ASSIGNMENT III

Probability and statistics

Parneet Kaur
RA2311056010122

Students	1	2	3	4	5	6	7	8	9	10
(X) Marks in maths	78	36	98	25	75	82	90	62	65	39
(Y) Marks in physics	84	51	91	60	68	62	86	58	63	47

To find:- Rank correlation coefficient

X	Y	Rank X (x)	Rank Y (y)	d = x − y	d^2
78	84	4	3	1	1
36	51	9	9	0	0
98	91	1	1	0	0
25	60	10	7	3	9
75	68	5	4	1	1
82	62	3	6	−3	9
90	86	2	2	0	0
62	58	7	8	−1	1
65	63	6	5	1	1
39	47	8	10	−2	4
					26

$$\rho = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 26}{10(10^2-1)}$$

$$= 0.843$$

Rank correlation coefficient is 0.843.

2) b)

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

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
		$\Sigma u = 0$	$\Sigma v = 0$	$\Sigma u^2 = 36$	$\Sigma v^2 = 44$	$\Sigma uv = 24$

$$\bar{X} = \frac{\Sigma x}{n} = \frac{544}{8} = 68$$

$$\bar{Y} = \frac{\Sigma y}{n} = \frac{552}{8} = 69$$

$$r_{xy} = r_{uv} = \frac{Cov(u,v)}{\sigma_u \sigma_v}$$

$$Cov(u,v) = \frac{\Sigma uv}{n} - \left(\frac{\Sigma u}{n}\right)\left(\frac{\Sigma v}{n}\right) = \frac{24}{8} - \left(\frac{0}{8}\right)\left(\frac{0}{8}\right) \Rightarrow 3.$$

$$\sigma_u = \sqrt{\frac{\Sigma u^2}{n} - \left(\frac{\Sigma u}{n}\right)^2} = \sqrt{\frac{36}{8} - (0)^2} = 2.12$$

$$\sigma_v = \sqrt{\frac{\Sigma v^2}{n} - \left(\frac{\Sigma v}{n}\right)^2} = \sqrt{\frac{44}{8} - (0)^2} = 2.345$$

$$r_{uv} = \frac{3}{2.12 \times 2.345} = 0.6036$$

X on Y

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

$$x - (\bar{u} + 68) = r_{uv} \frac{\sigma_u}{\sigma_v} (Y - (\bar{v} + 69))$$

$\because \bar{u} = \frac{\Sigma u}{n}$

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

$$x - 68 = 0.6 \times (Y - 69) \times 0.9$$

$$x - 68 = (0.6Y - 41.4) \times 0.9$$

~~$x - 0.6Y = -41.4 + 68$~~ $x - 68 = 0.54Y - 37.26$

~~$x - 0.6Y = 26.6$~~ $\boxed{x = 0.54Y + 30.74}$

~~$\boxed{x = 0.6Y + 26.6}$~~

Y on X

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

$$y - (\bar{v} + 69) = r_{uv} \frac{\sigma_v}{\sigma_u} (x - (\bar{x} + 68))$$

$$y - 69 = 0.6 \times \frac{2.345}{2.12} \times (x - 68)$$

$$y - 69 = 0.66 \times (x - 68)$$

$$y - 69 = 0.66x - 44.88$$

$$\boxed{y = 0.66x + 24.12}$$

Value of X when Y = 70

Find x:- X on Y

$$X = 0.54Y + 30.74$$

$$x = 0.54 \times 70 + 30.74$$

$$\boxed{x = 68.54}$$

Value of Y when X = 71

$$y = 0.66 \times 71 + 24.12$$

$$\boxed{y = 70.98}$$

2 way

Makes

A	B	C
5	8	7
6	10	3
8	11	5
9	12	4
7	4	1

H_0: There is no significant difference

H_1: There is a significant difference

x_1	x_2	x_3	Total	x_1^2	x_2^2	x_3^2
5	8	7	20	25	64	49
6	10	3	19	36	100	9
8	11	5	24	64	121	25
9	12	4	25	81	144	16
7	4	1	12	49	16	1
$\Sigma x_1 = 35$	$\Sigma x_2 = 45$	$\Sigma x_3 = 20$	$T = 100$	$\Sigma x_1^2 = 255$	$\Sigma x_2^2 = 445$	$\Sigma x_3^2 = 100$

N = Total no. of items in a given data $\Rightarrow 15$

T = Sum of all items $= 100$

$\frac{T^2}{N}$ = Correction factor (CF) $= \frac{(100)^2}{15} = 666.66$

$SS = (\Sigma x_1) \quad \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$

$255 + 445 + 100 - 666.66$

$= 133.34$

$SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N}$

$= \frac{(35)^2}{5} + \frac{(45)^2}{5} + \frac{(20)^2}{5} - 666.66$

$245 + 405 + 80 - 666.66$

63.34

$SSE = TSS - SSC$

$= 133.34 - 63.34$

$= 70$

C = No. of samples = 3, $C-1 = 3-1 = 2$

$N - C = 15 - 3 = 12$

$$MSC = \frac{SSC}{C-1} = \frac{63.34}{2} = 31.67$$

$$MSE = \frac{SSE}{N-C} = \frac{70}{12} = 5.83$$

ANOVA TABLE

Source of Variation	sum of squares	Degree of freedom	Mean
B/w treatments	SSE = 63.34	C−1 = 2	MSC = 31.67
errors	SSE = 70	N−C = 12	MSE = 5.83

Variance Ratio

$F = \frac{MSC}{MSE} = 5.43$

Table value at 5% significance

3.89

Calculated F > Tabulated f,

H_0 Rejected.

24b

Workers	A	B	C	D
	44	38	47	36
	46	40	52	43
	34	36	44	42
	43	38	46	33
	38	42	49	39

a) H_0: All workers have the same mean productivity.

H_1: There is a significant diffrence in mean productivity.

A	B	C	D	Row sum	Row Mean
44	38	47	36	165	41.25
46	40	52	43	181	45.25
34	36	44	32	146	36.5
43	38	46	33	160	40
38	42	49	39	168	42
$\Sigma A = 205$	$\Sigma B = 194$	$\Sigma C = 238$	$\Sigma D = 183$		
41	38.8	47.6	36.6		

Mean

$\bar{A} = 41$

Mean production for each worker

Worker 1: $\frac{44+38+47+36}{4} = 41.25$

Worker 2: $\frac{46+40+52+43}{4} = 45.25$

Worker 3: $\frac{34+36+44+33}{4} = 36.75$

Worker 4: $\frac{43+38+46+33}{4} = 40$

Worker 5: $\frac{38+42+49+39}{4} = 42$

Grand Mean (GM)

GM = $\frac{\text{Sum of all observations}}{20} = \frac{205+194+238+183}{20}$

$= 41.0$

Calculate SSW:

$4(41.25-41)^2 + 4(45.25-41)^2 + 4(36.25-41)^2 + 4(40-41)^2 + 4(42-41)^2$

$4(0.0625 + 20.25 + 18.0625 + 1 + 1)$

$= 143.5 + 18$

$= 161.5$

Now, Mean production for each machine

Machine A:- $\frac{44+46+34+43+38}{5} = 41$

Machine B:- $\frac{38+40+36+38+42}{5} = 38.8$

Machine C:- $\frac{47+52+44+46+49}{5} = 47.6$

Machine D:- $\frac{36+43+33+33+39}{5} = 36.8$

Calculate SSM:-

$5(41-41)^2 + 5(38.8-41)^2 + 5(47.6-41)^2 + 5(36.8-41)^2$

$= 322 + 18.8 = 338.8$

Calculating SST:-

$(44-41)^2+(46-41)^2+(34-41)^2+(43-41)^2+(38-41)^2+$
$(38-41)^2+(40-41)^2+(36-41)^2+(38-41)^2+(42-41)^2+$
$(47-41)^2+(52-41)^2+(44-41)^2+(46-41)^2+(49-41)^2+$
$(36-41)^2+(43-41)^2+(32-41)^2+(33-41)^2+(39-41)^2$

$9+25+49+4+9+9+1+25+9+1+36+121+$
$9+25+40\ 64+25+4+81+64+4$
$= 574.$

Sum of squares for Rows (SSR) - worker effect

$$SSR = \frac{1}{c}\sum_{i=1}^{r}(R_i - GM)^2$$

$SSE \Rightarrow 574 - 161.5 - 338.8 = 73.7$

Degree of freedom for workers

$df_w = 5-1 = 4$

For machines $df_m = 4-1 = 3$

For error $df_e = (4)(3) = 12$

Total $df_t = 5\times 4 - 1 = 19$

Mean squares

$MS_R \Rightarrow \frac{161.5}{4} = 40.375$

$MS_C \Rightarrow \frac{338.8}{4} = 112.933$

$MS_E = \frac{73.7}{12} = 6.1417$

Updated F Values $\Rightarrow \frac{MS_R}{MS_E} = 6.57$ $\quad F_C = \frac{112.93}{6.1417} = 18.39$

$F_R = 6.57 > 3.26$ we reject $H_0 \rightarrow$ significant diff in productivity

$F_C = 18.39 > 3.49$ reject $H_0 \rightarrow$ diff. in machine productivity