UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA
SECCIONAL TUNJA

VIGILADA MINEDUCACIÓN - SNIES 1732

Regresión por mínimos cuadrados

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

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

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

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

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

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

$(\sum x_i)^2 =$

¡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

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

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

$\left(\sum x_i\right)^2 = 784$

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

$\bar{x} = 4$

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

¡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

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

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

$(\sum x_i)^2 = 784$

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

$\bar{x} = 4$

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

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

¡Siempre hacia lo alto!

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:

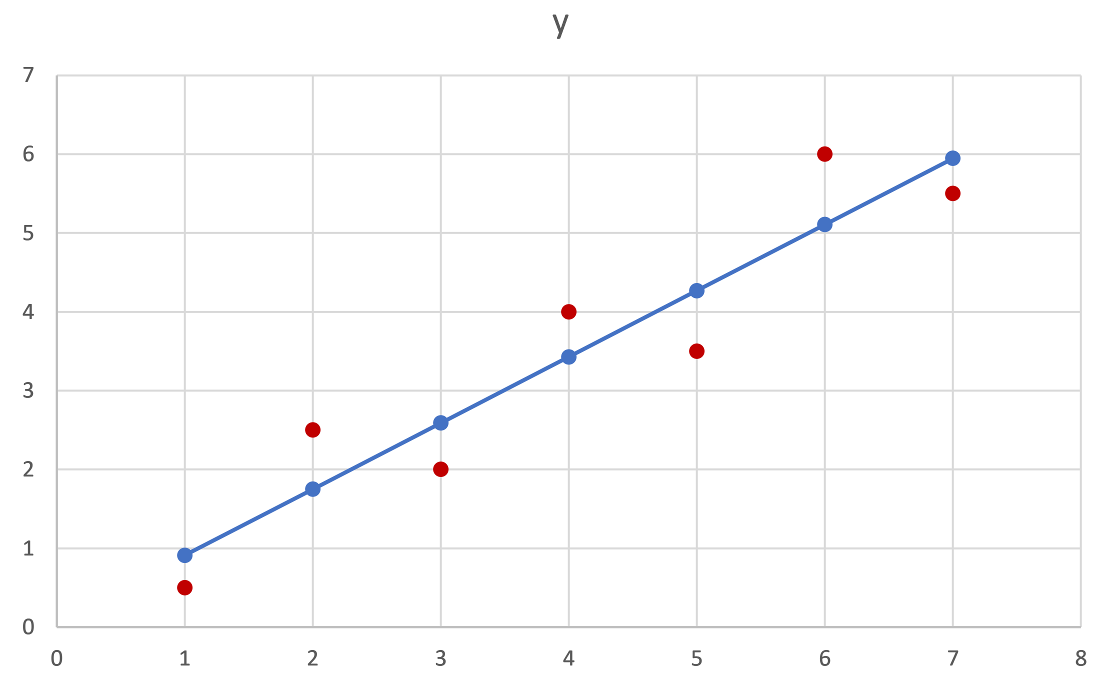
$$y = \ a_0 + a_1 x$$

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

Regresión lineal por mínimos cuadrados

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

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



Regresión lineal por mínimos cuadrados

Ejercicio: medición del error

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

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

x	y	a0	a1xi	e^2
1	0,5	0,07136	0,839285	0,168629316
2	2,5	0,07136	1,67857	0,562605005
3	2	0,07136	2,517855	0,347174316
4	4	0,07136	3,35714	0,32661225
5	3,5	0,07136	4,196425	0,589493806
6	6	0,07136	5,03571	0,797323985
7	5,5	0,07136	5,874995	0,199232786
				2,991071464

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

Error estándar estimado

Grado de asociación lineal entre dos variables

Grado de dispersión

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

Error estándar estimado

Grado de asociación lineal entre dos variables

Grado de dispersión

Desviación de 0 y correlación de 1 significa que los valores se ajustan al 100% a la recta hallada.

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

Error estándar estimado

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

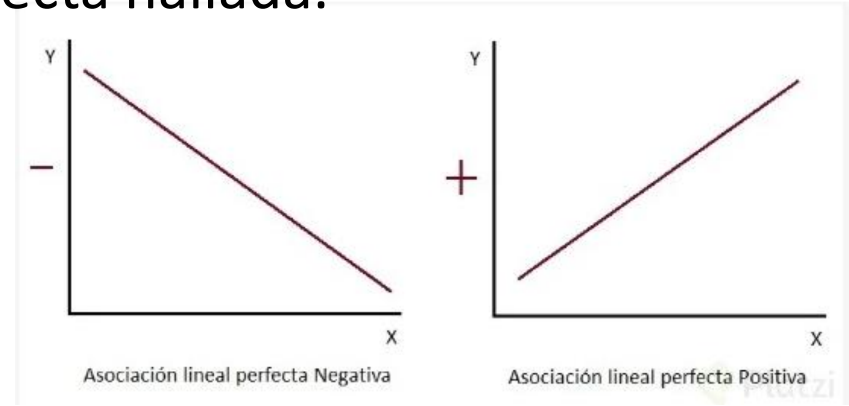
Grado de asociación lineal entre dos variables

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

Grado de dispersión

$$S_t = \Sigma(y_i - \bar{y})^2$$

Desviación de 0 y correlación de 1 significa que los valores se ajustan al 100% a la recta hallada.



Tomada de: httpshttps://static.platzi.com/media/user_upload/ejemplo3-ac75f3c4-5ccf-4c0c-a559-13aaa757173e.jpg

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

x	y	yi-mediay	(yi-mediay)^2
1	0,5	-2,92857	8,576530612
2	2,5	-0,92857	0,862244898
3	2	-1,42857	2,040816327
4	4	0,571429	0,326530612
5	3,5	0,071429	0,005102041
6	6	2,571429	6,612244898
7	5,5	2,071429	4,290816327
28	24		22,71428571

$$S_y = \sqrt{\frac{S_t}{n-1}} =$$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

x	y	yi-mediay	(yi-mediay)^2
1	0,5	-2,92857	8,576530612
2	2,5	-0,92857	0,862244898
3	2	-1,42857	2,040816327
4	4	0,571429	0,326530612
5	3,5	0,071429	0,005102041
6	6	2,571429	6,612244898
7	5,5	2,071429	4,290816327
28	24		22,71428571

$$S_y = \sqrt{\frac{S_t}{n-1}} = 1{,}945691$$

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}} =$$

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

x	y	yi-mediay	(yi-mediay)^2
1	0,5	-2,92857	8,576530612
2	2,5	-0,92857	0,862244898
3	2	-1,42857	2,040816327
4	4	0,571429	0,326530612
5	3,5	0,071429	0,005102041
6	6	2,571429	6,612244898
7	5,5	2,071429	4,290816327
28	24		22,71428571

$$S_y = \sqrt{\frac{S_t}{n-1}} = 1{,}945691$$

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}} = 0{,}773443$$

$S_{y/x} < S_y \Rightarrow$ Modelo de regresión adecuado

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

x	y	yi-mediay	$(\text{yi-mediay})^2$
1	0,5	-2,92857	8,576530612
2	2,5	-0,92857	0,862244898
3	2	-1,42857	2,040816327
4	4	0,571429	0,326530612
5	3,5	0,071429	0,005102041
6	6	2,571429	6,612244898
7	5,5	2,071429	4,290816327
28	24		22,71428571

$$S_y = \sqrt{\frac{S_t}{n-1}} = 1{,}945691$$

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}} = 0{,}773443$$

$$r^2 = \frac{22{,}714285 - 2{,}991071}{22{,}714285}$$

$$r^2 = \frac{S_t - S_r}{S_t}$$

Coeficiente de determinación

$S_{y/x} < S_y \Rightarrow$ Modelo de regresión adecuado

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

x	y	yi-mediay	$(\text{yi-mediay})^2$
1	0,5	-2,92857	8,576530612
2	2,5	-0,92857	0,862244898
3	2	-1,42857	2,040816327
4	4	0,571429	0,326530612
5	3,5	0,071429	0,005102041
6	6	2,571429	6,612244898
7	5,5	2,071429	4,290816327
28	24		22,71428571

$$S_y = \sqrt{\frac{S_t}{n-1}} = 1{,}945691$$

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}} = 0{,}773443$$

$$r^2 = \frac{22{,}714285 - 2{,}991071}{22{,}714285} = 0{,}868317$$

$$r^2 = \frac{S_t - S_r}{S_t}$$

Coeficiente de determinación

$S_{y/x} < S_y \Rightarrow$ Modelo de regresión adecuado

¡Siempre hacia lo alto!

Regresión lineal por mínimos cuadrados

Ejercicio: desviación estándar, error estándar estimado y coeficiente de correlación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

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

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \Sigma(y_i - \bar{y})^2$$

x	y	yi-mediay	(yi-mediay)^2
1	0,5	-2,92857	8,576530612
2	2,5	-0,92857	0,862244898
3	2	-1,42857	2,040816327
4	4	0,571429	0,326530612
5	3,5	0,071429	0,005102041
6	6	2,571429	6,612244898
7	5,5	2,071429	4,290816327
28	24		22,71428571

$$S_y = \sqrt{\frac{S_t}{n-1}} = 1{,}945691$$

$$r^2 = \frac{S_t - S_r}{S_t}$$

Coeficiente de determinación

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}} = 0{,}773443$$

$$r^2 = \frac{22{,}714285 - 2{,}991071}{22{,}714285} = 0{,}868317$$

$$r = 0{,}931835$$

$S_{y/x} < S_y \Rightarrow$ Modelo de regresión adecuado

Los resultados indican que el modelo lineal explicó el 86.8% de la incertidumbre original.

Regresión lineal por mínimos cuadrados

Ejercicio 2: encontrar una recta que se ajuste a los siguientes datos. Calcule la desviación estándar, el coeficiente de correlación, el error estándar estimado y el coeficiente de determinación.

x	0	2	4	6	9	11	12	15	17	19
y	5	6	7	6	9	8	7	10	12	12

Referencias

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

¡Siempre hacia lo alto!

UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA
SECCIONAL TUNJA

VIGILADA MINEDUCACIÓN - SNIES 1732

¡Siempre hacia lo alto!

USTATUNJA.EDU.CO

@santotomastunja