Regression Cont…

Regression equation using Regression Coefficients:

1. Using Actual values of X and Y.
2. Using deviation from Actual means.
3. Using deviation from Assumed means.
4. Using r, σ_x,σ_y values

Regression using Actual Values:

Regression equation Y on X :

$$\left(Y-\overline{Y}\right)=b_{yx}\left(X-\overline{X}\right)$$

$$b_{yx}=\left(\frac{N\sum XY-\sum X\sum Y}{N\sum X^2-(\sum X)^2}\right)$$

Regression line X on Y:

$$\left(X-\overline{X}\right)=b_{xy}\left(Y-\overline{Y}\right)$$

$$b_{xy}=\left(\frac{N\sum XY-\sum X\sum Y}{N\sum Y^2-(\sum Y)^2}\right)$$

1. Calculate the regression equation of X on Y using actual values of X and Y.

X:	1	2	3	4	5
Y:	2	5	3	8	7

X	Y	XY	Y^2	X^2
1	2	2	4	1
2	5	10	25	4
3	3	9	9	9
4	8	32	64	16
5	7	35	49	25
15	25	88	151	55

Regression equation X on Y :

$$\overline{Y} = \frac{\Sigma Y}{N} = 5;\ \overline{X} = \frac{\Sigma X}{N} = 3$$

$$\left(X - \overline{X}\right) = b_{xy}\left(Y - \overline{Y}\right)$$

$$b_{xy} = \left(\frac{N\sum XY - \sum X \sum Y}{N\sum Y^2 - (\sum Y)^2}\right)$$

$$b_{yx} = \left(\frac{5 \times 88 - 15 \times 25}{5 \times 151 - 25^2}\right) = 0.5$$

$$\left(X - 3\right) = 0.5\left(Y - 5\right)$$

$$X = 0.5Y + 0.5$$

$$\text{N} = 5,\ \sum X = 15,\ \sum X^2 = 55, \sum Y = 25, \sum Y^2 = 151,\ \sum XY = 88$$

Regression equation Y on X :

$$\overline{Y} = \frac{\Sigma Y}{N} = 5;\ \overline{X} = \frac{\Sigma X}{N} = 3$$

$$\left(Y - \overline{Y}\right) = b_{yx}\left(X - \overline{X}\right)$$

$$b_{yx} = \left(\frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2}\right)$$

$$b_{yx} = \left(\frac{5 \times 88 - 15 \times 25}{5 \times 55 - 15^2}\right) = 1.3$$

$$\left(Y - 5\right) = 1.3\left(X - 3\right)$$

$$Y = 1.3X + 1.1$$

2. **Calculate the regression equation of Yon X and X on Y, Using deviation from Actual means**.

X	:	2	4	6	8	10	12
Y	:	4	2	5	10	3	6

$$\left(Y - \overline{Y}\right) = b_{yx}\left(X - \overline{X}\right)$$

$$b_{yx} = \left(\frac{\sum\left(X - \overline{X}\right)\left(Y - \overline{Y}\right)}{\Sigma\left(X - \overline{X}\right)^2}\right)$$

X	Y	$(X-\overline{X})=(X-7)$	$(Y-\overline{Y})=(Y-5)$	$(X-\overline{X})(\mathrm{Y}-\overline{\mathrm{Y}})=(X-7)(\mathrm{Y}-5)$	$(X-\overline{X})^2=(X-7)^2$	$(Y-\overline{Y})^2=(Y-5)^2$
2	4	-5	-1	5	25	1
4	2	-3	-3	9	9	9
6	5	-1	0	0	1	0
8	10	1	5	5	1	25
10	3	3	-2	-6	9	4
12	6	5	1	5	25	1
42	30	0	0	18	70	40

$$\overline{X}=\frac{\Sigma X}{N}=\frac{42}{6}=7; \quad \overline{Y}=\frac{\Sigma Y}{N}=\frac{30}{6}=5$$

$$(Y-5)=b_{yx}(X-7)$$

$$b_{yx}=\left(\frac{\Sigma(X-7)(Y-5)}{\Sigma(X-7)^2}\right)=\frac{18}{70}=0.257$$

$$(Y-5)=0.257\,(X-7)$$

$$Y=0.287X+3.201$$

$\text{Re}\,gression\ X on\ Y$

$$(X-5)=b_{xy}(Y-7)$$

$$b_{xy}=\left(\frac{\Sigma(X-7)(Y-5)}{\Sigma(Y-5)^2}\right)=\frac{18}{40}=0.45$$

$$(X-7)=0.45\,(Y-5)$$

$$X=0.45X+4.75$$

Regression equation using Regression Coefficients, Using deviation from Assumed means.

Regression equation Y on X :

$$(Y-\overline{Y})=b_{yx}(X-\overline{X})$$

$$b_{yx}=\left(\frac{N\sum dX\,dY-\sum dX\sum dY}{N\sum dX^2-(\sum dX)^2}\right)$$

$where\quad dX=X-A_x,\ and\quad A_x\ \ is\ the\ assumed\ mean\ of\ X.$

$dY=X-A_y,\ and\quad A_y\ \ is\ the\ assumed\ mean\ of\ Y$

Regression line X on Y

$$(X-\overline{X})=b_{xy}(Y-\overline{Y})$$

$$b_{xy}=\left(\frac{N\sum dX\,dY-\sum dX\sum dY}{N\sum dY^2-(\sum dY)^2}\right)$$

$where\quad dX=X-A_x,\ and\quad A_x\ \ is\ the\ assumed\ mean\ of\ X.$

$dY=X-A_y,\ and\quad A_y\ \ is\ the\ assumed\ mean\ of\ Y$

:

3. Calculate the regression of Y on X and X on Y by using deviations from assumed mean.

X	:	78	89	97	69	59	79	68	61
Y	:	125	137	156	112	107	136	124	108

X	Y	dX= X - A_x = X - 69	dY=Y- A_y =Y - 112	dX^2	dy^2	dX dY
78	125	9	13	81	169	117
89	137	20	25	400	625	500
97	156	28	44	784	1936	1232
69	112	0	0	0	0	0
59	107	-10	-5	100	25	50
79	136	10	24	100	576	240
68	124	-1	12	1	144	-12
61	108	-8	-4	64	16	32
		48	109	1530	3491	2159

$$\overline{X} = A_X + \frac{\Sigma dX}{N} = 69 + \frac{48}{8} = 75; \qquad \overline{Y} = A_Y + \frac{\Sigma dY}{N} = 112 + \frac{109}{8} = 125.625$$

$$b_{yx} = \left(\frac{8\times 2159 - 48\times 109}{8\times 1530 - 48^2}\right) = 1.21$$

$$(Y - 125.625) = 1.21(X - 75)$$

$$Y = 1.21X + 34.87$$

$$b_{yx} = \left(\frac{8\times 2159 - 48\times 109}{8\times 3491 - 109^2}\right) = 0.75$$

$$(X - 75) = 0.75(Y - 125.625)$$

$$Y = 1.21X + 34.87$$

Regression Equation using Regression coefficients , using Standard deviation

$$\text{Regression Equation } \; Y \; on \; X:$$

$$\left(Y - \overline{Y}\right) = b_{yx}\left(X - \overline{X}\right) \qquad where$$

$$\text{Regression Equation } \; X \; on \; Y:$$

$$\left(X - \overline{X}\right) = b_{xy}\left(Y - \overline{Y}\right) \qquad where \qquad b_{xy} = r\frac{\sigma_x}{\sigma_y}$$

4. Estimate the value of Y when X = 9

	X	Y
Arithmetic Mean	5	12
Standard Deviation	2.6	3.6
Correlation Coefficient	0.7	

$\text{Regression Equation}\quad Y \ \ on \ \ X:$

$$\left(Y-\overline{Y}\right)=b_{yx}\left(X-\overline{X}\right) \qquad where \qquad b_{yx}=r\frac{\sigma_y}{\sigma_x}$$

$$b_{yx}=0.7\frac{3.6}{2.6}=0.9692$$

$$\therefore \left(Y-12\right)=0.9692\left(X-5\right)$$

$$Y=0.9692X+7.15$$

$$When \quad x=9$$

$$Y=0.9692\times 9+7.5=15.87$$

5. If the average of X and Y are 25 and 120 and b_{xy}=2.Estimate the value of X when Y=30.

$$\left(X-\overline{X}\right)=b_{xy}\left(Y-\overline{Y}\right)$$

$$\left(X-25\right)=2\left(Y-120\right) \qquad (1),$$

$$When\ Y=130, using\ in(1),\ X=45$$

6.Given two regression lines 3X + 4 Y = 44; 5X+ 8Y = 8, Variance of X is 30. Find mean

of X and Y and also r and σ_y

$$3X+4Y=44, \text{Regression line } X \text{ on } Y$$

$$so\ b_{xy}=-\frac{4}{3}$$

$$5X+8Y=80, \text{ Regression line } Y \text{ on } X$$

$$so\ b_{yx}=-\frac{5}{8}$$

Since $(\overline{X}\ \overline{Y})$ *is a point lies on the line. It should satisfy both the equations. Solving the equations the solution is* $(\overline{X},\overline{Y})$ *and* $\overline{X}=8,\ \overline{Y}=5$

$$r^2=b_{xy}\times b_{yx}=-\frac{4}{3}\times-\frac{5}{8}=0.82$$

$$\therefore\ r=\sqrt{0.82}.$$

$$b_{xy}=-\frac{4}{3}=\frac{r\sigma_x}{\sigma_y}\Rightarrow\sigma_y=3.73$$