

## A2-Regresión Múltiple

Eliezer Cavazos

2024-09-17

```
oCorte = read.csv("C:\\Users\\eliez\\OneDrive\\Desktop\\Clases\\AlCorte.csv")  
#Leer la base de datos
```

### 1. Haz un análisis descriptivo de los datos: medidas principales y gráficos

```
oModelo = lm(Resistencia~., data=oCorte)  
oPasos = step(oModelo, direction="both", trace=1)  
  
## Start: AIC=102.96  
## Resistencia ~ Fuerza + Potencia + Temperatura + Tiempo  
##  
##           Df Sum of Sq    RSS    AIC  
## - Fuerza    1    26.88  692.00 102.15  
## - Tiempo    1    40.04  705.16 102.72  
## <none>                                665.12 102.96  
## - Temperatura 1    252.20  917.32 110.61  
## - Potencia    1   1341.01 2006.13 134.08  
##  
## Step: AIC=102.15  
## Resistencia ~ Potencia + Temperatura + Tiempo  
##  
##           Df Sum of Sq    RSS    AIC  
## - Tiempo    1    40.04  732.04 101.84  
## <none>                                692.00 102.15  
## + Fuerza    1    26.88  665.12 102.96  
## - Temperatura 1    252.20  944.20 109.47  
## - Potencia    1   1341.02 2033.02 132.48  
##  
## Step: AIC=101.84  
## Resistencia ~ Potencia + Temperatura  
##  
##           Df Sum of Sq    RSS    AIC  
## <none>                                732.04 101.84  
## + Tiempo    1    40.04  692.00 102.15  
## + Fuerza    1    26.88  705.16 102.72  
## - Temperatura 1    252.20  984.24 108.72  
## - Potencia    1   1341.01 2073.06 131.07
```

## 2. Encuentra el mejor modelo de regresión que explique la variable Resistencia. Analiza el modelo basándote en:

Significancia del modelo: Economía de las variables Significación global (Prueba para el modelo) Significación individual (Prueba para cada  $\beta_i$ ) Variación explicada por el modelo

```
oModeloNulo = lm(Resistencia~1, data=oCorte)
oPasos2 = step(oModeloNulo, scope=list(lower=oModeloNulo, upper = oModelo),
direction="forward")

## Start:  AIC=132.51
## Resistencia ~ 1
##
##              Df Sum of Sq    RSS    AIC
## + Potencia    1   1341.01   984.24 108.72
## + Temperatura 1    252.20 2073.06 131.07
## <none>                        2325.26 132.51
## + Tiempo      1     40.04 2285.22 133.99
## + Fuerza      1     26.88 2298.38 134.16
##
## Step:  AIC=108.72
## Resistencia ~ Potencia
##
##              Df Sum of Sq    RSS    AIC
## + Temperatura 1    252.202 732.04 101.84
## <none>                        984.24 108.72
## + Tiempo      1     40.042 944.20 109.47
## + Fuerza      1     26.882 957.36 109.89
##
## Step:  AIC=101.84
## Resistencia ~ Potencia + Temperatura
##
##              Df Sum of Sq    RSS    AIC
## <none>                        732.04 101.84
## + Tiempo    1     40.042 692.00 102.15
## + Fuerza    1     26.882 705.16 102.72

#oModelo2 = lm(Resistencia ~ Potencia + Temperatura, data=oModelo)
oModelo2 = summary(oPasos2)

oN = length(oCorte$Resistencia)
oPasos3=step(oModelo,direction="both",k=log(oN))

## Start:  AIC=109.97
## Resistencia ~ Fuerza + Potencia + Temperatura + Tiempo
##
##              Df Sum of Sq    RSS    AIC
## - Fuerza      1     26.88 692.00 107.76
## - Tiempo      1     40.04 705.16 108.32
## <none>                        665.12 109.97
```

```

## - Temperatura 1 252.20 917.32 116.21
## - Potencia 1 1341.01 2006.13 139.69
##
## Step: AIC=107.76
## Resistencia ~ Potencia + Temperatura + Tiempo
##
##           Df Sum of Sq    RSS    AIC
## - Tiempo    1    40.04  732.04 106.04
## <none>                        692.00 107.76
## + Fuerza    1    26.88  665.12 109.97
## - Temperatura 1    252.20  944.20 113.68
## - Potencia  1   1341.02 2033.02 136.69
##
## Step: AIC=106.04
## Resistencia ~ Potencia + Temperatura
##
##           Df Sum of Sq    RSS    AIC
## <none>                        732.04 106.04
## + Tiempo    1    40.04  692.00 107.76
## + Fuerza    1    26.88  705.16 108.32
## - Temperatura 1    252.20  984.24 111.52
## - Potencia  1   1341.01 2073.06 133.87

summary(oPasos3)

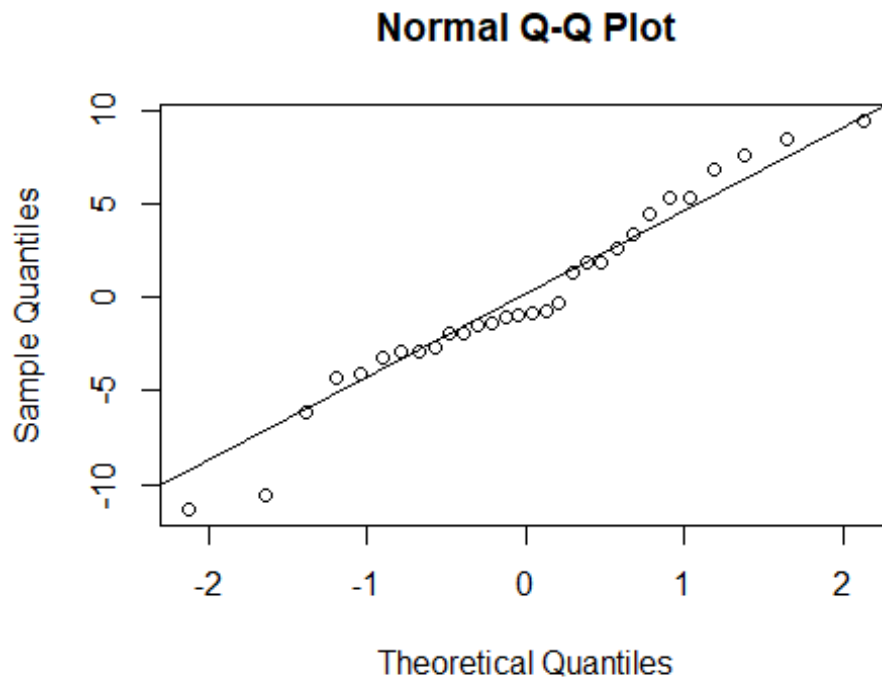
##
## Call:
## lm(formula = Resistencia ~ Potencia + Temperatura, data = oCorte)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.3233  -2.8067  -0.8483   3.1892   9.4600
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -24.90167   10.07207  -2.472  0.02001 *
## Potencia     0.49833    0.07086   7.033 1.47e-07 ***
## Temperatura  0.12967    0.04251   3.050 0.00508 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.207 on 27 degrees of freedom
## Multiple R-squared:  0.6852, Adjusted R-squared:  0.6619
## F-statistic: 29.38 on 2 and 27 DF, p-value: 1.674e-07

```

### 3. Analiza la validez del modelo encontrado:

Análisis de residuos (homocedasticidad, independencia, etc)

```
qqnorm(oPasos3$residuals)
qqline(oPasos3$residuals)
```



```
t.test(oPasos3$residuals)

##
##  One Sample t-test
##
## data:  oPasos3$residuals
## t = 8.8667e-17, df = 29, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -1.876076  1.876076
## sample estimates:
##  mean of x
## 8.133323e-17
```

Homocedasticidad

```
library(lmtest)

## Loading required package: zoo

##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':  
##  
##    as.Date, as.Date.numeric  
  
bptest(oPasos3)  
  
##  
## studentized Breusch-Pagan test  
##  
## data: oPasos3  
## BP = 4.0043, df = 2, p-value = 0.135
```

Linealidad

```
dwtest(oPasos3)  
  
##  
## Durbin-Watson test  
##  
## data: oPasos3  
## DW = 2.3511, p-value = 0.8267  
## alternative hypothesis: true autocorrelation is greater than 0
```

No multicolinealidad de Xi

```
library(car)  
  
## Loading required package: carData  
  
vif(oPasos3)  
  
##    Potencia Temperatura  
##          1          1
```

#### 4. Emite conclusiones sobre el modelo final encontrado e interpreta en el contexto del problema el efecto de las variables predictoras en la variable respuesta

La conclusion final que se puede identificar es que las variables que más influyen a la Resistencia es Potencia y Temperatura, los valores de Multicolinealidad de estas variables es de 1, lo que se puede entender que no existe multicolinealidad en el modelo, con esto se puede identificar que en el contexto del problema hay menos resistencia cuando se emplea más fuerza y/o hay mayor temperatura al momento de cortar.