Método de mínimos Cuadrados

Introducción

Una línea del mejor ajuste es una recta que muestra la mejor aproximación del conjunto dado de datos dispersos. Se utiliza para estudiar la naturaleza de la relación entre dos variables.

Una línea del mejor ajuste puede determinarse usando un método de “simple vista” dibujando una línea recta en un diagrama de dispersión de modo que el número de puntos por encima y por debajo de la línea sea aproximadamente igual.

Una forma más precisa de encontrar la línea del mejor ajuste es el método de mínimos cuadrdados.

Ajuste por mínimos cuadrados

Los siguientes pasos sirven para encontrar la ecuación lineal del mejor ajuste para un conjunto de pares ordenados *(x1, y1), (x2, y2), ..., (xn, yn).*

Paso 1: Calcular la media de los valores *x* y la media de los valores *y*.

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\overline{Y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

Continuación

Paso 2: La siguiente fórmula da la pendiente de la línea de mejor ajuste:

$$m = \frac{\sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})}{\sum_{i=1}^{n}(x_i - \overline{X})^2}$$

Paso 3: Calcule la intercepción en *y* de la línea usando la fórmula:

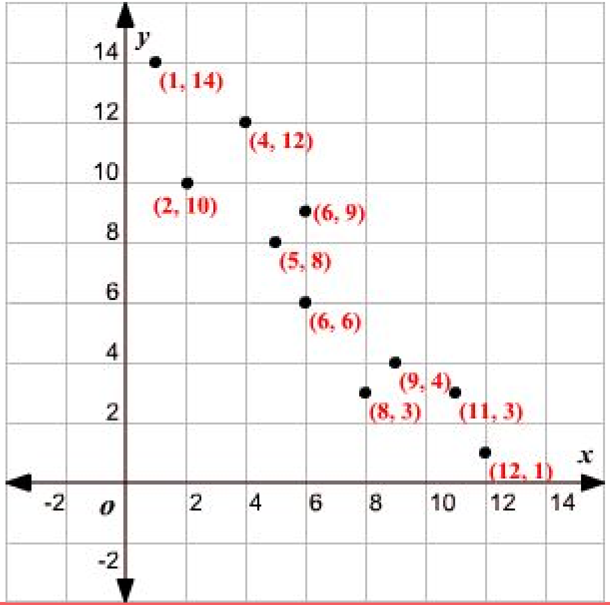
$$b = \overline{Y} - m\overline{X}$$

Paso 4: Utilice la pendiente *m* y la intercepción en *y* *b* para formar la ecuación de la recta.

Ejemplo

Utilice el método de mínimos cuadrados para determinar la ecuación de la línea del mejor ajuste para los datos. Luego trace la línea.

x	8	2	11	6	5	4	12	9	6	1
y	3	10	3	6	8	12	1	4	9	14



Solución

1. Calcular las medias de los valores-*x* y los valores-*y*.

$$\overline{X} = \frac{8+2+11+6+5+4+12+9+6+1}{10} = 6.4$$

$$\overline{Y} = \frac{3+10+3+6+8+12+1+4+9+14}{10} = 7$$

Ahora, calcular $x_i - \overline{X}$, $y_i - \overline{Y}$, $(x_i - \overline{X})(y_i - \overline{Y})$, y $(x_i - \overline{X})^2$ por cada par de datos *i*.

i	x_i	y_i	$x_i - \overline{X}$	$y_i - \overline{Y}$	$(x_i - \overline{X})(y_i - \overline{Y})$	$(x_i - \overline{X})^2$
1	8	3	1.6	−4	−6.4	2.56
2	2	10	−4.4	3	−13.2	19.36
3	11	3	4.6	−4	−18.4	21.16
4	6	6	−0.4	−1	0.4	0.16
5	5	8	−1.4	1	−1.4	1.96
6	4	12	−2.4	5	−12	5.76
7	12	1	5.6	−6	−33.6	31.36
8	9	4	2.6	−3	−7.8	6.76
9	6	9	−0.4	2	−0.8	0.16
10	1	14	−5.4	7	−37.8	29.16
					$\sum_{i=1}^{n} (x_i - \overline{X})(y_i - \overline{Y}) = -131$	$\sum_{i=1}^{n} (x_i - \overline{X})^2 = 118.4$

Solución (continuación)

2. Calcular la pendiente.

$$m = \frac{\sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})}{\sum_{i=1}^{n}(x_i - \overline{X})^2} = \frac{-131}{118.4} \approx -1.1$$

3. Calcular la intercepción de y.

$$\begin{aligned} b &= \overline{Y} - m\overline{X} \\ &= 7 - (-1.1 \times 6.4) \\ &= 7 + 7.04 \\ &\approx 14.0 \end{aligned}$$

Solución (continuación)

4. Utilice la pendiente y la intersección en y para formar la ecuación de la línea del mejor ajuste.

La pendiente de la recta es -1.1 y la intersección

en y es 14.0. Por lo tanto, la ecuación es:

$$y = -1.1x + 14.0$$