UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA

SECCIONAL TUNJA

VIGILADA MINEDUCACIÓN - SNIES 1732

Acreditación Institucional Internacional
OTORGADA POR EL IAC CINDA ACUERDO 55 DEL 9 DE MAYO-VIGENCIA 5 AÑOS

ACREDITACIÓN INSTITUCIONAL DE ALTA CALIDAD MULTICAMPUS
Res. MEN No. 01456 del 29 de enero de 2016
Vigencia por seis años

QS STARS RATED FOR EXCELLENCE

ISO 9001 icontec SC4289-1

CERTIFIED IQNet MANAGEMENT SYSTEM

UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA
SECCIONAL TUNJA

VIGILADA MINEDUCACIÓN - SNIES 1732

Regresión por mínimos cuadrados

Regresión por mínimos cuadrados

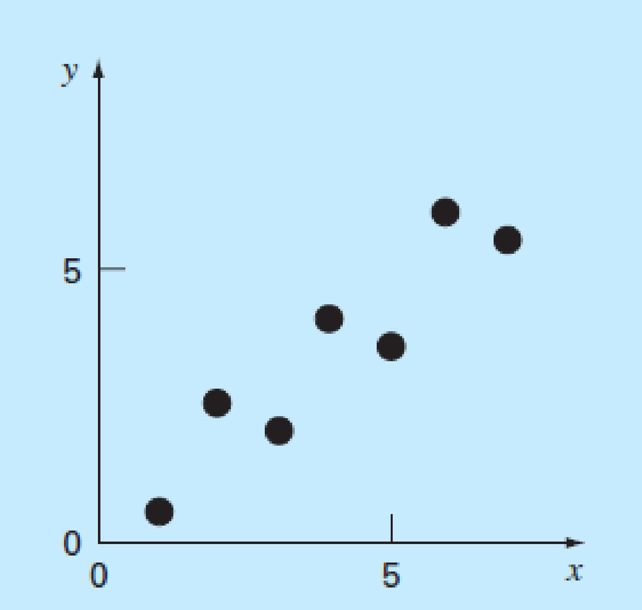
Cuando se tiene una serie de valores tabulados, extraídos experimentalmente, se necesita muchas veces predecir o aproximar valores para datos no medidos.

Regresión por mínimos cuadrados

Cuando se tiene una serie de valores tabulados, extraídos experimentalmente, se necesita muchas veces predecir o aproximar valores para datos no medidos.

Para ello se estudia la tendencia de los datos obtenidos y se busca encontrar una ecuación que se ajuste a éstos.

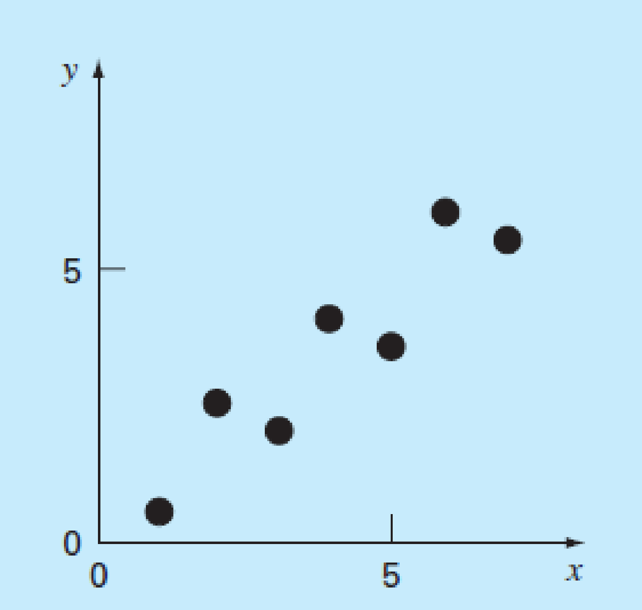
Regresión por mínimos cuadrados



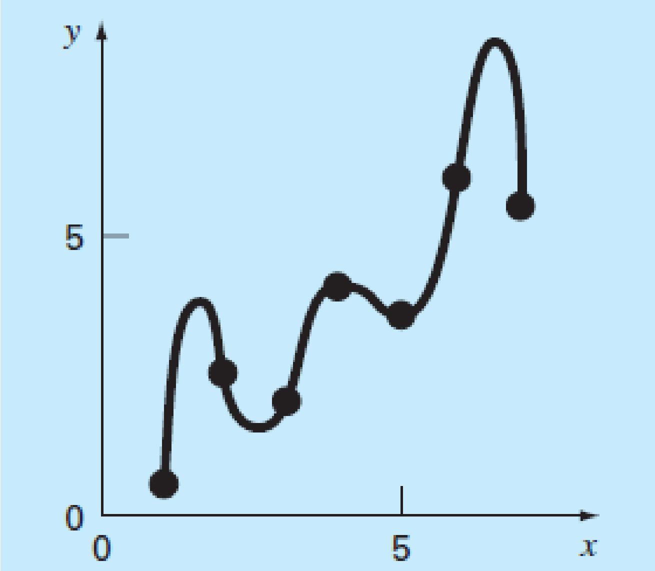
a)

Datos obtenidos experimentalmente

Regresión por mínimos cuadrados



a)

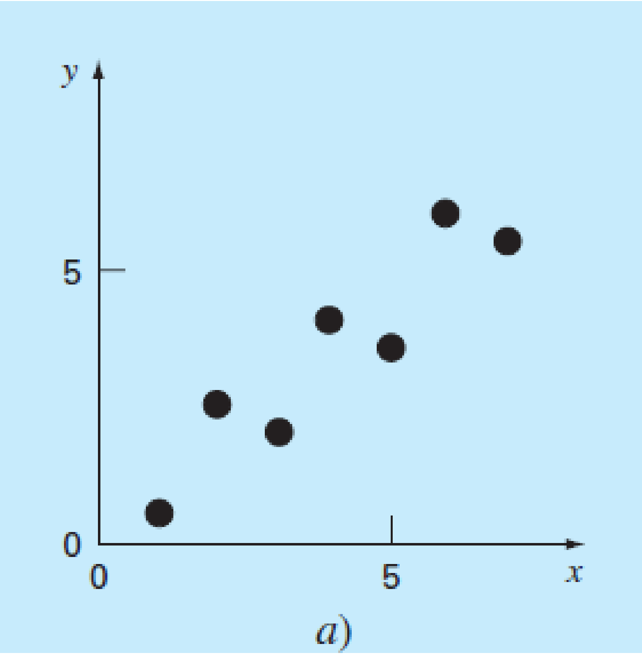


b)

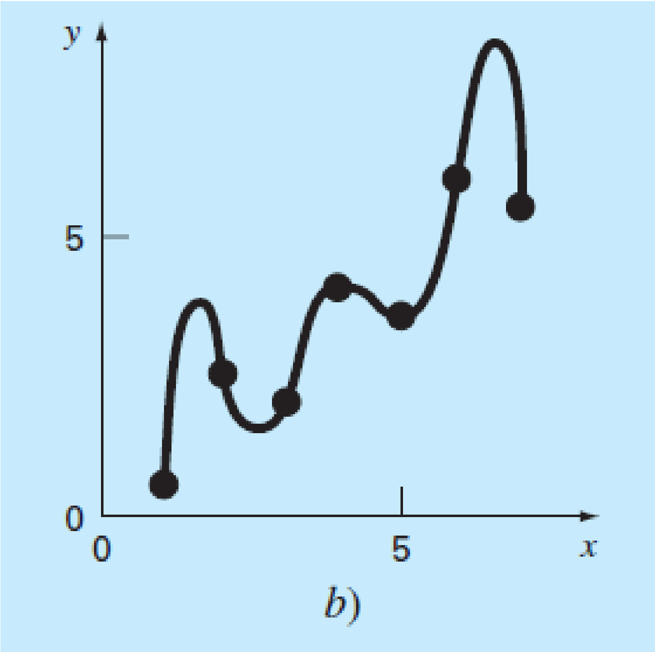
Datos obtenidos experimentalmente

Ajuste polinomial

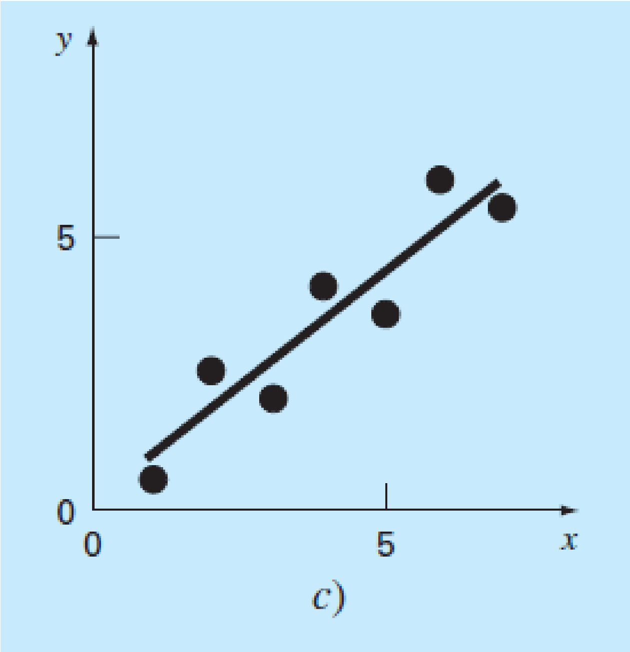
Regresión por mínimos cuadrados



Datos obtenidos experimentalmente



Ajuste polinomial



Ajuste por mínimos cuadrados

¡Siempre hacia lo alto!

Regresión lineal

El proceso de aproximación por mínimos cuadrados consiste, en su forma más simple, en ajustar una línea recta a un conjunto de datos que están definidos por los puntos ($x1$, $y1$), ($x2$, $y2$),…, (xn, yn).

Regresión lineal

El proceso de aproximación por mínimos cuadrados consiste, en su forma más simple, en ajustar una línea recta a un conjunto de datos que están definidos por los puntos ($x1$, $y1$), ($x2$, $y2$),…, (xn, yn).

La expresión matemática que define la recta es:

$$y = a_0 + a_1x + e$$

Donde:

- a_0 corresponde a la intersección con el eje y.
- a_1 corresponde a la pendiente.
- e corresponde al error (diferencia entre el modelo y las observaciones).

Regresión lineal

El proceso de aproximación por mínimos cuadrados consiste, en su forma más simple, en ajustar una línea recta a un conjunto de datos que están definidos por los puntos ($x1$, $y1$), ($x2$, $y2$),…, (xn, yn).

La expresión matemática que define la recta es:

$$y = a_0 + a_1 x + e$$

Donde:

- a_0 corresponde a la intersección con el eje y.
- a_1 corresponde a la pendiente.
- e corresponde al error (diferencia entre el modelo y las observaciones).

$$e = y - (a_0 + a_1 x)$$

Regresión lineal

Criterio para el mejor ajuste consiste en minimizar la sumatoria de los errores:

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_{i,\text{medida}} - y_{i,\text{modelo}})^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$$

n corresponde al total de puntos obtenidos.

Regresión lineal

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_{i,\text{medida}} - y_{i,\text{modelo}})^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$$

Para hallar los valores de a_0 y a_1, se deriva la anterior ecuación respecto a cada coeficiente:

Regresión lineal

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_{i,\text{medida}} - y_{i,\text{modelo}})^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$$

Para hallar los valores de a_0 y a_1, se deriva la anterior ecuación respecto a cada coeficiente:

$$\frac{\partial S_r}{\partial a_0} = -2\sum (y_i - a_0 - a_1 x_i)$$

$$\frac{\partial S_r}{\partial a_1} = -2\sum [(y_i - a_0 - a_1 x_i)x_i]$$

Regresión lineal

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_{i,\text{medida}} - y_{i,\text{modelo}})^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$$

Para hallar los valores de a_0 y a_1, se deriva la anterior ecuación respecto a cada coeficiente:

$$\frac{\partial S_r}{\partial a_0} = -2\sum (y_i - a_0 - a_1 x_i)$$

$$\frac{\partial S_r}{\partial a_1} = -2\sum [(y_i - a_0 - a_1 x_i)x_i]$$

Al igual a cero las derivadas parciales, se hará S_r mínimo.

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

¡Siempre hacia lo alto!

Regresión lineal

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

¡Siempre hacia lo alto!

Regresión lineal

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

Si se tiene que $\sum a_0 = na_0$, entonces:

Regresión lineal

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

Si se tiene que $\sum a_0 = n a_0$, entonces:

$$n a_0 + \left(\sum x_i\right) a_1 = \sum y_i$$

$$\left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 = \sum x_i y_i$$

¡Siempre hacia lo alto!

Regresión lineal

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

Si se tiene que $\sum a_0 = na_0$, entonces:

$$na_0 + \left(\sum x_i\right) a_1 = \sum y_i$$

$$\left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 = \sum x_i y_i$$

Despejando a_1, se tiene que:

$$a_1 = \frac{\sum y_i - na_0}{\sum x_i}$$

$$a_1 = \frac{\sum x_i y_i - (\sum x_i) a_0}{(\sum x_i{}^2)}$$

Regresión lineal

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

Si se tiene que $\sum a_0 = na_0$, entonces:

$$na_0 + \left(\sum x_i\right) a_1 = \sum y_i$$

$$\left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 = \sum x_i y_i$$

Despejando a_1, se tiene que:

$$a_1 = \frac{\sum y_i - na_0}{\sum x_i} \qquad a_1 = \frac{\sum x_i y_i - (\sum x_i) a_0}{(\sum x_i{}^2)}$$

Resolviendo, se tiene que:

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

¡Siempre hacia lo alto!

Regresión lineal

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

Si se tiene que $\sum a_0 = n a_0$, entonces:

$$n a_0 + \left(\sum x_i\right) a_1 = \sum y_i$$

$$\left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 = \sum x_i y_i$$

Despejando a_1, se tiene que:

$$a_1 = \frac{\sum y_i - n a_0}{\sum x_i} \qquad a_1 = \frac{\sum x_i y_i - (\sum x_i) a_0}{(\sum x_i{}^2)}$$

Resolviendo, se tiene que:

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

Además, de la primera ecuación:

$$a_0 = \bar{y} - a_1 \bar{x}$$

$\bar{y}$ y $\bar{x}$ son las medias de y y x

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$n = 7$

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$n = 7$

	x	y	xiyi	xi^2
	1	0,5	0,5	1
	2	2,5	5	4
	3	2	6	9
	4	4	16	16
	5	3,5	17,5	25
	6	6	36	36
	7	5,5	38,5	49
Sumatoria	28	24	119,5	140

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$n = 7$

	x	y	xiyi	xi^2
	1	0,5	0,5	1
	2	2,5	5	4
	3	2	6	9
	4	4	16	16
	5	3,5	17,5	25
	6	6	36	36
	7	5,5	38,5	49
Sumatoria	28	24	119,5	140

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

$(\sum x_i)^2 =$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$n = 7$

	x	y	xiyi	xi^2
	1	0,5	0,5	1
	2	2,5	5	4
	3	2	6	9
	4	4	16	16
	5	3,5	17,5	25
	6	6	36	36
	7	5,5	38,5	49
Sumatoria	28	24	119,5	140

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

$(\sum x_i)^2 = 784$

$\bar{y} =$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$n = 7$

	x	y	xiyi	xi^2
	1	0,5	0,5	1
	2	2,5	5	4
	3	2	6	9
	4	4	16	16
	5	3,5	17,5	25
	6	6	36	36
	7	5,5	38,5	49
Sumatoria	28	24	119,5	140

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

$(\sum x_i)^2 = 784$

$\bar{y} = 3{,}4285$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

n= 7

	x	y	xiyi	xi^2
	1	0,5	0,5	1
	2	2,5	5	4
	3	2	6	9
	4	4	16	16
	5	3,5	17,5	25
	6	6	36	36
	7	5,5	38,5	49
Sumatoria	28	24	119,5	140

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

$(\sum x_i)^2$= 784

$\bar{y}$= 3,4285

$\bar{x}$= 4

$$a_1 = \frac{7(119{,}5) - (28)(24)}{7(140) - (784)} =$$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$n = 7$

	x	y	xiyi	xi^2
	1	0,5	0,5	1
	2	2,5	5	4
	3	2	6	9
	4	4	16	16
	5	3,5	17,5	25
	6	6	36	36
	7	5,5	38,5	49
Sumatoria	28	24	119,5	140

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

$(\sum x_i)^2 = 784$

$\bar{y} = 3,4285$

$\bar{x} = 4$

$$a_1 = \frac{7(119,5) - (28)(24)}{7(140) - (784)} = \frac{836,5 - 672}{980 - 784} = \frac{164,5}{196} = 0,839285$$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5.5
Σ	24.0

$n = 7$

	x	y	xiyi	xi^2
	1	0,5	0,5	1
	2	2,5	5	4
	3	2	6	9
	4	4	16	16
	5	3,5	17,5	25
	6	6	36	36
	7	5,5	38,5	49
Sumatoria	28	24	119,5	140

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$a_0 = \bar{y} - a_1\bar{x}$$

$(\sum x_i)^2 = 784$

$\bar{y} = 3,4285$

$\bar{x} = 4$

$$a_1 = \frac{7(119,5) - (28)(24)}{7(140) - (784)} = \frac{836,5 - 672}{980 - 784} = \frac{164,5}{196} = 0,839285$$

$$a_0 = 3,4285 - 0,839285(4) = 0,07136$$

Regresión lineal por mínimos cuadrados

Ejercicio: ajustar la recta de los valores de la tabla dada.

$$a_1 = \frac{7(119{,}5) - (28)(24)}{7(140) - (784)} = \frac{836{,}5 \ - 672}{980 - 784} = \frac{164{,}5}{196} = 0{,}839285$$

$$a_0 = 3{,}4285 \ - 0{,}839285(4) = 0{,}07136$$

Por tanto, la recta que se ajusta a los valores dados es:

$$y = \ a_0 + a_1 x$$

$$y = 0{,}07136 + 0{,}839285x$$

Referencias

Chapra, S. C., & Canale, R. P. (2007). Métodos numéricos para ingenieros. McGraw-Hill,.

¡Siempre hacia lo alto!