

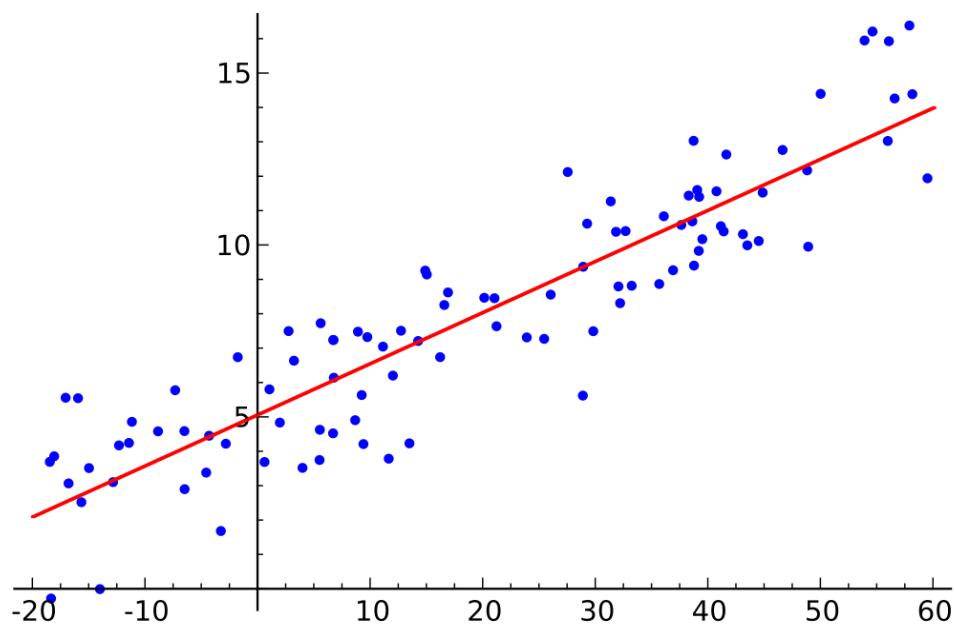
## PROGRAMA “REGRESIÓN LINEAL”

Ing. Brian García Sarmina

### Descripción del problema

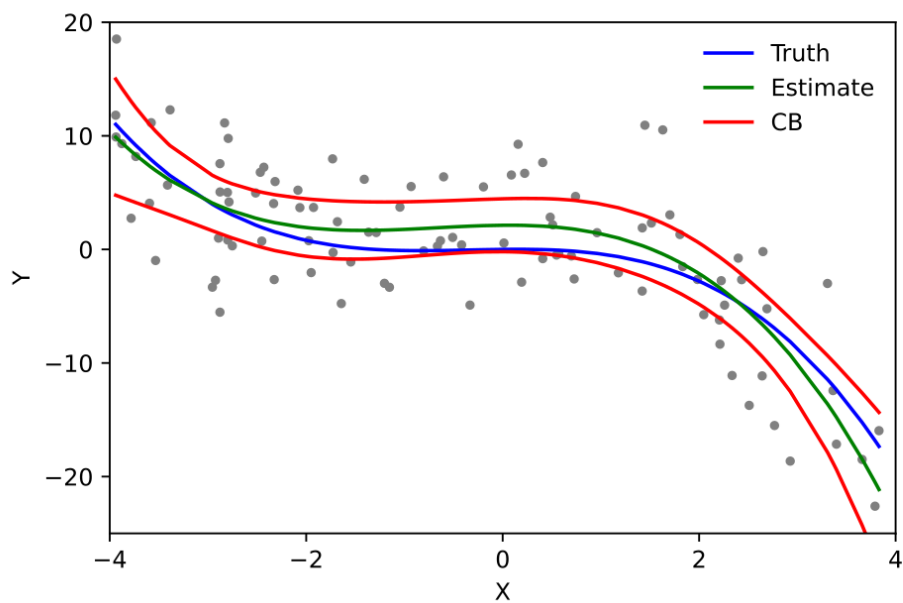
Se busca analizar el modelo simple de “Regresión Lineal”, este modelo es ocupado en procesos de inteligencia artificial (con la obviedad del aumento de complejidad dependiendo del problema abordado). La regresión lineal busca predecir una respuesta a partir de las características de la información usada, en este caso, al tratarse del modelo simple solo se contempla una característica.

Este modelo asume que **LAS DOS VARIABLES ESTÁN RELACIONADAS LINEALMENTE**, y se busca que la función lineal generada sea capaz de predecir un valor respuesta **Y** de la forma más precisa, todo a partir de la variable independiente **X**.



Fuente:

[https://en.wikipedia.org/wiki/Linear\\_regression#/media/File:Linear\\_regression.svg](https://en.wikipedia.org/wiki/Linear_regression#/media/File:Linear_regression.svg)



Fuente:

[https://en.wikipedia.org/wiki/Linear\\_regression#/media/File:Polyreg\\_scheffe.svg](https://en.wikipedia.org/wiki/Linear_regression#/media/File:Polyreg_scheffe.svg)

Sea:

- $f(x_i)$  es el valor de predicción según la función de regresión.
- $\beta_0$  corresponde al valor del intercepto de  $y$ .
- $\beta_1$  corresponde a la pendiente de la línea de regresión.
- $\epsilon_i$  es un error asociado.
- $n$  es el número de elementos empleados de información.
- $\bar{x}$  promedio de valores de coordenada  $x$ .
- $\bar{y}$  promedio de valores de coordenada  $y$ .
- $SDC_{xy}$  "Suma de Desviaciones Cruzadas" entre  $x$  y  $y$ .
- $SDC_{xx}$  "Suma de Desviaciones Cuadradas" de  $x$ .

Función de recta de regresión:

$$f(x_i) = \beta_0 + \beta_1 x_i + \epsilon_i$$

Suma de "desviación cruzada" entre  $x$  y  $y$ :

$$SDC_{xy} = \sum_{i=1}^n x_i y_i - n \bar{x} \cdot \bar{y}$$

Suma de "desviación cuadrada" de  $x$ .

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

Cálculo de  $\beta_1$  y  $\beta_0$ :

$$\beta_1 = \frac{SDC_{xy}}{SDC_{xx}} \quad \beta_0 = \bar{y} - \beta_1 \bar{x}$$

Aplicar una metodología para resolver el problema

1. Generar puntos aleatorios (muestra) con una distribución lineal.
2. Procesar los puntos para calcular
  - a. Promedio de x.
  - b. Promedio de y.
  - c.  $SDC_{xy}$
  - d.  $SDC_{xx}$
  - e.  $\beta_1$  y  $\beta_0$
  - f. Calcular la ecuación de la recta de regresión.
  - g. Graficar los puntos generados y la recta de regresión.