

Regresion_Lineal

Arturo

2023-08-29

Autor: Arturo Garza Campuzano

Matrícula: A00828096

Regresión lineal

1. Recta de mejor ajuste

Instalación y carga de paquetes

```
if (!require(nortest) || !require(tinytex)) {  
  install.packages('nortest')  
  install.packages("tinytex")  
}
```

```
## Loading required package: nortest
```

```
## Loading required package: tinytex
```

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

```
M = read.csv("Estatura-peso_HyM.csv")  
MM = subset(M,M$Sexo=="M")  
MH = subset(M,M$Sexo=="H")  
M1=data.frame(MH$Estatura,MH$Peso,MM$Estatura,MM$Peso)  
cor(M1)
```

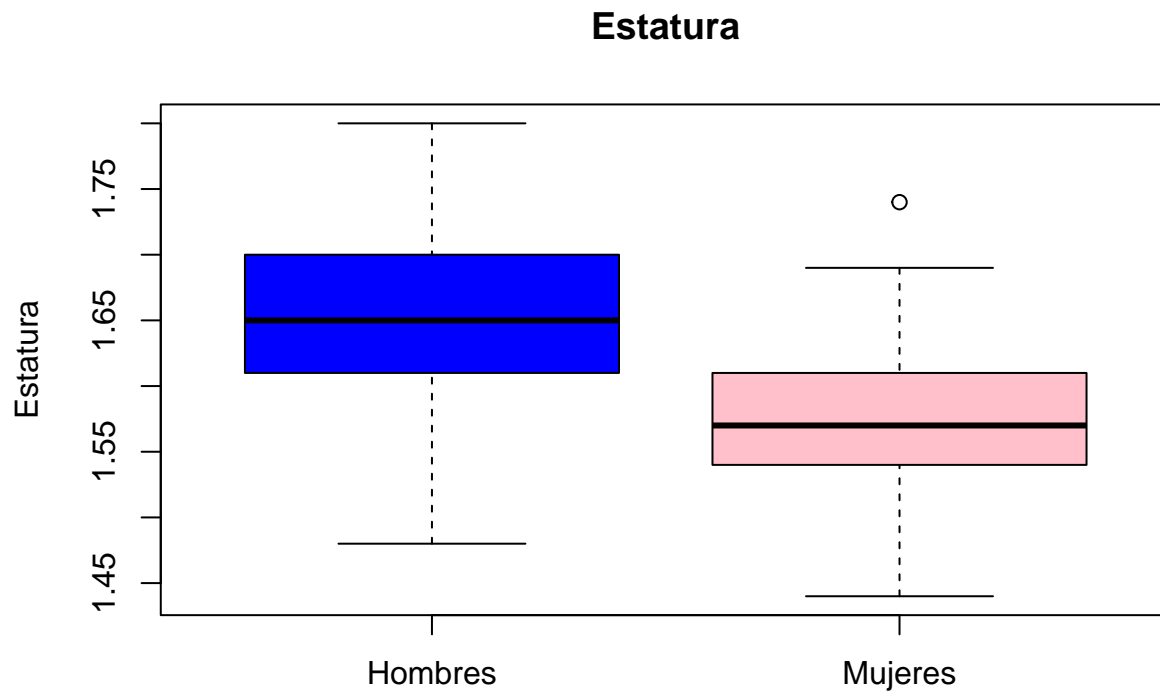
```
##           MH.Estatura    MH.Peso  MM.Estatura    MM.Peso  
## MH.Estatura 1.0000000000 0.846834792 0.0005540612 0.04724872  
## MH.Peso      0.846834792 1.000000000 0.0035132246 0.02154907  
## MM.Estatura 0.0005540612 0.003513225 1.0000000000 0.52449621  
## MM.Peso      0.0472487231 0.021549075 0.5244962115 1.00000000
```

De acuerdo a las aproximaciones de Cohen, el coeficiente de correlación de la muestra para hombres corresponde a una correlación **muy fuerte** entre las variables estatura y peso, considerando el valor del coeficiente como **0.85**.

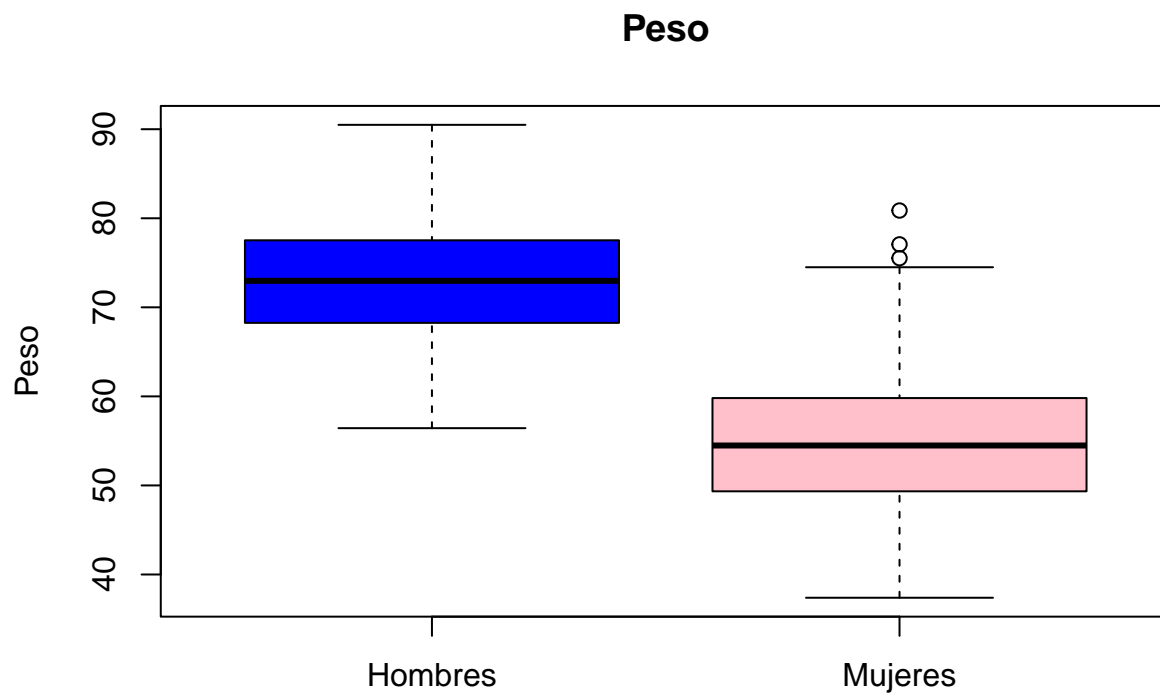
Por otro lado, el coeficiente de correlación de la muestra para mujeres corresponde a una correlación **fuerte** entre las variables estatura y peso, considerando el valor del coeficiente como **0.52**.

1.2 Obtén medidas (media, desviación estándar, etc) que te ayuden a analizar los datos.

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



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



```
n=4 # Numero 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("Mínimo", "Q1", "Mediana", "Media", "Q3", "Máximo", "DesvEst")
m
```

##		Minimo	Q1	Mediana	Media	Q3	Máximo	DesvEst
##	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

Algunas observaciones de los estadísticos presentados:

- Los hombres tienden a estaturas ligeramente más altas que las mujeres, como se puede observar en las medias y medianas de ambas categorías.
- Los hombres tienen un rango mayor de peso, como se puede ver por la diferencia entre los valores máximo y mínimo.
- Las mujeres tienden a menos variabilidad en cuanto a estatura, pero más variabilidad en cuanto a peso. Esto se puede contemplar en el estadístico de desviación estándar.

1.3. Encuentra la ecuación de regresión de mejor ajuste.

1.3.1 Realiza la regresión entre las variables involucradas.

Modelo A: Modelo con 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 = A$coefficients[1]
b1_a = A$coefficients[2]
b2_a = A$coefficients[3]
```

```
cat("Peso =", b0_a, "+", b1_a, "Estatura", b2_a, "SexoM")
```

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

```
# Ecuación del modelo
```

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

```
## Para mujeres
```

```
cat("Peso =", b0_a+b2_a, "+", b1_a, "Estatura", "\n")
```

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

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

```
## Para hombres
```

```
cat("Peso =", b0_a, "+", b1_a, "Estatura", "\n")
```

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

Modelo B: Modelo con Interacción

```
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$SexoM
##          -83.68             94.66             11.12             -13.51
b0_b = B$coefficients[1]
b1_b = B$coefficients[2]
b2_b = B$coefficients[3]
```

1.3.2 Verifica el modelo.

Modelo A: Modelo con 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
```

Significancia global

Paso 1. Definir las hipótesis

Se establecen la hipótesis nula y la alternativa:

- $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$
- $H_1 : \text{al menos una } \beta_k \neq 0$

Paso 2. Regla de decisión

Considerando que $\alpha = 0.03$ entonces la hipótesis nula se rechaza:

- Si F^* es mayor a $|F_0|$
- Si valor $p < \alpha$

Paso 3. Análisis de resultado

Gracias al resumen del modelo se tiene F^* , el valor de p y los grados de libertad. Solo falta calcular F_0 .

```
alfa <- 0.03
k <- 2
grados_libertad_ma <- 437

F0 <- qf(1 - alfa, k, grados_libertad_ma)
print(F0)

## [1] 3.534846
```

Paso 4. Conclusiones

Después de haber realizados los cálculos se pueden hacer las siguientes observaciones:

- Como F^* (791.5) es mayor a F_0 (3.53) entonces se rechaza H_0 .
- Como el valor p (2.2e-16) es menor que alfa (0.03) entonces se rechaza H_0 .

Por lo tanto, el modelo lineal entre la variable respuesta (M\$Peso) y alguna de las variables independientes sí explica a Y. Se cumple que para al menos una de las X:

$$Y = \hat{\beta}_0 + \hat{\beta}_i X + e$$

Entonces sí existe una relación significativa entre la variable Y y al menos una de las variables X_i , lo cual parece indicar que la regresión lineal puede ser apropiada para el modelo.

Significancia individual

Paso 1. Definir las hipótesis

Se establecen la hipótesis nula y la alternativa:

- $H_0 : \beta_i = 0$
- $H_1 : \beta_i \neq 0$

Paso 2. Regla de decisión

Considerando que $\alpha = 0.03$ entonces la hipótesis nula se rechaza:

- Si t^* es mayor a $|t_0|$
- Si valor p $< \alpha$

Paso 3. Análisis de resultado

Gracias al resumen del modelo se tienen todas las t^* , el valor de p de cada una y los grados de libertad. Solo falta calcular t_0 .

```
n <- grados_libertad_ma + k + 1

t0 <- qt(1 - alfa/2, df = n - 2)

print(t0)

## [1] 2.177184
```

Paso 4. Conclusiones

Después de haber realizados los cálculos se pueden hacer las siguientes observaciones:

- Para la estatura, como t^* es mayor a t_0 entonces se rechaza H_0 . En los betas restantes no se rechaza H_0 .
- Como el valor p (2.2e-16) es menor que alfa (0.03) para todos los betas entonces se rechaza H_0 .

Por lo tanto, hay relación estadística entre la variable respuesta y las variables significativas. Además, se puede observar que $\Pr(>|t|)$ para todos los beta es cercano a cero. Todo esto indica al parecer la regresión lineal se ajusta a este caso.

Coefficiente de detrmínación

Este valor corresponde a r cuadrada ajustado del resumen del modelo A, ya que se esta trabajando con más de una variable. Su valor es de 0.7827, considerablemente cerca del uno.

Modelo B: Modelo con Interacción

```
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
```

Significancia global

Paso 1. Definir las hipótesis

Se establecen la hipótesis nula y la alternativa:

- $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$
- $H_1 : \text{al menos una } \beta_k \neq 0$

Paso 2. Regla de decisión

Considerando que $\alpha = 0.03$ entonces la hipótesis nula se rechaza:

- Si F^* es mayor a $|F_0|$
- Si valor $p < \alpha$

Paso 3. Análisis de resultado

Gracias al resumen del modelo se tiene F^* , el valor de p y los grados de libertad. También F_0 ya se calculó al hacer la prueba de hipótesis de la significancia global del modelo anterior.

Paso 4. Conclusiones

Se rechaza a H_0 ya que se cumplen las dos reglas de decisión. Se observa además que el modelo A está muchísimo más alejado del valor frontera, lo cual puede indicar que el modelo A es más apropiado en este caso.

Significancia individual

Paso 1. Definir las hipótesis

Se establecen la hipótesis nula y la alternativa:

- $H_0 : \beta_i = 0$
- $H_1 : \beta_i \neq 0$

Paso 2. Regla de decisión

Considerando que $\alpha = 0.03$ entonces la hipótesis nula se rechaza:

- Si t^* es mayor a $|t_0|$
- Si valor $p < \alpha$

Paso 3. Análisis de resultado

Gracias al resumen del modelo se tienen todas las t^* , el valor de p de cada una y los grados de libertad. También t_0 ya se calculó al hacer la prueba de hipótesis de la significancia individual del modelo anterior.

Paso 4. Conclusiones

Después de haber realizados los cálculos se pueden hacer las siguientes observaciones:

- Para la estatura, como t^* es mayor a t_0 entonces se rechaza H_0 . En los betas restantes no se rechaza H_0 .
- Tanto para la intersección como la estatura, como el valor p ($2.2e-16$) es menor que alfa (0.03) los betas entonces se rechaza H_0 . No obstante, para las betas restantes H_0 no se rechaza.

Por lo tanto, hay relación estadística entre la variable respuesta y las variables significativas. Sin embargo, se puede observar que $\Pr(>|t|)$ para algunos beta rebasa el alfa, lo cual parece ser señal de que la regresión lineal no se ajusta a este caso.

Coefficiente de determinación

Este valor corresponde a r cuadrada ajustado del resumen del modelo B, ya que se está trabajando con más de una variable. Su valor es de 0.7832, considerablemente cerca del uno.

¿Qué modelo se debería utilizar?

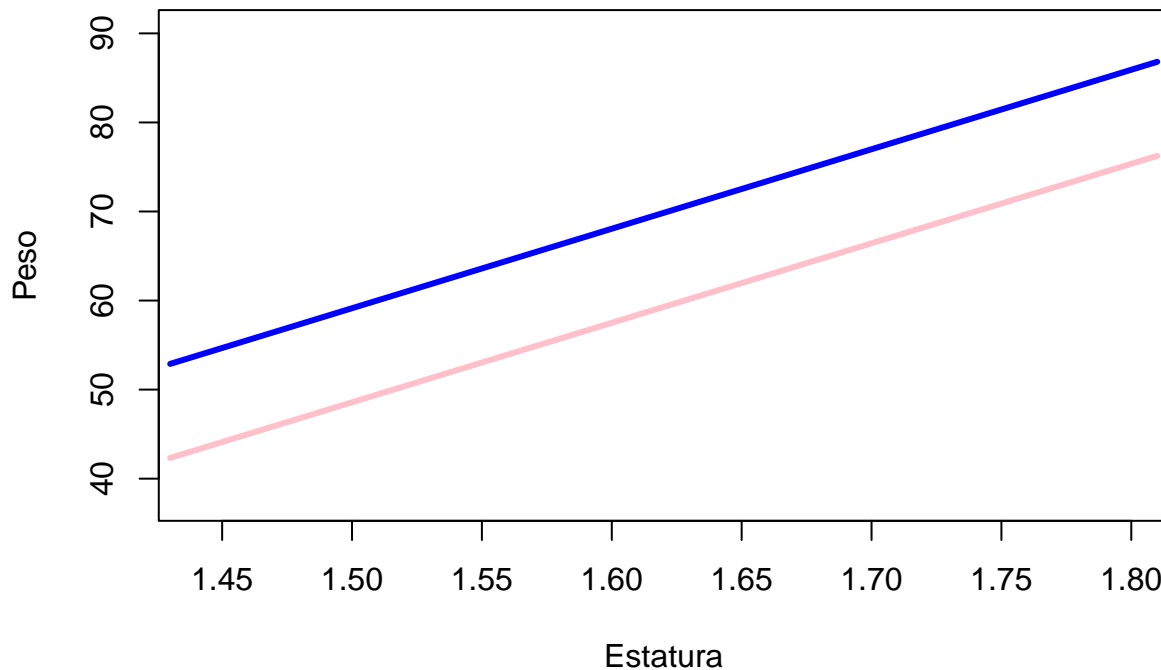
El modelo A, ya que de su verificación se puede sugerir con más firmeza que la regresión lineal es el modelo más adecuado.

1.4 Dibuja el diagrama de dispersión de los datos y la recta de mejor ajuste.

```
Ym = function(x){b0_a+b2_a+b1_a*x}
Yh = function(x){b0_a+b1_a*x}

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

Relación de Peso vs. Estatura



1.5 Interpreta en el contexto del problema

1.5.1 ¿Qué información proporciona $\hat{\beta}_0$ sobre la relación entre la estatura y el peso de hombres y mujeres?

El coeficiente $\hat{\beta}_0$ es la intersección de la línea de regresión con el eje vertical. En el contexto del problema, $\hat{\beta}_0$ proporciona información sobre el valor del peso esperado cuando la estatura es cero. Sin embargo, en este contexto, no tiene un significado debido a que no tiene sentido que un individuo tenga una estatura de cero. Por lo tanto, la información proporcionada por este coeficiente no es relevante.

1.5.2 ¿Cómo interpretas $\hat{\beta}_1$ en la relación entre la estatura y el peso de hombres y mujeres?

El coeficiente $\hat{\beta}_1$ es la pendiente de la línea de regresión. En el contexto del problema, $\hat{\beta}_1$ indica la dirección y magnitud del cambio en el peso en respuesta a un cambio en la estatura. Como se puede observar, este cambio es el mismo para los hombres que para las mujeres.

2. Valida el modelo

2.1 Analiza si el (los) modelo(s) obtenidos son apropiados para el conjunto de datos. Realiza el análisis de los residuos:

2.1.1 Normalidad de los residuos

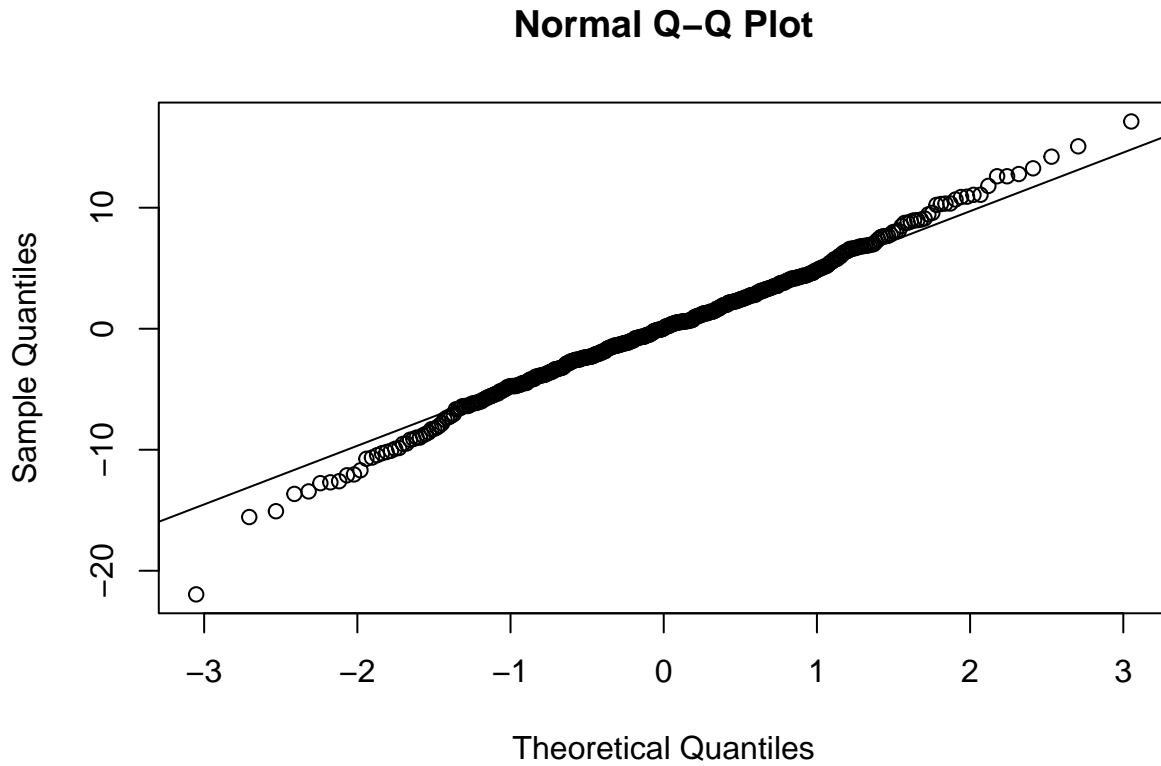
Prueba de normalidad Anderson - Darling

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

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


QQ plot

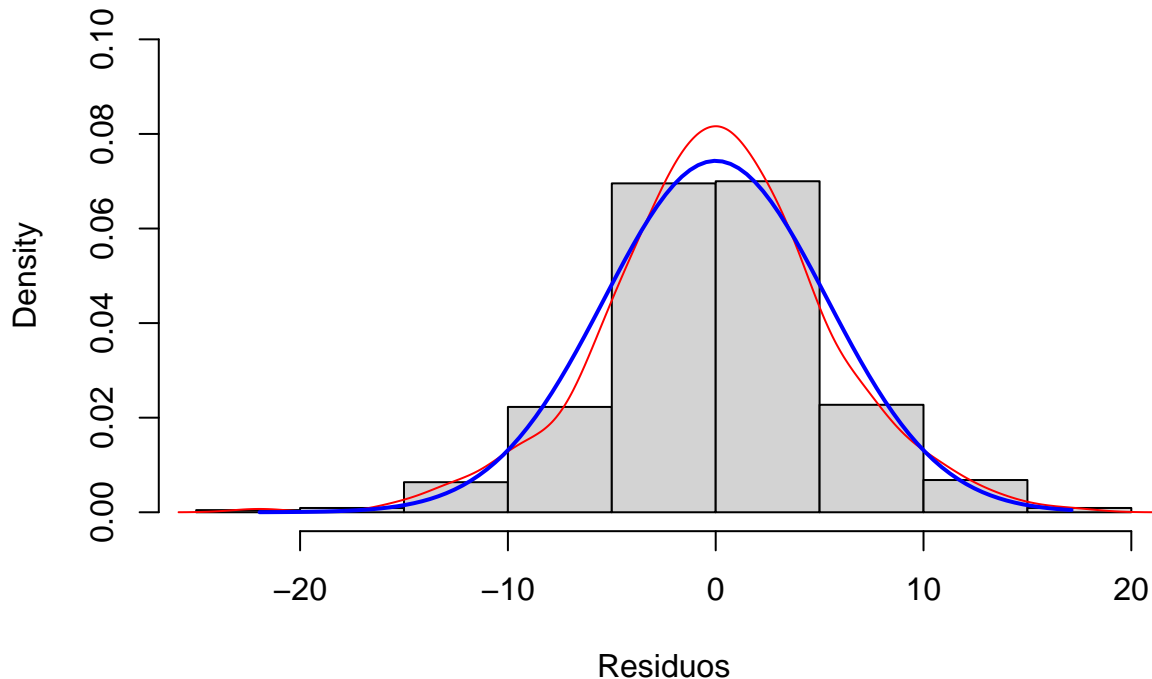
```
qqnorm(A$residuals)
qqline(A$residuals)
```



Histograma de Residuos

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

Histograma de Residuos



Interpretación en el contexto del problema:

- **Prueba de normalidad Anderson - Daling:** Considerando que alfa tiene un valor de 0.03 entonces se puede afirmar que los datos provienen de una población normal; la hipótesis nula no es rechazada.
- **QQ plot:** Como se puede contemplar en la gráfica de normalidad, la relación del porcentaje acumulado y la X se acerca a una línea recta; lo cual sugiere que hay normalidad.
- **Histograma de Residuos:** La distribución de los residuos se asemeja a una distribución normal, contando con una forma ligeramente leptocúrtica. Por lo tanto, a un nivel práctico se puede asumir normalidad.

Conclusión: Considerando las tres pruebas de normalidad que se realizaron, la distribución de los residuos se puede asumir como normal.

2.1.2 Verificación de media cero

Paso 1. Definir las hipótesis.

$$H_0 : \mu_e = 0$$

$$H_1 : \mu_e \neq 0$$

Estimador: \bar{x}_e

$$\text{Estadístico de prueba: } t^* = \frac{\bar{x}_e - \mu_e}{s_{\bar{x}_e}}$$

Distribución del estadístico: t de Student con $gl = n - 1$

Paso 2. Regla de decisión.

Nivel de confianza = 0.97

$$\alpha = 0.03$$

```
alfa <- 0.03
n <- length(A$residuals)
```

```
t0 <- qt(alfa/2,n-1)
cat("t0 =",t0)
```

```
## t0 = -2.177168
```

H_0 se rechaza si:

- $|t^*| > 2.18$
- valor $p < 0.03$

Paso 3. Análisis de resultado.

```
result <- t.test(A$residuals, alternative = "two.sided", mu = 0, conf.level = 0.97)
print(result)
```

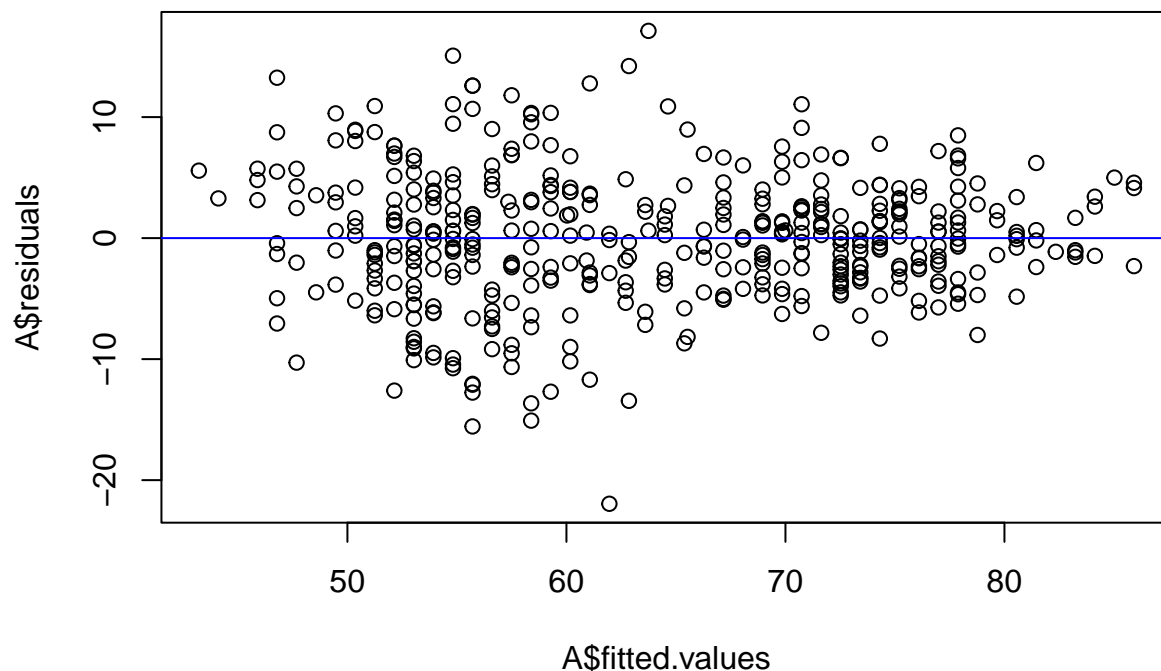
```
##
## One Sample t-test
##
## data: A$residuals
## t = 6.941e-16, df = 439, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 97 percent confidence interval:
## -0.5571867 0.5571867
## sample estimates:
## mean of x
## 1.776357e-16
```

Paso 4. Conclusiones.

Interpretación en el contexto del problema: Como la regla clásica y la regla de valor p no se cumplen entonces no se rechaza la hipótesis nula. Es decir, la media de los residuos es de cero.

2.1.3 Homocedasticidad e independencia

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



Interpretación en el contexto del problema: Como se puede observar en la gráfica, Los residuos cuentan una simetría y homocedasticidad.

3. Conclusión final sobre el análisis de regresión lineal

La ecuación de regresión de mejor ajuste que mejor describe el comportamiento de los datos es la ecuación proporcionada por el Modelo A (Modelo con Sexo):

- Para mujeres: $y = -85.32 + 89.26x$
- Para hombres: $y = -74.75 + 89.26x$

Este modelo fue elegido después de haber sido verificado con la significancia global, significancia individual y el coeficiente de determinación.

Después se realizó la validez del modelo que consistió en un análisis de residuos cuyos resultados fueron:

- Los residuos se distribuyen como una normal.
- La media de los residuos no es significativamente diferente de cero.
- Tiene la misma varianza en todos los valores del dominio de x.

4. Intervalos de confianza y predicción

4.1 Estimación por intervalo de beta 1

```
confint(A, level= 1-alfa)
```

```
##              1.5 %      98.5 %
## (Intercept) -91.20451 -58.304689
## M$Estatura   79.32465  99.196052
## M$SexoM      -11.93983  -9.189113
```

Interpretación: esto significa que con un nivel de confianza del 98.5% se puede estar seguro de que el verdadero valor del coeficiente de estatura caerá dentro del rango establecido.

4.2 Estimación por intervalo y

```
xe= data.frame(x = c(1.65, 1.57, 1.61, 1.74, 1.80, 1.73))
predict(object=A, newdata=xe, interval="prediction", level=1-alfa)
```

```
##      fit      lwr      upr
## 1  68.95457 57.20540 80.70374
## 2  68.95457 57.20540 80.70374
## 3  76.98800 65.23787 88.73813
## 4  72.52498 60.78380 84.26617
## 5  78.77321 67.01363 90.53279
## 6  70.73978 58.99628 82.48327
## 7  82.34362 70.55511 94.13214
## 8  74.31019 62.56795 86.05243
## 9  74.31019 62.56795 86.05243
## 10 72.52498 60.78380 84.26617
## 11 70.73978 58.99628 82.48327
## 12 76.98800 65.23787 88.73813
## 13 76.09540 64.34874 87.84206
## 14 67.16936 55.41117 78.92755
## 15 77.88061 66.12617 89.63504
## 16 73.41759 61.67629 85.15888
## 17 72.52498 60.78380 84.26617
```

18 67.16936 55.41117 78.92755
19 67.16936 55.41117 78.92755
20 74.31019 62.56795 86.05243
21 77.88061 66.12617 89.63504
22 75.20279 63.45876 86.94683
23 67.16936 55.41117 78.92755
24 76.98800 65.23787 88.73813
25 75.20279 63.45876 86.94683
26 68.95457 57.20540 80.70374
27 76.98800 65.23787 88.73813
28 76.98800 65.23787 88.73813
29 70.73978 58.99628 82.48327
30 78.77321 67.01363 90.53279
31 69.84717 58.10126 81.59308
32 76.09540 64.34874 87.84206
33 66.27676 54.51280 78.04072
34 75.20279 63.45876 86.94683
35 69.84717 58.10126 81.59308
36 72.52498 60.78380 84.26617
37 66.27676 54.51280 78.04072
38 75.20279 63.45876 86.94683
39 71.63238 59.89046 83.37430
40 80.55842 68.78604 92.33079
41 70.73978 58.99628 82.48327
42 68.06197 56.30871 79.81523
43 68.95457 57.20540 80.70374
44 72.52498 60.78380 84.26617
45 76.09540 64.34874 87.84206
46 74.31019 62.56795 86.05243
47 68.95457 57.20540 80.70374
48 64.49155 52.71355 76.26955
49 70.73978 58.99628 82.48327
50 67.16936 55.41117 78.92755
51 66.27676 54.51280 78.04072
52 73.41759 61.67629 85.15888
53 63.59895 51.81267 75.38522
54 77.88061 66.12617 89.63504
55 72.52498 60.78380 84.26617
56 72.52498 60.78380 84.26617
57 77.88061 66.12617 89.63504
58 85.91404 74.08330 97.74477
59 67.16936 55.41117 78.92755
60 64.49155 52.71355 76.26955
61 64.49155 52.71355 76.26955
62 77.88061 66.12617 89.63504
63 83.23623 71.43840 95.03405
64 78.77321 67.01363 90.53279
65 75.20279 63.45876 86.94683
66 71.63238 59.89046 83.37430
67 77.88061 66.12617 89.63504
68 77.88061 66.12617 89.63504
69 63.59895 51.81267 75.38522
70 77.88061 66.12617 89.63504
71 70.73978 58.99628 82.48327

72 73.41759 61.67629 85.15888
73 70.73978 58.99628 82.48327
74 80.55842 68.78604 92.33079
75 62.70635 50.91097 74.50172
76 73.41759 61.67629 85.15888
77 75.20279 63.45876 86.94683
78 77.88061 66.12617 89.63504
79 76.98800 65.23787 88.73813
80 67.16936 55.41117 78.92755
81 64.49155 52.71355 76.26955
82 71.63238 59.89046 83.37430
83 81.45102 69.67099 93.23105
84 83.23623 71.43840 95.03405
85 76.98800 65.23787 88.73813
86 68.06197 56.30871 79.81523
87 62.70635 50.91097 74.50172
88 75.20279 63.45876 86.94683
89 67.16936 55.41117 78.92755
90 72.52498 60.78380 84.26617
91 76.98800 65.23787 88.73813
92 70.73978 58.99628 82.48327
93 77.88061 66.12617 89.63504
94 77.88061 66.12617 89.63504
95 68.95457 57.20540 80.70374
96 76.09540 64.34874 87.84206
97 78.77321 67.01363 90.53279
98 69.84717 58.10126 81.59308
99 79.66581 67.90025 91.43137
100 71.63238 59.89046 83.37430
101 74.31019 62.56795 86.05243
102 77.88061 66.12617 89.63504
103 72.52498 60.78380 84.26617
104 74.31019 62.56795 86.05243
105 69.84717 58.10126 81.59308
106 71.63238 59.89046 83.37430
107 80.55842 68.78604 92.33079
108 67.16936 55.41117 78.92755
109 83.23623 71.43840 95.03405
110 65.38416 53.61359 77.15472
111 72.52498 60.78380 84.26617
112 69.84717 58.10126 81.59308
113 75.20279 63.45876 86.94683
114 73.41759 61.67629 85.15888
115 68.95457 57.20540 80.70374
116 75.20279 63.45876 86.94683
117 76.98800 65.23787 88.73813
118 77.88061 66.12617 89.63504
119 68.06197 56.30871 79.81523
120 76.09540 64.34874 87.84206
121 72.52498 60.78380 84.26617
122 69.84717 58.10126 81.59308
123 68.95457 57.20540 80.70374
124 73.41759 61.67629 85.15888
125 77.88061 66.12617 89.63504

126 69.84717 58.10126 81.59308
127 69.84717 58.10126 81.59308
128 75.20279 63.45876 86.94683
129 64.49155 52.71355 76.26955
130 63.59895 51.81267 75.38522
131 69.84717 58.10126 81.59308
132 74.31019 62.56795 86.05243
133 66.27676 54.51280 78.04072
134 76.98800 65.23787 88.73813
135 63.59895 51.81267 75.38522
136 65.38416 53.61359 77.15472
137 70.73978 58.99628 82.48327
138 73.41759 61.67629 85.15888
139 81.45102 69.67099 93.23105
140 79.66581 67.90025 91.43137
141 60.92114 49.10507 72.73721
142 84.12883 72.32086 95.93680
143 77.88061 66.12617 89.63504
144 80.55842 68.78604 92.33079
145 76.98800 65.23787 88.73813
146 84.12883 72.32086 95.93680
147 71.63238 59.89046 83.37430
148 76.09540 64.34874 87.84206
149 76.09540 64.34874 87.84206
150 71.63238 59.89046 83.37430
151 74.31019 62.56795 86.05243
152 62.70635 50.91097 74.50172
153 74.31019 62.56795 86.05243
154 65.38416 53.61359 77.15472
155 69.84717 58.10126 81.59308
156 67.16936 55.41117 78.92755
157 76.09540 64.34874 87.84206
158 75.20279 63.45876 86.94683
159 72.52498 60.78380 84.26617
160 76.98800 65.23787 88.73813
161 77.88061 66.12617 89.63504
162 72.52498 60.78380 84.26617
163 69.84717 58.10126 81.59308
164 70.73978 58.99628 82.48327
165 71.63238 59.89046 83.37430
166 70.73978 58.99628 82.48327
167 73.41759 61.67629 85.15888
168 69.84717 58.10126 81.59308
169 75.20279 63.45876 86.94683
170 75.20279 63.45876 86.94683
171 73.41759 61.67629 85.15888
172 68.95457 57.20540 80.70374
173 57.35072 45.48339 69.21806
174 72.52498 60.78380 84.26617
175 83.23623 71.43840 95.03405
176 73.41759 61.67629 85.15888
177 68.06197 56.30871 79.81523
178 74.31019 62.56795 86.05243
179 68.95457 57.20540 80.70374

180 73.41759 61.67629 85.15888
181 68.06197 56.30871 79.81523
182 74.31019 62.56795 86.05243
183 80.55842 68.78604 92.33079
184 74.31019 62.56795 86.05243
185 72.52498 60.78380 84.26617
186 62.70635 50.91097 74.50172
187 70.73978 58.99628 82.48327
188 70.73978 58.99628 82.48327
189 72.52498 60.78380 84.26617
190 68.95457 57.20540 80.70374
191 71.63238 59.89046 83.37430
192 70.73978 58.99628 82.48327
193 74.31019 62.56795 86.05243
194 79.66581 67.90025 91.43137
195 85.91404 74.08330 97.74477
196 85.91404 74.08330 97.74477
197 80.55842 68.78604 92.33079
198 68.95457 57.20540 80.70374
199 74.31019 62.56795 86.05243
200 60.02853 48.20088 71.85619
201 65.38416 53.61359 77.15472
202 70.73978 58.99628 82.48327
203 73.41759 61.67629 85.15888
204 78.77321 67.01363 90.53279
205 76.09540 64.34874 87.84206
206 66.27676 54.51280 78.04072
207 60.92114 49.10507 72.73721
208 84.12883 72.32086 95.93680
209 81.45102 69.67099 93.23105
210 64.49155 52.71355 76.26955
211 71.63238 59.89046 83.37430
212 73.41759 61.67629 85.15888
213 68.95457 57.20540 80.70374
214 67.16936 55.41117 78.92755
215 85.02143 73.20249 96.84037
216 62.70635 50.91097 74.50172
217 81.45102 69.67099 93.23105
218 71.63238 59.89046 83.37430
219 66.27676 54.51280 78.04072
220 72.52498 60.78380 84.26617
221 51.24927 39.50039 62.99815
222 57.49750 45.75329 69.24170
223 52.14187 40.39618 63.88757
224 55.71229 43.97095 67.45363
225 58.39010 46.64320 70.13700
226 54.81969 43.07852 66.56085
227 58.39010 46.64320 70.13700
228 50.35667 38.60376 62.10958
229 59.28270 47.53227 71.03313
230 60.17531 48.42050 71.93011
231 53.03448 41.29113 64.77782
232 57.49750 45.75329 69.24170
233 49.46406 37.70628 61.22184

234 56.60489 44.86254 68.34724
235 51.24927 39.50039 62.99815
236 63.74572 51.96507 75.52638
237 53.92708 42.18525 65.66892
238 61.96051 50.19446 73.72657
239 50.35667 38.60376 62.10958
240 58.39010 46.64320 70.13700
241 61.96051 50.19446 73.72657
242 58.39010 46.64320 70.13700
243 54.81969 43.07852 66.56085
244 60.17531 48.42050 71.93011
245 65.53093 53.73235 77.32951
246 52.14187 40.39618 63.88757
247 56.60489 44.86254 68.34724
248 51.24927 39.50039 62.99815
249 52.14187 40.39618 63.88757
250 54.81969 43.07852 66.56085
251 53.03448 41.29113 64.77782
252 50.35667 38.60376 62.10958
253 58.39010 46.64320 70.13700
254 53.92708 42.18525 65.66892
255 55.71229 43.97095 67.45363
256 58.39010 46.64320 70.13700
257 53.92708 42.18525 65.66892
258 56.60489 44.86254 68.34724
259 53.03448 41.29113 64.77782
260 53.92708 42.18525 65.66892
261 54.81969 43.07852 66.56085
262 59.28270 47.53227 71.03313
263 60.17531 48.42050 71.93011
264 54.81969 43.07852 66.56085
265 52.14187 40.39618 63.88757
266 54.81969 43.07852 66.56085
267 61.06791 49.30790 72.82792
268 53.92708 42.18525 65.66892
269 53.03448 41.29113 64.77782
270 54.81969 43.07852 66.56085
271 46.78625 35.00886 58.56365
272 59.28270 47.53227 71.03313
273 51.24927 39.50039 62.99815
274 53.92708 42.18525 65.66892
275 54.81969 43.07852 66.56085
276 61.06791 49.30790 72.82792
277 53.03448 41.29113 64.77782
278 53.03448 41.29113 64.77782
279 62.85312 51.08018 74.62606
280 51.24927 39.50039 62.99815
281 64.63832 52.84912 76.42753
282 44.10844 32.30393 55.91296
283 58.39010 46.64320 70.13700
284 58.39010 46.64320 70.13700
285 59.28270 47.53227 71.03313
286 53.03448 41.29113 64.77782
287 52.14187 40.39618 63.88757

288 55.71229 43.97095 67.45363
289 46.78625 35.00886 58.56365
290 69.99395 58.13609 81.85180
291 59.28270 47.53227 71.03313
292 57.49750 45.75329 69.24170
293 61.06791 49.30790 72.82792
294 59.28270 47.53227 71.03313
295 58.39010 46.64320 70.13700
296 45.89365 34.10804 57.67926
297 60.17531 48.42050 71.93011
298 57.49750 45.75329 69.24170
299 50.35667 38.60376 62.10958
300 51.24927 39.50039 62.99815
301 61.96051 50.19446 73.72657
302 54.81969 43.07852 66.56085
303 51.24927 39.50039 62.99815
304 54.81969 43.07852 66.56085
305 50.35667 38.60376 62.10958
306 61.06791 49.30790 72.82792
307 60.17531 48.42050 71.93011
308 52.14187 40.39618 63.88757
309 52.14187 40.39618 63.88757
310 55.71229 43.97095 67.45363
311 45.89365 34.10804 57.67926
312 56.60489 44.86254 68.34724
313 55.71229 43.97095 67.45363
314 60.17531 48.42050 71.93011
315 59.28270 47.53227 71.03313
316 57.49750 45.75329 69.24170
317 53.03448 41.29113 64.77782
318 57.49750 45.75329 69.24170
319 56.60489 44.86254 68.34724
320 47.67886 35.90883 59.44888
321 55.71229 43.97095 67.45363
322 52.14187 40.39618 63.88757
323 56.60489 44.86254 68.34724
324 53.92708 42.18525 65.66892
325 50.35667 38.60376 62.10958
326 53.92708 42.18525 65.66892
327 62.85312 51.08018 74.62606
328 49.46406 37.70628 61.22184
329 46.78625 35.00886 58.56365
330 58.39010 46.64320 70.13700
331 60.17531 48.42050 71.93011
332 47.67886 35.90883 59.44888
333 48.57146 36.80798 60.33494
334 59.28270 47.53227 71.03313
335 53.03448 41.29113 64.77782
336 49.46406 37.70628 61.22184
337 55.71229 43.97095 67.45363
338 53.03448 41.29113 64.77782
339 53.03448 41.29113 64.77782
340 54.81969 43.07852 66.56085
341 55.71229 43.97095 67.45363

342 49.46406 37.70628 61.22184
343 53.03448 41.29113 64.77782
344 45.89365 34.10804 57.67926
345 52.14187 40.39618 63.88757
346 55.71229 43.97095 67.45363
347 47.67886 35.90883 59.44888
348 58.39010 46.64320 70.13700
349 53.92708 42.18525 65.66892
350 50.35667 38.60376 62.10958
351 61.06791 49.30790 72.82792
352 61.06791 49.30790 72.82792
353 53.92708 42.18525 65.66892
354 54.81969 43.07852 66.56085
355 55.71229 43.97095 67.45363
356 51.24927 39.50039 62.99815
357 53.92708 42.18525 65.66892
358 53.03448 41.29113 64.77782
359 56.60489 44.86254 68.34724
360 43.21584 31.40063 55.03105
361 51.24927 39.50039 62.99815
362 57.49750 45.75329 69.24170
363 59.28270 47.53227 71.03313
364 55.71229 43.97095 67.45363
365 58.39010 46.64320 70.13700
366 51.24927 39.50039 62.99815
367 53.03448 41.29113 64.77782
368 53.03448 41.29113 64.77782
369 53.92708 42.18525 65.66892
370 55.71229 43.97095 67.45363
371 53.03448 41.29113 64.77782
372 47.67886 35.90883 59.44888
373 61.06791 49.30790 72.82792
374 57.49750 45.75329 69.24170
375 57.49750 45.75329 69.24170
376 56.60489 44.86254 68.34724
377 46.78625 35.00886 58.56365
378 60.17531 48.42050 71.93011
379 52.14187 40.39618 63.88757
380 53.03448 41.29113 64.77782
381 59.28270 47.53227 71.03313
382 61.96051 50.19446 73.72657
383 57.49750 45.75329 69.24170
384 46.78625 35.00886 58.56365
385 59.28270 47.53227 71.03313
386 49.46406 37.70628 61.22184
387 46.78625 35.00886 58.56365
388 52.14187 40.39618 63.88757
389 47.67886 35.90883 59.44888
390 55.71229 43.97095 67.45363
391 54.81969 43.07852 66.56085
392 54.81969 43.07852 66.56085
393 55.71229 43.97095 67.45363
394 55.71229 43.97095 67.45363
395 59.28270 47.53227 71.03313

```
## 396 58.39010 46.64320 70.13700
## 397 60.17531 48.42050 71.93011
## 398 52.14187 40.39618 63.88757
## 399 56.60489 44.86254 68.34724
## 400 48.57146 36.80798 60.33494
## 401 56.60489 44.86254 68.34724
## 402 46.78625 35.00886 58.56365
## 403 53.03448 41.29113 64.77782
## 404 53.92708 42.18525 65.66892
## 405 53.92708 42.18525 65.66892
## 406 52.14187 40.39618 63.88757
## 407 50.35667 38.60376 62.10958
## 408 57.49750 45.75329 69.24170
## 409 53.92708 42.18525 65.66892
## 410 49.46406 37.70628 61.22184
## 411 53.03448 41.29113 64.77782
## 412 57.49750 45.75329 69.24170
## 413 53.03448 41.29113 64.77782
## 414 52.14187 40.39618 63.88757
## 415 60.17531 48.42050 71.93011
## 416 53.03448 41.29113 64.77782
## 417 53.03448 41.29113 64.77782
## 418 57.49750 45.75329 69.24170
## 419 51.24927 39.50039 62.99815
## 420 62.85312 51.08018 74.62606
## 421 54.81969 43.07852 66.56085
## 422 62.85312 51.08018 74.62606
## 423 64.63832 52.84912 76.42753
## 424 49.46406 37.70628 61.22184
## 425 61.06791 49.30790 72.82792
## 426 52.14187 40.39618 63.88757
## 427 53.03448 41.29113 64.77782
## 428 54.81969 43.07852 66.56085
## 429 56.60489 44.86254 68.34724
## 430 55.71229 43.97095 67.45363
## 431 65.53093 53.73235 77.32951
## 432 54.81969 43.07852 66.56085
## 433 56.60489 44.86254 68.34724
## 434 54.81969 43.07852 66.56085
## 435 61.06791 49.30790 72.82792
## 436 55.71229 43.97095 67.45363
## 437 54.81969 43.07852 66.56085
## 438 53.92708 42.18525 65.66892
## 439 58.39010 46.64320 70.13700
## 440 63.74572 51.96507 75.52638
```

Interpretación: cada una de las filas del resultado proporcionado representa una observación distinta para un valor dado de estatura. El intervalo de predicción para cada observación se calcula con un nivel de confianza del 98.5% y se presenta en tres columnas: el valor de la predicción, el límite inferior y el límite superior.

4.3 Intervalo de predicción y

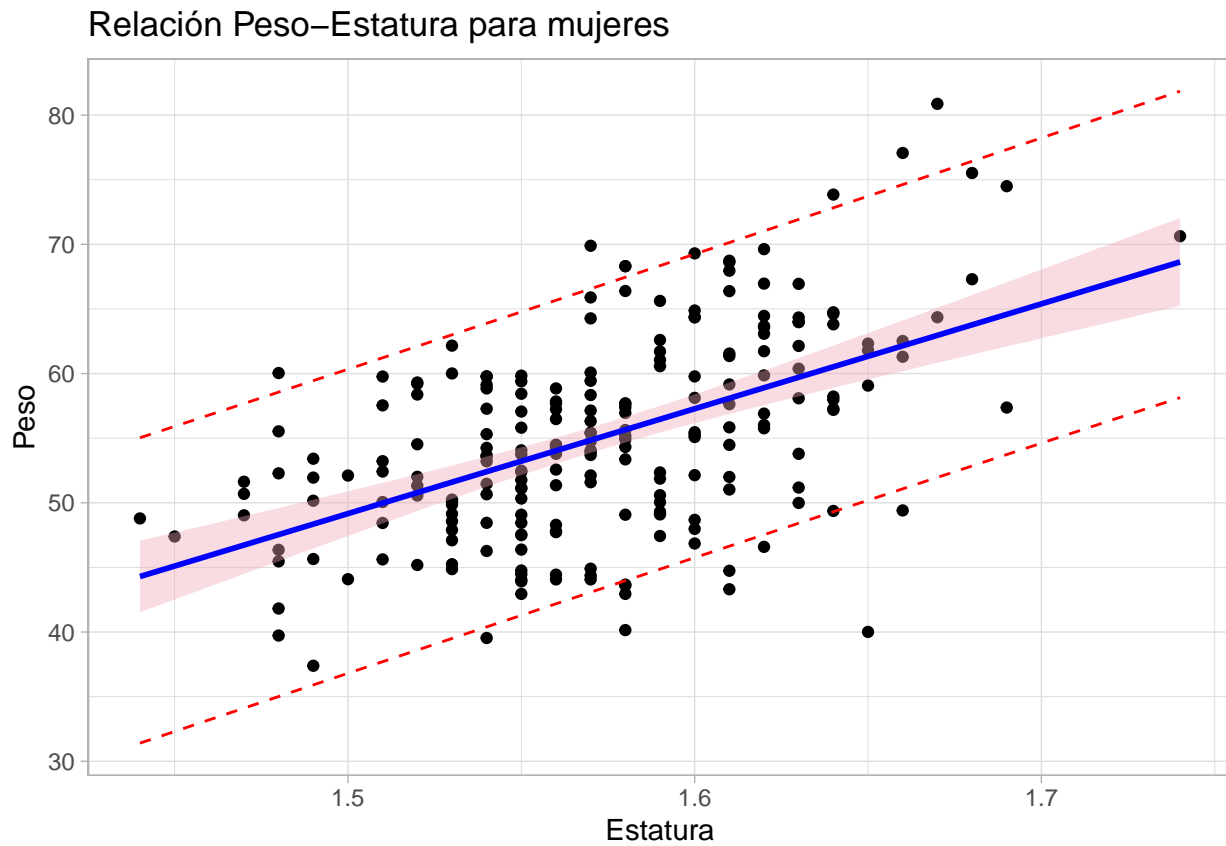
```
Ip=predict(object=A,interval="prediction",level=0.97)
datos1=cbind(M,Ip)
```

```

M2m = subset(datos1, Sexo == "M")
M2h = subset(datos1, Sexo == "H")

library(ggplot2)
ggplot(M2m, aes(x=Estatura, y=Peso)) +
  ggtitle("Relación 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="pink2") +
  theme_light()

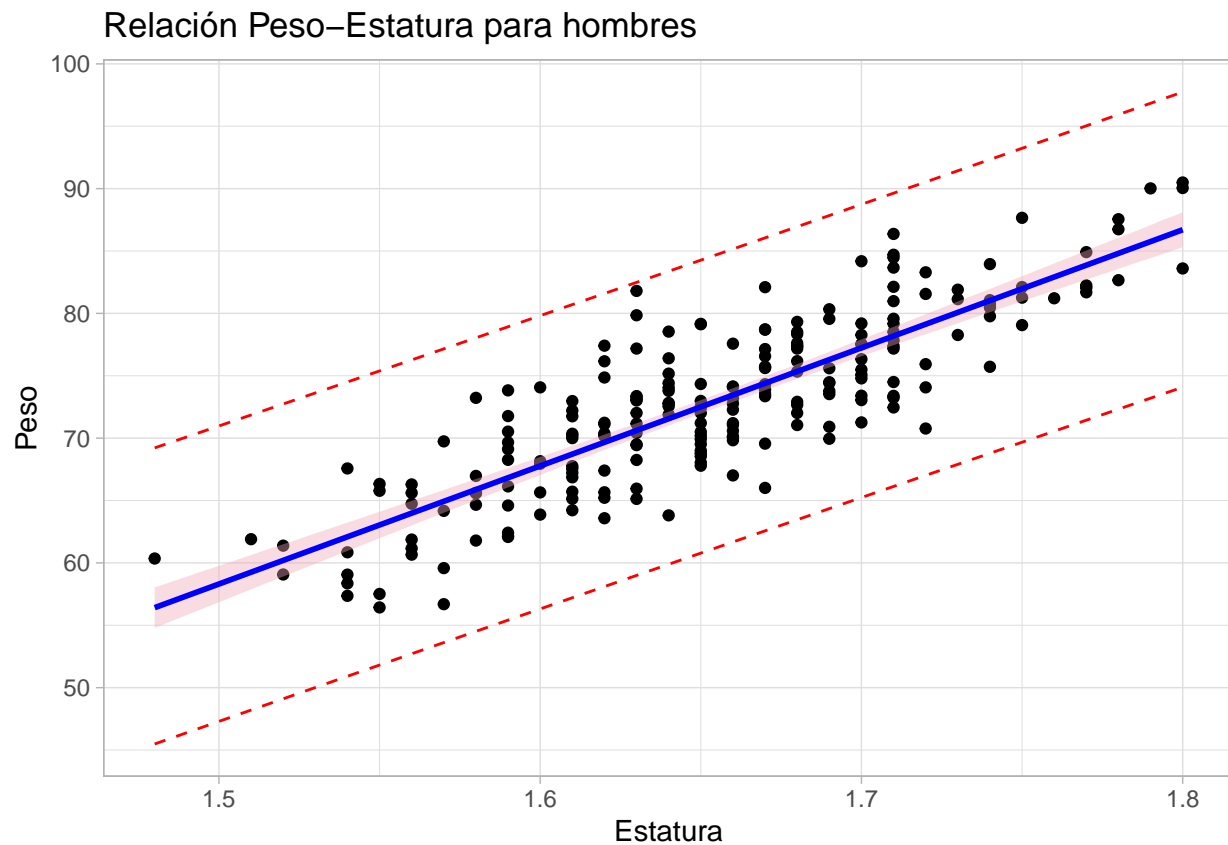
```



```

library(ggplot2)
ggplot(M2h, aes(x=Estatura, y=Peso)) +
  ggtitle("Relación 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()

```



Interpretación: Se observa que estos intervalos de predicción se tienen que graficar por separado debido a que se tiene que incluir la variable Sexo; por lo tanto, hay que limitarlo a dos variables para que los intervalos sean correctos. Una vez realizada esta tarea se pueden observar los gráficos anteriormente presentados por sexo.