**The principle of least squares:**

The principle of least squares (LS) consists of determining the values of the unknown parameters that will minimize the sum of squares of errors (residuals).

**Error**

Error are defined as the differences between observed values and the estimated value by the fitted Model equation.

**Regression Equation:**

$\hat{Y}$**=a+b**$X\_{i}$**+**$ ε$

$$to find a ?$$

$a=\overbar{Y}$**-b**$\overbar{X}$

$to find b?$ **(y on x)**

$$ $$

$b\_{yx}$**=**$\frac{n\sum\_{}^{}xy-\left(\sum\_{}^{}x\right)\left(\sum\_{}^{}y\right)}{n\sum\_{}^{}x^{2}-\left(\sum\_{}^{}x\right)^{2}}$

$to find b\_{xy}$**(x on y)**

**bxy=**$ \frac{n\sum\_{}^{}xy-\left(\sum\_{}^{}x\right)\left(\sum\_{}^{}y\right)}{n\sum\_{}^{}y^{2}-\left(\sum\_{}^{}y\right)^{2}}$

**Example NO.1**

**Compute the least squares regression equation of Y on X for the following .what is the regression co efficient and what does it mean?**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **x** | **5** | **6** | **8** | **10** | **12** | **13** | **15** | **16** | **17** |
| **y** | **16** | **19** | **23** | **28** | **36** | **41** | **44** | **45** | **50** |

**Solution**

The estimated regression line is Y on X is

$\hat{Y}$**=a+bX**

**Necessary calculations are:**

|  |  |  |  |
| --- | --- | --- | --- |
| **x** | **Y** | **Xy** | $$x^{2}$$ |
| 5 | 16 | 80 | 25 |
| 6 | 19 | 114 | 36 |
| 8 | 23 | 184 | 64 |
| 10 | 28 | 280 | 100 |
| 12 | 36 | 432 | 144 |
| 13 | 41 | 533 | 169 |
| 15 | 44 | 660 | 225 |
| 16 | 45 | 720 | 256 |
| 17 | 50 | 850 | 289 |
| $\sum\_{}^{}x$**=102** | $\sum\_{}^{}y$**=302** | $\sum\_{}^{}xy$**=3853** | **1308** |

Here y on x so

**byx =**$\frac{n\sum\_{}^{}xy-\left(\sum\_{}^{}x\right)\left(\sum\_{}^{}y\right)}{n\sum\_{}^{}x^{2}-\left(\sum\_{}^{}x\right)^{2}}$

 n=9

$b\_{yx}$=$\frac{9\left(3853\right)-(102)(302)}{9\left(1308\right)-\left(102\right)^{2}}$

b= 2.831

now we have to find a so

**a**$=\overbar{Y}$**-b**$\overbar{X}$

$\overbar{Y}=\frac{\sum\_{}^{}y}{n}$ **=**$\frac{302}{9}$**=33.56**

$\overbar{X}=\frac{\sum\_{}^{}x}{n}$**=**$\frac{102}{9}$**=11.33**

**a**=33.56-2.831(11.33)

**a=1.47**

**Hence the desire estimated regression equation is**

$\hat{Y}$**=a+b**$X\_{i}$**+**$ ε$

$$\hat{Y}=1.47+2.831X$$

**Example no.2**

in an experiment **to measure** the stiffness of a spring ,the length of the spring under different loads was measured as follows

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **X=loads**  | **3** | **5** | **6** | **9** | **10** | **12** | **15** | **20** | **22** | **28** |
| **Y=length** | **10** | **12** | **15** | **18** | **20** | **22** | **27** | **30** | **32** | **34** |

Find the regression equation

1. **The length,given the weight on the spring (y on x)**
2. **The weight ,given the length of the spring (x on y)**

**Solution:**

**Second part x on y**

$\hat{X}$**=a+bY**

|  |  |  |  |
| --- | --- | --- | --- |
| **X** | **Y** | $$y^{2}$$ | **Xy** |
| 3 | 10 | 100 | 30 |
| 5 | 12 | 144 | 60 |
| 6 | 15 | 225 | 90 |
| 9 | 18 | 324 | 162 |
| 10 | 20 | 400 | 200 |
| 12 | 22 | 484 | 264 |
| 15 | 27 | 729 | 405 |
| 20 | 30 | 900 | 600 |
| 22 | 32 | 1024 | 704 |
| 28 | 34 | 1156 | 932 |
| $\sum\_{}^{}x$**=130** | $\sum\_{}^{}y$**=220** | **5486** | **3467** |

**bxy=**$\frac{n\sum\_{}^{}xy-\left(\sum\_{}^{}x\right)\left(\sum\_{}^{}y\right)}{n\sum\_{}^{}y^{2}-\left(\sum\_{}^{}y\right)^{2}}$

$b\_{xy}$=$\frac{10\left(3467\right)-(130)(220)}{10\left(5486\right)-\left(220\right)^{2}}$=0.94

now we have to find a so

**a**$=\overbar{X}$**-b**$\overbar{Y}$

$\overbar{Y}=\frac{\sum\_{}^{}y}{n}$ **=**$\frac{220}{10}$**=22**

$\overbar{X}=\frac{\sum\_{}^{}x}{n}$**=**$\frac{130}{10}$**=13**

**a=xbar-b(ybar)**

**a**=13-0.94(22)

**a=-7.68**

**Hence the desire estimated regression equation is**

$\hat{X}$**=a+ bY+**$ ε$

$$\hat{Y}=-7.68+0.94X+ε$$

 **Part (i) is for home assignment.**

$$ $$