Estadística y Diseño de Experimentos
Ejercicios 2

10

Nombre de lxs alumnxs:

•Rodrigo Pacheco Otero.

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

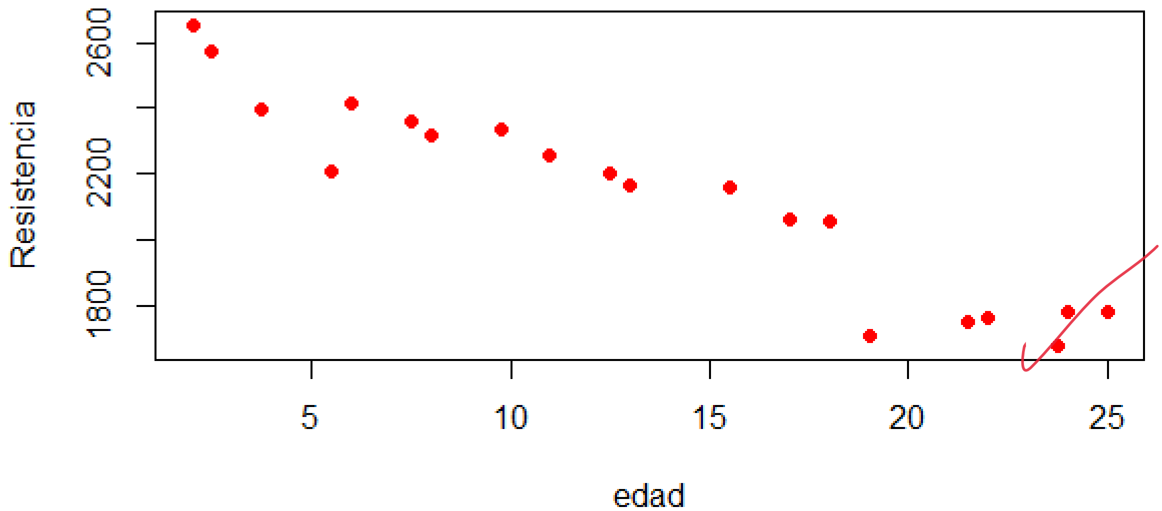
Cuadro 1: Resistencia obtenida según la edad del propulsor tipo 1.

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

R= La edad es la variable regresora X, y la resistencia la variable Y de resistencia. Respuesta

- Y = b0 + b1X.
- Resistencia = b0 + b1 * edad del lote tipo 1. ✓

2. **Grafique el diagrama de dispersión de los datos.**



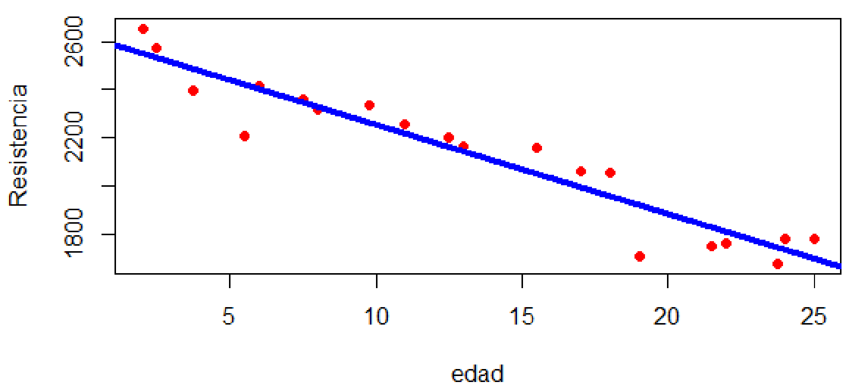
3. **Obtenga los estimadores para β0 y β1 y escriba la ecuación de la recta ajustada.**

```
> RL$coefficients
(Intercept)        edad
 2627.82236   -37.15359
```

R= Ecuación -> Resistencia = 2627.822 – 37.153 * edad

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

Da una sensación de que los datos están más inclinados hacia arriba de la recta, sin embargo, los poco que están por debajo están más alejados de la recta lo que compensaría ese “desfase”, por lo que yo diría que es un buen ajuste.



5. **Efectué la prueba de significancia de la regresión para un nivel α=0.05. Escriba el valor del p − valor. ¿Qué conclusiones puede hacer sobre β1?**

   R= p-valor = 1.64^(-10)

   Por hipótesis b1 es distinto de 0, y además podemos decir que la edad influye en la resistencia.

```
> alpha<-0.05
> p.valor<-1.64^(-10)
> p.valor<alpha #si es true se reachaza H0
[1] TRUE
```

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)?**

   Según nuestro modelo

   Resistencia = 2627.822 – 37.153 * 5 = 2442.054

   Resistencia = 2627.822 – 37.153 * 10 = 2256.286

   Resistencia = 2627.822 – 37.153 * 15 = 2070.518

```
> predict(RL, list(edad=c(5)), interval = 'prediction')
       fit      lwr      upr
1 2442.054 2229.021 2655.087
> predict(RL, list(edad=c(10)), interval = 'prediction')
       fit      lwr      upr
1 2256.286 2048.385 2464.188
> predict(RL, list(edad=c(15)), interval = 'prediction')
       fit      lwr      upr
1 2070.518 1863.382 2277.655
```

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

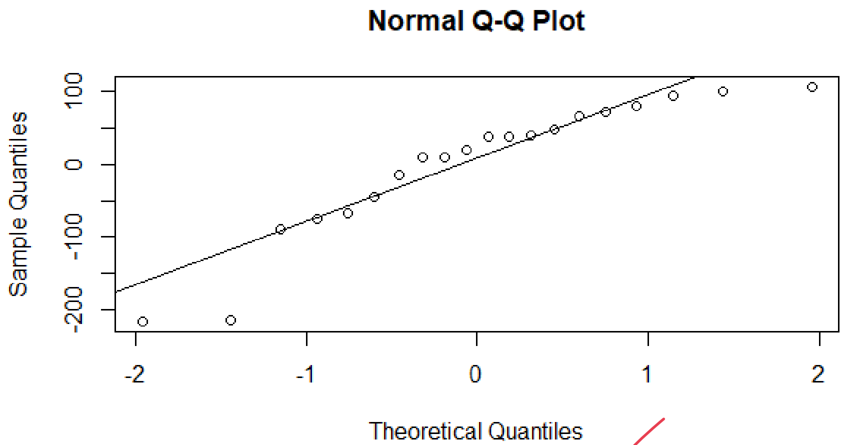
   R= R^2 = 0.9018 `Multiple R-squared:  0.9018,`

   Como nuestro valor es cercano a cero, podemos decir que la dependencia es baja, explica poco la variabilidad de la resistencia.

8. VERIFICACION DE SUPUESTOS DEL MODELO. (Obtenga primero los residuales)

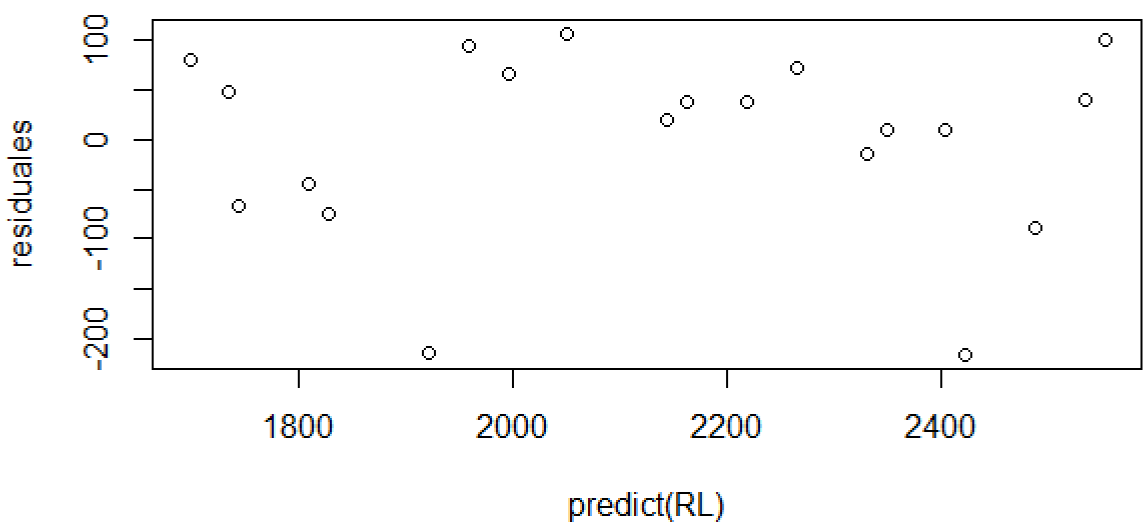
a) NORMALIDAD. Grafique los residuales contra los cuantiles de una normal (qqnorm, qqline). ¿Se satisface este supuesto?

R= Residuales yi -y´i



Como la mayoría de los datos siguen la recta y además la muestra es pequeña, podríamos decir que cumple el supuesto de normalidad, a pesar de los datos que no la siguen.

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



¿Observa alguna anomalía?

R=No se ve un patrón aparente, sin embargo, hay una cierta tendencia por encima del cero.

9. PUNTOS ATÍPICOS E INFLUYENTES.

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

R=Podemos tener 2 datos atípicos por estar muy alejados y tener una tendencia distinta a los de los demás, esos datos son los que están por debajo del -2.

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

Según la distancia de Cook, no hay puntos influyentes.

```
> cooks.distance(RL)>1 #Si es true, entonces es un punto influyente.
    1     2     3     4     5     6     7     8     9    10    11    12    13    14 
FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE 
   15    16    17    18    19    20 
FALSE FALSE FALSE FALSE FALSE FALSE 
```

10. Escriba una conclusión general para este problema.

Podemos observar que el modelo de regresión no es directamente proporcional respecto a las variables, ya que mientras más aumente la edad, menor resistencia tendremos. Además de tener una dependencia pequeña entre variables ya que la edad no explica bien la variabilidad de la resistencia en el modelo.

Es un modelo aceptable para los pocos datos que posee, ya que no aparenta datos atípicos ni influyentes, lo que nos serviría para aproximaciones en predicciones, sin embargo, aplicarlo en un caso real quizá no sea lo mejor por la ausencia de otras variables.