

Correlación y regresión lineal simple

Estudiantes:

Juan Sebastián López Hiquita
Jorge Andrés Zapata López

Asignatura:
Álgebra lineal

Profesor:
Jorge Luis Mejía Galeano

Grupo:
054 - Martes - Jueves: 10-12

Problema - situación planteada

El departamento de investigaciones económicas de una empresa desea realizar un estudio sobre los precios y la demanda de su principal producto. Para ello cuenta con la siguiente información:

- Variable x : Precio (miles de pesos)
- Variable y : Demanda (número de unidades)

x	5	7	9	12	17	23	30	10	?	3	?	?	120
y	100	90	86	72	60	55	43	?	50	?	25	80	?

a) Objetivos del estudio con base a dichas variables.

El objetivo del estudio es establecer un modelo de regresión lineal entre las dos variables analizadas y determinar el grado de correlación entre las mismas.

b) Elaborar el diagrama de dispersión (nube de puntos). ¿Que tendencia se visualiza en el gráfico?.

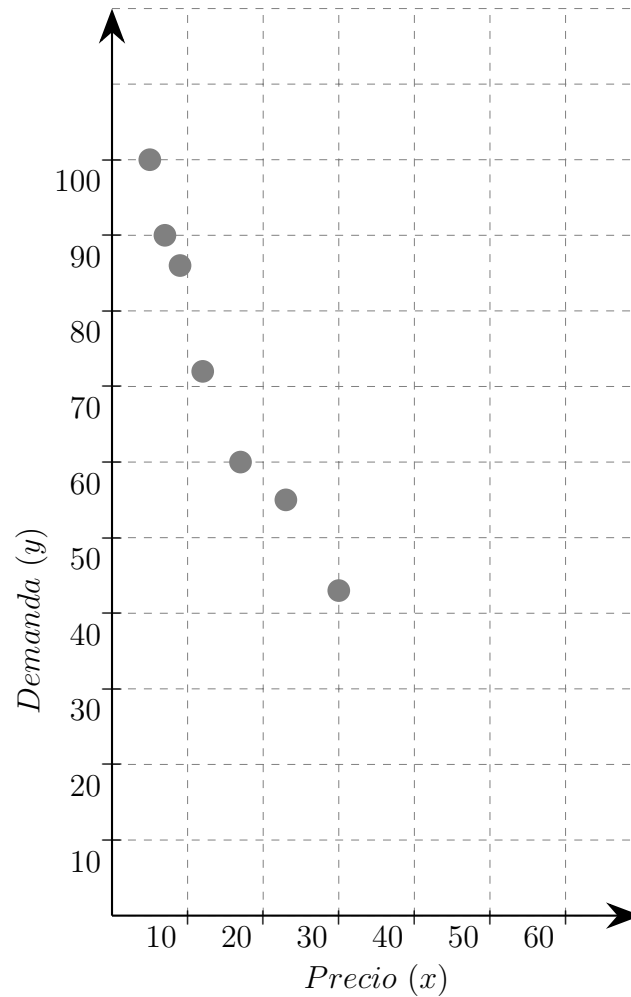


Figure 1: Diagrama de dispersión

Se aprecia que hay una correlación lineal inversa bastante marcada entre los datos suministrados para ambas variables.

c) Calcular el coeficiente de correlación Υ analizarlo e interpretarlo.

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
5	100	-9.71	27.72	94.28	768.39	-269.16
7	90	-7.71	17.72	59.44	313.99	-136.62
9	86	-5.71	13.72	32.60	188.23	-78.34
12	72	-2.71	-0.28	7.3	0.07	0.75
17	60	2.29	-12.28	5.24	150.79	-28.12
23	55	8.29	-17.28	68.72	298.59	-143.25
30	43	15.29	-29.28	233.78	857.31	-447.69
$\Sigma 103$	$\Sigma 506$	$\Sigma 0.03$	$\Sigma 0.04$	$\Sigma 501.36$	$\Sigma 2577.37$	$\Sigma -1102.43$

$$\Upsilon = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{(\sum (x - \bar{x})^2)(\sum (y - \bar{y})^2)}}$$

$$\bar{x} = \frac{\sum x}{n} = \frac{103}{7} \approx 14.71$$

$$\bar{y} = \frac{\sum y}{n} = \frac{506}{7} \approx 72.28$$

$$\sum (x - \bar{x})^2 = 501.36$$

$$\sum (y - \bar{y})^2 = 2577.37$$

$$\sum (x - \bar{x})(y - \bar{y}) = -1102.43$$

$$\Upsilon = \frac{-1102.43}{\sqrt{501.36 * 2577.37}} = -0.969$$

Interpretación: El valor negativo indica que la correlación es inversa (a mayor precio, menores unidades demandadas); y el valor cercano a -1 indica que el modelo de regresión lineal se ajusta bien con los datos suministrados.

d) Calcular la función de ajuste (Ecuación de Regresión) y graficarla sobre el diagrama. Método de los mínimos cuadrados.

x	y	xy	x^2
5	100	500	25
7	90	630	49
9	86	774	81
12	72	864	144
17	60	1020	289
23	55	1265	529
30	43	1290	900
$\Sigma 103$	$\Sigma 506$	$\Sigma 6343$	$\Sigma 2017$

$$y = mx + b, \quad m = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}, \quad n = 7$$

$$m = \frac{6343 - \frac{103 \cdot 506}{7}}{2017 - \frac{(103)^2}{7}} = \frac{6343 - 7445.428}{2017 - 1515.571} = \frac{-1102.428}{501.429} = -2.198$$

$$b = \bar{y} - m\bar{x} = 72.28 - (-2.198 * 14.71) = 104.612$$

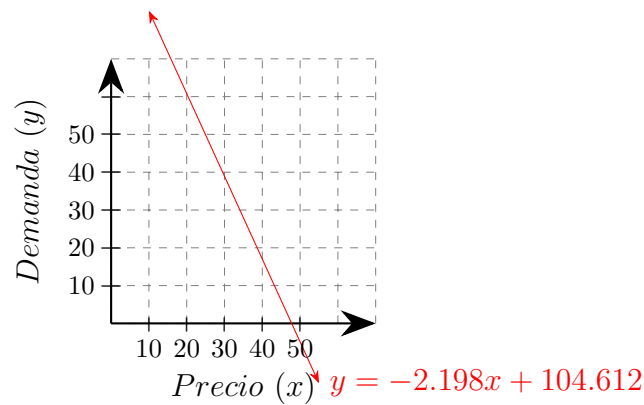


Figure 2: Ecuación de regresión

e) Pronósticar el precio de venta para 50, 25 y 80 unidades demandadas.

$$y = -2.198x + 104.612$$

$$y - 104.612 = -2.198x$$

$$x = \frac{y-104.612}{-2.198}$$

Para $y = 50$:

$$x = \frac{50-104.612}{-2.198} = \frac{-54.612}{-2.198} = 24.846$$

Para $y = 25$:

$$x = \frac{25-104.612}{-2.198} = \frac{-79.612}{-2.198} = 36.220$$

Para $y = 80$:

$$x = \frac{80-104.612}{-2.198} = \frac{-24.612}{-2.198} = 11.197$$

f) Pronósticar las cantidades demandadas para un precio de \$3000 y \$120000.

Para $x = 10$:

$$y = -2.198 * 10 + 104.612 = 82.632$$

Para $x = 3$:

$$y = -2.198 * 3 + 104.612 = 98.018$$

Para $x = 120$:

$$y = -2.198 * 120 + 104.612 = -159.148$$

g) Calcular el coeficiente de determinación e interpretar los resultados.

$$\Upsilon^2 = (-0.969)^2 = 0.938$$

Interpretación de resultados: El valor del coeficiente de determinación es cercano a 1. Indicando que el modelo si puede explicar los resultados observados.

h) Sacar conclusiones y sugerencias a partir de los resultados obtenidos a travez de este modelo.

El modelo de regresión se ajusta bien con los datos históricos y se considera un modelo válido para hacer pronósticos.

Gráfico completo

Diagrama de dispersión y gráfica de la ecuación de regresión.

x	5	7	9	12	17	23	30	10	24.846	3	36.220	11.97	120
y	100	90	86	72	60	55	43	82.632	50	98.018	25	80	-159.148

