

Ejemplo.

La siguiente tabla muestra datos sobre el desgaste de acero dulce y y la viscosidad del aceite x .

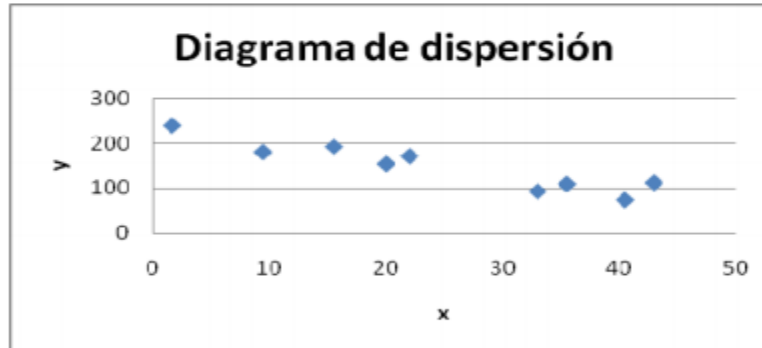
y [10^{-4} mm ³]	240	181	193	155	172	110	113	75	94
x	1.6	9.4	15.5	20	22	35.5	43	40.5	33

Las sumas relevantes, se muestran a continuación:

Sumas	
y	1333
x	220.5
xy	26864.4
x^2	7053.67
y^2	220549

- Trazar la gráfica de la dispersión de los datos. ¿Parece conveniente el uso de un modelo de regresión lineal?
- Ajustar un modelo de regresión lineal simple.
- Determinar el valor que se espera del desgaste cuando la viscosidad es 30
- ¿Puede considerarse válido el modelo? Justificar su respuesta.

a)



b) El modelo es:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^9 x_i y_i - \frac{\sum_{i=1}^9 x_i \sum_{i=1}^9 y_i}{9}}{\sum_{i=1}^9 x_i^2 - \frac{\left(\sum_{i=1}^9 x_i\right)^2}{9}} = \frac{26864.4 - \frac{(220.5)(1333)}{9}}{7053.67 - \frac{(220.5)^2}{9}} \approx -3.5086$$

y

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_0 = \frac{1333}{9} - (-3.5086) \left(\frac{220.5}{9} \right) \approx 234.0718$$

El modelo queda:

$$\hat{y} = 234.0718 - 3.5086x$$

e) Para el valor que se espera del desgaste cuando la viscosidad es 30, sustituyendo en el modelo:
 $\hat{y}(30) = 234.0718 - 3.5086(30) \approx 128.8138$

d) Para determinar si el modelo es válido debe obtenerse el coeficiente de determinación.
El coeficiente de correlación, está definido por:

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$$

$$SS_{xx} = \sum_{i=1}^9 x_i^2 - \frac{\left(\sum_{i=1}^9 x_i\right)^2}{9} = 7053.67 - \frac{(220.5)^2}{9} = 1651.42$$

$$SS_{yy} = \sum_{i=1}^9 y_i^2 - \frac{\left(\sum_{i=1}^9 y_i\right)^2}{9} = 220549 - \frac{(1333)^2}{9} \approx 23116.8889$$

$$SS_{xy} = \sum_{i=1}^9 x_i y_i - \frac{\sum_{i=1}^9 x_i \sum_{i=1}^9 y_i}{9} = 26864.4 - \frac{(220.5)(1333)}{9} = -5794.1$$

sustituyendo:

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}} = \frac{-5794.1}{\sqrt{(1651.42)(23116.8889)}} \approx -0.9378$$

Entonces el coeficiente de determinación será:

$$r^2 = R^2 = \frac{SS_{xy}^2}{SS_{xx} SS_{yy}} = \frac{(-5794.1)^2}{(1651.42)(23116.8889)} \approx 0.8794$$

El ajuste es regular y puede considerarse válido el modelo dependiendo del error que se esté dispuesto a cometer.