

# Análisis de Regresión Lineal Simple

Aplicaciones Prácticas en Diferentes Contextos

Angello Zamora Valencia

Escuela Profesional de Estadística e Informática

28 de mayo de 2025

## 1. Introducción

La regresión lineal simple es una de las técnicas estadísticas más fundamentales para el análisis de relaciones entre variables. Esta metodología permite modelar la relación entre una variable dependiente ( $Y$ ) y una variable independiente ( $X$ ) mediante una función lineal de la forma:

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (1)$$

donde  $\beta_0$  es el intercepto,  $\beta_1$  es la pendiente, y  $\epsilon$  representa el error aleatorio.

Los coeficientes se estiman utilizando el método de mínimos cuadrados ordinarios:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (3)$$

El presente documento analiza tres casos prácticos de regresión lineal en diferentes contextos: educativo, comercial y laboral.

## 2. Análisis de Casos

### 2.1. Caso 1: Horas de Estudio vs Calificación

Este primer análisis examina la relación entre las horas dedicadas al estudio y las calificaciones obtenidas por los estudiantes.

**Datos utilizados:**

Horas de Estudio	Calificación
1	50
2	55
3	65
4	70
5	75
6	85
7	88
8	92

**Resultados del análisis:**

- Ecuación de regresión:  $\hat{y} = 5,36x + 43,75$
- Error cuadrático medio (MSE): 12.65
- Coeficiente de determinación:  $R^2 = 0,973$

**Interpretación:**

- Por cada hora adicional de estudio, la calificación aumenta en promedio 5.36 puntos.
- Sin estudiar (0 horas), la calificación base esperada sería de 43.75 puntos.
- El modelo explica el 97.3 % de la variación en las calificaciones, indicando un ajuste excelente.

**2.2. Caso 2: Temperatura vs Ventas de Helado**

Este análisis estudia cómo la temperatura ambiente afecta las ventas de helado en una heladería local.

**Datos utilizados:**

Temperatura (°C)	Ventas (unidades)
20	10
22	15
25	25
28	35
30	45
32	55
35	70
37	80

**Resultados del análisis:**

- Ecuación de regresión:  $\hat{y} = 4,13x - 72,14$
- Error cuadrático medio (MSE): 19.64
- Coeficiente de determinación:  $R^2 = 0,988$

**Interpretación:**

- Cada grado adicional de temperatura aumenta las ventas en 4.13 unidades de helado.
- A 0°C, las ventas estimadas serían -72.14 unidades (extrapolación no realista).
- El modelo explica el 98.8 % de la variación en ventas, mostrando una relación muy fuerte.

### 2.3. Caso 3: Años de Experiencia vs Salario

Este tercer caso analiza la relación entre los años de experiencia laboral y el salario anual de los empleados.

**Datos utilizados:**

Experiencia (años)	Salario (USD)
0	25,000
1	28,000
2	32,000
3	36,000
5	42,000
7	48,000
10	58,000
12	65,000

**Resultados del análisis:**

- Ecuación de regresión:  $\hat{y} = 3333,33x + 25000,00$
- Error cuadrático medio (MSE): 4,722,222.22
- Coeficiente de determinación:  $R^2 = 0,989$

**Interpretación:**

- Cada año adicional de experiencia incrementa el salario en \$3,333.33.
- El salario base (sin experiencia) sería de \$25,000.
- El modelo explica el 98.9% de la variación salarial, indicando una relación muy sólida.

## 3. Predicciones

Utilizando los modelos desarrollados, se realizaron las siguientes predicciones:

- **Estudio:** Con 5.5 horas de estudio, se predice una calificación de 73.2 puntos.
- **Helados:** A 33°C, se estiman ventas de 64.3 unidades.
- **Salario:** Con 6 años de experiencia, el salario estimado sería \$45,000.

## 4. Resumen Comparativo

Modelo	$R^2$	Calidad del Ajuste	Aplicabilidad
Estudio $\rightarrow$ Nota	0.973	Excelente	Alta
Temperatura $\rightarrow$ Ventas	0.988	Excelente	Alta
Experiencia $\rightarrow$ Salario	0.989	Excelente	Alta

## 5. Conclusiones

El análisis de regresión lineal simple aplicado a los tres casos de estudio demuestra la versatilidad y efectividad de esta técnica estadística:

1. **Alta capacidad explicativa:** Los tres modelos presentan coeficientes de determinación superiores a 0.97, indicando que explican más del 97 % de la variabilidad en las variables dependientes.
2. **Interpretación práctica:** Cada modelo proporciona información valiosa para la toma de decisiones en sus respectivos contextos.
3. **Validez predictiva:** Los modelos permiten realizar predicciones confiables dentro del rango de datos observados.
4. **Aplicabilidad universal:** La regresión lineal simple es aplicable en diversos campos, desde la educación hasta los negocios y recursos humanos.

Sin embargo, es importante recordar las limitaciones del análisis:

- Las predicciones son más confiables dentro del rango de datos observados.
- La relación lineal asumida puede no ser válida en todos los casos.
- Otros factores no incluidos en el modelo pueden influir en las variables dependientes.

En conclusión, la regresión lineal simple constituye una herramienta fundamental en el análisis estadístico, proporcionando insights valiosos sobre las relaciones entre variables en diferentes contextos profesionales y académicos.