**Lab. Meteorological**

**Data analysis ........ Fourth stage**

**(The second Semester)**

**Department of Atmospheric Sciences**

**2024 – 2025**

**Lecturers :**

**Dr.** **Iqbal khalaf , Dr.** **Khawla Zaki , Dr. Farah Haseeb ,**

 **L. Ruaa mazin,L. Luma mahdi ,L. Salwa salman,**

 **L. Yasamin qusay, A.L.Zahraa abdul razzaq, A.L.Zahraa araf**

 ***Preparing by: Dr. Farah Haseeb , L. Ruaa mazin , L. Luma Mahdi***

**(( Fourth Lecture ))**

**The Coefficient of Determination (**$R^{2}$**):**

The coefficient of determination R^2 is used to find out the number of points on the drawn regression line, which shows the relationship between the independent variable and the dependent variable whose value ranges between 0 and 1.

If the coefficient is 0.80, then 80% of the points should fall within the regression line. Values of 1 or 0 would indicate the regression line represents all or none of the data, respectively. A higher coefficient is an indicator of a better goodness of fit for the observations.



or

$R^{2}$ = $[\frac{n(\sum\_{}^{}xy)-(\sum\_{}^{}x)(\sum\_{}^{}y)}{\sqrt{\left[n\sum\_{}^{}x^{2 }-(\sum\_{}^{}x)^{2}\right]\left[n\sum\_{}^{}y^{2}-\left(\sum\_{}^{}y)^{2}\right)\right]}}]\^2$

**Example**: Find the regression equation and the determination coefficient for the following data :

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **x** | **6** | **8** | **9** | **8** | **7** | **6** | **5** | **6** | **5** | **5** |
| **y** | **10** | **13** | **15** | **14** | **9** | **7** | **6** | **6** | **5** | **5** |

**Solution :-**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **y** | **x** | **xy** | **x^2** | **y^2** |
| 6 | 10 | 60 | 100 | 36 |
| 8 | 13 | 104 | 169 | 64 |
| 9 | 15 | 135 | 225 | 81 |
| 8 | 14 | 112 | 196 | 64 |
| 7 | 9 | 63 | 81 | 49 |
| 6 | 7 | 42 | 49 | 36 |
| 5 | 6 | 30 | 36 | 25 |
| 6 | 6 | 36 | 36 | 36 |
| 5 | 5 | 25 | 25 | 25 |
| 5 | 5 | 25 | 25 | 25 |
| **∑=65** | **∑=90** | **∑=632** | **∑=942** | **∑=441** |

b= 0.356 , a=3.26 , n=10



**Standard Error of Estimate:**

The standard error coefficient of the calculated or estimated values (Se) This parameter calculates the vertical distance between the points scattered around the regression line and between the regression line and the lower its value, the more accurate the calculated values and vice versa. it is used to check the accuracy of predictions made with the regression line.



This law requires that we calculate the speculative value for each value of y and this requires that we calculate a and b and the solution becomes complex.

In order to get rid of the complexity in the law, the standard error coefficient of the calculated values can be calculated through the equation below

Referring to the above example, we extract the value of syy

Syy=441-

Syy=18.5

Sxy =632-

Sxy=47



**Se= 0.47**

**H.W./** Find the determination coefficient and Standard Error for a group of people of different ages and different weights , as shown in the table below :

|  |  |  |
| --- | --- | --- |
| The number of people | Age | Weight |
| 1 | 40 | 78 |
| 2 | 21 | 70 |
| 3 | 25 | 60 |
| 4 | 31 | 55 |
| 5 | 38 | 80 |
| 6 | 47 | 66 |