REGRESIÓN LINEAL SIMPLE

SISTEMAS EMBEBIDOS

REGRESIÓN LINEAL SIMPLE

Una de las aplicaciones mas importantes de la estadística implica la estimación del valor medio de una variable de respuesta y o la predicción de algún valor futuro de y con base el conocimiento de un conjunto de variables independientes relacionadas, $x_1, x_2, . . . x_k$.

REGRESIÓN LINEAL SIMPLE

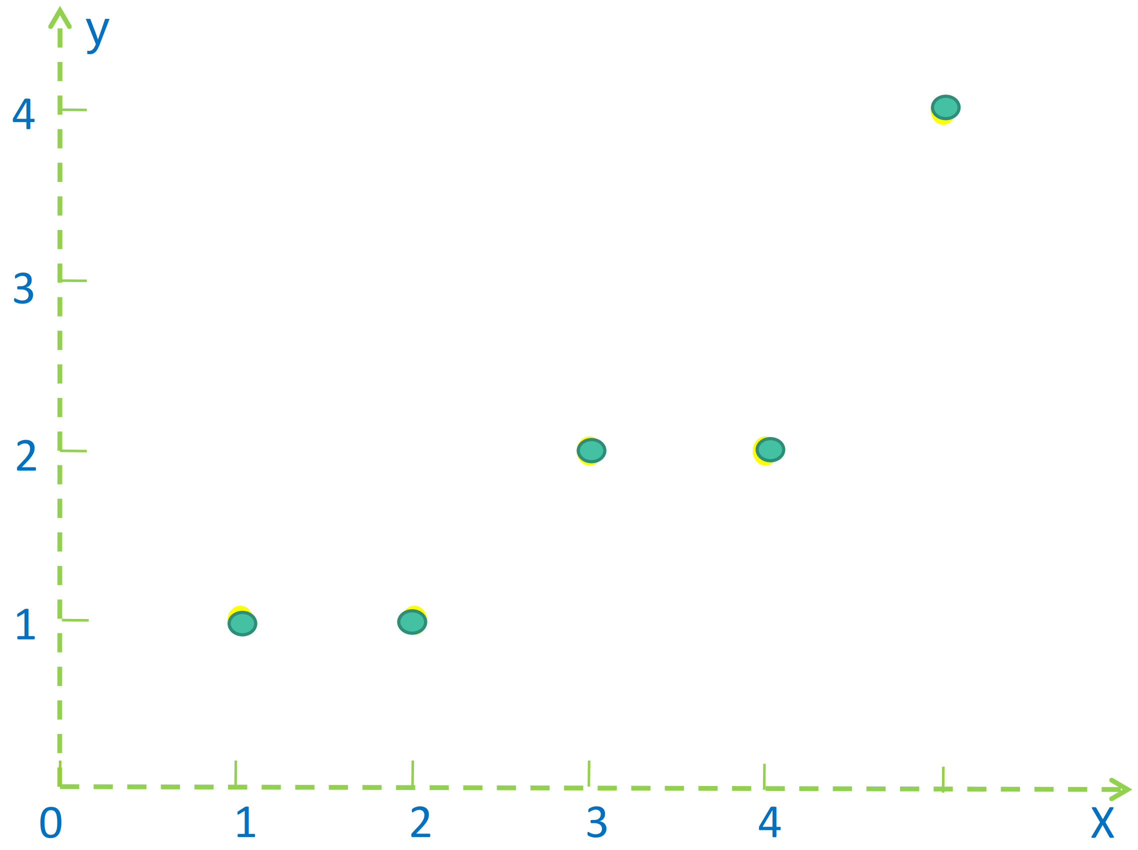
Los modelos que se emplean para relacionar una variable dependiente y con las variables independientes $x_1, x_2, \dots x_k$ se denominan *modelos de regresión* o *modelos estadísticos lineales* porque expresan el valor medio de y para valores dados de $x_1, x_2, \dots x_k$ como una función lineal de un conjunto de parámetros desconocidos.

REGRESIÓN LINEAL SIMPLE

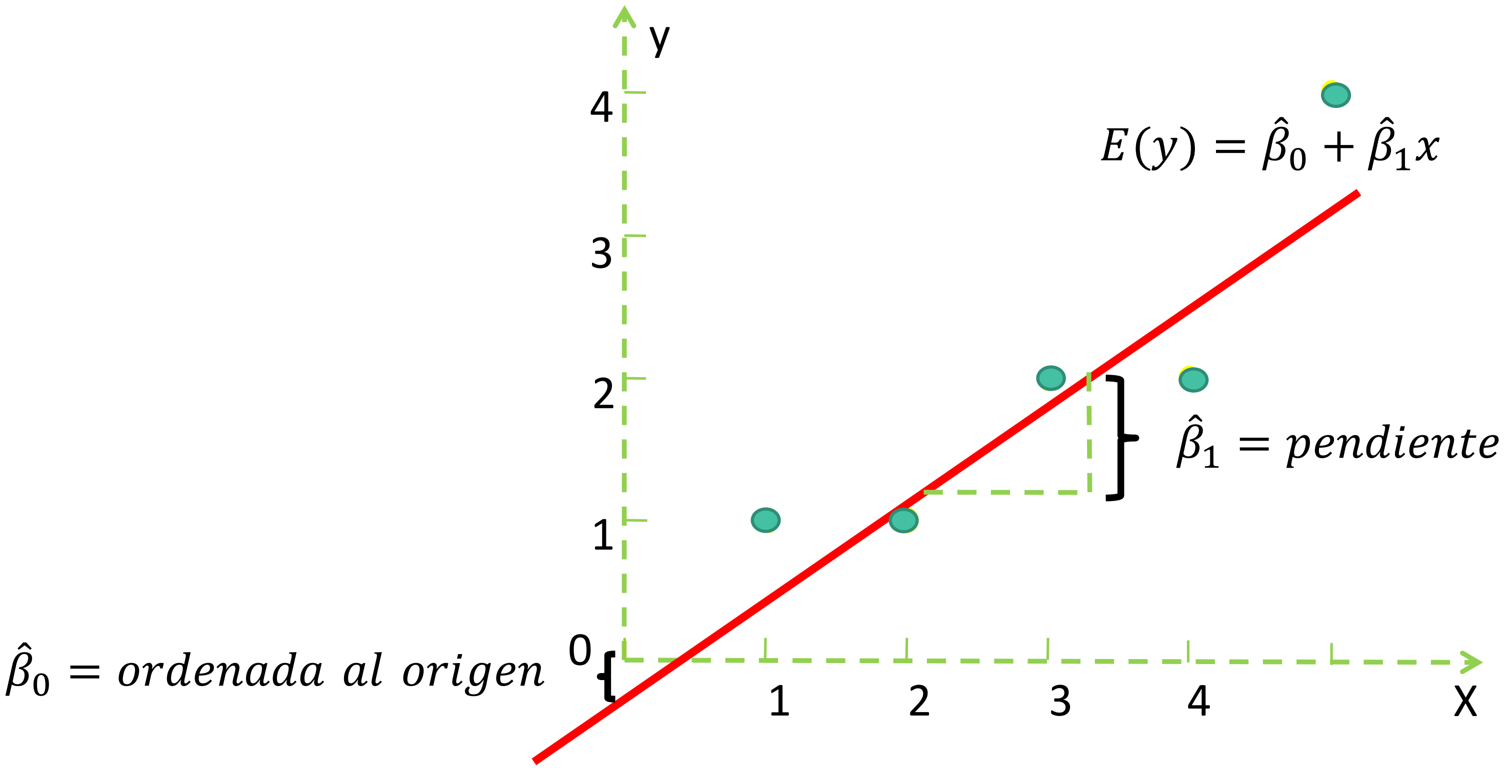
Los conceptos de análisis de regresión se presentan empleando un modelo de regresión muy sencillo, uno que relaciona y con una sola variable x. Aprenderemos a ajustar este modelo a un conjunto de datos mediante el *método de los mínimos cuadrados.*

REGRESIÓN LINEAL SIMPLE

Figura 1.



REGRESIÓN LINEAL SIMPLE



REGRESIÓN LINEAL SIMPLE

Donde:

y = variable dependiente

x = variable independiente

$$y = \beta_0 + \beta_1 x + \varepsilon$$

$E(y)=\beta_0+\beta_1 x$ es el componente determinístico (la ecuación de una línea recta)

ε = componente de error aleatorio

$\hat{\beta}_0$ = punto en que la línea corta el eje y

$\hat{\beta}_1$ = pendiente de la línea

REGRESIÓN LINEAL SIMPLE

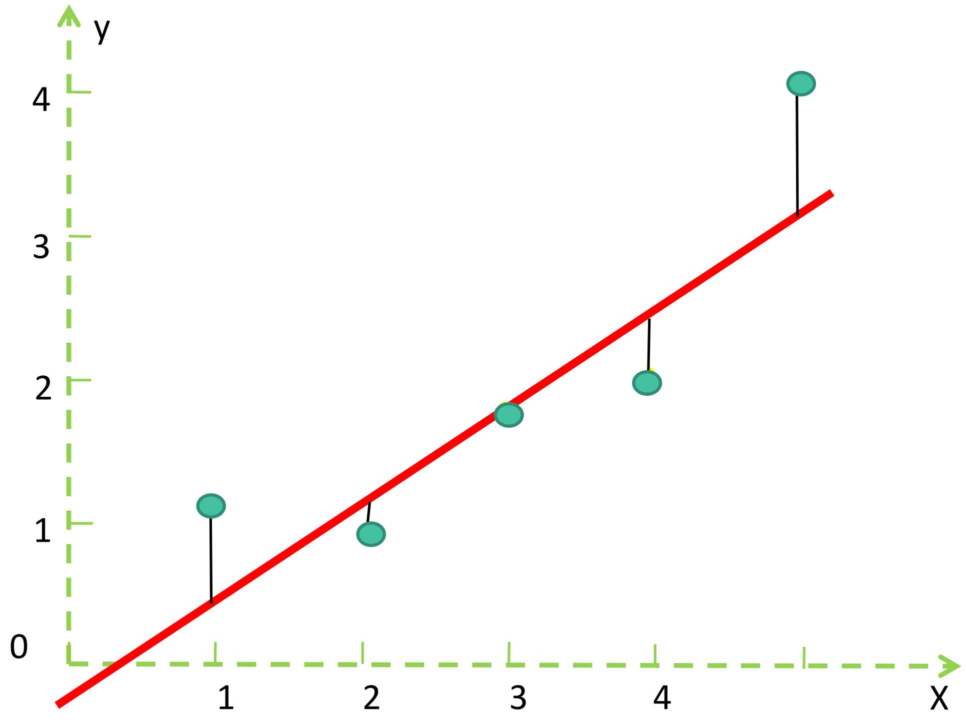
$y = \beta_0 + \beta_1 x + \varepsilon$

$y = \beta_0 + mx$

β_0 = (Σx^2Σy-ΣxΣxy)/[nΣx^2 –(Σx)^2]

m = (nΣxy-Σxy)/[nΣx^2 –(Σx)^2]

REGRESIÓN LINEAL SIMPLE



REGRESIÓN LINEAL SIMPLE

Pendiente: $$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$$

Ordenada al origen: $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

Donde: $$SS_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$SS_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

R CUADRADO

$$R^2 = \frac{\sum_{t=1}^{T}(\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^{T}(Y_t - \bar{Y})^2}$$

El coeficiente de determinación se define como la proporción de la varianza total de la variable explicada por la regresión. El coeficiente de determinación, también llamado R cuadrado, refleja la bondad del ajuste de un modelo a la variable que pretender explicar.

Es importante saber que el resultado del coeficiente de determinación oscila entre 0 y 1. Cuanto más cerca de 1 se sitúe su valor, mayor será el ajuste del modelo a la variable que estamos intentando explicar. De forma inversa, cuanto más cerca de cero, menos ajustado estará el modelo y, por tanto, menos fiable será.