

Predicción peso de personas

Rubén Pizarro Gurrola

2025-01-16

Objetivo

Predicir y evaluar un modelo de regresión lineal simple.

Descripción

Se va a predecir usando un modelo de regresión lineal simple el peso de una persona en función de la estatura.

Librerías

```
library(readr)
library(ggplot2)
library(dplyr)
```

Cargar los datos

```
datos <- read.csv("https://raw.githubusercontent.com/rpizarro/Ciencia-de-los-Datos-Descriptivo-Predictivo/main/data/weight.csv")
```

```
##   estatura peso
## 1      170   75
## 2      180   80
## 3      190   88
## 4      170   65
## 5      178   80
## 6      192  110
## 7      150   80
## 8      184   84
## 9      170   72
## 10     180   90
## 11     175   75
## 12     180   82
## 13     190   94
## 14     170   74
```

Construir el modelo de regresión lineal simple

Se utiliza la función `lm($y \sim x$, data = datos)` ajusta un modelo de regresión lineal simple, donde y se predice a partir de x .

- x es el peso
- y la estatura

La expresión $\text{peso} \sim \text{estatura}$ significa que vamos a predecir el peso en función de la estatura.

La expresión `summary(modelo)` extrae estadísticos y los coeficientes a y b que necesitamos para $Y = a + bx$.

```
modelo <- lm(data = datos,
             formula = peso ~ estatura)
summary(modelo)

##
## Call:
## lm(formula = peso ~ estatura, data = datos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.357  -3.857  -2.690   1.976  17.976
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -35.9735    38.7335  -0.929   0.371
## estatura      0.6667     0.2184   3.053   0.010 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.649 on 12 degrees of freedom
## Multiple R-squared:  0.4372, Adjusted R-squared:  0.3903
## F-statistic: 9.321 on 1 and 12 DF,  p-value: 0.01003
```

Valores de los Coeficientes

```
a <- modelo$coefficients[1]
b <- modelo$coefficients[2]
paste("Valor de a = ", round(a, 4))

## [1] "Valor de a = -35.9735"

paste("Valor de b=", round(b, 4))

## [1] "Valor de b= 0.6667"
```

Significa que una predicción de un peso de persona deberá tener $Y = a + b \cdot x$ o sea $Y = -35.9735 + 0.6667 \cdot x$ y si la persona mide:

- 178
- 182
- 160

¿cuánto debe pesar?

Las expresiones que estan en signos de pesos se llama código latex que es lenguaje matemático.

- $\text{peso} = -35.9735 + 0.6667 \cdot 178;$
- $\text{peso} = -35.9735 + 0.6667 \cdot 182;$
- $\text{peso} = -35.9735 + 0.6667 \cdot 160$

```
pred1 <- a + b * 178
pred2 <- a + b * 182
pred3 <- a + b * 160
```

```
pred1
```

```
## (Intercept)
##      82.69046
```

```
pred2
```

```
## (Intercept)
##      85.35707
```

```
pred3
```

```
## (Intercept)
##      70.69074
```

Por medio de la función *predict()*

```
a_predecir <- data.frame(estatura = c(178, 182, 160, 200, 0))
# Predicciones
predicciones <- predict(modelo, newdata = a_predecir)
predicciones
```

```
##           1           2           3           4           5
##  82.69046  85.35707  70.69074  97.35679 -35.97350
```

Visualmente con una linea de tendencia

```
ggplot(datos, aes(x = estatura, y = peso)) +
  geom_point(color = "blue", alpha = 0.7) + # Puntos de datos
  geom_smooth(method = "lm", se = TRUE, color = "red") + # Línea de regresión
  labs(
    title = "Regresión Lineal Simple",
    x = "Variable Independiente (x)",
    y = "Variable Dependiente (y)"
  ) +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

Regresión Lineal Simple

