

## Actividad 10

Saúl Francisco Vázquez del Río

2024-08-30

### Primer problema

Analiza la base de datos de estatura y peso Download datos de estatura y peso de los hombres y mujeres en México y obten el mejor modelo de regresión para esos datos

```
M=read.csv("C:\\Users\\saulv\\OneDrive\\Escritorio\\Septimo semestre\\Estatura-peso_HyM.csv") #Leer la base de datos
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
```

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

### Matriz de correlacion

Obtén la matriz de correlación de los datos que se te proporcionan. Interpreta.

```
cor(M1)

##           MH.Estatura    MH.Peso  MM.Estatura    MM.Peso
## MH.Estatura 1.0000000000 0.846834792 0.0005540612 0.04724872
## MH.Peso      0.8468347920 1.0000000000 0.0035132246 0.02154907
## MM.Estatura 0.0005540612 0.003513225 1.0000000000 0.52449621
## MM.Peso      0.0472487231 0.021549075 0.5244962115 1.00000000
```

La matriz de correlacion establece las relaciones entre peso y estatura de hombres y mujeres.

```
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

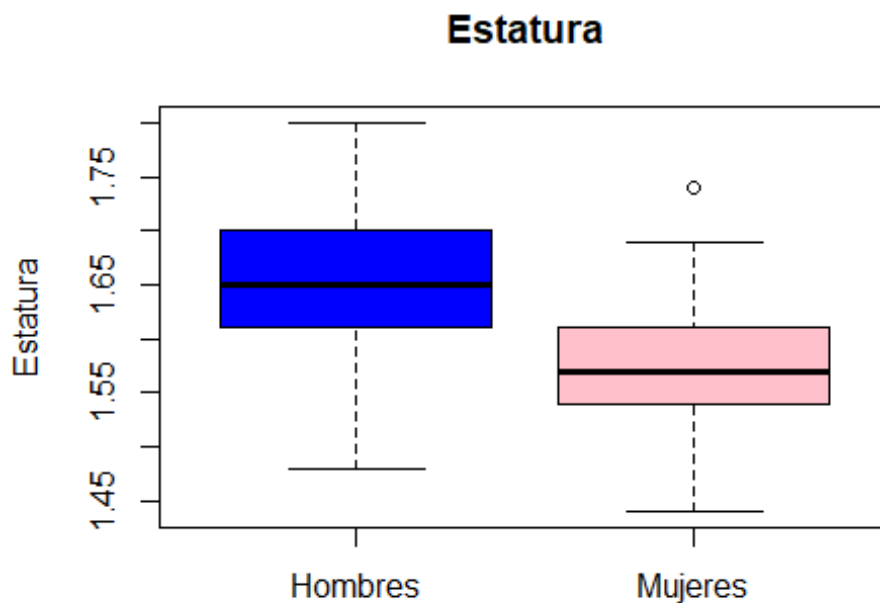
Establecemos la variables como las mediana y la media para entedner mejor los datos

### Box plot

```

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

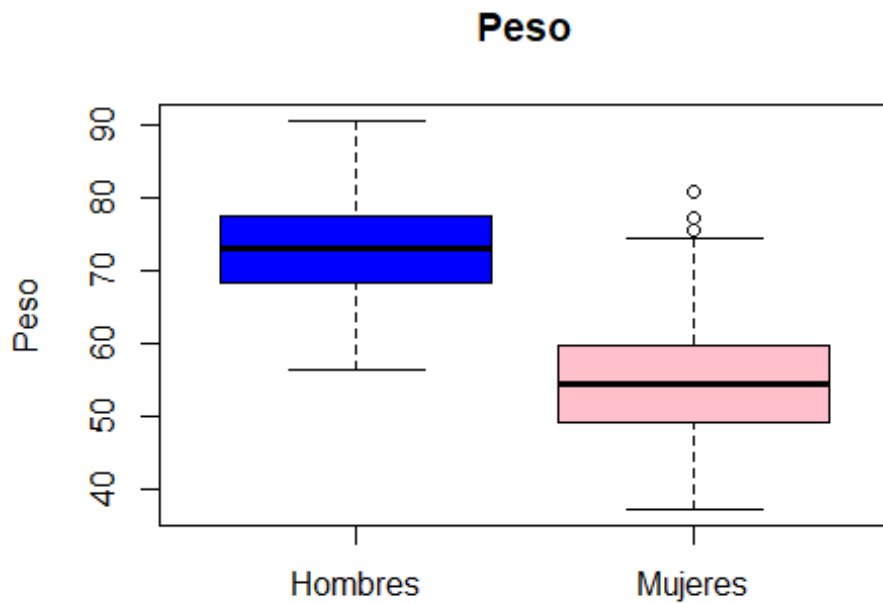
```



```

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

```



En el box plot se muestran la relacion que tienen los hombres y las mujeres con la estatura y con el peso

### Hombres

```
Modelo1H = lm(Peso~Estatura, data = MH)
```

```
Modelo1H
```

```
##
## Call:
## lm(formula = Peso ~ Estatura, data = MH)
##
## Coefficients:
## (Intercept)      Estatura
##      -83.68       94.66
```

```
Modelo1M = lm(Peso~Estatura, data = MM)
```

```
Modelo1M
```

```
##
## Call:
## lm(formula = Peso ~ Estatura, data = MM)
##
## Coefficients:
## (Intercept)      Estatura
##      -72.56       81.15
```

Se hacen los modelos de hombres y mujeres para ver la relacion entre el peso y la estatura

Hipotesis

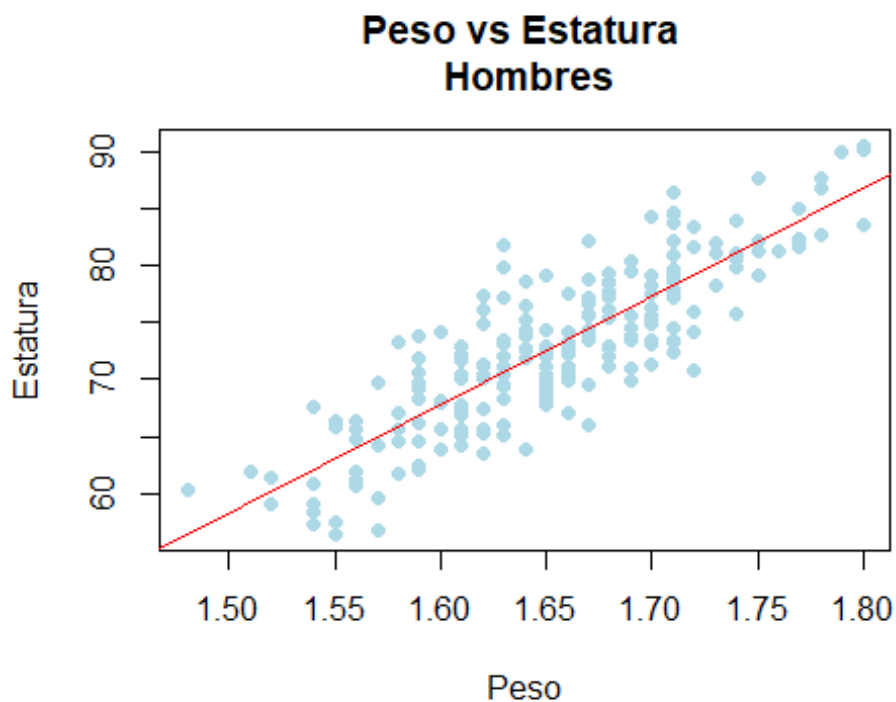
$$H_0: \beta_1 = 0$$

```
summary(Modelo1H)
```

```
##
## Call:
## lm(formula = Peso ~ Estatura, data = MH)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.3881 -2.6073 -0.0665  2.4421 11.1883
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -83.685      6.663   -12.56  <2e-16 ***
## Estatura      94.660      4.027    23.51  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.678 on 218 degrees of freedom
## Multiple R-squared:  0.7171, Adjusted R-squared:  0.7158
## F-statistic: 552.7 on 1 and 218 DF,  p-value: < 2.2e-16
```

### Grafico de dispersion para hombres

```
plot(MH$Estatura, MH$Peso, col='lightblue', main="Peso vs Estatura \n Hom
bres", ylab="Estatura", xlab="Peso", pch = 19 )
abline(Modelo1H, col="red", lwd=1.5)
```



La grafica muestra la relacion entre el peso y la estatura de los hombres

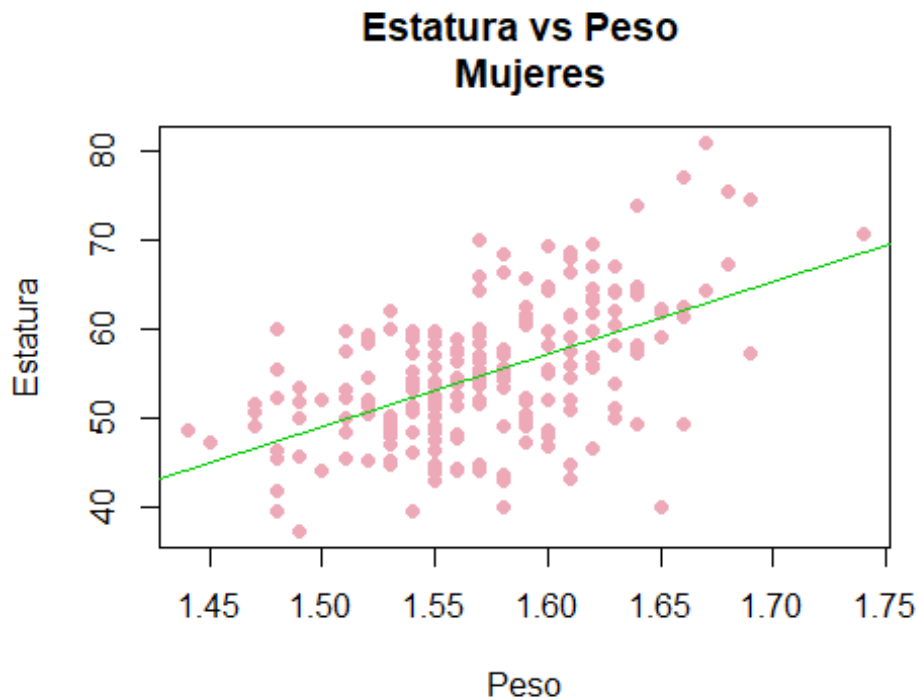
### Mujeres

`summary(Modelo1M)`

```
##
## Call:
## lm(formula = Peso ~ Estatura, data = MM)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.3256  -4.1942   0.4004   4.2724  17.9114
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -72.560     14.041  -5.168 5.34e-07 ***
## Estatura      81.149       8.922   9.096 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.65 on 218 degrees of freedom
## Multiple R-squared:  0.2751, Adjusted R-squared:  0.2718
## F-statistic: 82.73 on 1 and 218 DF, p-value: < 2.2e-16
```

### Grafico para mujeres

```
plot(MM$Estatura, MM$Peso, col='pink2', main="Estatura vs Peso \n Mujeres", ylab="Estatura", xlab="Peso", pch = 19 )  
abline(Modelo1M, col="Green3", lwd=1.5)
```



La grafica muestra la relacion entre el peso y la estatura de las mujeres

### Un solo modelo

```
Modelo2 = lm(Peso~Estatura + Sexo, M)  
Modelo2  
  
##  
## Call:  
## lm(formula = Peso ~ Estatura + Sexo, data = M)  
##  
## Coefficients:  
## (Intercept)      Estatura      SexoM  
##      -74.75         89.26      -10.56
```

```
summary(Modelo2)  
  
##  
## Call:  
## lm(formula = Peso ~ Estatura + Sexo, data = M)  
##  
## 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 ***
## Estatura 89.2604 4.5635 19.560 <2e-16 ***
## 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
```

Se crea un modelo en el cual se pueden ver las diferencias del peso y la estatura entre hombres y mujeres

A 0.05 sí es significativo y los modelos quedarían

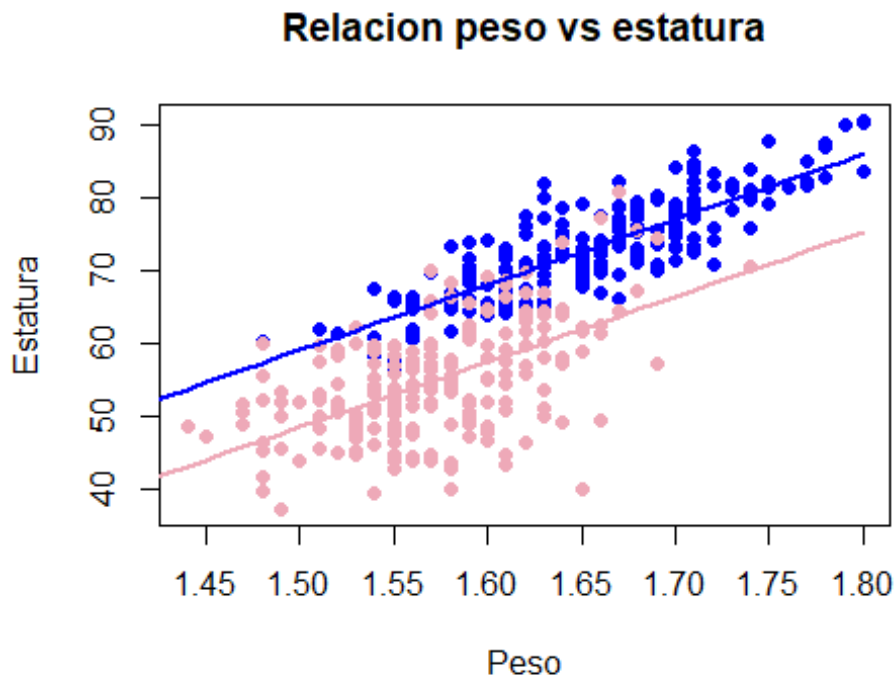
Hombres:  $\text{Estatura} = 1.3862 + 0.00399P$   
 $\text{Estatura} = 1.2727097 + 0.0052296P + 0.0121799$

Mujeres:  $\text{Estatura} = 1.3862 + 0.00399P$   
 $\text{Estatura} = 1.2727097 + 0.0052296P + 0.0121799$

```
b0 = Modelo2$coefficients[1]
b1 = Modelo2$coefficients[2]
b2 = Modelo2$coefficients[3]

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

colores = c("blue", "pink2")
plot(M$Estatura, M$Peso, col=colores[factor(M$Sexo)], pch=19, ylab="Estatura",
     xlab="Peso", main="Relacion peso vs estatura")
x = seq(1.40, 1.80, 0.01)
lines(x, Ym(x), col = "pink2", lwd = 2)
lines(x, Yh(x), col="blue", lwd = 2)
```



La grafica

muestra las diferencias de peso y estatura de hombres y mujeres

Interpreta en el contexto del problema: ¿Qué información proporciona  $\hat{\beta}_0$  sobre la relación entre la estatura y el peso de hombres y mujeres? Los modelos presentan la relación que los hombres y las mujeres tienen con su peso y estatura, además que se nota un control de peso por la estatura.

¿Cómo interpretas  $\hat{\beta}_1$  en la relación entre la estatura y el peso de hombres y mujeres? La relación se puede interpretar que esta varía dependiendo del sexo que se está tratando ya que los hombres tienden a ser más altos lo que conlleva a que estos pesen más, al igual en el lado de las mujeres estas tienden a ser más bajas y por esto tienden a pesar menos.

#### Otro Modelo (segunda entrega)

```
Modelo3 = lm(Peso ~ Estatura * Sexo, data = M)
Modelo3
```

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

Formula de hombres =  $-83.68454 + 11.12409 + 94.66024$  Formula de mujeres =  $-83.68454 + 94.66024$



```
summary(Modelo3)

##
## Call:
## lm(formula = Peso ~ Estatura * Sexo, data = M)
##
## 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 ***
## Estatura       94.660      5.882   16.092  <2e-16 ***
## SexoM          11.124     14.950    0.744    0.457
## Estatura: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

anova(Modelo3)

## Analysis of Variance Table
##
## Response: Peso
##              Df Sum Sq Mean Sq    F value Pr(>F)
## Estatura      1  37731   37731  1306.5938 <2e-16 ***
## Sexo          1   8097    8097   280.3892 <2e-16 ***
## Estatura:Sexo  1     61      61    2.1085 0.1472
## Residuals    436  12590      29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

rsq_adj <- summary(Modelo3)$adj.r.squared
print(paste("R-squared ajustado: ", rsq_adj))

## [1] "R-squared ajustado: 0.783219679867522"

# Asignar Los coeficientes del modelo con interacción
b0_A <- Modelo3$coefficients[1]
b1_A <- Modelo3$coefficients[2]
b2_A <- Modelo3$coefficients[3]
b3_A <- Modelo3$coefficients[4]

# Imprimir Los coeficientes
print(b0_A)
```

```

## (Intercept)
## -83.68454

print(b1_A)

## Estatura
## 94.66024

print(b2_A)

## SexoM
## 11.12409

print(b3_A)

## Estatura:SexoM
## -13.51113

# Funciones para Las líneas de regresión
Ym <- function(x) { b0_A + b2_A + (b1_A + b3_A) * x } # Línea para mujeres
Yh <- function(x) { b0_A + b1_A * x } # Línea para hombres

# Colores para Los puntos y Las líneas
colores <- c("#66BD63", "#FDAE61")

# Crear La gráfica de dispersión
plot(M$Estatura, M$Peso, col=colores[factor(M$Sexo)], pch=19, ylab="Peso",
, xlab="Estatura", main="Relación entre estatura y peso")

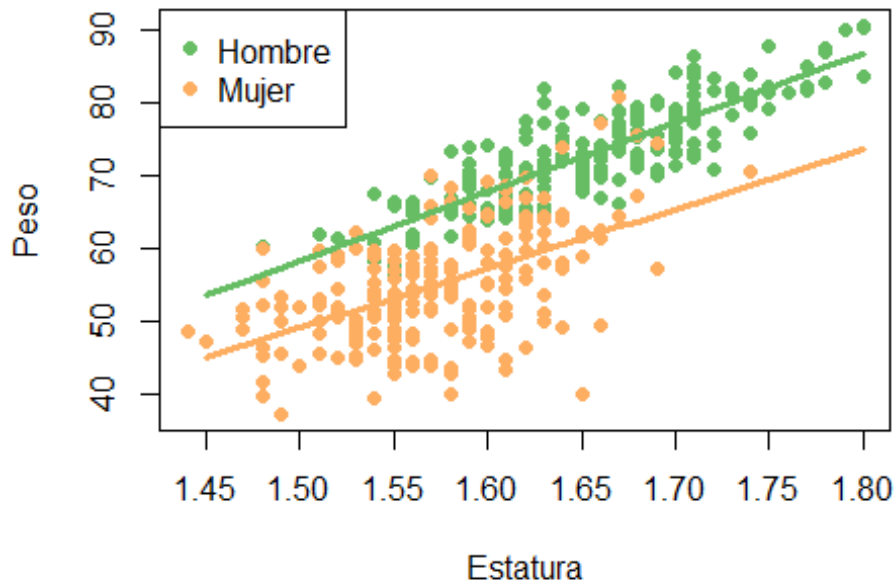
# Generar secuencia de valores para Las líneas de regresión
x <- seq(1.45, 1.80, 0.01)

# Añadir Las líneas de regresión al gráfico
lines(x, Ym(x), col="#FDAE61", lwd=3) # Línea para mujeres
lines(x, Yh(x), col="#66BD63", lwd=3) # Línea para hombres

# Añadir Leyenda al gráfico
legend("topleft", legend=c("Hombre", "Mujer"), pch=19, col=c("#66BD63", "#FDAE61"))

```

## Relación entre estatura y peso



B\_0: respresenta el promedio del peso cuando la estarua de los hombres 0, siendo esto el punto de inicio / referencia para emprezar a comprar las diferencias de sexo.

B\_1: Muestra el cambio de peso en los hombres

B\_2: Muestra el cambio de peso en las mujeres, pero comparado con el de los hombres

B\_3: Muestra la relacion entre la estatura de las mujeres comparado con la de los hombres

Para concluir el modelo3 no aporta un valor adicional a la relacion entre estatura y peso, ya que tiene falta de significancia ademas de una complejidad inecesaria para la interpretacion de los datos.