

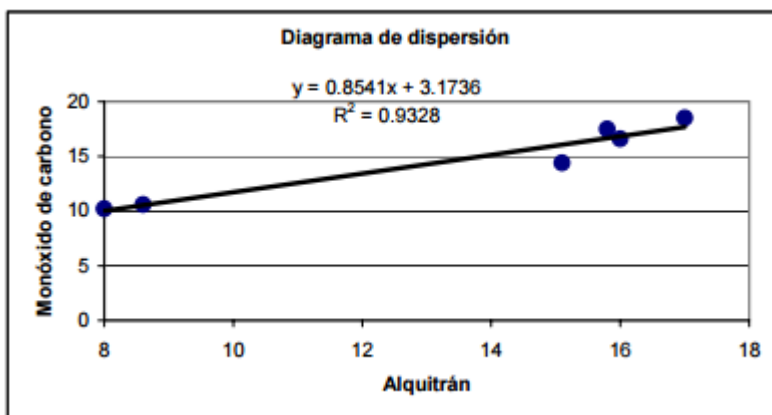
Ejemplo.

La Procuraduría del Consumidor evalúa anualmente distintas marcas de cigarros de acuerdo con el contenido de alquitrán, nicotina y monóxido de carbono (CO). La asociación de médicos considera peligrosas cada una de estas sustancias en la salud del fumador. Estudios anteriores han demostrado que un aumento en el contenido de alquitrán y nicotina de un cigarro está acompañado de un incremento en el monóxido de carbono emitido en el humo del cigarro. La tabla siguiente muestra los valores para seis marcas de cigarras.

Marca	Alquitrán [mg]	CO [mg]
Benson&Hedges	16,0	16,6
Camel Lights	8,0	10,2
Marlboro	15,1	14,4
Raleigh	15,8	17,5
Montana	17,0	18,5
Viceroy Light	8,6	10,6

- Elaborar un diagrama de dispersión entre el contenido de alquitrán (x), y CO (y)
- Obtener la recta de regresión y trazarla en el diagrama de dispersión del inciso (a)
- ¿Corroboran los resultados el hecho de que un aumento de alquitrán conlleva un aumento de monóxido de carbono?
- ¿Considera que el modelo proporcionado es bueno? ¿Qué porcentaje del monóxido de carbono emitido en el humo del cigarro queda explicado por el modelo?

a) El diagrama de dispersión es:



- b) Los parámetros y el modelo, son:

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

$$\hat{\beta}_1 = \frac{\sum_{i=1}^6 x_i y_i - \frac{\sum_{i=1}^6 x_i \sum_{i=1}^6 y_i}{6}}{\sum_{i=1}^6 x_i^2 - \frac{\left(\sum_{i=1}^6 x_i\right)^2}{6}} = \frac{1246.80 - \frac{(80.50)(87.80)}{6}}{1160.61 - \frac{(80.50)^2}{6}} = \frac{68.817}{80.568} \approx 0.8541$$

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

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 14.633 - (0.854)(13.417) \approx 3.1741$$

Por lo tanto el modelo está dado por:

$$\hat{y} = 0.8541x + 3.1741$$

- b) De acuerdo a la gráfica, si se corrobora el hecho de que un aumento de alquitrán conlleva un aumento de monóxido de carbono.
- c) Si es bueno el modelo que se obtuvo, ya que, el porcentaje del monóxido de carbono emitido en el humo del cigarro queda explicado por el modelo de manera muy aceptable como se observa en lo que sigue.

Para determinar si el modelo es válido debe obtenerse el coeficiente de determinación.

El coeficiente de correlación, está definido por:

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

$$SS_{xx} = \sum_{i=1}^6 x_i^2 - \frac{\left(\sum_{i=1}^6 x_i\right)^2}{6} = 1160.61 - \frac{(80.50)^2}{6} = 80.568$$

$$SS_{yy} = \sum_{i=1}^6 y_i^2 - \frac{\left(\sum_{i=1}^6 y_i\right)^2}{6} = 1347.82 - \frac{(87.80)^2}{6} = 63.013$$

$$SS_{xy} = \sum_{i=1}^6 x_i y_i - \frac{\sum_{i=1}^6 x_i \sum_{i=1}^6 y_i}{6} = 1246.80 - \frac{(80.50)(87.80)}{6} = 68.817$$

sustituyendo:

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}} = \frac{68.817}{\sqrt{(80.568)(63.013)}} \approx 0.9658$$

Entonces el coeficiente de determinación es:

$$r^2 = R^2 = \frac{SS_{xy}^2}{SS_{xx} SS_{yy}} = \frac{(68.817)^2}{(80.568)(63.013)} \approx 0.9328$$