Act 7

Ana Lucía Cárdenas Pérez A01284090

2023-08-29

```
M = read.csv("Estatura-peso_HyM.csv")
install.packages('nortest')
```

```
## Installing package into 'C:/Users/anaca/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)
```

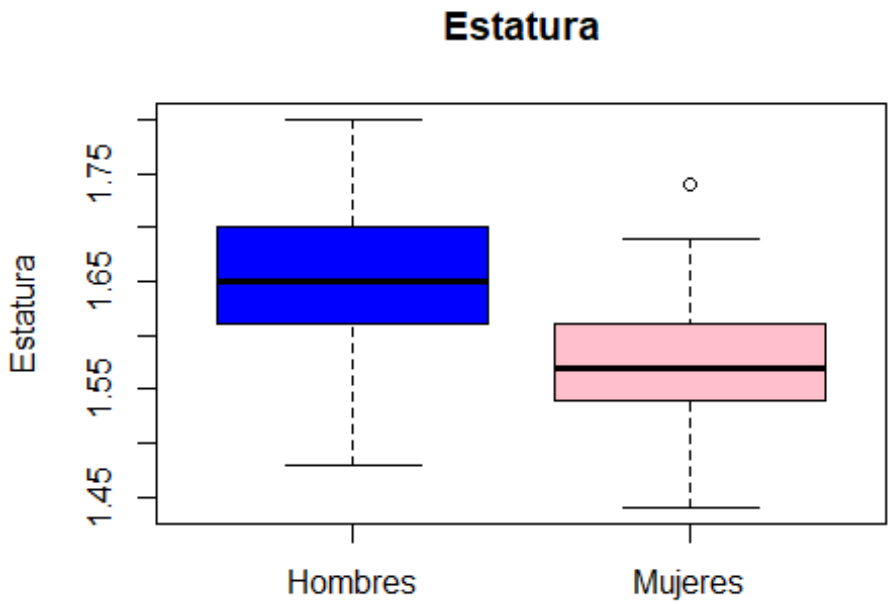
```
## package 'nortest' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\anaca\AppData\Local\Temp\RtmpWqRbmA\downloaded_packages
```

```
library(nortest)
```

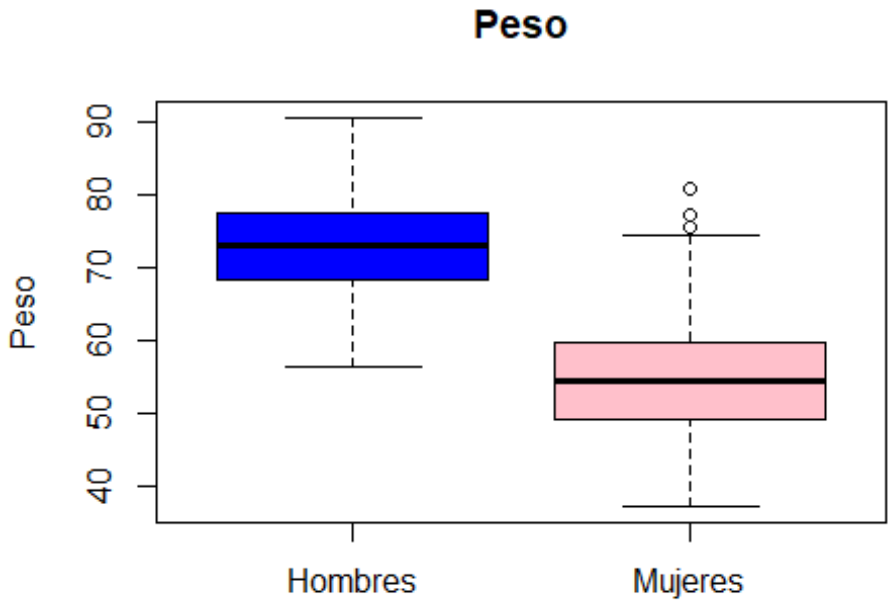
```
MM = subset(M,M$Sexo=="M")
MH = subset(M,M$Sexo=="H")
M1=data.frame(MH$Estatura,MH$Peso,MM$Estatura,MM$Peso)

n=4 #número de variables
d=matrix(NA,ncol=7,nrow=n)
for(i in 1:n){
  d[i,]<-c(as.numeric(summary(M1[,i])),sd(M1[ ,i]))
}
m=as.data.frame(d)

row.names(m)=c("H-Estatura","H-Peso","M-Estatura","M-Peso")
names(m)=c("Minimo","Q1","Mediana","Media","Q3","Máximo","Desv Est")
m
```

```
##            Minimo      Q1 Mediana     Media      Q3 Máximo   Desv Est
## H-Estatura   1.48  1.6100   1.650  1.653727  1.7000   1.80 0.06173088
## H-Peso      56.43 68.2575  72.975 72.857682 77.5225  90.49 6.90035408
## M-Estatura   1.44  1.5400   1.570  1.572955  1.6100   1.74 0.05036758
## M-Peso      37.39 49.3550  54.485 55.083409 59.7950  80.87 7.79278074
```

```
boxplot(M$Estatura~M$Sexo, ylab="Estatura", xlab="", col=c("blue","pink"),
names=c("Hombres", "Mujeres"), main="Estatura")
```




```r
boxplot(M$Peso~M$Sexo, ylab="Peso",xlab="", names=c("Hombres", "Mujeres"),
col=c("blue","pink"), main="Peso")
```



Regresion lineal

Preparacion de datos

verificar como esta definida la variable Sexo: caracter (variable categorica)

Modelo con Variable Sexo

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

```
##
## Call:
## lm(formula = M$Peso ~ M$Estatura + M$Sexo)
##
## Coefficients:
## (Intercept)   M$Estatura      M$SexoM
##      -74.75        89.26       -10.56
```

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

cat("Peso = ", b0, "+", b1, "Estatura", b2, "SexoM")
```

```
## Peso =  -74.7546 + 89.26035 Estatura -10.56447 SexoM
```

Verificacion del modelo

- Significancia Gobal
- Significancia Individual
- Porcentaje de variacion explicada por el modelo

```
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
```

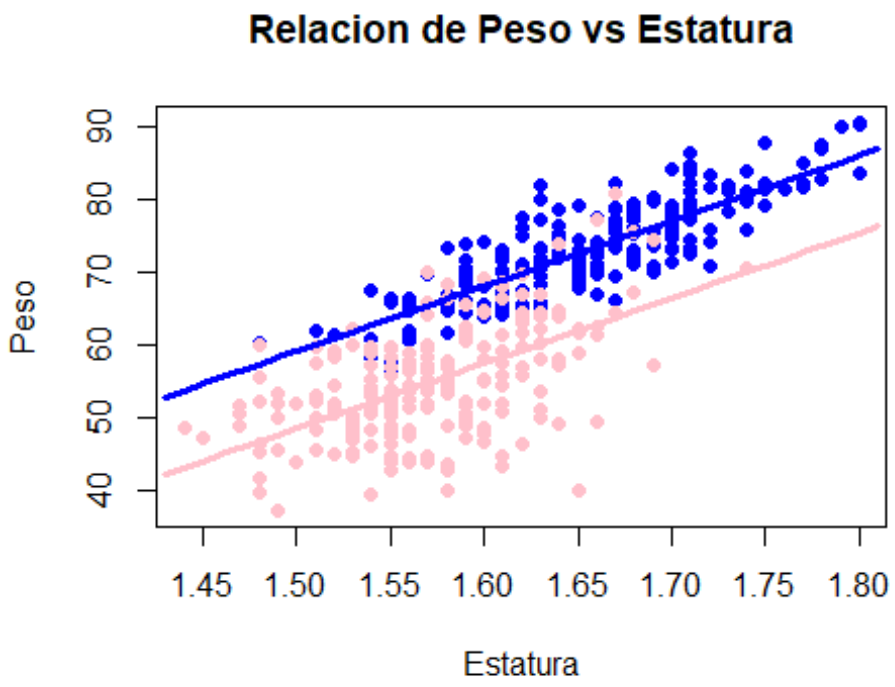
```
## Multiple R-squared:  0.7837, Adjusted R-squared:  0.7827
## F-statistic: 791.5 on 2 and 437 DF,  p-value: < 2.2e-16
```

Ecuacion del modelo

```
#Para mujeres (SexoM = 1)
cat("Para Mujeres", "\n")

## Para Mujeres

cat("Peso =",b0+b2, "+", b1, "Estatura", "\n")

## Peso = -85.31907 + 89.26035 Estatura

#Para hombres (SexoM = 0)
cat("Para Hombres", "\n")

## Para Hombres

cat("Peso = ", b0, "+", b1, "Estatura", "\n")

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

Grafica

```
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", main="Relacion de Peso vs 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)
```



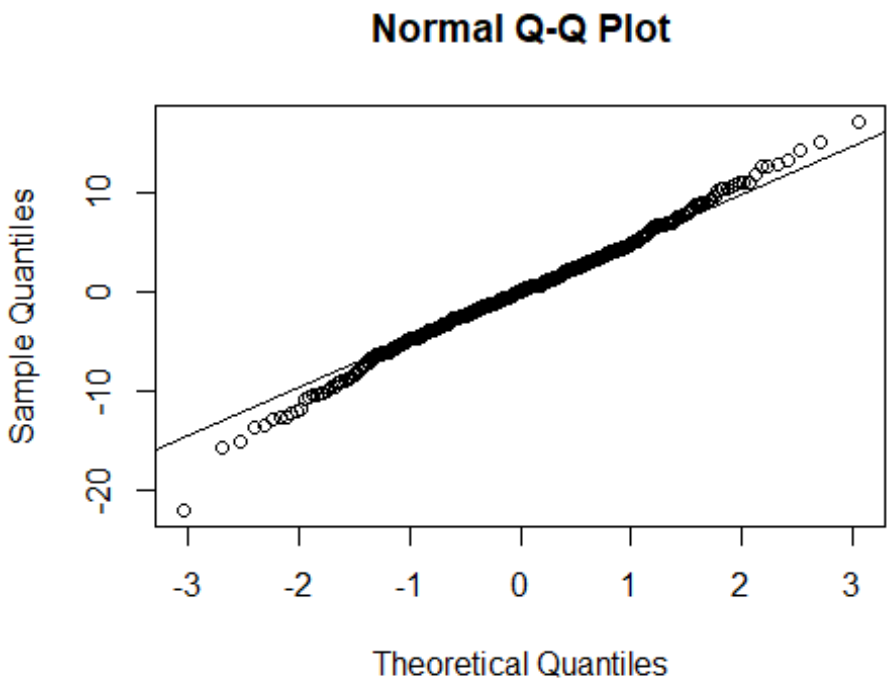

```
B = lm(M$Peso~M$Estatura*M$Sexo)
B

##
## Call:
## lm(formula = M$Peso ~ M$Estatura * M$Sexo)
##
## Coefficients:
##        (Intercept)          M$Estatura             M$SexoM
M$Estatura:M$SexoM
##             -83.68               94.66                11.12              -
13.51

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
```

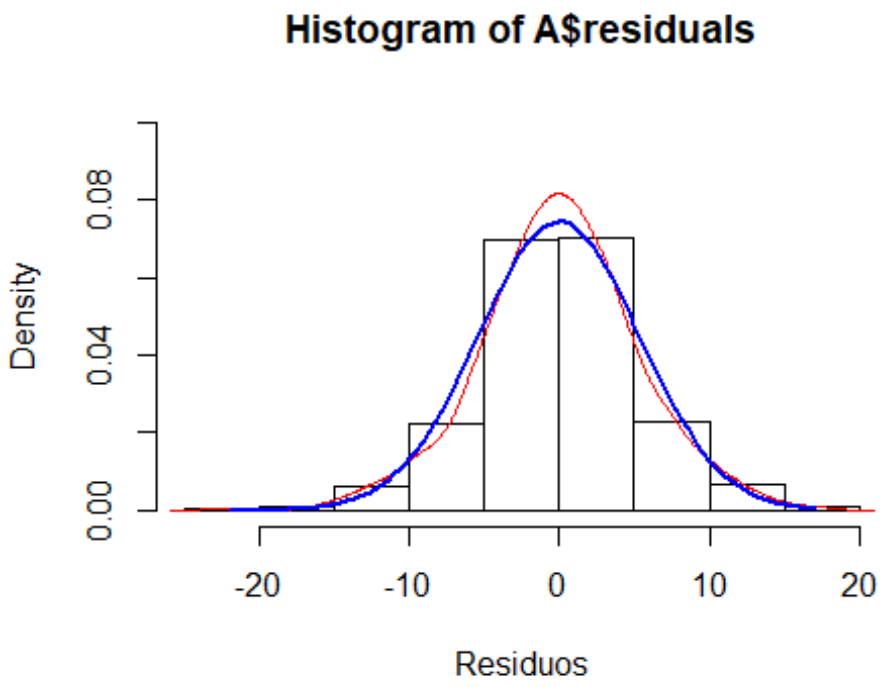
```
ad.test(A$residuals)
```

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

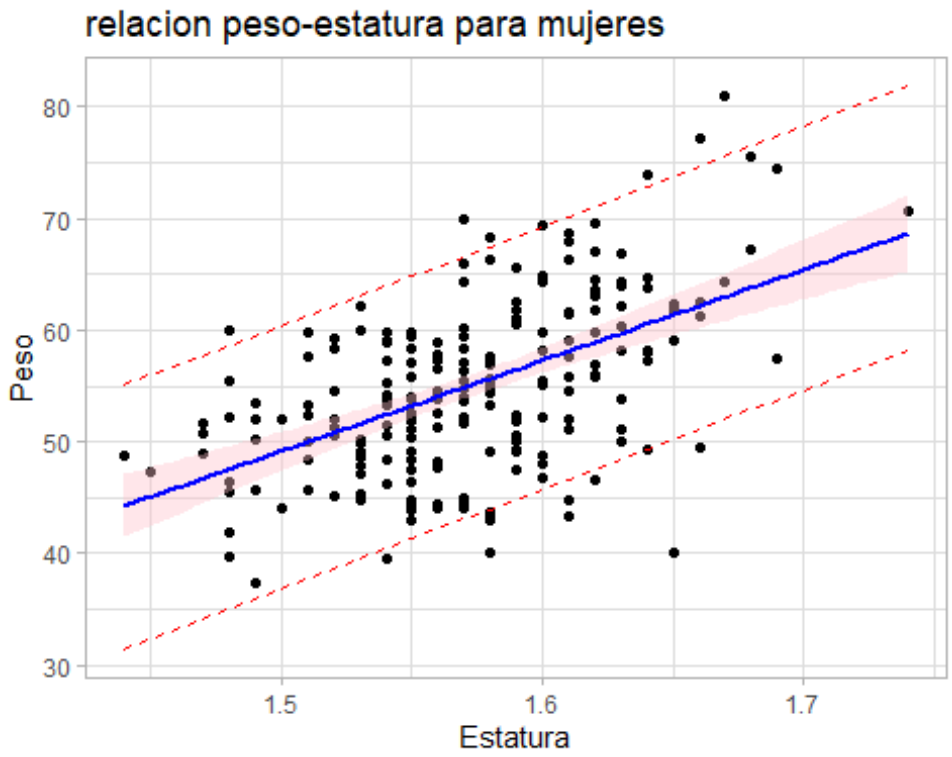
```
qqnorm(A$residuals)
qqline(A$residuals)
```




```
hist(A$residuals, freq = FALSE,ylim = c(0,0.1), xlab = "Residuos", col = 0)
lines(density(A$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)
```

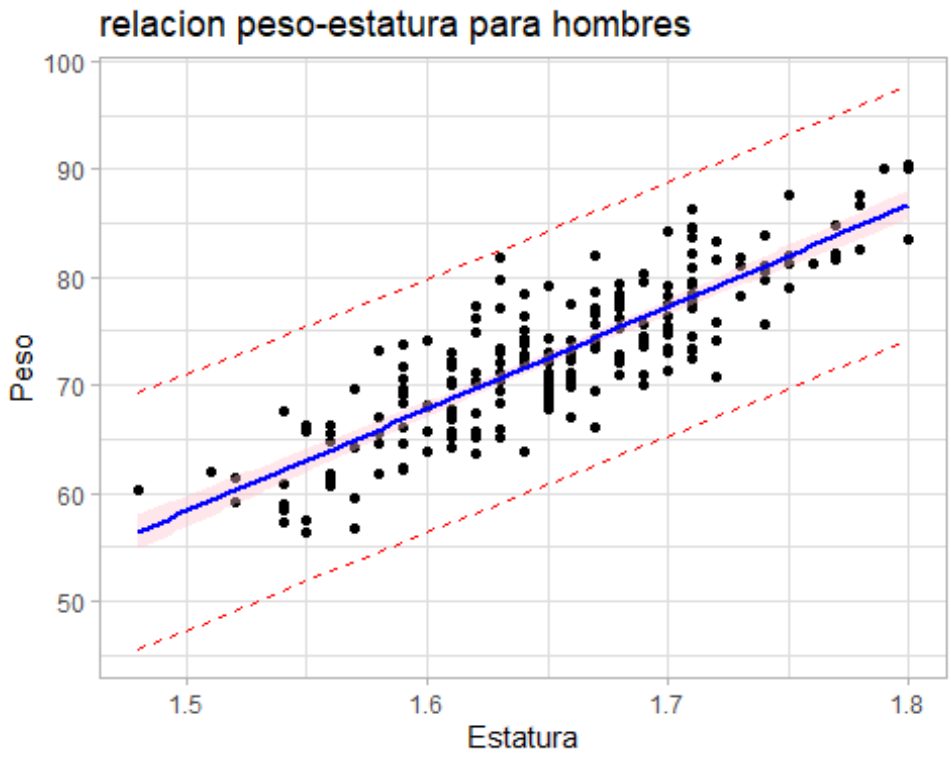



```
Ip = predict(object = A, interval = "prediction", level = 0.97)
M2 = cbind(M,Ip)
M2m = subset(M2, Sexo == "M")
M2h = subset(M2, Sexo == "H")
library(ggplot2)
ggplot(M2m, aes(x=Estatura, y=Peso)) +
  ggtitle("relacion peso-estatura para mujeres")+
  geom_point()+
  geom_line(aes(y=lwr), color = "red", linetype = "dashed") +
  geom_line(aes(y=upr), color = "red", linetype = "dashed") +
  geom_smooth(method=lm, formula=y~x, se=TRUE, level = 0.97, col = "blue",
fill = "pink") +
  theme_light()
```

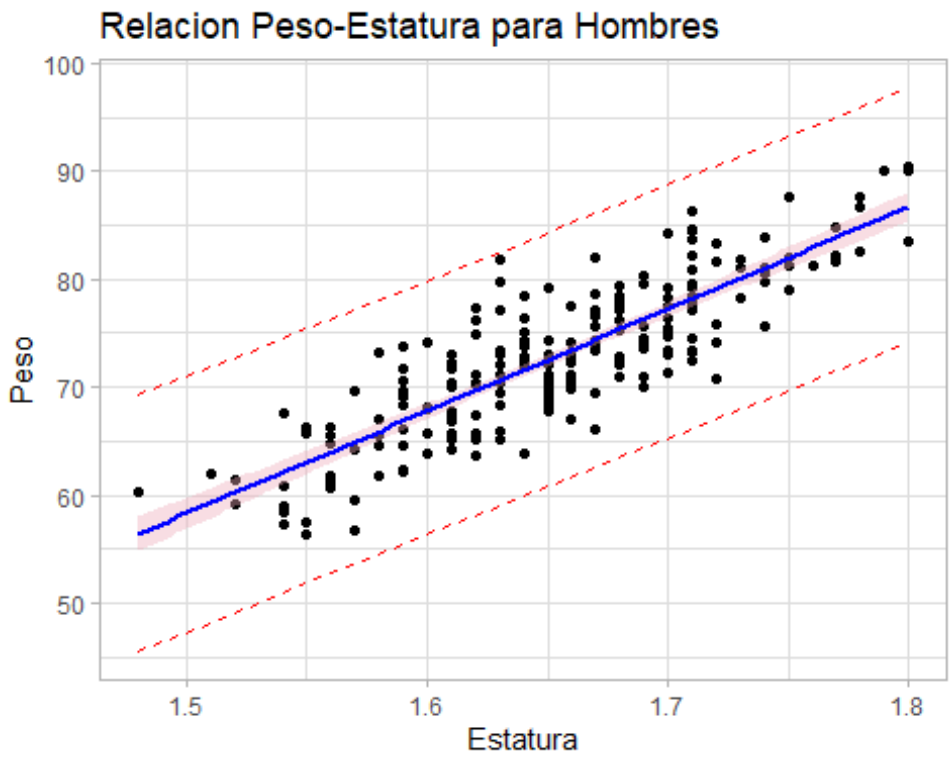


Intervalos de Confianza y prediccion para Y.

```
Ip = predict(object = A, interval = "prediction", level = 0.97)
M2 = cbind(M,Ip)
M2m = subset(M2, Sexo == "M")
M2h = subset(M2, Sexo == "H")
library(ggplot2)
ggplot(M2h, aes(x=Estatura, y=Peso)) +
  ggtitle("relacion peso-estatura para hombres")+
  geom_point()+
  geom_line(aes(y=lwr), color = "red", linetype = "dashed") +
  geom_line(aes(y=upr), color = "red", linetype = "dashed") +
  geom_smooth(method=lm, formula=y~x, se=TRUE, level = 0.97, col = "blue",
fill = "pink") +
  theme_light()
```




```
ggplot(M2h,aes(x=Estatura,y=Peso))+
  ggtitle("Relacion Peso-Estatura para Hombres")+
  geom_point()+
  geom_line(aes(y=lwr), color="red", linetype="dashed")+
  geom_line(aes(y=upr), color="red", linetype="dashed")+
  geom_smooth(method=lm, formula=y~x, se=TRUE, level=0.97, col="blue",
fill="pink2")+
  theme_light()
```


Relacion Peso-Estatura para Hombres
100
90
80
70
60
50
Peso
1.5
1.6
1.7
1.8
Estatura