Regresión lineal simple

Isaac Cortés Olmos

Universidad de Atacama

May 16, 2025

Esquema

Análisis de regresión simple

Análisis de regresión simple

Análisis de regresión simple

- Estudia la relación entre 2 variables.
- Ejemplo: Ingresos y Ahorros.
- En este caso, la variable Ingresos se denomina variable independiente representada mediante x y la cantidad Ahorros se llama dependiente representada por y.

Ejemplo:

El gerente de un banco desea saber si puede considerarse que el ahorro de las familias (variable y) depende de sus ingresos (variable x). En la Tabla se muestran los resultados obtenidos para una muestra de 10 familias.

Familia	Ingresos (x)	Ahorro (y)
1	11	0.5
2	14	1.1
3	12	0.9
4	9	0.6
5	13	1.2
6	13	0.9
7	15	1.5
8	17	1.3
9	15	1.1
10	13	0.7

Table 1: Ingreso y ahorro mensual de 10 familias (en miles de pesos).

Análisis de regresión simple

Ejemplo:

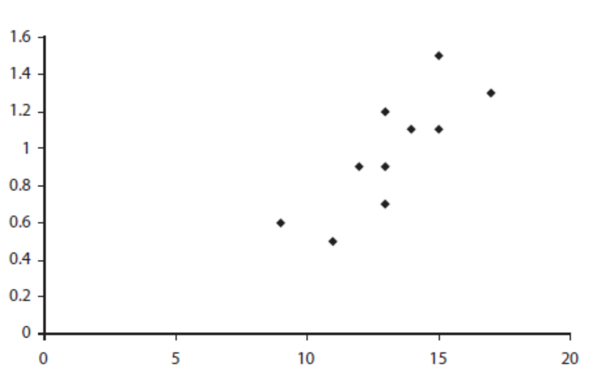

Figure 1: Diagrama de dispersión para los datos del ejemplo.

Tipo de relaciones:

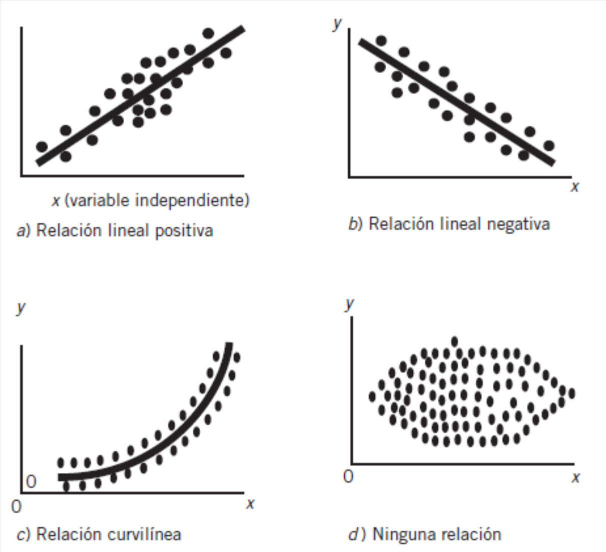

Figure 2: Ejemplos de diagramas de dispersión para datos con 2 variables.

Análisis de regresión simple

Método de mínimos cuadrados

- Reduce al mínimo el cuadrado de las distancias verticales entre cada uno de los puntos y la recta ajustada.

Análisis de regresión simple

Ejemplo:

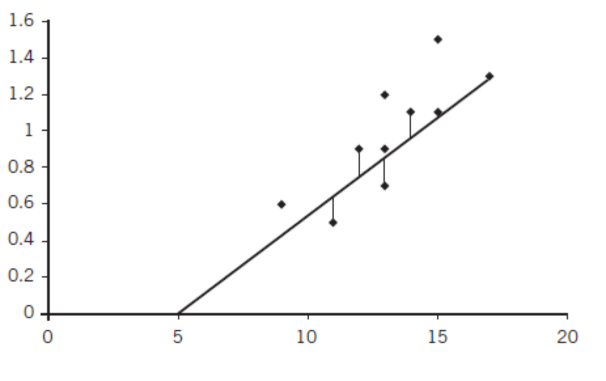

Figure 3: Distancias entre la ordenada de cada punto y la ordenada de la recta de regresión: $y_i - y_c$.

Análisis de regresión simple

Ecuaciones normales

- Conjunto de 2 ecuaciones que resueltas simultáneamente producen el valor de la pendiente y el de la ordenada al origen que son los parámetros de la ecuación lineal que arroja los mínimos cuadrados.

Análisis de regresión simple

Familia	Ingresos (x)	Ahorro (y)	xy	x^2
1	11	0.5	5.5	121
2	14	1.1	15.4	196
3	12	0.9	10.8	144
4	9	0.6	5.4	81
5	13	1.2	15.6	169
6	13	0.9	11.7	169
7	15	1.5	22.5	225
8	17	1.3	22.1	289
9	15	1.1	16.5	225
10	13	0.7	9.1	169
Totales	132	9.8	134.6	1788

Table 2: Ingreso y ahorro mensual de 10 familias (en miles de pesos).

Análisis de regresión simple

Método de mínimos cuadrados

$$
\begin{aligned}
b &= \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}} \\
&= \frac{134.6 - \frac{132(9.8)}{10}}{1788 - \frac{132^2}{10}} \\
&= 0.1149 \\
a &= \overline{y} - b\overline{x} \\
&= \frac{9.8}{10} - 0.1149\left(\frac{132}{10}\right) \\
&= -0.5367 \\
\widehat{y} &= -0.5367 + 0.1149x.
\end{aligned}
$$

De una muestra de 15 embarques de láminas de aluminio se registró la distancia en km al lugar de entrega así como el tiempo en horas para cada carga.

Embarque	Distancia en km	Tiempo de entrega en horas
1	400	7
2	800	12
3	120	2
4	340	5
5	520	8
6	300	4
7	100	1.4
8	85	1
9	589	9.6
10	1115	13
11	265	3
12	670	11
13	1215	15
14	550	9
15	215	2.6

Referencias I

- Devore, J. L. (2009). *Probabilidad y estadística para ingeniería y ciencias*. Cengage Learning Editores.
- Díaz Mata, A. (2013). *Estadística aplicada a la Administración y Economía*. McGraw–Hill.
- Ross, S. M. (2002). *Probabilidad y estadística para ingeniería y ciencias*. McGraw–Hill.
- Spiegel, M. R., & Stephens, L. J. (2009). *Estadística*. McGraw–Hill.