Practical 11

Calculation of pearson correlation coefficient in MS – Excel

Pearson's r measures the strength and direction of the linear relationship between two continuous variables.

The **requirements** when considering the use of Pearson's correlation coefficient are:

1. Scale of measurement should be interval or ratio.

2. Variables should be approximately normally distributed.

3. The association should be linear.

4. There should be no outliers in the data.

Pearson's r is calculated using the formula:

$$r = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_i (x_i - \overline{x})^2}\sqrt{\sum_i (y_i - \overline{y})^2}}$$

r can take a range of values from +1 to -1

● A value of 0 indicates that there is no association between the two variables.

● A value greater than 0 indicates a positive association; that is, as the value of one variable increases, so does the value of the other variable.

● A value less than 0 indicates a negative association; that is, as the value of one variable increases, the value of the other variable decreases.

Example 1

In the example below of 6 people with different ages and different weight, let us try calculating the value of the Pearson r.

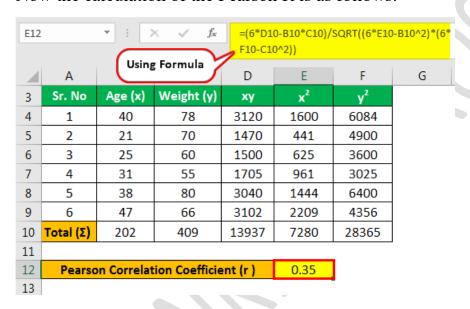| Sr. No	Age (x)	Weight (y)
1	40	78
2	21	70
3	25	60
4	31	55
5	38	80
6	47	66

Solution:

For the Calculation of the Pearson Correlation Coefficient, we will first calculate the following values:

Sr. No	Age (x)	Weight (y)	xy	x^2	y^2
1	40	78	3120	1600	6084
2	21	70	1470	441	4900
3	25	60	1500	625	3600
4	31	55	1705	961	3025
5	38	80	3040	1444	6400
6	47	66	3102	2209	4356
Total (Σ)	202	409	13937	7280	28365

Here the total number of people is 6 so, **n=6**

Now the calculation of the Pearson R is as follows:

r = (n (Σxy)- (Σx)(Σy))/(√ [n Σx2-(Σx)2][n Σy2– (Σy)2)

r = (6 * (13937)- (202)(409)) / (√ [6 *7280 -(202)2] * [6 * 28365- (409)2)

r = (6 * (13937)- (202) * (409))/(√ [6 *7280 -(202)2] * [6 * 28365- (409)2)

r = (83622- 82618)/(√ [43680 -40804] * [170190- 167281)

r = 1004/(√ [2876] * [2909)

r = 1004 / (√ 8366284)

r = 1004 / 2892.452938

r = 0.35

The value of the Pearson correlation coefficient is 0.35