Actividad 7 Regresion

Elías Garza A01284041

29/8/2023

```
M = read.csv("C://Users//elias//Downloads//Estatura-peso_HyM.csv")
head(M)
```

```
##   Estatura  Peso Sexo
## 1     1.61 72.21    H
## 2     1.61 65.71    H
## 3     1.70 75.08    H
## 4     1.65 68.55    H
## 5     1.72 70.77    H
## 6     1.63 77.18    H
```

```
summary(M)
```

```
##     Estatura          Peso           Sexo
##  Min.   :1.440   Min.   :37.39   Length:440
##  1st Qu.:1.560   1st Qu.:54.49   Class :character
##  Median :1.610   Median :64.53   Mode  :character
##  Mean   :1.613   Mean   :63.97
##  3rd Qu.:1.660   3rd Qu.:73.22
##  Max.   :1.800   Max.   :90.49
```

Modelo A:

```
A = lm(M$Peso ~ M$Estatura + M$Sexo)
summary(A)
```

```
##
## Call:
## lm(formula = M$Peso ~ M$Estatura + M$Sexo)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.9505  -3.2491   0.0489   3.2880  17.1243
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -74.7546     7.5555  -9.894   <2e-16 ***
## M$Estatura   89.2604     4.5635  19.560   <2e-16 ***
```

```
## M$SexoM     -10.5645     0.6317 -16.724   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.381 on 437 degrees of freedom
## Multiple R-squared:  0.7837, Adjusted R-squared:  0.7827
## F-statistic: 791.5 on 2 and 437 DF,  p-value: < 2.2e-16
```

```
b0 = A$coefficients[1]
b1 = A$coefficients[2]
b2 = A$coefficients[3]
```

```
cat('Para mujeres', '\n')
```

```
## Para mujeres
```

```
cat('Peso = ', b0, '+', b1, 'estatura')
```

```
## Peso =  -74.7546 + 89.26035 estatura
```

```
cat('Para hombres', '\n')
```

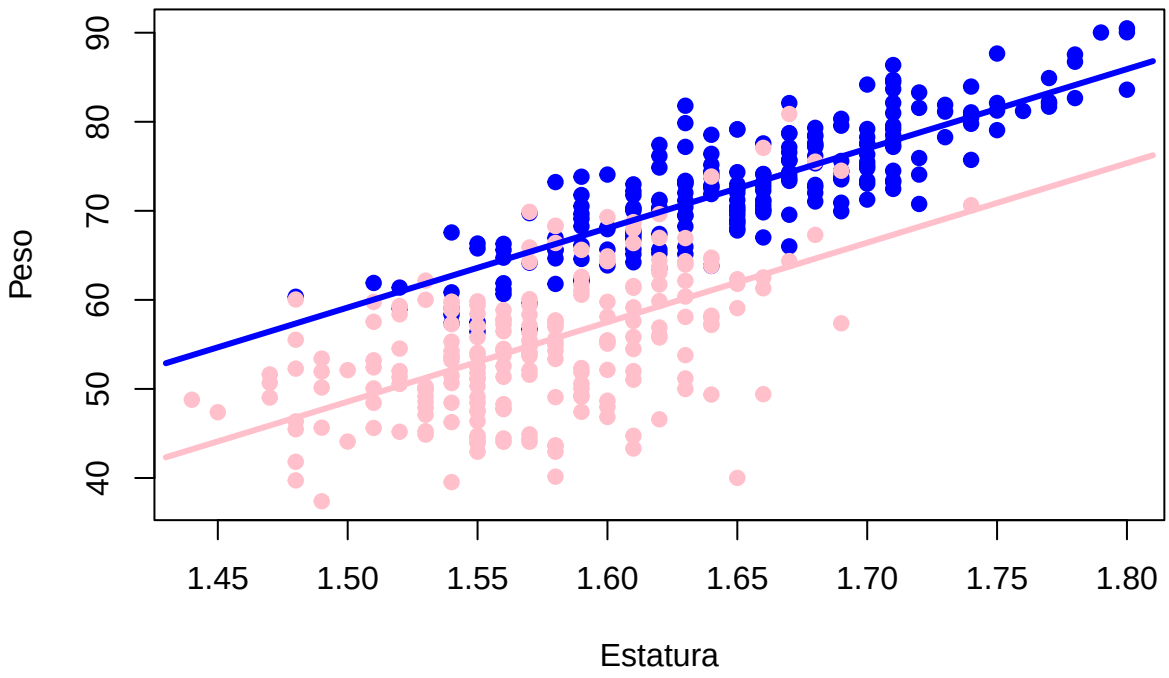
```
## Para hombres
```

```
cat('Peso = ', b0 + b2, '+', b1, 'estatura')
```

```
## Peso =  -85.31907 + 89.26035 estatura
```

```
Ym = function(x){b0 +b2 +b1*x}
Yh = function(x){b0 +b1*x}

colores = c('blue', 'pink')
plot(M$Estatura, M$Peso, col=colores[factor(M$Sexo)], pch=19, ylab='Peso', xlab='Estatura')
x = seq(1.43, 1.81, 0.01)
lines(x, Ym(x), col='pink', lwd=3)
lines(x, Yh(x), col='blue', lwd=3)
```



Como las pendientes son paralelas podemos decir que no tenemos interacción entre los grupos. Ahora generaremos para interaccion de variables

```
B = lm(M$Peso ~M$Estatura*M$Sexo)
summary(B)
```

```
##
## Call:
## lm(formula = M$Peso ~ M$Estatura * M$Sexo)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.3256  -3.1107   0.0204   3.2691  17.9114
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          -83.685      9.735  -8.597   <2e-16 ***
## M$Estatura            94.660      5.882  16.092   <2e-16 ***
## M$SexoM               11.124     14.950   0.744    0.457
## M$Estatura:M$SexoM  -13.511      9.305  -1.452    0.147
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.374 on 436 degrees of freedom
## Multiple R-squared:  0.7847, Adjusted R-squared:  0.7832
## F-statistic: 529.7 on 3 and 436 DF,  p-value: < 2.2e-16
```

Como vemos que el p-value es mayor a 0.05 no consideramos que haya interaccion.

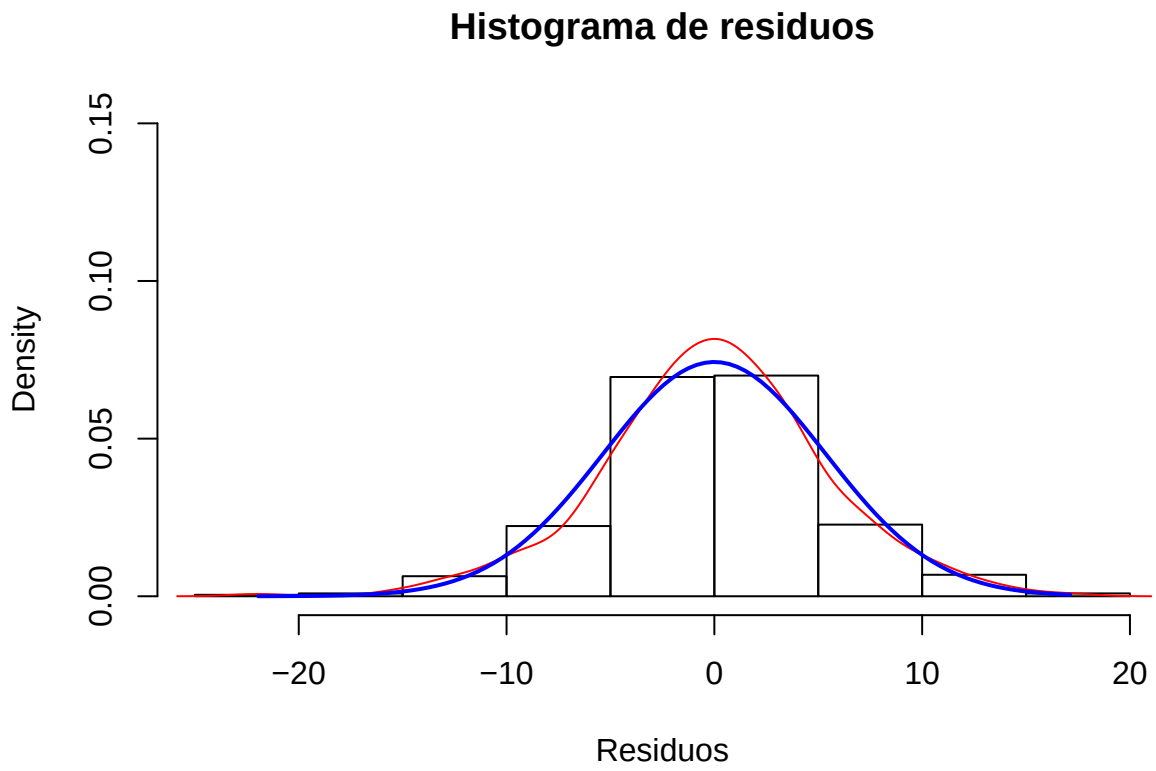
Ahora realizaremos el analisis de residuos.

```
library(nortest)
ad.test(A$residuals)
```

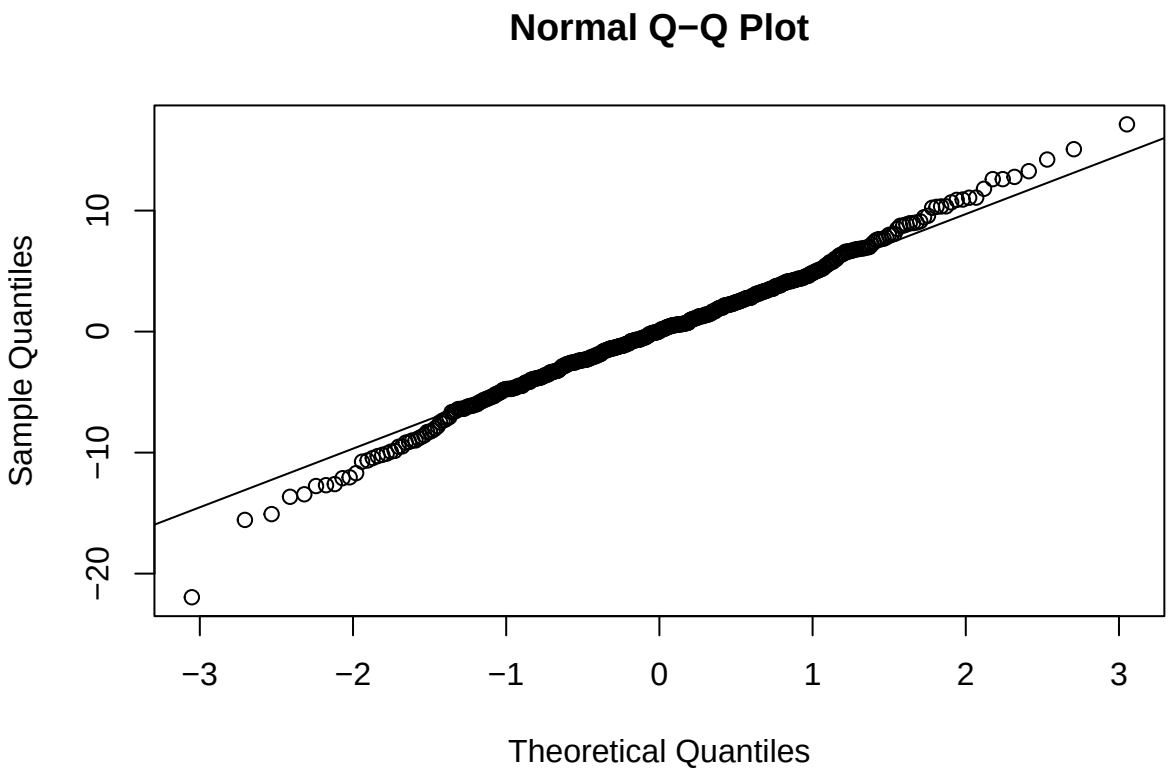
```
##
##  Anderson-Darling normality test
##
## data:  A$residuals
## A = 0.79651, p-value = 0.03879
```

Como tenemos un p-value mayor a 0.03 podemos decir que los residuos siguen una distribucion normal.

```
hist(A$residuals, freq = FALSE, ylim = c(0,0.15), xlab = "Residuos", col = 0, main = "Histograma de resi
lines(density(A$residuals), col = "red")
curve(dnorm(x, mean = mean(A$residuals), sd = sd(A$residuals)),
from = min(A$residuals), to = max(A$residuals), add = TRUE,
col = "blue", lwd =2)
```



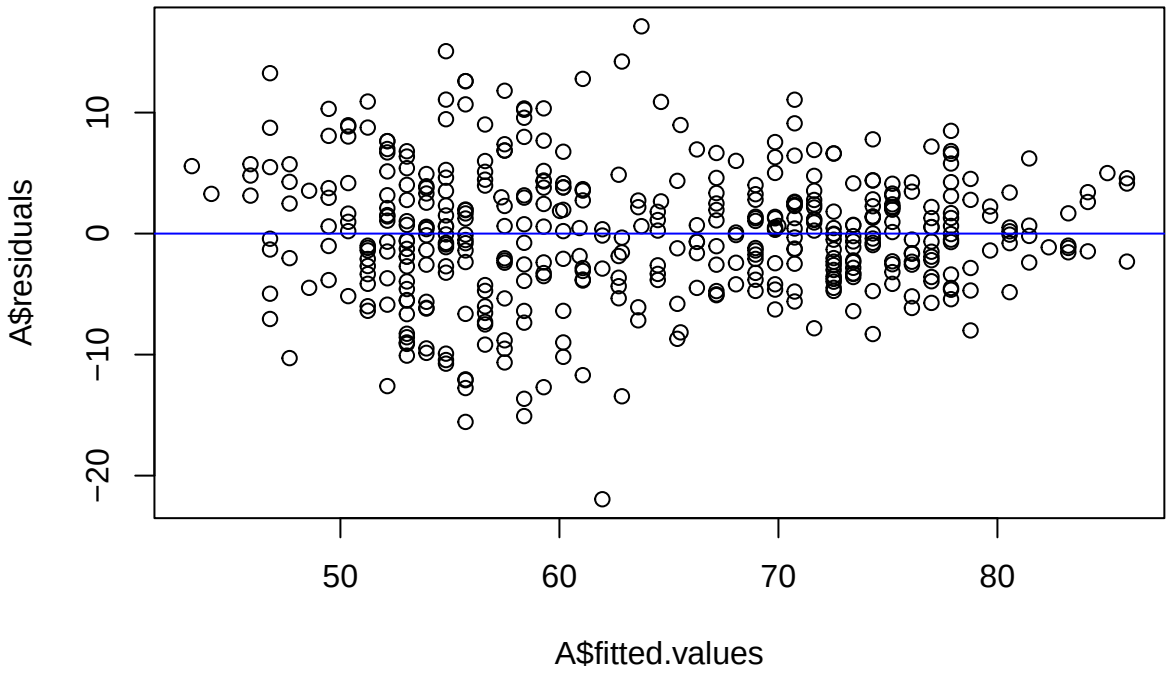

```
qqnorm(A$residuals)
qqline(A$residuals)
```



Como podemos ver, esto confirma que los residuos siguen una distribución normal.

Para ver homoestisidad

```
plot(A$fitted.values,A$residuals)
abline(h=0, col="blue")
```


A$residuals
10
0
−10
−20
50
60
70
80
A$fitted.values