Se observan las edades y las estaturas de un grupo de niños como se muestra en la siguiente tabla:

	Edad	3,3	4,3	5,2	5,6	6,3	7,0	8,1	8,6	9,3	10,1
x											
y	Estatura (cm)	94	112	115	123	117	117	125	139	134	140

$\} n = 10$

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n} = \bar{y} - b\bar{x}$$

Proponga un modelo de regresión lineal simple para los datos de la tabla anterior y conteste:

$n = 10$

$\Sigma xy = 8502{,}8 \quad \Sigma x = 67{,}8 \quad \Sigma y = 1216 \quad \Sigma x^2 = 509{,}59$

$$b = \frac{10 \cdot 8502{,}8 - 67{,}8 \cdot 1216}{10 \cdot 509{,}59 - 67{,}8^2} = 5{,}76$$

$$a = \frac{1216 - 5{,}76 \cdot 67{,}8}{10} = 82{,}55$$

$$\boxed{y = 82{,}55 + 5{,}76x}$$

a. ¿Cuál es la estatura esperada para un niño de 5 años?

$y = 82{,}55 + 5{,}76 \cdot 5 = \boxed{111{,}35}$

b. ¿Cuál es la edad estimada de un niño que mide 125cm de estatura?

$$125 = 82{,}55 + 5{,}76x \rightarrow x = \frac{125 - 82{,}55}{5{,}76} = \boxed{7{,}37}$$

c. ¿Cuál sería la estatura esperada para un niño a los 25 años?

$y = 82{,}55 + 5{,}76 \cdot 25 = \boxed{226{,}55}$

En un experimento donde se quería estudiar la asociación entre consumo de sal y la presión arterial, se asignó aleatoriamente a algunos individuos una cantidad diaria constante de sal en su dieta, y al cabo de un mes se les midió la tensión arterial media. Algunos resultados fueron los siguientes:

X(sal en g)	Y (presión en mm de Hg)
1.8	100
2.2	98
3.5	110
4.0	110
4.3	112
5.0	120

Determine la ecuación de la recta de regresión lineal para los datos de la tabla.

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n} = \bar{y} - b\bar{x}$$

$n = 6 \quad \Sigma xy = 2302{,}2 \quad \Sigma x = 20{,}8 \quad \Sigma y = 650 \quad \Sigma x^2 = 79{,}82$

$$b = \frac{6 \cdot 2302{,}2 - 20{,}8 \cdot 650}{6 \cdot 79{,}82 - 20{,}8^2} = 6{,}34$$

$$a = \frac{650 - 6{,}34 \cdot 20{,}8}{6} = 86{,}35$$

$$\boxed{y = 86{,}35 + 6{,}34x}$$

Ejemplo 88 *Un estudiante de computación, que tiene una pequeña empresa de elaboración de software, llevó a cabo un estudio para determinar la relación entre el tiempo que tarda elaborando un software en días (X) y el ingreso obtenido por su venta en dólares (Y). Se recolectaron los valores de estas variables para 14 software, obteniendo los siguientes datos:*

X:	2	3	5	9	10	16	19	20	24	27	32	41	55	60
Y:	100	200	250	400	500	850	930	900	1300	1360	1500	2050	2800	2900

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n} = \bar{y} - b\bar{x}$$

1. *Encuentre la ecuación de regresión lineal para el ingreso como función del tiempo de elaboración de un software*

$n = 14 \quad \Sigma xy = 587890 \quad \Sigma x = 323 \quad \Sigma y = 16040 \quad \Sigma x^2 = 11871$

$$b = \frac{14 \cdot 587890 - 323 \cdot 16040}{14 \cdot 11871 - 323^2} = 49{,}29$$

$$a = \frac{16040 - 49{,}29 \cdot 323}{14} = 8{,}52$$

$$\boxed{y = 8{,}52 + 49{,}29x}$$

2. *¿Aproximadamente, al elaborar un software, cuánto es el valor promedio del ingreso fijo (ingreso que no depende del tiempo de elaboración)?*

8,52

3. *¿Aproximadamente, al elaborar un software, cuánto es el aumento promedio del ingreso por día de elabolación?*

49,29

Un vendedor de teléfonos celulares quiere poner a prueba cierta marca. Para esto, toma 7 teléfonos del mismo modelo y los carga hasta cierto porcentaje de la batería, y los deja proyectando el mismo video hasta que el celular se apague. Los resultados fueron los siguientes:

% de carga (X)	2	6	8	10	12	14	16
minutos antes de apagarse (Y)	1	5	8	11	15	20	25

1. Determine la ecuación de regresión lineal para la cantidad de minutos que dura en apagarse el celular proyectando el video en función del porcentaje de la carga.

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n} = \bar{y} - b\bar{x}$$

$n = 7 \quad \Sigma xy = 1066 \quad \Sigma x = 68 \quad \Sigma y = 85 \quad \Sigma x^2 = 800$

$$b = \frac{7 \cdot 1066 - 68 \cdot 85}{7 \cdot 800 - 68^2} = 1{,}73$$

$$\boxed{y = -4{,}66 + 1{,}73x}$$

$$a = \frac{85 - 1{,}73 \cdot 68}{7} = -4{,}66$$

Con la llegada de las lluvias aumenta la cantidad de accidentes en carretera. Durante los primeros días de la estación lluviosa se han registrado las horas de lluvia en algunos sectores del GAM y la cantidad de accidentes reportados. Con la información se construyó la tabla

y	**Accidentes**	32	30	34	26	30	35	43	33
x	**Horas**	3	2	4	2	3	4	5	3

$\}\ n = 8$

Utilice el método de mínimos cuadrados para estimar el parámetro β en la ecuación de regresión $y = 20 + \beta x$, en la que el número de accidentes reportados depende de las horas de lluvia.

Utilice el método de mínimos cuadrados para estimar el parámetro β en la ecuación de regresión $y = 20 + \beta x$, en la que el número de accidentes reportados depende de las horas de lluvia.

$y = 20 + \beta x \longrightarrow y - 20 - \beta x$

$\Sigma (y - 20 - \beta x)^2$

$\Sigma 2(y - 20 - \beta x) \cdot -x = 0$

$\Sigma xy - 20\Sigma x - \beta \Sigma x^2 = 0$

$$\beta = \frac{\Sigma xy - 20\Sigma x}{\Sigma x^2} \qquad \Sigma xy = 888 \quad \Sigma x = 26 \quad \Sigma x^2 = 92$$

$$\beta = \frac{888 - 20 \cdot 26}{92} = 4 \longrightarrow \boxed{y = 20 + 4x}$$

Ejemplo 89 *Considere los datos de la siguiente tabla:*

X :	2	3	5	9	13
Y :	8	10	18	30	40

A partir de estos datos, estime el coeficiente β *de la ecuación de regresión* $\mu_{Yx} = 2 + \beta x$ *utilizando el método de mínimos cuadrados.*

$$\sum (y - 2 - \beta x)^2$$

$$\sum 2(y - 2 - \beta x) \cdot -x = 0$$

$$-\sum xy + 2\sum x + \beta \sum x^2 = 0$$

$$\beta = \frac{\sum xy - 2\sum x}{\sum x^2} \qquad \sum xy = 926 \quad \sum x = 32 \quad \sum x^2 = 288$$

$$\beta = \frac{926 - 2 \cdot 32}{288} = 2{,}99$$

$$\boxed{y = 2 + 2{,}99x}$$

Ejercicio 44 *Considere los datos de la siguiente tabla:*

x :	1	5	10	20
y :	4	12	21	43

A partir de estos datos, estime el coeficiente β *de la ecuación de regresión* $\mu_{Yx} = \beta + \beta x$ *utilizando el método de mínimos cuadrados.* $R/ \quad b = \frac{567}{281}$

$$\sum (y - \beta - \beta x)^2$$

$$\sum 2(y - \beta - \beta x)(-1 - x) = 0$$

$$\sum (y - \beta - \beta x)(-1 - x) = 0$$

$$\sum (-y + \beta + \beta x - xy + \beta x + \beta x^2) = 0$$

$$\sum (-y + \beta + 2\beta x + \beta x^2 - xy) = 0$$

$$\sum(-y + \beta + 2\beta x + \beta x^2 - xy) = 0$$

$$-\sum y + \beta \sum 1 + 2\beta \sum x + \beta \sum x^2 - \sum xy = 0$$

$$\beta \sum 1 + 2\beta \sum x + \beta \sum x^2 = \sum xy + \sum y$$

$$\beta\left(\sum 1 + 2\sum x + \sum x^2\right) = \sum xy + \sum y$$

$$\beta = \frac{\sum xy + \sum y}{\sum 1 + 2\sum x + \sum x^2}$$

$\sum xy = 1134$

$\sum y = 80$

$\sum 1 = 1 \cdot 4 = 4$

$\sum x = 36$

$\sum x^2 = 526$

$$\beta = \frac{1134 + 80}{4 + 2 \cdot 36 + 526}$$

$$\beta = 2{,}0166$$

$$\boxed{y = 2{,}0166 + 2{,}0166x}$$

Ejemplo 90 *Un estudiante de computación, que tiene una pequeña empresa de elaboración de software, llevó a cabo un estudio para determinar la relación entre el tiempo que tarda elaborando un software en días* (X) *y el ingreso obtenido por su venta en dólares* (Y). *Se recolectaron los valores de estas variables para 14 software, obteniendo los siguientes datos:*

$$\sum x = 323 \qquad \sum y = 16040 \qquad \sum x^2 = 11871$$
$$\sum y^2 = 29162000 \qquad \sum xy = 587890$$

Asuma las hipótesis de regresión. En un ejemplo anterior se determinó que la ecuación de regresión lineal para el ingreso como función del tiempo de elaboración de un software es

$$\widehat{y} = 49.2935x + 8.44282$$

1. *Determine un intervalo de confianza del 90% para el valor promedio del ingreso fijo (ingreso que no depende del tiempo de elaboración)* $\to a$

$a = 8,44282 \quad b = 49,2935 \quad n = 14$

$$a \pm t_{\delta/2,\nu} s \sqrt{\frac{\sum x^2}{n S_{xx}}} \text{ con } \nu = n-2$$

$$S_{xx} = \sum(x_i - \bar{x})^2 = (n-1)s_x^2 = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$S_{yy} = \sum(y_i - \bar{y})^2 = (n-1)s_y^2 = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

$$s = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2} = \sqrt{\frac{S_{yy} - bS_{xy}}{n-2}}$$

$$S_{xx} = 11871 - \frac{323^2}{14} = \frac{61865}{14}$$

$$S_{yy} = 29162000 - \frac{16040}{14} = 1078474 2,86$$

$$S_{xy} = 587890 - \frac{323 \cdot 16040}{14} = \frac{1524470}{7}$$

$$s = \sqrt{\frac{10784742,86 - 49,2935 \cdot \frac{1524470}{7}}{14-2}} = 62,86323259$$

$$\alpha = 0,10 \to \frac{\alpha}{2} = 0,05 \to t_{0,05,12} = \pm 1,78229 \quad V = 12$$

$$\text{Lim inf} = 8,44282 - 1,78229 \cdot 62,86323259 \sqrt{\frac{11871}{14 \cdot \frac{61865}{14}}}$$
$$= -40,63623863$$

$$\text{Lim sup} = 8,44282 + 1,78229 \cdot 62,86323259 \sqrt{\frac{11871}{14 \cdot \frac{61865}{14}}}$$
$$= 57,52187863$$

$$\boxed{\text{IC }]-40.63623863,\ 57.52187863[}$$

2. *Determine un intervalo de confianza del 90% para el $\underbrace{\text{aumento promedio del ingreso por día de}}_{b}$ elabolación*

$b = 49{,}2935$

IC para β: $b \pm t_{\delta/2,\nu} s \sqrt{\frac{1}{S_{xx}}}$ con $\nu = n - 2$

$t_{0.05,12} = \pm 1{,}78229$ $\quad S_{xx} = \frac{61865}{14}$

$s = 62{,}86323259$

$49{,}2935 - 1{,}78229 \cdot 62{,}86323259 \cdot \sqrt{\frac{1}{\frac{61865}{14}}} = 47{,}60804$

$49{,}2935 + 1{,}78229 \cdot 62{,}86323259 \cdot \sqrt{\frac{1}{\frac{61865}{14}}} = 50{,}97895$

$\boxed{]47.60804, 50.97895[}$

Determine el IC de 90% para el ingreso esperado por venta de software con un tiempo de elaboración de 30 días

$y = 49{,}2935 \cdot 30 + 8{,}44282$ $\quad n = 14$

$y = 1487{,}24782$ $\quad v = 12$

$\hat{y} \pm t_{\delta/2,\nu} s \sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}}$ con $\nu = n - 2$ y $\hat{y} = a + bx$

$t_{0.05,12} = \pm 1{,}78229$ $\quad s = 62{,}86323259$

$S_{xx} = \frac{61865}{14}$ $\quad x = 30$

$\bar{x} = \frac{\Sigma x}{n} = \frac{323}{14}$

Lim inf:

$1487{,}24782 - 1{,}78229 \cdot \sqrt{\frac{1}{14} + \frac{\left(30 - \frac{323}{14}\right)^2}{\frac{61865}{14}}}$

$= 1486{,}736542$

Lim Sup:

$$1487{,}24782 - 1{,}78229 \cdot 62{,}86323259 \cdot \sqrt{\frac{1}{14} + \frac{\left(30 - \frac{323}{14}\right)^2}{\frac{63865}{14}}}$$

$$= 1519{,}38843$$

$$\boxed{\left]1486{,}736542 \,,\ 1519{,}38843\right[}$$

Determine un intervalo de predicción de 90% para el ingreso por venta de software con un tiempo de elaboración de 30 días

$$y = 49{,}2935 \cdot 30 + 8{,}44282 \qquad n = 14$$

$$y = 1487{,}24782 \qquad v = 12$$

Intervalo de predicción para $Y|x$ cuando se cumplen las hipótesis de regresión

Extremos del intervalo: $\hat{y} \pm t_{\delta/2,\nu}\, s \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}}$

con $\nu = n - 2$ y $\hat{y} = a + bx$

$$t_{0.05,12} = {}^{+}1{,}78229 \qquad S = 62{,}86323259$$

$$S_{xx} = \frac{63865}{14} \qquad x = 30$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{323}{14}$$

Lim inf:

$$1487{,}24782 - 1{,}78229 \cdot 62{,}86323259 \cdot \sqrt{1 + \frac{1}{14} + \frac{\left(30 - \frac{323}{14}\right)^2}{\frac{63865}{14}}}$$

$$= 1370{,}688415$$

$$1487{,}24782 + 1{,}78229 \cdot 62{,}86323259 \cdot \sqrt{1 + \frac{1}{14} + \frac{\left(30 - \frac{323}{14}\right)^2}{\frac{63865}{14}}}$$

$$= 1603{,}807225$$

$$\boxed{\left]1370{,}688415 \,,\ 1603{,}807225\right[}$$

Con un nivel de signicancia del 10%, ¿Hay evidencia de que el aumento promedio del ingreso por día de elabolación sea de 50 dólares por día?

$H_0: \beta = 50$ $\quad b = 49{,}2935$ $\quad S_{xx} = \frac{63865}{19}$ $\quad n = 14$

$H_1: \beta \neq 50$ $\quad \beta_0 = 50$ $\quad v = 12$

$S = 62{,}86323259$ $\quad T = (B - \beta_0)\frac{\sqrt{S_{xx}}}{S}$ con $\nu = n - 2$

$$t_{obs} = \frac{(49{,}2935 - 50) \cdot \sqrt{\frac{63865}{19}}}{62{,}86323259} = -0{,}74709$$

$\alpha = 0{,}1 \rightarrow \frac{\alpha}{2} = 0{,}05 \rightarrow t_{0.05,12} = \pm 1{,}78229$

$\boxed{\text{R/ } -1{,}78229 < -0{,}74709 < 1{,}78229, \text{ No se rechaza } H_0}$

Con valor P

$2P(t > -0{,}74709) = 0{,}46940$ $\quad \alpha = 0{,}10$

$\boxed{\text{R/ Como } P = 0{,}46940 > \alpha = 0{,}10, \text{ No se rechaza } H_0}$

Ejemplo

Una cooperativa de ahorro ha sufrido un cierre técnico debido a una deficiente administración. Esta entidad financiera estima que el tiempo (en días) para la devolución de dinero depende linealmente de la cantidad de dinero (millones de colones) que el cliente tiene en su cuenta. Los siguientes datos se refieren a información de 10 clientes sobre el dinero en su cuenta (x) y el tiempo que han esperado para su devolución (y).

$$\sum_{i=1}^{10} x_i^2 = 433 \qquad \sum_{i=1}^{10} x_i = 51 \qquad \sum_{i=1}^{10} y_i^2 = 1545 \qquad \sum_{i=1}^{10} y_i = 95 \qquad \sum_{i=1}^{10} x_i y_i = 808$$

a) Según la recta de regresión lineal, ¿cuántos días más, en promedio, debe esperar un cliente por la devolución de su dinero por cada millón de colones en su cuenta? **(4 puntos)**

$$b = \frac{10 \cdot 808 - 51 \cdot 95}{10 \cdot 833 - 51^2} = 1{,}871$$

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$n = 10$

c) Construya un intervalo de predicción del 95 % para el tiempo que debe esperar un cliente que tiene 5 millones de colones en su cuenta. **(5 puntos)**

$$a = \frac{95 - 1{,}871 \cdot 51}{10} = -0{,}0421$$

$$a = \frac{\sum y - b\sum x}{n} = \bar{y} - b\bar{x}$$

$$\hat{y} \pm t_{\delta/2,\nu} s\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}}$$

con $\nu = n-2$ y $\hat{y} = a + bx$

$y = -0{,}0421 + 1{,}871x \qquad x = 5$

$y = -0{,}0421 + 1{,}871 \cdot 5 = 9{,}3129 \qquad n = 10 \quad \nu = 8$

$\alpha = 0{,}05 \to \frac{\alpha}{2} = 0{,}025 \to t_{0.025,8} = 2{,}306$

$$S_{xx} = 433 - \frac{51^2}{10} = 172{,}9$$

$$S_{xx} = \sum(x_i - \bar{x})^2 = (n-1)s_x^2 = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$S_{yy} = \sum(y_i - \bar{y})^2 = (n-1)s_y^2 = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

$$s = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2} = \sqrt{\frac{S_{yy} - bS_{xy}}{n-2}}$$

$$S_{yy} = 1545 - \frac{95^2}{10} = 642{,}5$$

$$S_{xy} = 808 - \frac{51 \cdot 95}{10} = 323{,}5$$

$$s = \sqrt{\frac{642{,}5 - 1{,}871 \cdot 323{,}5}{10-2}} = 2{,}157$$

$y = 9{,}3129$ $\qquad n = 10$ $\qquad \hat{y} \pm t_{\delta/2,\nu} s \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}}$

con $\nu = n - 2$ y $\hat{y} = a + bx$

$t_{0.025,8} = 2{,}306$ $\qquad x = 5$

$S = 2{,}157$ $\qquad \bar{x} = \frac{51}{10} = 5{,}1$

$S_{xx} = 172{,}9$

$$9{,}3129 - 2{,}306 \cdot 2{,}157 \cdot \sqrt{1 + \frac{1}{10} + \frac{(5 - 5{,}1)^2}{172{,}9}}$$

$$9{,}3129 + 2{,}306 \cdot 2{,}157 \cdot \sqrt{1 + \frac{1}{10} + \frac{(5 - 5{,}1)^2}{172{,}9}}$$

$$\boxed{]4{,}09594,\ 14{,}52985[}$$

Ejemplo

Una cooperativa de ahorro ha sufrido un cierre técnico debido a una deficiente administración. Esta entidad financiera estima que el tiempo (en días) para la devolución de dinero depende linealmente de la cantidad de dinero (millones de colones) que el cliente tiene en su cuenta. Los siguientes datos se refieren a información de 10 clientes sobre el dinero en su cuenta (x) y el tiempo que han esperado para su devolución (y).

$$\sum_{i=1}^{10} x_i^2 = 433 \qquad \sum_{i=1}^{10} x_i = 51 \qquad \sum_{i=1}^{10} y_i^2 = 1545 \qquad \sum_{i=1}^{10} y_i = 95 \qquad \sum_{i=1}^{10} x_i y_i = 808$$

$\rightarrow r^2$

b) ¿Qué porcentaje de la variación en los días de espera para la devolución del dinero depende de factores diferentes a la cantidad de dinero en la cuenta del cliente? **(3 puntos)**

$$b = \frac{10 \cdot 808 - 51 \cdot 95}{10 \cdot 433 - 51^2} = 1{,}87 \qquad r^2 = b\frac{s_x}{s_y} = \left(b\sqrt{\frac{S_{xx}}{S_{yy}}}\right)^2$$

$$S_{xx} = 433 - \frac{51^2}{10} = 172{,}9 \qquad b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$S_{yy} = 1545 - \frac{95^2}{10} = 642{,}5$$

$$r^2 = \left(1{,}87 \cdot \sqrt{\frac{172{,}9}{692{,}5}}\right)^2 = 0{,}941$$

$$r^2 = b\frac{s_x}{s_y} = \left(b\sqrt{\frac{S_{xx}}{S_{yy}}}\right)^2$$

$$1 - 0{,}941 = 0{,}059 \approx 0{,}06$$

Al rededor del 6% de variacion

- Una fabrica recolectó la siguiente información de 8 de sus trabajadores sobre el número de minutos que llegaron tarde al trabajo (X) y el número de piezas defectuosas que fabricaron ese día (Y):

X	2	5	10	15	20	25	30	40
Y	2	4	9	12	15	20	24	30

Suponiendo que se cumplen las hipótesis de regresión.

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n} = \bar{y} - b\bar{x}$$

- Encuentre la ecuación de regresión lineal para el número de piezas defectuosas que fabrica un empleado por día como función del número de minutos que llega tarde al trabajo.

$n = 8 \quad \Sigma xy = 3014 \quad \Sigma x = 147 \quad \Sigma y = 116 \quad \Sigma x^2 = 3879$

$$b = \frac{8 \cdot 3014 - 147 \cdot 116}{8 \cdot 3879 - 147^2} = 0{,}749230606$$

$$a = \frac{116 - 0{,}749230606 \cdot 147}{8} = 0{,}732887615$$

$$\boxed{R/\ y = 0{,}732887615 + 0{,}749230606}$$

- Aproximadamente, ¿qué porcentaje de $\overbrace{\text{variación}}^{r^2}$ de en el número de piezas defectuosas que elabora un empleado en un día se debe a otros factores distintos al número de minutos que llega tarde a trabajar?

Sin importar $x \to 1-r^2$

$$r^2 = b\frac{s_x}{s_y} = \left(b\sqrt{\frac{S_{xx}}{S_{yy}}}\right)^2$$

$n = 8$ $\quad \Sigma x^2 = 3879$ $\quad \Sigma y^2 = 2346$

$b = 0{,}749230606$ $\quad \Sigma x = 147$ $\quad \Sigma y = 116$

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$S_{xx} = 3879 - \frac{147^2}{8} = 1177{,}875$$

$$S_{yy} = 2346 - \frac{116^2}{8} = 664$$

$$r^2 = b\frac{s_x}{s_y} = \left(b\sqrt{\frac{S_{xx}}{S_{yy}}}\right)^2$$

$$1 - r^2 = 1 - \left(0{,}749230606 \cdot \sqrt{\frac{1177{,}875}{664}}\right)^2 = \boxed{0{,}004222877}$$

- Una fabrica recolectó la siguiente información de 8 de sus trabajadores sobre el número de minutos que llegaron tarde al trabajo (X) y el número de piezas defectuosas que fabricaron ese día (Y):

X	2	5	10	15	20	25	30	40
Y	2	4	9	12	15	20	24	30

Suponiendo que se cumplen las hipótesis de regresión.

- A un nivel de significancia del 5%, ¿existe evidencia de que el aumento promedio en el número de piezas defectuosas que realiza un empleado por cada minuto que llega tarde al trabajo es mayor a 0.7?

$H_0: \beta = 0,7 \ (\leq)$

$H_1: \beta > 0,7$

$n = 8$

$b = 0,749230606$

$\sum x^2 = 3879$

$\sum x = 147$

$\sum xy = 3014$

$\sum y^2 = 2396$

$\sum y = 116$

$\alpha = 0,05$

$v = 6$

$\beta_0 = 0,7$

$$T = (B - \beta_0)\frac{\sqrt{S_{xx}}}{S} \text{ con } \nu = n - 2$$

Como es 1 cola se usa directa sin dividir

$S_{xx} = 3879 \quad \frac{147^2}{8} = 1177,875$

$S_{yy} = 2396 - \frac{116^2}{8} = 664$

$S_{xy} = 3014 - \frac{147 \cdot 116}{8} = 882,5$

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

$$s = \sqrt{\frac{S_{yy} - bS_{xy}}{n-2}}$$

$$S = \sqrt{\frac{664 - 0,749230606 \cdot 882,5}{8-2}} = 0,68361663$$

$$T = (B - \beta_0)\frac{\sqrt{S_{xx}}}{S} \text{ con } \nu = n - 2$$

$$t_{obs} = (0,749230606 - 0,7) \cdot \frac{\sqrt{1177,875}}{0,68361663} = 2,47157$$

$t_c = t_{0.05,6} = 1,94318$

R/ Como $t_{obs} = 2,47157 > t_c = 1,94318$ se rechaza H_0

$P(t > 2,47157) = 0,02418 < \alpha = 0,05$ se rechaza H_0