Page :
Date : / /

Rank Correlation

Now we move to another concept of statistics that is known as Rank correlation.

The correlation obtained from the table contains ranks as the variables x and y is known as rank correlation.

That means when we obtain correlation then the values of the variables x and y are the the numbers that may be marks of the student, rupees of the employees, Intelligence quotients, etc. but when we obtain rank-correlation then the varibles x and y be represented by just ranks. That means by 1, 2, 3, 4, 5 6, 7, etc.

Now you understand it clearly that Rank correlation is nothing but a special type of correlation.

Obviously you can obtain the coefficient of rank-correlation by the use of formula

$$\gamma = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$ which is well known

formula to obtain the coefficient of correlation. But there is also a particular formula for this which we can obtain by simplifying above one.

We know

$$\gamma = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$ where $X = x - \bar{x}$ and $Y = y - \bar{y}$

$\bar{x}$ and $\bar{y}$ are the means of x-series and y-series respectively

Date: / /

Let we have n items with n different ranks.

$\therefore$ x-series and y-series both contains the numbers from 1 to n but this is not necessary that the first and second column ~~are~~ of x and y be same.

$$\therefore \bar{x} = \frac{1+2+3+\cdots+n}{n} = \bar{y}$$

$$\therefore \bar{x} = \frac{n(n+1)}{2} \Big/ n = \bar{y}$$

$$\therefore \bar{x} = \frac{n+1}{2} = \bar{y}$$

$$\sum_{i=1}^{n} x = 1+2+3+\cdots+n$$

$$\sum_{i=1}^{n} y = 1+2+3+\cdots+n$$

$$\therefore \sum X^2 = \sum (x-\bar{x})^2$$

$$= \sum (x^2 - 2\bar{x}x + \bar{x}^2)$$

$$= \sum_{i=1}^{n} x^2 - 2\bar{x}\sum_{i=1}^{n} x + \bar{x}^2 \sum_{i=1}^{n} 1$$

$$= (1^2+2^2+3^2+\cdots+n^2) - 2\frac{(n+1)}{2}\cdot(1+2+3+\cdots+n) + \left(\frac{n+1}{2}\right)^2 (\underbrace{1+1+\cdots+1}_{n\ times})$$

$$= \frac{n(n+1)(2n+1)}{6} - (n+1)\cdot\frac{n(n+1)}{2} + \left(\frac{n+1}{2}\right)^2 \cdot n$$

$$= \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)^2}{2} + \frac{n(n+1)^2}{4}$$

$$= \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)^2}{4}$$

$$= \frac{n(n+1)}{2}\left[\frac{2n+1}{3} - \frac{n+1}{2}\right]$$

$$= \frac{n(n+1)}{2}\left(\frac{4n+2-3n-3}{6}\right)$$

$$= \frac{n(n+1)(n-1)}{2\times 6}$$

$$= \frac{n(n^2-1)}{12}$$

Similarly $\Sigma Y^2 = \frac{n(n^2-1)}{2\times 6}$.

Now $\Sigma(X-Y)^2 = \Sigma(X^2+Y^2-2XY)$

$\Rightarrow \Sigma\{x-\bar{x}-(y-\bar{y})\}^2 = \Sigma X^2 + \Sigma Y^2 - 2\Sigma XY$

$\Rightarrow \Sigma(x-y)^2 = \frac{n(n^2-1)}{2\times 6} + \frac{n(n^2-1)}{2\times 6} - 2\Sigma XY$

$= \frac{n(n^2-1)}{6} - 2\Sigma XY$ $\quad \because \bar{x} = \frac{n+1}{2} = \bar{y}$

$\Rightarrow 2\Sigma XY = \frac{n(n^2-1)}{6} - \Sigma(x-y)^2$

$\therefore \Sigma XY = \frac{n(n^2-1)}{12} - \frac{\Sigma(x-y)^2}{2}$

Substituting the ΣX^2, ΣY^2 and ΣXY in the original r. we get,

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

$$= \frac{\frac{n(n^2-1)}{12} - \frac{\Sigma(x-y)^2}{2}}{\sqrt{\frac{n(n^2-1)}{2\times 6} \cdot \frac{n(n^2-1)}{2\times 6}}}$$

$$= \frac{\frac{n(n^2-1)}{12} - \frac{\Sigma(x-y)^2}{2}}{\frac{n(n^2-1)}{12}}$$

$$= 1 - \frac{\Sigma(x-y)^2}{2} \times \frac{12}{n(n^2-1)}$$

$$r = 1 - \frac{6\Sigma(x-y)^2}{n(n^2-1)}$$

$$\boxed{\therefore r = 1 - \frac{6\Sigma d^2}{n(n^2-1)}}$$ [If $x-y$ is denoted by d.]

It is known as Spearman's coefficient of rank correlation and it is denoted by ρ(rho) particularly. So

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

Rank correlation coefficient always lies between -1 & 1.
There is an example now,

Q. Compute Spearman's rank correlation co-efficient for the following data:

Person	A	B	C	D	E	F	G	H	I	J
Rank in 1st Subject	9	10	6	5	7	2	4	8	1	3
Rank in 2nd Subject	1	2	3	4	5	6	7	8	9	10

Ans → Here total no. of persons = 10

$\therefore n = 10.$

Person	x	y	$d=x-y$	d^2
A	9	1	8	64
B	10	2	8	64
C	6	3	3	9
D	5	4	1	1
E	7	5	2	4
F	2	6	-4	16
G	4	7	-3	9
H	8	8	0	0
I	1	9	-8	64
J	3	10	-7	49

$\therefore \Sigma d^2 = 280$

$\therefore$ Coefficient of rank correlation from

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

$$= 1 - \frac{6\times 280}{10(10^2-1)}$$

$$= 1 - 1.697$$

$$= -0.697 \text{ Ans}$$

Scanned by CamScanner

Date : / /

Q. Three judges A, B, C give the following ranks. Find which pair of judges has Common approach

A:	1	6	5	10	3	2	4	9	7	8
B:	3	5	8	4	7	10	2	1	6	9
C:	6	4	9	8	1	2	3	10	5	7

Ans:→

Here $n = 10$.

To find the common approach of pair of judges we have to find the coefficient of rank correlation three times because there are three pair of judges (A, B), (B, C) and (C, A).

∴ (I) To find $\rho(A,B)$ (II) To find $\rho(B,C)$

$x(A)$	$y(B)$	$d=x-y$	d^2
1	3	-2	4
6	5	1	1
5	8	-3	9
10	4	6	36
3	7	-4	16
2	10	-8	64
4	2	2	4
9	1	8	64
7	6	1	1
8	9	-1	1

$\Sigma d^2 = 200$

$x(B)$	$y(C)$	$d=x-y$	d^2
3	6	-3	9
5	4	1	1
8	9	-1	1
4	8	-4	16
7	1	6	36
10	2	8	64
2	3	-1	1
1	10	-9	81
6	5	1	1
9	7	2	4

$\Sigma d^2 = 214$

$$\therefore \rho(A,B) = 1 - \frac{6\Sigma d^2}{n(n^2-1)} \quad \text{and} \quad \rho(B,C) = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6\times 200}{10(100-1)} \qquad = 1 - \frac{6\times 214}{10(100-1)}$$

$$= -0.2 \qquad = -0.3$$

Date : / /

Similarly to find $\rho(C,A)$

$x(C)$	$y(A)$	$d=x-y$	d^2
6	1	5	25
4	6	-2	4
9	5	4	16
8	10	-2	4
1	3	-2	4
2	2	0	0
3	4	-1	1
10	9	1	1
5	7	-2	4
7	8	-1	1

$\Sigma d^2 = 60$

$$\therefore \rho(C,A) = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6\times 60}{10(100-1)}$$

$$= 0.6$$

Here $\rho(C,A)$ is maximum out of all the three. Hence the judges A and C have the common approach.

Ans

[we see here just the magnitude not the + and − sign.]

Scanned by CamScanner