

# Día 1: Introducción a la Regresión Logística

Carlos

July 6, 2024

## 1 Conceptos Básicos de la Regresión Logística

La regresión logística es una técnica de modelado estadístico utilizada para predecir la probabilidad de un evento binario (es decir, un evento que tiene dos posibles resultados) en función de una o más variables independientes. A diferencia de la regresión lineal, que se utiliza para predecir valores continuos, la regresión logística se usa cuando la variable dependiente es categórica.

## 2 Diferencias entre Regresión Lineal y Logística

### 2.1 Regresión Lineal

La regresión lineal busca modelar la relación entre una variable dependiente continua  $Y$  y una o más variables independientes  $X_1, X_2, \dots, X_n$  mediante una ecuación de la forma:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon$$

donde  $\beta_0, \beta_1, \dots, \beta_n$  son los coeficientes del modelo y  $\epsilon$  es el término de error.

### 2.2 Regresión Logística

La regresión logística, en cambio, modela la probabilidad de que un evento ocurra (por ejemplo, éxito vs. fracaso) utilizando la función logística. La variable dependiente  $Y$  es binaria, tomando valores de 0 o 1. La ecuación de la regresión logística es:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

donde  $p$  es la probabilidad de que  $Y = 1$ . La función logística es:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}}$$

### 3 Casos de Uso de la Regresión Logística

La regresión logística se utiliza en una variedad de campos para problemas de clasificación binaria, tales como:

- **Medicina:** Predicción de la presencia o ausencia de una enfermedad.
- **Marketing:** Determinación de la probabilidad de que un cliente compre un producto.
- **Finanzas:** Evaluación del riesgo de crédito, es decir, si un cliente va a incumplir o no con un préstamo.
- **Seguridad:** Detección de fraudes o intrusiones.

### 4 Implementación Básica en R

Para implementar una regresión logística en R, primero es necesario instalar y cargar los paquetes necesarios. Aquí se muestra un ejemplo básico de implementación:

#### 4.1 Instalación y Configuración de R y RStudio

- Descargue e instale R desde <https://cran.r-project.org/>.
- Descargue e instale RStudio desde <https://rstudio.com/products/rstudio/download/>.

#### 4.2 Introducción Básica a R

- Sintaxis básica de R.
- Operaciones básicas: asignación, operaciones aritméticas, funciones básicas.

#### 4.3 Ejemplo de Regresión Logística en R

```
# Instalación del paquete necesario
install.packages("stats")

# Carga del paquete
library(stats)

# Ejemplo de conjunto de datos
data <- data.frame(
  outcome = c(1, 0, 1, 0, 1, 1, 0, 1, 0, 0),
  predictor = c(2.3, 1.9, 3.1, 2.8, 3.6, 2.4, 2.1, 3.3, 2.2, 1.7)
)

# Ajuste del modelo de regresión logística
```

```
model <- glm(outcome ~ predictor, data = data, family = binomial)

# Resumen del modelo
summary(model)
```

En este ejemplo, se utiliza el conjunto de datos ‘data’ que contiene una variable de resultado binaria ‘outcome’ y una variable predictora continua ‘predictor’. El modelo de regresión logística se ajusta utilizando la función `glm` con la familia binomial.