

---

## Raíz de Error cuadrático medio al cuadrado (RMSE)

Error de raíz cuadrada media ( $RMSE$ ) es la desviación estándar de los valores residuales (errores de predicción). Los valores residuales son una medida de la distancia de los puntos de datos de la línea de regresión;  $RMSE$  es una medida de cual es el nivel de dispersión de estos valores residuales.

$$RMSE = \sqrt{E[(\hat{X}_i - X_i)^2]}$$
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2}$$

```
library(FNN)
library(scales)
library(lattice)
library(caret)

## Loading required package: ggplot2

library(fastDummies) #Para variables ficticias
setwd("C:\\Users\\81799\\OneDrive\\Documentos\\ESFM_CLASES\\Servicio Social ARTF\\Machine Learning")
Data <- read.csv("rmse.csv")
head(Data , 5)

##   price  pred
## 1  24.0 21.05
## 2  21.6 22.76
## 3  34.7 38.64
## 4  33.4 35.47
## 5  36.2 34.75
```

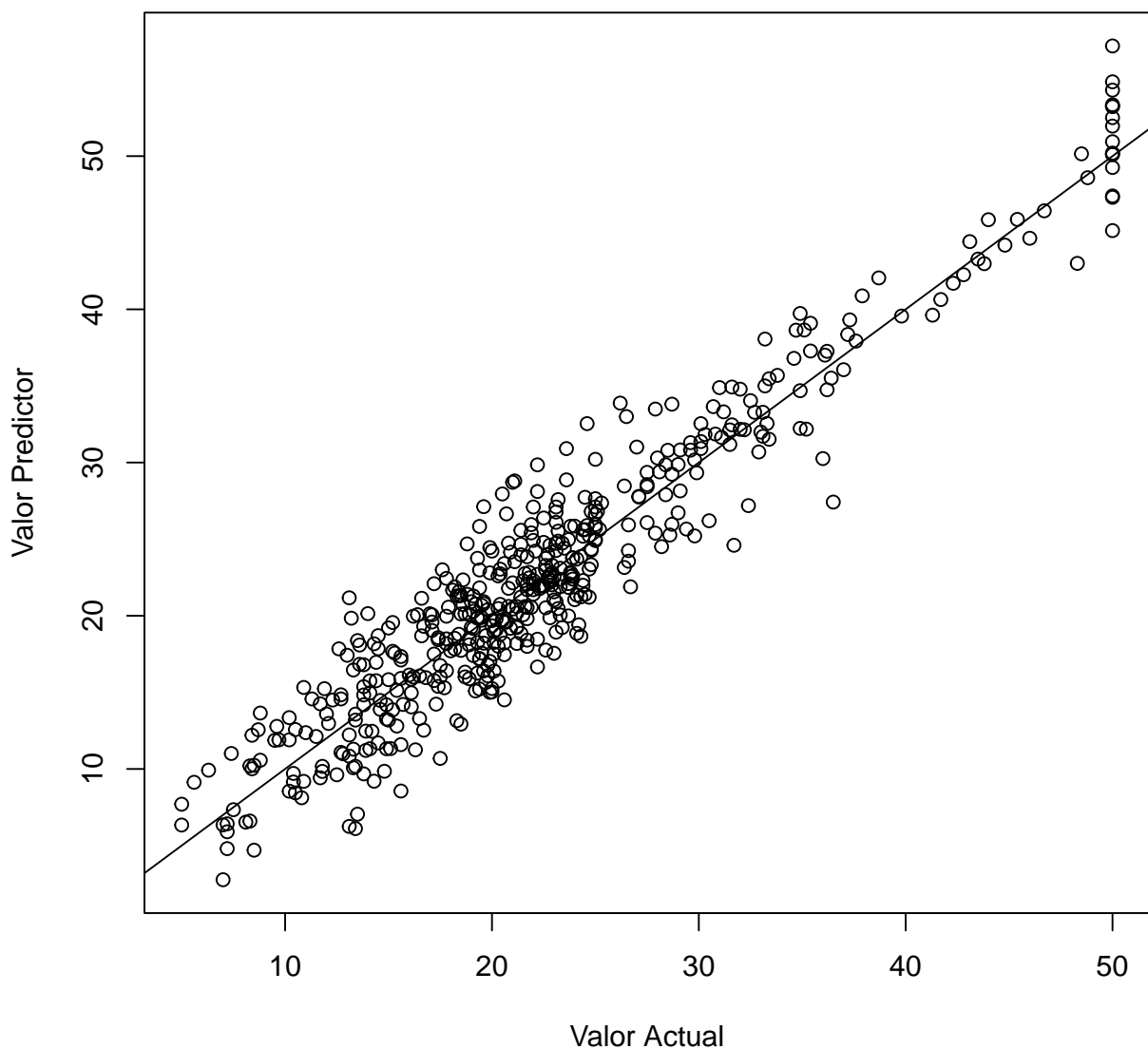
Aplicando la formula de  $RMSE$  tendremos:

```
RMSE <- sqrt(mean((Data$pred - Data$price )^2) )
RMSE

## [1] 2.934995
```

Esto quiere decir que en promedio nos equivocamos  $\pm 2.935$  en la predicción con respecto a los datos originales.

```
plot(Data$price , Data$pred , xlab = "Valor Actual", ylab = "Valor Predictor")
abline(0,1)
```



El objetivo es que los valores originales y los valores predictores al gráficarlos tengamos una linea recta con pendiente igual a 1. Haciendo una función para calcular lo anterior, tendremos lo siguiente:

```
RMSE_funcion <- function(prediccion, original){  
  return( sqrt(mean((prediccion- original )^2) ))  
}  
RMSE_funcion(Data$pred, Data$price) #Para no estar escribiendo la formula  
## [1] 2.934995
```