M- III (P&S)

UNIT- H Assignment Questions

① Define correlation, Types of correlation, Regression and Regression Equations.

Correlation

⇒ A relation between two or more variables is called as co-relation

⇒ Correlation is a statistical measure for finding out the degree of association between two variables.

(or)

⇒ The relationship between two variables such that a change in one variable results in positive or the negative change in the other is called correlation

(or)

⇒ A greater change in one variable resulting in a corresponding greater or smaller change in the other

variable is also called correlation

eg:- Rainfall & crop yield.

Types of correlation

① positive (+), negative (-) correlation

→ When two variables tends to move together in the same direction then the correlation is called positive (+) correlation other wise negative (-) correlation

eg:- ① income & expenditure

② height & weigh

③ age & IQ

④ Demand & supply

② Simple & Multiple

→ When we study only two variable say x, y then relationship is desired as simple co-relation.

→ If we study more than two variable's simultaniously then the co-relation is called Multiple correlation.

Partial & Total

→ If the study of the variable excludes some other variable's then it is called partial correlation

→ If all-facts are taken into the accounts then the co-relation is called total co-relation.

Regression

⇒ In regression we can estimate the value of one variable with the value of another variable which is know

⇒ The statical method which estimate the unknow value of one variable from the know value of the related variable is called regression.

Regression Equations

Regression Equation x on y

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

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

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

Regression Equation y on x

$$y-\bar{y} = b_{yx}(x-\bar{x})$$

$$b_{yx} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

② Find coefficient of correlation and regression equation from the following data.

x	92	89	87	86	83	77	71	63	53	50
y	86	88	91	77	68	85	52	82	37	57

Sol coefficient of correlation given by

$$r = \frac{\Sigma dx\, dy}{\sqrt{dx^2 dy^2}} \Rightarrow \frac{1843.7}{\sqrt{(2086.9)(2972.1)}} \Rightarrow 0.7402$$

$$\bar{x} = \frac{\Sigma x}{n} \Rightarrow \frac{751}{10} \Rightarrow \underline{75.1}$$

$$\bar{y} = \frac{\Sigma y}{n} \Rightarrow \frac{723}{10} \Rightarrow \underline{72.3}$$

computation Table

x	$dx = x - \bar{x}$	$dx^2 = (x-\bar{x})^2$	y	$dy = y - \bar{y}$	$dy^2 = (y-\bar{y})^2$	$dxdy = (x-\bar{x})(y-\bar{y})$
92	92 − 75.1 = 16.9	285.61	86	86 − 72.3 = 13.7	187.69	231.53
89	13.9	193.21	88	15.7	246.49	218.23
87	11.9	141.61	91	18.7	349.69	222.53
86	10.9	118.81	77	4.7	22.09	51.23
83	7.9	62.41	68	−4.3	18.49	−33.97
77	1.9	3.61	85	12.7	161.29	24.13
71	−4.1	16.81	52	−20.3	412.09	83.23
63	−12.1	146.41	82	9.7	94.09	−117.37
53	−22.1	488.41	37	−35.3	1246.09	780.13
50	−25.1	630.01	57	−15.3	234.09	~~780.13~~ 384.03

$\Sigma x = 751$ $\Sigma dx = 0$ $\Sigma dx^2 = 20869$ $\Sigma y = 723$ $\Sigma dy = 2972.1$ $\Sigma dxdy = 1843.7$

Regression Eq

x on y → $x - \bar{x} = b_{xy}(y - \bar{y})$ -①

$$b_{xy} = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{\Sigma (y-\bar{y})^2} \Rightarrow \frac{1843.7}{2972.1}$$

$$\boxed{b_{xy} = 0.6203}$$

Sub in ①

$x - 75.1 = 0.6203(y - 72.3)$

$x = 75.1 + 0.6203y - 44.84769$

$x = 0.6203y + 30.25231$

y on x

$\Rightarrow y - \bar{y} = b_{yx}(x - \bar{x})$

$y - 72.3 = b_{yx}(x - 75.1)$

$$b_{yx} = \frac{\Sigma (y-\bar{y})(x-\bar{x})}{\Sigma (x-\bar{x})^2} \Rightarrow \frac{1843.7}{2086.9}$$

$b_{yx} \Rightarrow 0.8834$

$y - 72.3 = 0.8834x - 66.34334$

$y = 72.3 + 0.8834x - 66.34334$

$y = 5.95666$ //

3. Find the rank correlation co-efficient from the following data

x	10.	15.	12.	17.	13.	16	24	14	22
y	30	42	45	46	33	34	40	35	39

Sol

Rank correlation co-efficient is given by

$$\rho = 1 - \frac{6 \leq D^2}{n(n^2-1)} \Rightarrow 1 - \frac{6(72)}{9(80)}$$

$\Rightarrow 1 - 0.6$

$\Rightarrow 0.4$,

X	Rank of X	Y	Rank of Y	D = X-Y	$D^2 = (X-Y)^2$
10	9	30	9	0	0
15	5	42	3	2	4
12	8	45	2	6	36
17	3	46	1	2	4
13	7	33	8	-1	1
16	4	34	7	-3	9
24	1	40	4	-3	9
14	6	35	6	0	0
22	2	39	5	-3	9

$\Sigma D^2 = 72$

(H) Using the method of least squares fit a curve of the form $y = a + bx + cx^2$ to the following data

x	-1	1	2	3	5
y	8	0	5	16	56

Sol Consider $y = a + bx + cx^2$

Whose Normal eq's are

$$\Sigma y = na + b\Sigma x + c\Sigma x^2 \quad -(1)$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3 \quad -(2)$$

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4 \quad -(3)$$

Computation table

x	y	x^2	x^3	x^4	xy	x^2y
-1	8	1	-1	1	-8	8
1	0	1	1	1	0	0
2	5	4	8	16	10	20
3	16	9	27	81	48	144
5	56	25	125	625	280	1400
10	85	40	160	724	330	1572

$85 = 5a + 10b + 40c$

$330 = 10a + 40b + 160c$

$1572 = 40a + 160b + 724c$

we get $a = 1$

Solving (1) & (2) & (3)

$b = -4, \; c = 3$

$$y = 1 - 4x + 3x^2$$

Q5 Fit a least square curve of the form $y = ax^b$ for the following data

x	61	26	7	26
y	350	400	500	600

Sol Consider $y = ab^x$ whose normal eq's are

$$\Sigma y = nA + B\Sigma x \quad -(1)$$

$$\Sigma XY = A\Sigma X + B\Sigma x^2 \quad -(2)$$

X	Y	$y = \log_{10} y$	X^2	XY
61	350	2.5440	3721	155.184
26	400	2.6020	676	67.652
7	500	2.6989	49	18.8923
26	600	2.7781	676	72.2306
120	1850	10.623	5122	313.9589

$$10.623 = 4A + 120B$$

$$313.9589 = 120A + 5122B$$

$A = 2.7490$ $\quad B = -3.108 \times 10^{-3} = -0.00310$

$a = 561.04$ $\quad b = 0.992$

$$y = (561.04)(0.992)^x$$