Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA
Unidad Iztapalapa

# UEA: Estadística y Diseño de Experimentos.

## Ejercicios 2.

Equipo:

- Velasco Islas Bryan Daniel.
- 2223009392
- Benavides Santiago Magaly.
- 2203007338

Docente: Montesinos Vázquez Joel.

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.

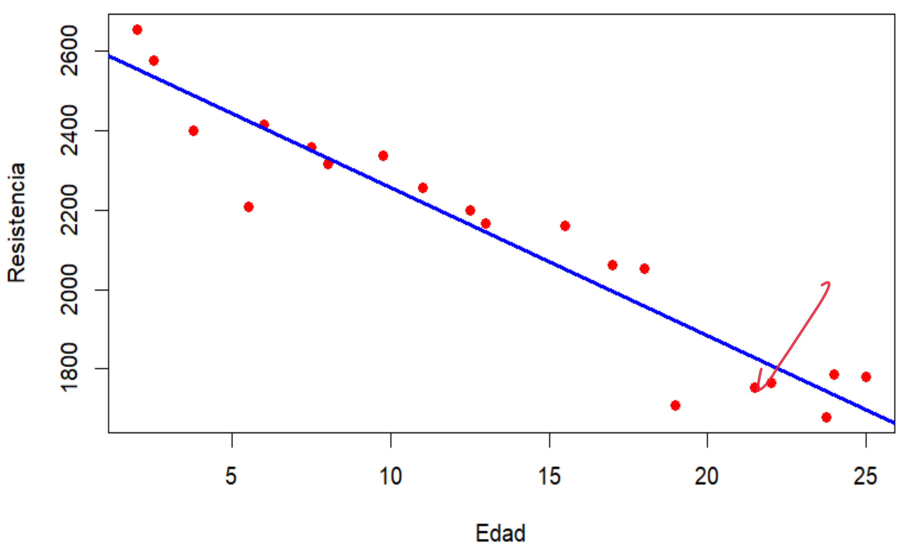
1.Identifique quién es la variable de respuesta Y y quién es la variable regresora X y escriba el modelo de regresión.

- - La variable de respuesta Y es la Resistencia.
- - La variable regresora X es la Edad.
- - Y=β0+β1X

```
> #1) Modelo de regresión
> Resistencia<-c(2158.70,1678.15,2316,2061.30,2207.5,1708.30,1784.70,2575,2357.90,2256.70,2165.20,2
399.95,1779.8,2336.75,1765.30,2053.5,2414.4,2200.5,2654.20,1753.7)
> Edad<-c(15.50,23.75,8,17,5.5,19,24,2.5,7.5,11,13,3.75,25,9.75,22,18,6,12.5,2,21.50)
```

2. Grafique el diagrama de dispersión de los datos.

```
#2) Dispersión
plot(Edad,Resistencia,pch=16,col='orange')
```



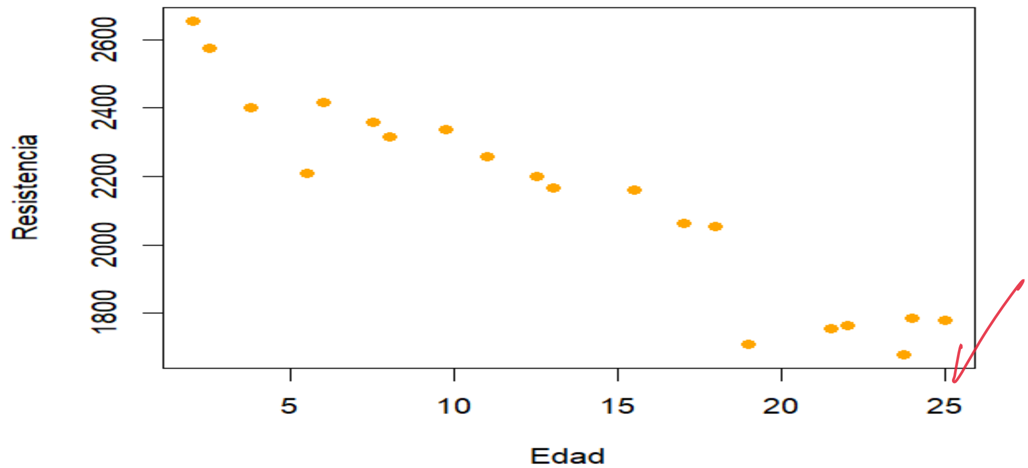

3. Obtenga los estimadores para β0 y β1 y escriba la ecuación de la recta ajustada.

*β^0 = 2627.88879*

*β^1 = -37.15707*

Y = β^ 0 + β^1X

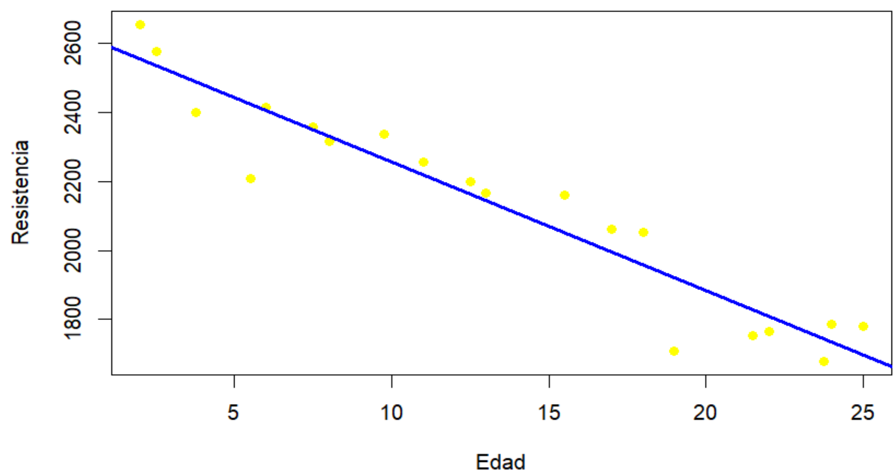
```
> #3) Ajuste de regresión betas
> RL<-lm(Resistencia~Edad)
> #Valores para beta0 y beta1
> RL$coefficients
(Intercept)        Edad
 2627.88879   -37.15707
```




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

Tiene tendencia negativa, es decir, a medida que la edad aumenta, la resistencia disminuye. Visualmente se observa que los puntos quedan un poco dispersos si hacemos una regresión con tendencia lineal.

```
> #4) Gráfica de regresión
> plot(Edad,Resistencia,pch=16,col='yellow')
> abline(RL,col='blue',lwd=3)
```




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 $\beta 1$?

Con $\alpha$= 0.05 rechazamos a H_0, es decir que beta es diferente de 1. La edad está asociada negativamente a la resistencia (pendiente negativa).

```
> #Nivel de significancia
> alpha<-0.05
> p.valor<-0.00014
> p.valor<alpha
[1] TRUE
> # Intervalos de confianza
> confint(RL)
                2.5 %      97.5 %
(Intercept) 2535.08166 2720.69592
Edad          -43.22556  -31.08857
>
```

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)?

Escribe los valores estimados

```
> #6) Predicción
> #i)Estimación para la resistencia si la Edad=5
> 2627.88879-37.15707*5
[1] 2442.103
> predict(RL,list(Edad=c(5)),interval='confidence')
       fit      lwr      upr
1 2442.103 2374.185 2510.022
> predict(RL,list(Edad=c(5)),interval='prediction')
       fit      lwr      upr
1 2442.103 2229.116 2655.091
> #ii)Estimación para la resistencia si Edad=10
> 2627.88879-37.15707*10
[1] 2256.318
> predict(RL,list(Edad=c(10)),interval='confidence')
       fit      lwr      upr
1 2256.318 2206.781 2305.855
> predict(RL,list(Edad=c(10)),interval='prediction')
       fit      lwr      upr
1 2256.318 2048.461 2464.176
> #iii)Estimación para la resistencia si la Edad=15
> 2627.88879-37.15707*15
[1] 2070.533
> predict(RL,list(Edad=c(15)),interval='confidence')
       fit      lwr      upr
1 2070.533 2024.313 2116.753
> predict(RL,list(Edad=c(15)),interval='prediction')
       fit      lwr      upr
1 2070.533 1863.441 2277.625
```

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

Nos está indicando que el 90.2% de la variabilidad de la resistencia se explica linealmente con la edad.

```
> #7) Coeficiente de determinación
> summary(RL)

Call:
lm(formula = Resistencia ~ Edad)

Residuals:
    Min      1Q  Median      3Q     Max 
-216.03  -50.66   28.71   66.60  106.75 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2627.889     44.174   59.49  < 2e-16 ***
Edad         -37.157      2.888  -12.86 1.64e-10 ***
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 96.09 on 18 degrees of freedom
Multiple R-squared:  0.9019,    Adjusted R-squared:  0.896
4
F-statistic: 165.5 on 1 and 18 DF,  p-value: 1.635e-10
```

8. VERIFICACIÓN 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?

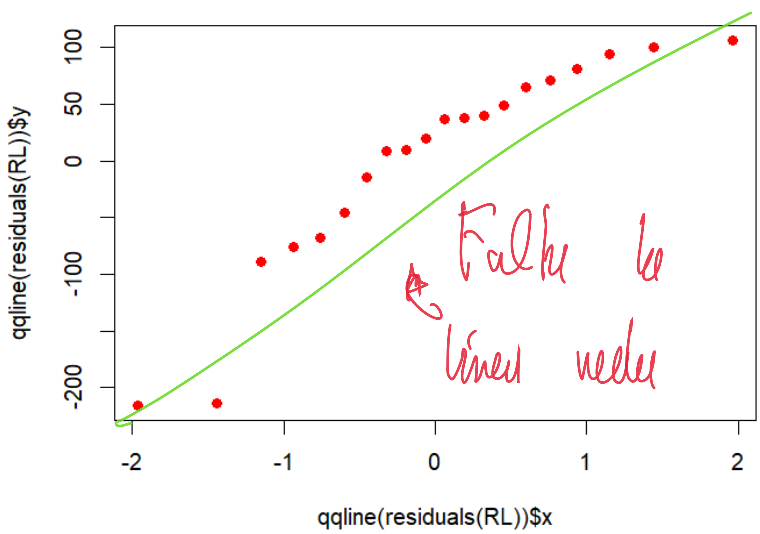
Los puntos en el gráfico Q-Q parecen seguir de cerca la línea recta (qqline), especialmente en el centro. Las desviaciones en los extremos no son extremas. Por lo tanto, se puede considerar que el supuesto de normalidad de los residuales se satisface razonablemente.

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

¿Observa alguna anomalía?

Se observan puntos que no siguen la tendencia lineal pero no afecta tanto a los resultados y se puede graficar sin alteraciones a los datos.

```
> #8) Verificación de supuestos
> #i)Normalidad
> qqnorm(residuals(RL))
> qqline(residuals(RL))
> plot(qqnorm(residuals(RL)),qqline(residuals(RL)),pch=16,col='red')
> #ii)media 0,v.constante e independecia
> plot(predict(RL),rstudent(RL))
```




9. PUNTOS ATÍPICOS E INFLUYENTES.

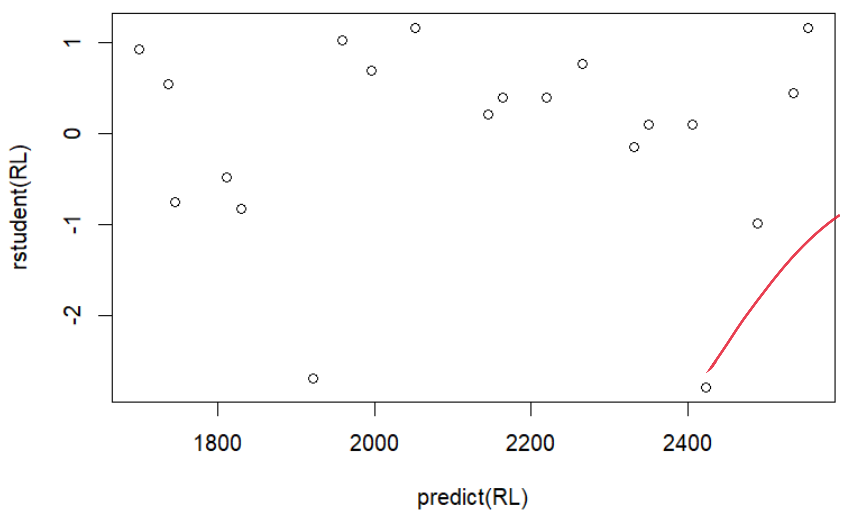
a) Utilizando la gráfica anterior, .se observan puntos que puedan considerarse como atípicos (outliers)?

Sí, se observan puntos que se desvían considerablemente de la línea teórica en los extremos de la gráfica, lo que sugiere la presencia de valores atípicos en los residuales.

¿Cuáles puntos?

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

```
> plot(predict(RL),rstudent(RL))
> #9) Puntos atipicos
> cooks.distance(RL)
           1            2            3            4            5            6 
0.0373353549 0.0497266372 0.0010319566 0.0161515191 0.3346663957 0.2291830248 
           7            8            9           10           11           12 
0.0270795864 0.0190854129 0.0003925013 0.0047043945 0.0012461632 0.0755914526 
          13           14           15           16           17           18 
0.0890722829 0.0192423778 0.0166299858 0.0387292119 0.0005901489 0.0041854588 
          19           20 
0.1316150188 0.0425809858 
> cooks.distance(RL)>1
    1     2     3     4     5     6     7     8     9    10    11    12    13 
FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE 
   14    15    16    17    18    19    20 
FALSE FALSE FALSE FALSE FALSE FALSE FALSE 
```




10. Escriba una conclusión general para este problema.

Se observó que existe una relación lineal entre ambas variables. La pendiente estimada del modelo indica cómo varía la resistencia conforme aumenta la edad de los propulsores, mostrando que, a mayor tiempo transcurrido, la resistencia tiende a disminuir.

El valor del coeficiente de determinación permitió evaluar la calidad del ajuste, indicando qué proporción de la variabilidad de la resistencia puede explicarse por la edad. Además, la prueba de significancia de la regresión mostró que el modelo es

estadísticamente válido para describir esta relación.
Finalmente, la verificación de los supuestos del modelo (normalidad, homocedasticidad e independencia de los residuales) no evidenció violaciones importantes, por lo que el modelo ajustado puede considerarse adecuado para realizar estimaciones y predicciones sobre la resistencia del material en función de la edad de los propulsores.