Regresión lineal

La correlación indica el grado de asociación entre las variables; mientras que la regresión es útil para averiguar la forma probable de relación entre ellos.

Para obtener la ecuación de la recta de regresión se empleará el principio de mínimos cuadrados, denominado así debido a que considera que la recta de regresión encontrada es la mejor, en sentido de que la suma de los cuadrados de las distancias verticales de los distintos puntos a la recta de regresión es mínima.

Así, la ecuación de la recta es: $\hat{y} = mx + b$

$$m = \frac{\sum_{i=1}^{n}(x_i - \tilde{x})(y_i - \check{y})}{\sum_{i=1}^{n}(x_i - \tilde{x})^2}$$

$$\boldsymbol{m = \frac{SC_{xy}}{SC_x}}$$

$$b = \frac{\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i}{n}$$

$$\boldsymbol{b = \bar{y} - m\bar{x}}$$

Donde:

$\check{y}$ representa el valor estimado de $y$.

$\bar{y}$ representa la media de $y$.

$\bar{x}$ representa la media de $x$.

$m$ representa la pendiente de la recta de regresión.

$b$ representa la ordenada en el origen de la recta de regresión.

Realiza las siguientes actividades de aprendizajes

| **Horas de estudios** | **16** | **18** | **20** | **22** | **23** | **27** | **32** | **34** |
|---|---|---|---|---|---|---|---|---|
| **Calificación en examen** | **61** | **72** | **64** | **77** | **70** | **88** | **92** | **84** |

| Horas de estudios $x_i$ | Calificación de examen $y_i$ | $x_i y_i$ | $x_i^2$ | $y_i^2$ |
|---|---|---|---|---|
| 16 | 61 | $16 * 61 = 976$ | $16^2 = 256$ | 3721 |
| 18 | 72 | 1296 | 324 | 5184 |
| 20 | 64 | 1280 | 400 | 4096 |
| 22 | 77 | 1694 | 484 | 5929 |
| 23 | 70 | 1610 | 529 | 4900 |
| 27 | 88 | 2376 | 729 | 7744 |
| 32 | 92 | 2944 | 1024 | 8464 |
| 34 | 84 | 2856 | 1156 | 7056 |
| $\sum_{i=1}^{n} x_i = 192$ | $\sum_{i=1}^{n} y_i = 608$ | $\sum_{i=1}^{n} x_i y_i = 15032$ | $\sum_{i=1}^{n} x_i^2 = 4902$ | $\sum_{i=1}^{n} y_i^2 = 47094$ |

$$m = \frac{SC_{xy}}{SC_x}$$

$$sc_x = \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} = 4902 - \frac{(192)^2}{8} = 4902 - 4608 = 294$$

$$sc_x = 294$$

$$sc_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n} \rightarrow sc_{xy} = 15032 - \frac{(192)(608)}{8} = 440$$

$$sc_{xy} = 440$$

**Para encontrar la pendiente.**

$$m = \frac{440}{294} = 1.49$$

$$m = 1.49$$

**Aplicamos las siguientes formulas para hallar la ordenada.**

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\bar{x} = \frac{192}{8} = 24$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$\bar{y} = \frac{608}{8} = 76$$

$$\boldsymbol{b = \bar{y} - m\bar{x}}$$

$$\boldsymbol{b = 76 - 1.49(24)}$$

$$\boldsymbol{b = 76 - 1.49(24)}$$

$$\boldsymbol{b = 40.24}$$

$$\hat{y} = mx + b$$

$$\hat{y} = \mathbf{1.49}x + \mathbf{40.24}$$

$$\hat{y} = \mathbf{1.49}(8) + \mathbf{40.24}$$

$$\hat{y} = 52.16$$

$$\hat{y} = \mathbf{1.49}(5) + \mathbf{40.24}$$

$$\hat{y} = 47.69$$

$$\hat{y} = \mathbf{1.49}(20) + \mathbf{40.24}$$

$$\hat{y} = 70.04$$

Calcules en coeficiente de correlación y regresión e interpreta el resultado de los datos que se proporcionan a continuación.

| **Semanas de experiencia.** | **7** | **9** | **6** | **14** | **8** | **12** | **10** | **4** | **2** | **11** | **1** | **8** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Números de rechazos | 26 | 20 | 28 | 16 | 23 | 18 | 24 | 26 | 38 | 22 | 32 | 25 |

| $x_i$ | $y_i$ | $x_i y_i$ | $x_i^2$ | $y_i^2$ |
|---|---|---|---|---|
| 7 | 26 | 182 | 49 | 676 |
| 9 | 20 | 180 | 81 | 400 |
| 6 | 28 | 168 | 36 | 784 |
| 14 | 16 | 224 | 196 | 256 |
| 8 | 23 | 184 | 64 | 529 |
| 12 | 18 | 216 | 144 | 324 |
| 10 | 24 | 240 | 100 | 576 |
| 4 | 26 | 104 | 16 | 676 |
| 2 | 38 | 76 | 4 | 1444 |
| 11 | 22 | 242 | 121 | 484 |
| 1 | 32 | 32 | 1 | 1024 |
| 8 | 25 | 200 | 64 | 625 |
| $\sum_{i=1}^{n} x_i = 92$ | $\sum_{i=1}^{n} y_i = 298$ | $\sum_{i=1}^{n} x_i y_i = 2048$ | $\sum_{i=1}^{n} x_i^2 = 876$ | $\sum_{i=1}^{n} y_i^2 = 7798$ |

$$sc_x = \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} = 876 - \frac{(92)^2}{12} = 876 - \frac{8464}{12}$$

$$= 876 - 705.33 = 170.66$$

$$sc_x = 170.66$$

$$sc_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n} = 2048 - \frac{92 * 298}{12} = -236.66$$

$$sc_{xy} = -236.66$$

$$m = \frac{-236.66}{170.66} = -1.38$$

$$\overline{x} = \frac{92}{12} = 7.66$$

$$\overline{y} = \frac{298}{12} = 24.83$$

$$b = \overline{y} - m\overline{x}$$

$$b = 24.83 - (-1.38)(7.66)$$

$$b = 14.25$$

$$\hat{y} = mx + b$$

$$\hat{y} = -1.38x + 14.25$$

$$r = \frac{SC_{xy}}{\sqrt{SC_x SC_y}} = \frac{-236.66}{\sqrt{170.66(\quad)}}$$