



Regresión

Es una técnica estadística utilizada para simular la relación existente entre dos o más variables.

Por lo tanto se puede emplear para construir un modelo que permita predecir el comportamiento de una variable dada.



Introducción

-ESTUDIO DE UNA VARIABLE AISLADA

-ESTUDIO DE UNA VARIABLE CUANDO ESTÁ RELACIONADA CON OTRA

Tipos de Regresión

REGRESIÓN	ECUACIÓN
Lineal	$y = A + Bx$
Logarítmica	$y = A + B\ln(x)$
Exponencial	$y = Ae^{(Bx)}$
Cuadrática	$y = A + Bx + Cx^2$

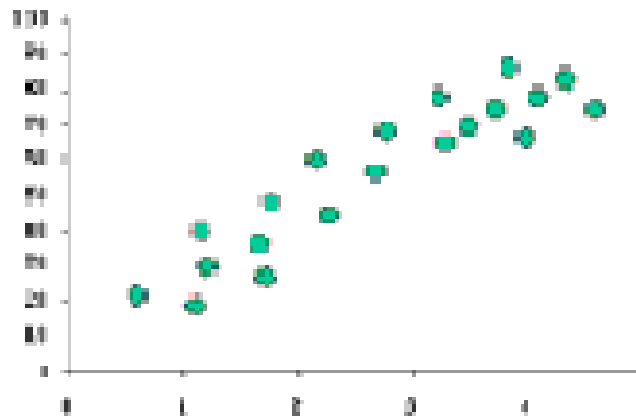


Utilidad

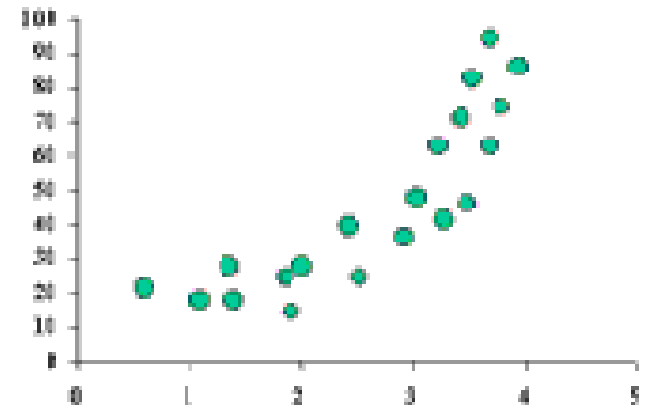
El análisis de regresión es una herramienta estadística que permite analizar y predecir o estimar observaciones futuras de dos o más variables relacionadas entre sí, es decir una herramienta útil para la planeación

RELACIONES

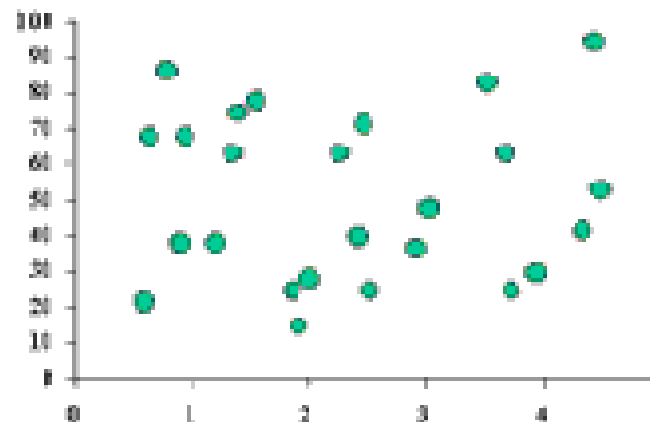
RELACION LINEAL



RELACION EXPONENCIAL



SIN RELACION



Regresión Lineal

Ajuste de una recta:

Pero de todos estos principios, el más simple es el de mínimos cuadrados: con los demás se llegaría a los cálculos más complicados.

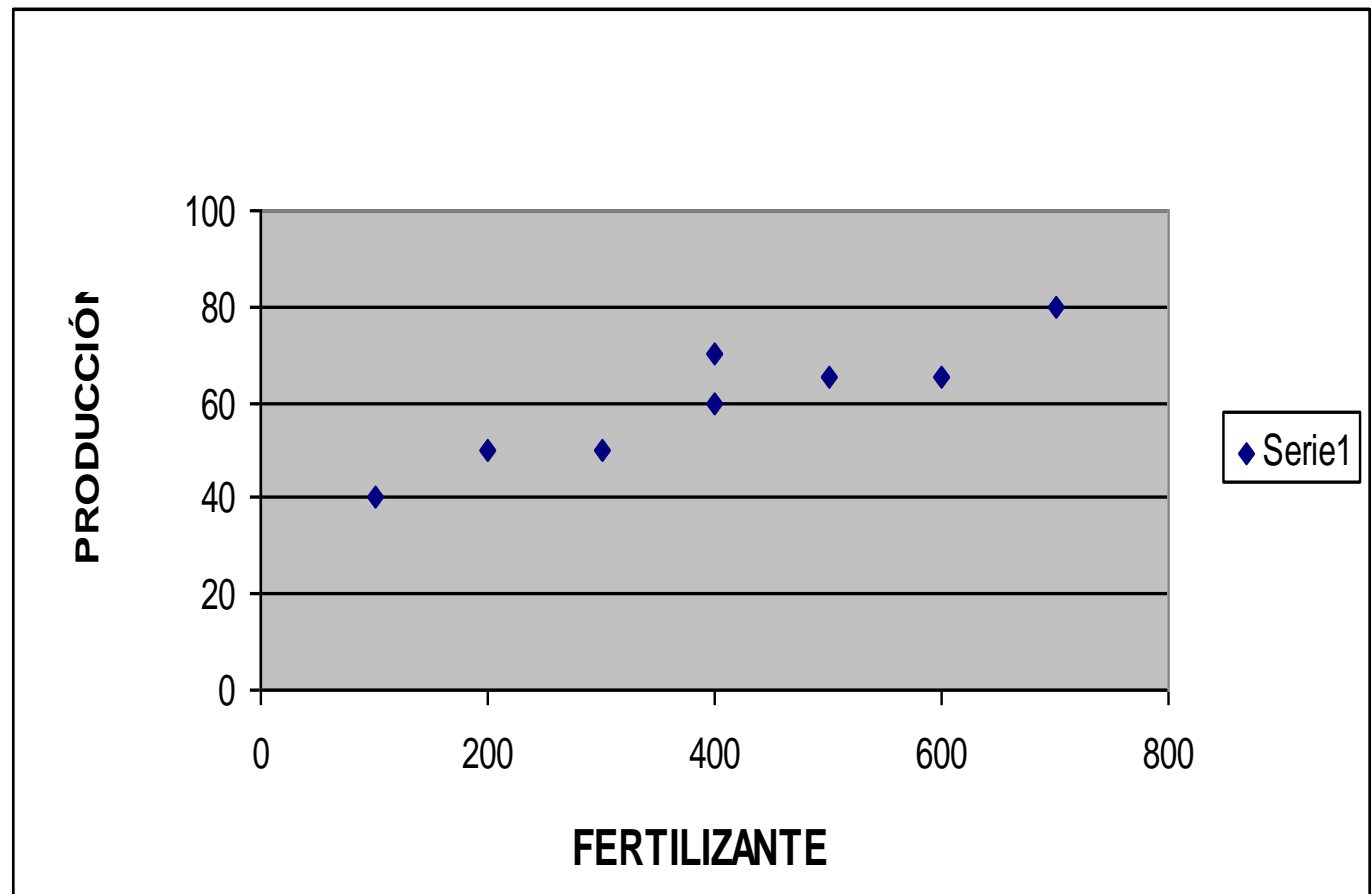
K.F. Gauss, 1809

EJEMPLO

- La producción en función de un fertilizante nos produce la siguiente tabla de relación

FERTILIZANTE	PRODUCCION
X	Y
100	40
200	50
300	50
400	70
500	65
600	65
700	80

GRAFICO DE LA RELACIÓN



MÍNIMOS CUADRADOS

Los cálculos consisten en encontrar una recta cuya ecuación es $Y' = a + bX$ que se ajuste a los puntos de la relación .

Es decir encontrar las fórmulas para encontrar la pendiente b y la ordenada al origen a .

Nótese que X es la misma para los puntos y la función, en cambio Y no necesariamente es igual a Y' a no ser que coincidan exactamente un punto cualquiera con la recta encontrada.

FÓRMULAS

$$b = \frac{\sum xy}{\sum x^2}$$

$$a = \bar{Y} - b\bar{X}$$

Con $x = X - \bar{X}$

$y = Y - \bar{Y}$

FÓRMULAS

$$Y = a_0 + a_1 X$$

$$\begin{cases} \sum Y = a_0 N + a_1 \sum X \\ \sum XY = a_0 \sum X + a_1 \sum X^2 \end{cases}$$

$$a_0 = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{N \sum X^2 - (\sum X)^2} \quad a_1 = \frac{N \sum XY - (\sum X)(\sum Y)}{N \sum X^2 - (\sum X)^2}$$

CÁLCULOS

X	Y	X ²	XY	Y ²
100	40	10000	4000	1600
200	50	40000	10000	2500
300	50	90000	15000	2500
400	70	160000	28000	4900
500	65	250000	32500	4225
600	65	360000	39000	4225
700	80	490000	56000	6400
$\sum X = 2800$	$\sum Y = 420$	$\sum X^2 = 1400000$	$\sum XY = 184500$	$\sum Y^2 = 26350$

ECUACIONES

$$\sum X = 2800 \quad \sum Y = 420 \quad \sum X^2 = 1400000 \quad \sum XY = 184500 \quad \sum Y^2 = 26350$$

$$\begin{cases} \sum Y = a_0 N + a_1 \sum X \\ \sum XY = a_0 \sum X + a_1 \sum X^2 \end{cases}$$

$$\begin{cases} 7a_0 + 2800a_1 = 420 \\ 2800a_0 + a_1 1400000 = 184500 \end{cases}$$

$$\begin{cases} a_0 = 36.4 \\ a_1 = 0.059 \end{cases}$$

RECTA ENCONTRADA

ECUACIONES

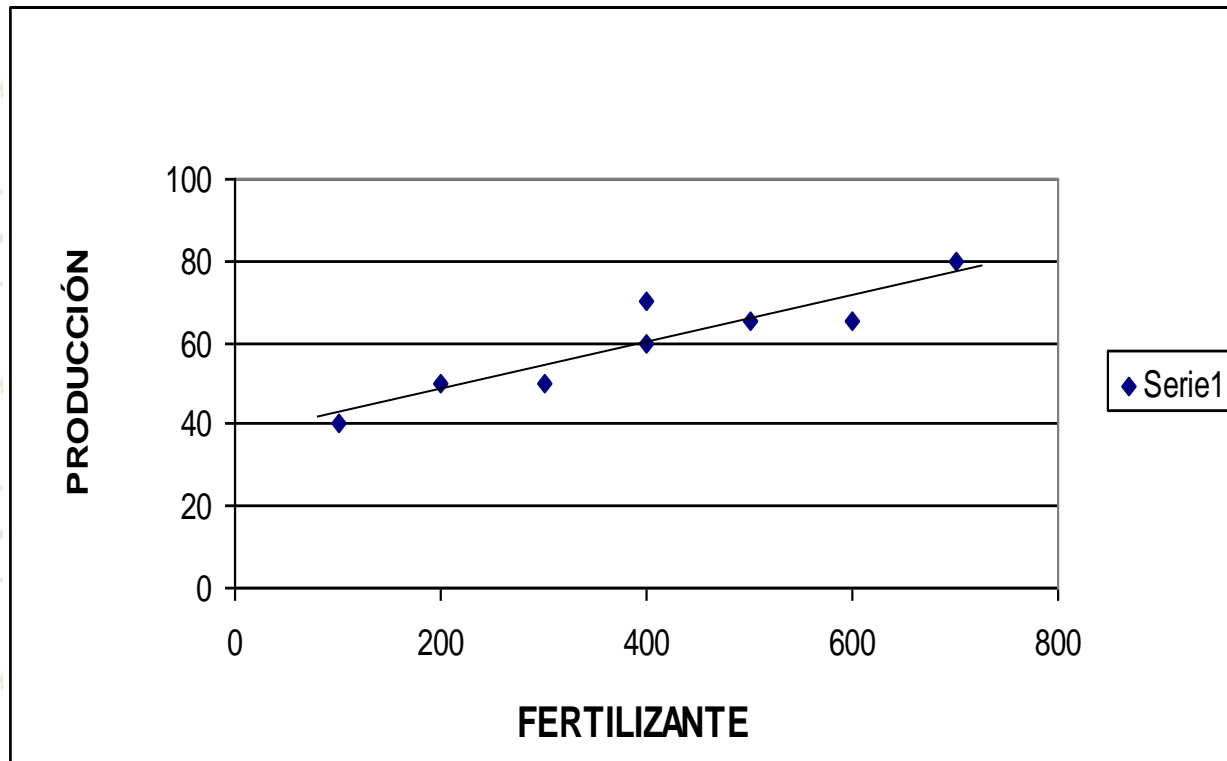
$$\begin{cases} a_0 = 36.4 \\ a_1 = 0.059 \end{cases}$$

$$Y = a_0 + a_1 X$$

$$Y = 36.4 + 0.059X$$

RECTA ENCONTRADA

GRÁFICO



40	42,32
50	48,21
50	54,11
70	60,00
65	65,89
65	71,79
80	77,68

PRONÓSTICO

✱ Para el valor de fertilizante 800 podemos predecir que se obtendrá una producción de :

$$✱ Y = 36.4 + 0.059(800) = 83.6$$

RELACIONES NO LINEALES

PARÁBOLA DE LOS MÍNIMOS CUADRADOS

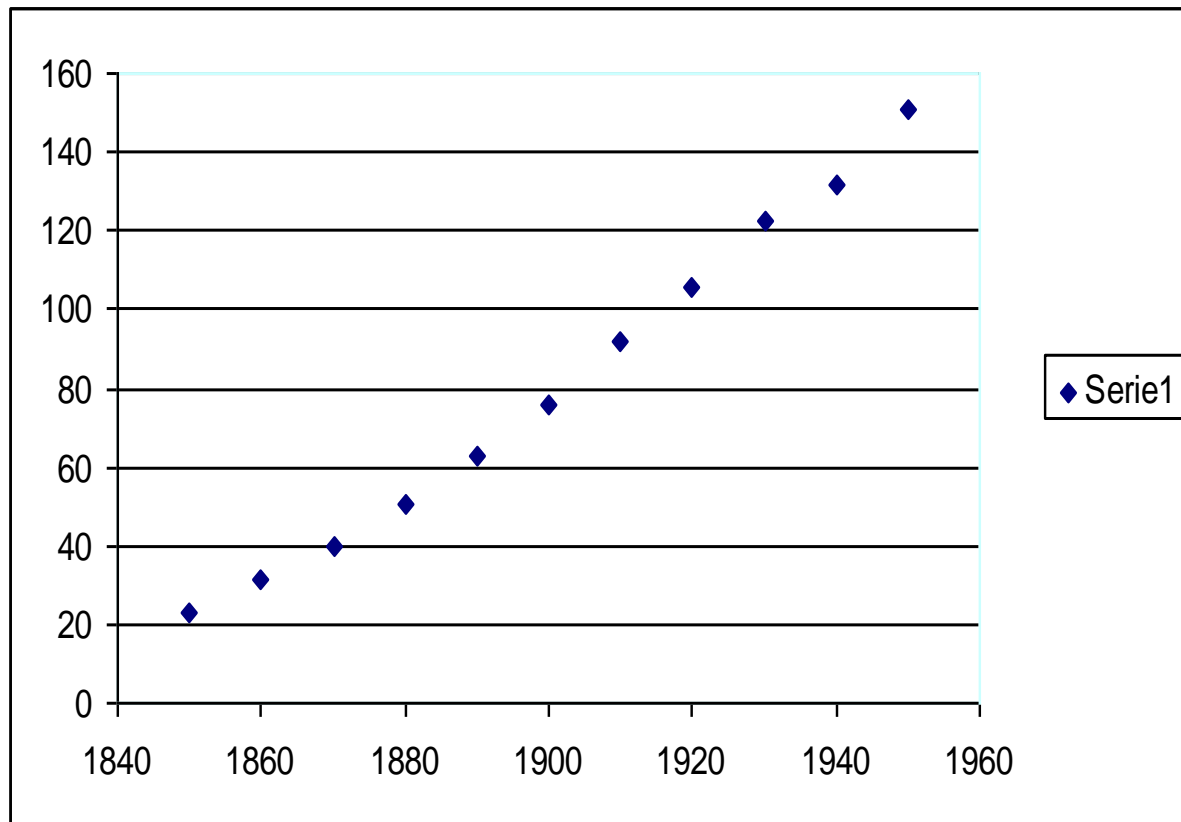
$$Y = a_0 + a_1 X + a_2 X^2$$

$$\begin{cases} \sum Y = a_0 N + a_1 \sum X + a_2 \sum X^2 \\ \sum XY = a_0 \sum X + a_1 \sum X^2 + a_2 \sum X^3 \\ \sum X^2 Y = a_0 \sum X^2 + a_1 \sum X^3 + a_2 \sum X^4 \end{cases}$$

EJEMPLO

AÑO	POBLACIÓN(MILLONES)
1850	23.2
1860	31.4
1870	39.8
1880	50.2
1890	62.9
1900	76.0
1910	92.0
1920	105.7
1930	122.8
1940	131.7
1950	151.1

GRÁFICO DE LOS PUNTOS



CÁLCULOS

AÑO	X	Y	X ²	X ³	X ⁴	XY	X ² Y
1850	-5	23,2	25	-125	625	-116	580
1860	-4	31,4	16	-64	256	-125,6	502,4
1870	-3	39,8	9	-27	81	-119,4	358,2
1880	-2	50,2	4	-8	16	-100,4	200,8
1890	-1	62,9	1	-1	1	-62,9	62,9
1900	0	76,0	0	0	0	0,0	0,0
1910	1	92,0	1	1	1	92,0	92,0
1920	2	105,7	4	8	16	211,4	422,8
1930	3	122,8	9	27	81	368,4	1105,2
1940	4	131,7	16	64	256	526,8	2107,2
1950	5	151,1	25	125	625	755,5	3777,5
	$\sum X = 0$	$\sum Y = 886,8$	$\sum X^2 = 110$	$\sum X^3 = 0$	$\sum X^4 = 1958$	$\sum XY = 1429,8$	$\sum X^2Y = 9209$

ECUACIONES

$$\sum X = 0 \quad \sum Y = 886,8 \quad \sum X^2 = 110 \quad \sum X^3 = 0 \quad \sum X^4 = 1958 \quad \sum XY = 1429,8 \quad \sum X^2Y = 9209$$

$$\begin{cases} \sum Y = a_0 N + a_1 \sum X + a_2 \sum X^2 \\ \sum XY = a_0 \sum X + a_1 \sum X^2 + a_2 \sum X^3 \\ \sum X^2Y = a_0 \sum X^2 + a_1 \sum X^3 + a_2 \sum X^4 \end{cases}$$

$$\begin{cases} 11a_0 + 110a_2 = 886.8 \\ 110a_1 = 1429.8 \\ 110a_0 + 1958a_2 = 9209.0 \end{cases}$$

ECUACIONES

$$\begin{cases} a_0 = 76.64 \\ a_1 = 13.00 \\ a_2 = 0.3974 \end{cases}$$

$$Y = a_0 + a_1X + a_2X^2$$

$$Y = 76.64 + 13.00X + 0.3974X^2$$

PRONÓSTICO

Pronóstico: De acuerdo a la curva encontrada la población para el año 1960 sería

$$Y = 76.64 + 13.00(6) + 0.3974(6)^2$$
$$= 168.95 \text{ millones}$$