

10

Ejercicio 2

Estadística y Diseño de Experimentos

2223011525
Diaz Aguilar Hans Uriel

Profesor: Joel
Montesinos Vázquez

Universidad Autónoma Metropolitana
Iztapalapa

13 de octubre de 2025

Problema y datos

En la fabricación de un motor se deben unir dos tipos de propulsores (tipo 1 y tipo 2). Se sospecha que la resistencia al corte de esta unión está relacionada con la edad (en semanas) del lote de propulsores del tipo 1. En la siguiente tabla se muestra la Resistencia al corte (medida en psi) y la Edad (en semanas) del lote del propulsor tipo 1.

```
Resistencia <- c(2158.70, 1678.15, 2316.00, 2061.30, 2207.50,  
1708.30, 1784.70, 2575.00, 2357.90, 2256.70, 2165.20, 2399.55, 1779.80,  
2336.75, 1765.30, 2053.50, 2414.40, 2200.50, 2654.20, 1753.70)  
  
Edad <- c(15.50, 23.75, 8.00, 17.00, 5.50, 19.00, 24.00, 2.50, 7.50,  
11.00, 13.00, 3.75, 25.00, 9.75, 22.00, 18.00, 6.00, 12.50, 2.00, 21.50)
```

1. Identifique quién es la variable de respuesta Y y quién es la variable regresora X y escriba el modelo de regresión

Variable de respuesta (Y):
Resistencia al corte.

Variable regresora (X):
Edad del propulsor tipo 1.

Modelo de regresión lineal

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Donde:

B_0 : Intercepto

B_1 : Pendiente

ε : Error aleatorio

2. Grafique el diagrama de dispersión de los datos

```
plot(Edad, Resistencia, pch = 16, col='red')
```

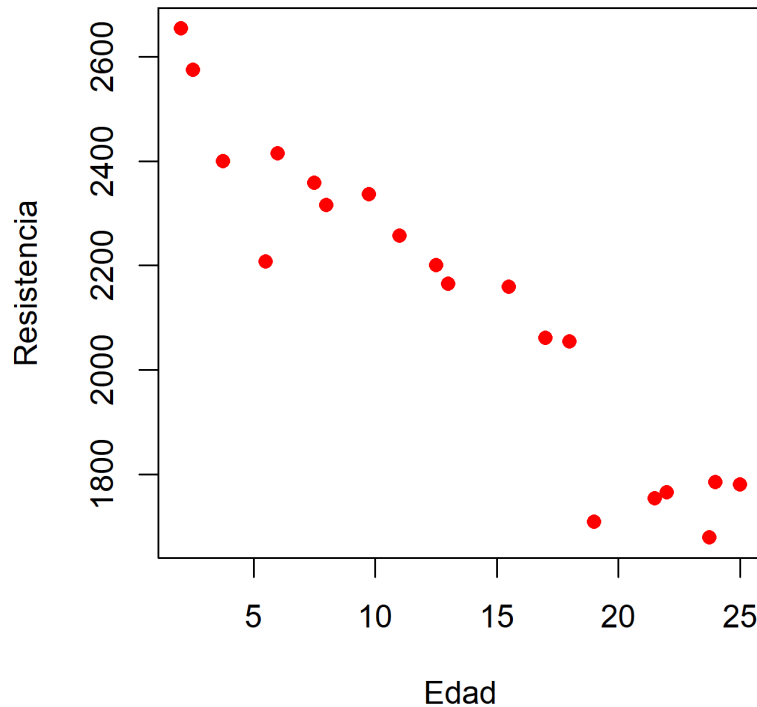


Figura 1: Diagrama de dispersión de Edad y Resistencia.

3. Obtenga los estimadores para B_0 y B_1 y escriba la ecuación de la recta ajustada

```
RL <- lm(Resistencia~Edad)
```

```
RL$coefficients
```

```
# (Intercept)      Edad
# 2627.82236   -37.15359
```

$$\hat{B}_0 = 2627.82236$$

$$\hat{B}_1 = -37.15359$$

Ecuación ajustada

$$\hat{Y} = 2627.82236 - 37.15359X$$

4. Grafique la recta de la regresión junto con los datos. ¿Qué tan bueno cree que es el ajuste?

```
plot(Edad, Resistencia, pch = 16, col='red')
abline(RL, col='lightgreen', lwd=3)
```

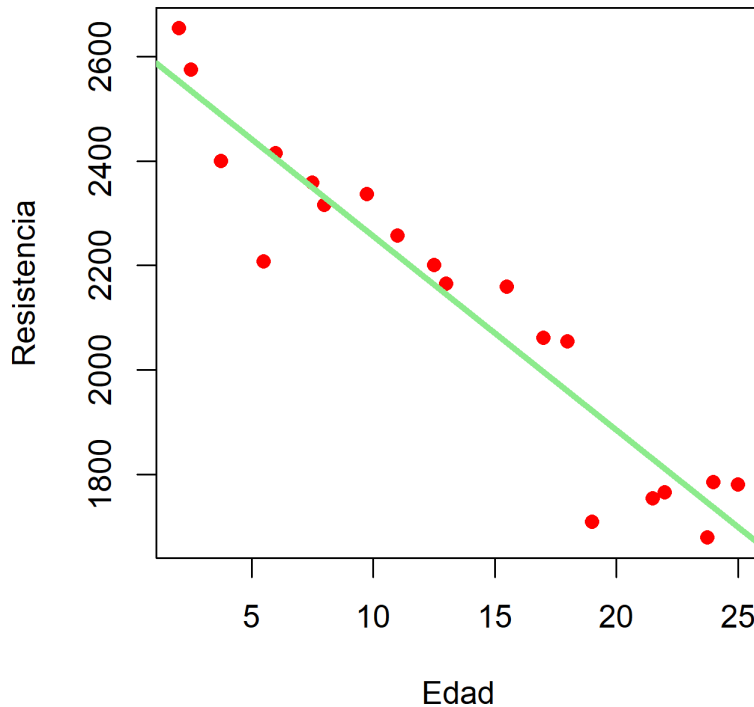


Figura 2: Recta de regresión sobre el diagrama de dispersión.

Visualmente, los puntos están bastante cerca de la recta. Personalmente creo que es un buen ajuste lineal, salvo algunos puntos que se notan más alejados.

5. Efectúe la prueba de significancia de la regresión para un nivel $\alpha = 0.05$. Escriba el valor del p – valor. ¿Qué conclusiones puede hacer sobre B_1 ?

```
alpha <- 0.05
p.valor <- summary(RL)$coefficients["Edad", "Pr(>|t|)"] # 1.643344e-10
p.valor < alpha # TRUE
```

Como $p\text{-valor} < 0.05$, se rechaza $H_0: B_1 = 0$. Así, se puede afirmar que la edad del propulsor sí tiene un efecto significativo sobre la resistencia.

6. Suponga que se tienen tres lotes del propulsor tipo 1, con 5, 10 y 15 semanas de edad respectivamente. ¿Cuál es la estimación para la resistencia según el modelo de regresión (para cada lote)?

```
predict(RL, list(Edad = c(5, 10, 15)))
#      1      2      3
# 2442.054 2256.286 2070.518
```

Con esto se puede observar que, a mayor edad, es menor la resistencia. ✓

7. Calcule el valor del coeficiente de determinación (Multiple R Squared). Según este coeficiente. ¿qué tan bueno es el ajuste de la regresión?

```
summary(RL)$r.squared # 0.9018414
```

$$R^2 = 0.9018414$$

El modelo puede explicar aproximadamente el 90.18% de la variabilidad en la resistencia mediante la edad, por lo que es un muy buen ajuste. ✓

8. VERIFICACIÓN DE SUPUESTOS DEL MODELO. (Obten- ga primero los residuales)

a) NORMALIDAD. Grafique los residuales contra los cuantiles de una normal (qqnorm, qqline). ¿Se satisface este supuesto?

```
qqnorm(residuals(RL))
qqline(residuals(RL))
```

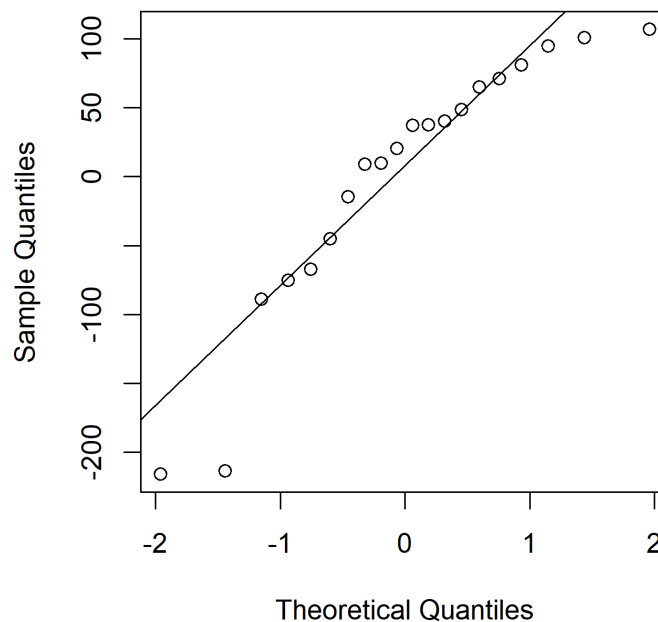


Figura 3: Gráfico de residuales del modelo vs una distribución normal teórica.

La mayoría de los puntos sigue la recta, excepto los 2 del lado inferior izquierdo y superior derecho. En general, no hay ninguna curvatura marcada, por lo que podemos decir que el supuesto de los errores con distribución normal se cumple.

b) MEDIA CERO, VARIANZA CONSTANTE E INDEPENDENCIA.
Grafique los residuales contra los predichos para verificar los tres supuestos (predict, rstudent). ¿Observa alguna anomalía?

```
plot(predict(RL), residuals(RL))  
abline(h = 0, col = "blue")  
  
plot(predict(RL), rstudent(RL))  
abline(h = 0, col = "blue")
```

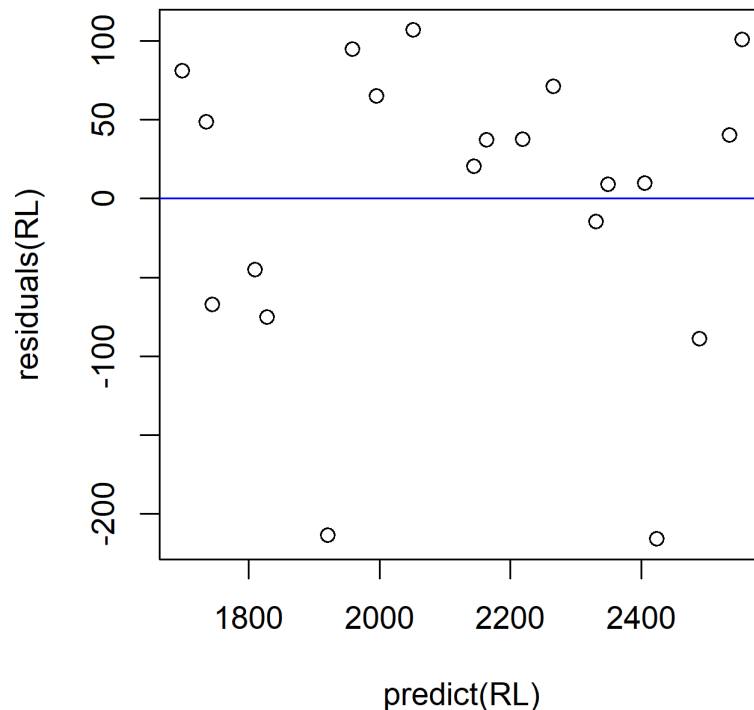


Figura 4: Gráfico de residuales vs valores ajustados.

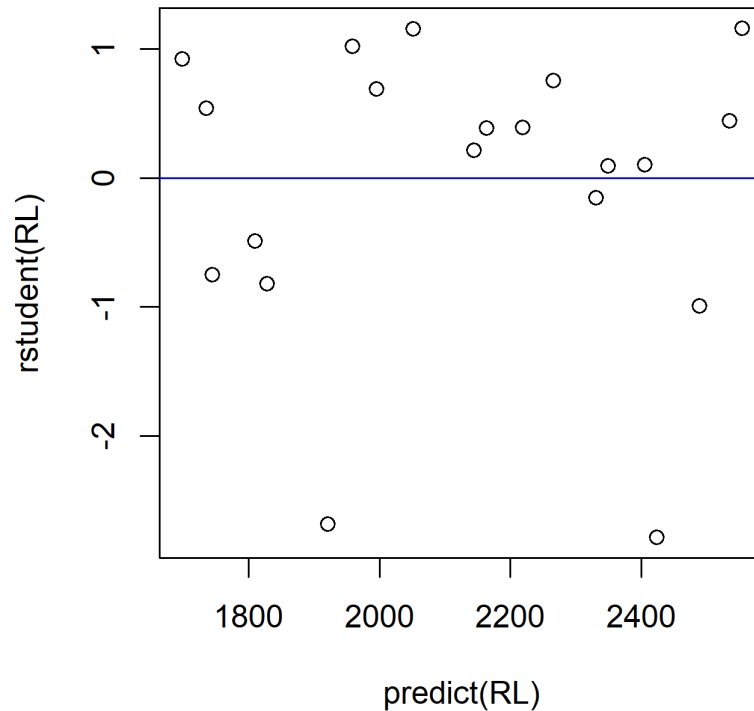


Figura 5: Gráfico de residuales estudentizados vs valores ajustados.

- Los puntos están dispersos de forma aleatoria alrededor de cero, sin notarse ningún patrón, por lo que se cumple el supuesto de media cero.
- La dispersión parece uniforme en todo el rango de valores (salvo 2 casos un poco más alejados) sin anomalías relevantes. Así, podemos decir que se cumple la varianza constante.
- No se observan agrupamientos o «estructuras», por lo que consideramos cierto que la independencia se cumple.

9. PUNTOS ATÍPICOS E INFLUYENTES

a) Utilizando la gráfica anterior, ¿se observan puntos que puedan considerarse como atípicos (outliers)?

Por debajo de -2 , utilizando la Figura 5, se muestran dos puntos un poco más alejados que los demás, por lo que podría tratarse de outliers.

b) Utilizando la distancia de Cook, verifique si hay puntos influyentes.

```
cook_values = cooks.distance(RL)
plot(cook_values, type = "h", ylim = c(0, max(cook_values, 1.2)))
abline(h = 1, col = "red")
```

R

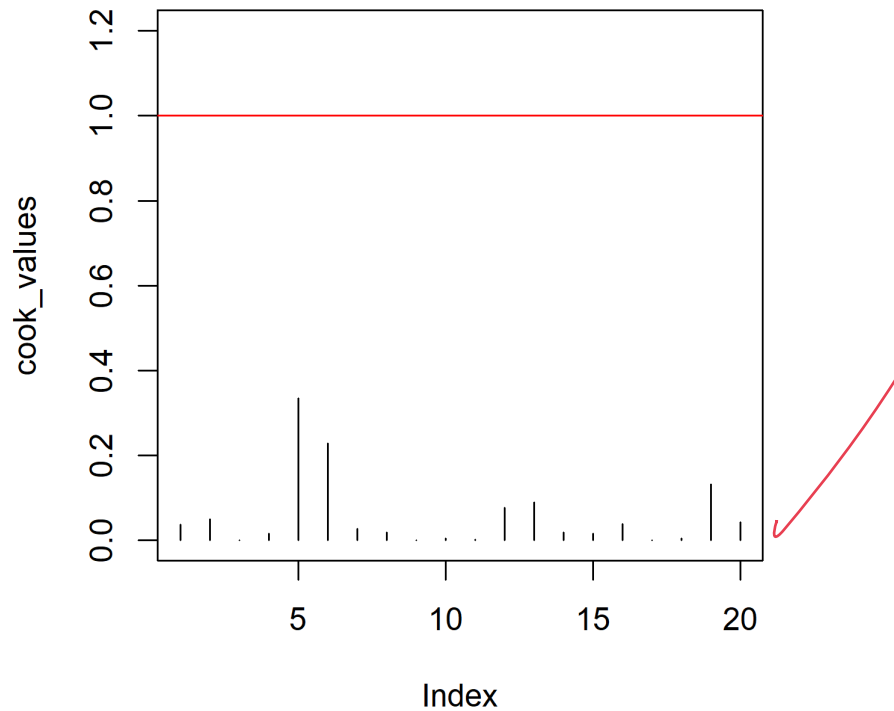


Figura 6: Gráfico para notar puntos influyentes. $D_i > 1$

No hay puntos con distancia de Cook mayor a 1, por tanto, no hay puntos influyentes.

10. Escriba una conclusión general para este problema

El análisis realizado muestra que existe una relación lineal negativa y significativa entre la edad del propulsor tipo 1 y la resistencia al corte. Mientras que aumenta la edad, la resistencia disminuye, como lo muestra el valor negativo de la pendiente de la recta.

El modelo de regresión lineal presenta un muy buen nivel de ajuste ($R^2 \approx 90\%$) y cumple con los supuestos de: normalidad, media cero, varianza constante e independencia de los errores. Y tampoco se detectaron puntos influyentes de acuerdo a la distancia de Cook. Así, concluimos que el modelo es bastante bueno para explicar y predecir la resistencia en función de la edad del propulsor.