CORRELATION:

By Prof. Karunesh Saxena
Faculty of Management Studies
Director IQAC Cell
Mohanlal Sukhadia University
Udaipur (Raj.)

Types of Correlation

- Positive and Negative Correlation
- Simple, Partial and Multiple Correlation
- Linear and Non- Linear Correlation

Methods of Studying the Correlation:

- **Scatter Diagram** – The simplest device for studying correlation between two variables is a special type of dot chart.

- **Karl Pearson's Coefficient of Correlation** - Karl Pearson's Coefficient of Correlation denoted by- 'r' The coefficient of correlation 'r' measure the degree of linear relationship between two variables say x & y.

- **Spearman's Rank Coefficient of Correlation**.

KARL PEARSON'S COEFFICIENT OF CORRELATION.

The standard formula used in the computation of Pearson's Product Moment Correlation Coefficient is as follows:

$$\frac{N\sum XY - \sum X \sum Y}{\sqrt{N\sum X^2 - (\sum X)^2 X\, N\sum Y^2 - (\sum Y)^2}}$$

EXAMPLE : Find the Pearson's Coefficient:

FIRM	1	2	3	4	5	6	7
PRICE	11	13	15	17	18	19	20
DEMAND	30	29	24	24	21	18	15

When to use Rank Correlation Coefficient:

Has two principles:

- The initial data are in the form of ranks.
- If N is fairly small rank methods is sometimes applied to interval data as an approximation to the more time consuming r.

Spearman's Rank Coefficient of Correlation

- When statistical series in which the variables under study are not capable of quantitative measurement but can be arranged in serial order, in such situation Pearson's correlation coefficient can not be used in such case Spearman Rank correlation can be used.
- $R = 1 - \frac{6\sum D^2}{N(N^2-1)}$
- R = Rank correlation coefficient
- D = Difference of rank between paired item in two series.
- N = Total number of observation.

Interpretation of Rank Correlation Coefficient (R)

- The value of Rank Correlation Coefficient, R ranges from -1 to +1.
- If R = +1, then there is complete agreement in the order of the ranks and the ranks are in the same direction.
- If R = -1, then there is complete agreement in the order of the ranks and the ranks are in the opposite direction.
- If R = 0, then there is no Correlation.

Rank Correlation Coefficient (R)

- **Problems where actual ranks are given.**

a. Calculate the difference ‘D’ of two Ranks i.e. (R1 – R2).

b. Square the difference & calculate the sum of the difference i.e. $\sum D$

c. Substitute the values obtained in the formula.

EXAMPLE: Two managers are asked to rank a group of employees in order of potential for eventually becoming top managers. The ranking are as follows:

Employees	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Ranking by Manager 1	10	2	1	4	3	6	5	8	7	9
Ranking by Manager 2	9	4	2	3	1	5	6	8	7	10

Solution:

Employees	Rank by Manager 1 (R1)	Rank by Manager 2 (R2)	$(R1-R2)^2$
A	10	9	1
B	2	4	4
C	1	2	1
D	4	3	1
E	3	1	4
F	6	5	1
G	5	6	1
H	8	8	0
I	7	7	0
J	9	10	1
N=10			**$\sum D^2 = 14$**

$$R_s = 1 - \frac{6\sum D^2}{N(N^2-1)}$$

$$= 1 - 6 \text{ X } 14 / 990$$

$$= 1 - 0.085 = 0.915$$

Thus, we find there is high degree of positive correlation in the ranks assigned by the two managers.

Rank Correlation Coefficient (R)

❑ Problems where Ranks are not given:

a. If the ranks are not given, then we need to assign ranks to the data series. The highest value in the series can be assigned rank 1. We need to follow the same scheme of ranking for the other series.
b. Then calculate the rank correlation coefficient in similar way as we do when the ranks are given.

EXAMPLE: An examination of ten applicants for a clerical post was taken by a firm. From the marks obtained by the applicants in the accountancy and statistics paper, compute rank coefficient of correlation.

Applicants	A	B	C	D	E	F	G	H	I	J
Marks in Statistics	92	89	87	86	83	77	71	63	53	50
Marks in Accountancy	86	83	91	77	68	85	52	82	37	57

Solution:

Applicants	Marks in Statistics	R1	Marks in Accountancy	R2	$(R1 - R2)^2$ D^2
A	92	10	86	9	1
B	89	9	83	7	4
C	87	8	91	10	4
D	86	7	77	5	4
E	83	6	68	4	4
F	77	5	85	8	9
G	71	4	52	2	4
H	63	3	82	6	9
I	53	2	37	1	1
J	50	1	57	3	4
N = 10					**$\sum D^2 = 44$**

$$R = 1 - \frac{6\sum D^2}{N(N^2 - 1)}$$

$$R = 1 - \frac{6 \text{ X } 44}{990}$$

$$= 1 - 0.267 = 0.733$$

MERITS OF RANK CORRELATION COEFFICIENT:

- Simpler to understand and easy to apply.
- Great advantage to the qualitative data (like honesty , efficiency, intelligence etc). For example, the two workers of same firm can be ranked in order of efficiency and degree of correlation.
- Only method that can be used on the given ranks and not on actual data.
- Ascertains rough Degree of Correlation on the actual data

DEMERITS OF RANK CORRELATION COEFFICIENT:

- Cannot be used to find the Correlation of Grouped Frequency Distribution.
- Method cannot be applied where N is exceeding 30 unless we are given the ranks not the actual value of the variables.