Una institución de investigación realiza un estudio para analizar los efectos de un nuevo tratamiento para controlar los niveles altos de ansiedad. Para eso consideran un puntaje (a mayor valor mayores niveles de ansiedad) y definen un conjunto experimental con 120 individuos que en ese puntaje presentan valores similares al inicio del estudio, 60 son hombres y 60 mujeres. En el mercado se sabe que hay otro tratamiento que se usa comúnmente para este fin, de manera que de forma aleatoria han divido a los 120 individuos en tres grupos: 40 a los que no se aplicó ningún tratamiento (control), 40 a los que se aplicó el tratamiento actual (Trat1) y 40 a los que se aplicó el nuevo tratamiento (Trat2); 20 hombres y 20 mujeres en cada grupo. Los datos se presentan en el archivo *Ex4A.csv*.

Los investigadores sospechan que para el nuevo tratamiento podría existir un efecto diferenciado de acuerdo con el sexo, por lo que consideran conveniente incluir esta variable en el análisis.

Inciso I (Análisis descriptivo).

Para realizar un análisis descriptivo de los datos primero tenemos que ver que tipos de datos son nuestras variables

```
## 'data.frame':    120 obs. of  3 variables:
##  $ Sexo   : chr  "Hombre" "Hombre" "Hombre" "Hombre" ...
##  $ Trat   : chr  "Control" "Control" "Control" "Control" ...
##  $ Puntaje: num  8.88 9.54 13.12 10.14 10.26 ...
```

Dado que necesitamos dos variables categóricas (*Sexo y Tratamiento*), vamos a convertirlas en factor

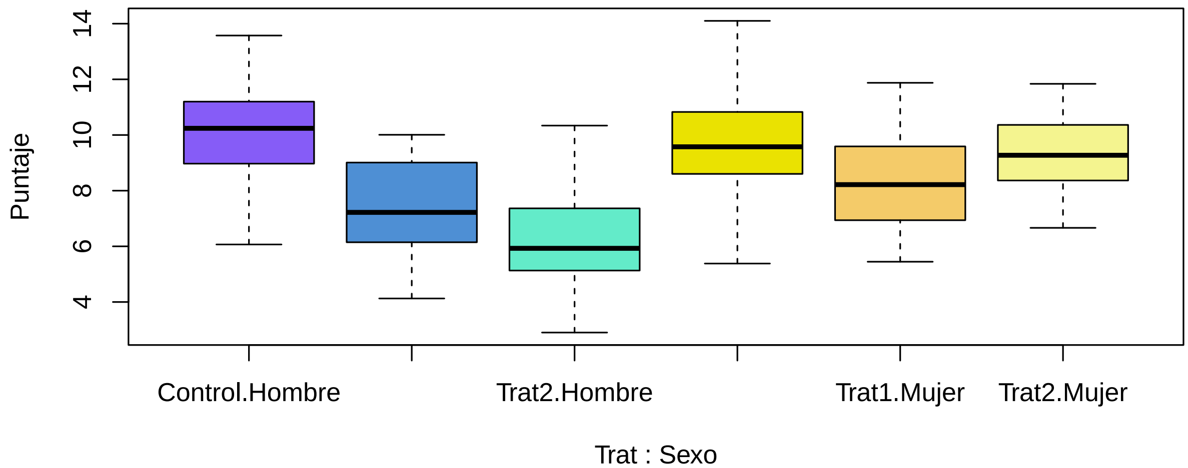
```
Datos4$Sexo = factor(Datos4$Sexo)
Datos4$Trat = factor(Datos4$Trat)


str(Datos4)
```

```
## 'data.frame':    120 obs. of  3 variables:
##  $ Sexo   : Factor w/ 2 levels "Hombre","Mujer": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Trat   : Factor w/ 3 levels "Control","Trat1",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ Puntaje: num  8.88 9.54 13.12 10.14 10.26 ...
```

Notemos que tenemos una variable categórica de 2 niveles (*Sexo*) y otra variable categórica de 3 niveles (*Tratamiento*)

Visualizamos estas variables



Esta vez estamos en un problema ANOVA de dos vias (ANOVA Factorial), de este modo tendremos dos factores que interaccionan con la variable dependiente $Puntaje$. Notamos que el factor $Trat$ y el factor $sexo$ tienen un efecto de manera conjunta en la variable dependiente $Puntaje$.

De acuerdo los boxplots, podemos decir que el Tratamiento 2 en los hombres es el que reduce más el puntaje, pues es el que tiene una media menor en los boxplots correspondientes a los hombres, esto implicará que el Tratamiento 2 es más efectivo para disminuir los niveles de ansiedad en los hombres.

Por otro lado, en las mujeres notamos que el Tratamiento 1 es el que tiene menor media a comparación del Tratamiento Control y el Tratamiento 2 de el lado de las mujeres, por lo que podemos decir que que el Tratamiento 1 es más efectivo para disminuir los niveles de ansiedad en las mujeres.

Inciso II (Expresión del puntaje promedio para cada valor)

Para dar la expresión del puntaje promedio para cada valor de las variables categóricas, primero tendremos que plantear nuestro modelo. Hay que considerar que tenemos dos variables categóricas, por lo tanto, nuestro nivel de referencia será igual al primer nivel de la variable $Trat : Trat = Control$. También hay que considerar que que la variable $Sexo$ es una variable categórica,y del mismo modo, solo tendremos el valor de $Mujer$ dentro de nuestro modelo. Aparte, tendremo nuestras interacciones (todo esto de acuerdo a la Nota que hay al final de las indicaciones del ejercicio)

$$y_i = \beta_0 + \beta_1 x_{1\text{Trat 1}} + \beta_2 x_{1\text{Trat 2}} + \beta_3 x_{2\text{Mujer}} + \beta_4 x_{1\text{Trat 1}} \cdot x_{2\text{Mujer}} + \beta_5 x_{1\text{Trat 2}} \cdot x_{2\text{Mujer}}$$

Por lo tanto, nuestra recta ajustada se verá del siguiente modo:

$$\widehat{\mathbb{E}}(Puntaje; Trat, Sexo) = \widehat{\beta_0} + \widehat{\beta_1} x_{1\text{Trat 1}} + \widehat{\beta_2} x_{1\text{Trat 2}} + \widehat{\beta_3} x_{2\text{Mujer}} + \widehat{\beta_4} x_{1\text{Trat 1}} \cdot x_{2\text{Mujer}} + \widehat{\beta_5} x_{1\text{Trat 2}} \cdot x_{2\text{Mujer}}$$

```
fit = lm(Puntaje ~ Trat*Sexo, data = Datos4)
summary(fit)
```

```
##
```

```
## Call:
## lm(formula = Puntaje ~ Trat * Sexo, data = Datos4)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3785 -1.1800 -0.0518  1.2159  4.3400
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          10.2832     0.3948  26.044  < 2e-16 ***
## TratTrat1            -2.8858     0.5584  -5.168 1.02e-06 ***
## TratTrat2            -4.0703     0.5584  -7.289 4.39e-11 ***
## SexoMujer            -0.5231     0.5584  -0.937   0.3509
## TratTrat1:SexoMujer   1.3758     0.7897   1.742   0.0842 .
## TratTrat2:SexoMujer   3.5914     0.7897   4.548 1.36e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.766 on 114 degrees of freedom
## Multiple R-squared:  0.4007, Adjusted R-squared:  0.3744
## F-statistic: 15.24 on 5 and 114 DF,  p-value: 1.873e-11
```

Los valores correspondientes a cada $\widehat{\beta_i}$ son los siguientes:

- $\widehat{\beta_0} = 10.2832$
- $\widehat{\beta_1} = -2.8858$
- $\widehat{\beta_2} = -4.0703$
- $\widehat{\beta_3} = -0.5231$
- $\widehat{\beta_4} = 1.3758$
- $\widehat{\beta_5} = 3.5914$

Por lo que nuestro modelo ajustado será:

$$\begin{aligned}\widehat{\mathbb{E}}(Puntaje; Trat, Sexo) = 10.2832 + (-2.8858)x_{1\text{Trat 1}} + (-4.0703)x_{1\text{Trat 2}} + (-0.5231)x_{2\text{Mujer}} \\ +(1.3758)x_{1\text{Trat 1}} \cdot x_{2\text{Mujer}} + (3.5914)x_{1\text{Trat 2}} \cdot x_{2\text{Mujer}}\end{aligned}$$

Para analizar el modelo hay que escribir todas las expresiones que corresponde a la esperanza de la variable dependiente *Puntaje* dado el valor de una de las variables categóricas dejando los demás fijos, entonces

- E(Puntaje; Trat = Control, Sexo = Hombre) $= \beta_0 + \beta_1(0) + \beta_2(0) + \beta_3(0) + \beta_4(0)(0) + \beta_5(0)(0) = \beta_0$
- E(Puntaje; Trat = Trat 1, Sexo = Hombre) $= \beta_0 + \beta_1(1) + \beta_2(0) + \beta_3(0) + \beta_4(1)(0) + \beta_5(0)(0) = \beta_0 + \beta_1$
- E(Puntaje; Trat = Trat 2, Sexo = Hombre) $= \beta_0 + \beta_1(0) + \beta_2(1) + \beta_3(0) + \beta_4(0)(0) + \beta_5(1)(0) = \beta_0 + \beta_2$
- E(Puntaje; Trat = Control, Sexo = Mujer) $= \beta_0 + \beta_1(0) + \beta_2(0) + \beta_3(1) + \beta_4(0)(1) + \beta_5(0)(1) = \beta_0 + \beta_3$

- $$\begin{aligned} \text{E(Puntaje; Trat = Trat 1, Sexo = Mujer)} &= \beta_0 + \beta_1(1) + \beta_2(0) + \beta_3(1) + \beta_4(1)(1) + \beta_5(0)(1) \\ &= \beta_0 + \beta_1 + \beta_3 + \beta_4 \end{aligned}$$

- $$\begin{aligned} \text{E(Puntaje; Trat = Trat 2, Sexo = Mujer)} &= \beta_0 + \beta_1(0) + \beta_2(1) + \beta_3(1) + \beta_4(0)(1) + \beta_5(1)(1) \\ &= \beta_0 + \beta_2 + \beta_3 + \beta_5 \end{aligned}$$

De este modo, la estimación puntual en cada caso será:

- E(Puntaje; Trat = Control, Sexo = Hombre) = 10.2832
- E(Puntaje; Trat = Trat 1, Sexo = Hombre) = $10.2832 + (-2.8858)$
- E(Puntaje; Trat = Trat 2, Sexo = Hombre) = $10.2832 + (-4.0703)$
- E(Puntaje; Trat = Control, Sexo = Mujer) = $10.2832 + (-0.5231)$
- E(Puntaje; Trat = Trat 1, Sexo = Mujer) = $10.2832 + (-2.8858) + (-0.5231) + 1.3758$
- E(Puntaje; Trat = Trat 2, Sexo = Mujer) = $10.2832 + (-4.0703) + (-0.5231) + 3.5914$

Otras estimaciones puntuales de acuerdo al archivo $EjemploRegLinMul1.R$

```r
vcov(fit)
```

```
##                     (Intercept)  TratTrat1  TratTrat2  SexoMujer
## (Intercept)           0.1559012 -0.1559012 -0.1559012 -0.1559012
## TratTrat1            -0.1559012  0.3118025  0.1559012  0.1559012
## TratTrat2            -0.1559012  0.1559012  0.3118025  0.1559012
## SexoMujer            -0.1559012  0.1559012  0.1559012  0.3118025
## TratTrat1:SexoMujer   0.1559012 -0.3118025 -0.1559012 -0.3118025
## TratTrat2:SexoMujer   0.1559012 -0.1559012 -0.3118025 -0.3118025
##                     TratTrat1:SexoMujer TratTrat2:SexoMujer
## (Intercept)                   0.1559012           0.1559012
## TratTrat1                    -0.3118025          -0.1559012
## TratTrat2                    -0.1559012          -0.3118025
## SexoMujer                    -0.3118025          -0.3118025
## TratTrat1:SexoMujer           0.6236049           0.3118025
## TratTrat2:SexoMujer           0.3118025           0.6236049
```

```r
sigma(fit)
```

```
## [1] 1.765793
```

Inciso III (tabla ANOVA).

Recordamos que en esta prueba no nos importa la redundancia, por lo que, de acuerdo a la tabla ANOVA tenemos la siguiente hipótesis:

$$H_0 : \beta_1 = 0 \wedge \beta_2 = 0 \wedge \beta_3 = 0 \wedge \beta_4 = 0 \wedge \beta_5 = 0 \quad vs \quad H_a : \beta_1 \neq 0 \vee \beta_2 \neq 0 \vee \beta_3 \neq 0 \vee \beta_4 \neq 0 \vee \beta_5 \neq 0$$

Con ayuda de la función *lm()* y *summary()* veremos si se rechaza o no se rechaza

```r
fit = lm(Puntaje ~ Trat*Sexo, data = Datos4)
summary(fit)
```

```
##
## Call:
## lm(formula = Puntaje ~ Trat * Sexo, data = Datos4)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3785 -1.1800 -0.0518  1.2159  4.3400
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          10.2832     0.3948  26.044  < 2e-16 ***
## TratTrat1            -2.8858     0.5584  -5.168 1.02e-06 ***
## TratTrat2            -4.0703     0.5584  -7.289 4.39e-11 ***
## SexoMujer            -0.5231     0.5584  -0.937   0.3509
## TratTrat1:SexoMujer   1.3758     0.7897   1.742   0.0842 .
## TratTrat2:SexoMujer   3.5914     0.7897   4.548 1.36e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.766 on 114 degrees of freedom
## Multiple R-squared:  0.4007, Adjusted R-squared:  0.3744
## F-statistic: 15.24 on 5 and 114 DF,  p-value: 1.873e-11
```

Notamos que el valor del último $p - value = 1.873e - 11$, como $p - value < 0.05$ podemos decir que se rechaza H_0, por lo que alguna de las β_i con $i = 1, ..., 5$ tiene significancia en nuestro modelo (*i.e.* alguna β puede ser diferente de cero).

Inciso IV (El sexo tiene un efecto en el puntaje?)

Para ver si el sexo tienen un efecto en el puntaje, nos podemos apoyar de la siguiente prueba de hipotesis simultanea (*La que el hint nos proporciona*):

$$H_0 : \text{E(Puntaje; Trat = k, Sexo = Hombre)} = \text{E(Puntaje; Trat = k, Sexo = Mujer)} \quad \forall k \in \{\text{Control,Trat1,Trat2}\}$$

$$vs$$

$$H_a : \text{E(Puntaje; Trat = k, Sexo = Hombre)} \neq \text{E(Puntaje; Trat = k, Sexo = Mujer)} \quad \forall k \in \{\text{Control,Trat1,Trat2}\}$$

Desarrollemos ambas hipotesis hasta tenerlas en terminos de β

$$\begin{gathered} H_0 : \text{E(Puntaje; Trat = Control, Sexo = Hombre)} = \text{E(Puntaje; Trat = Control, Sexo = Mujer)} \\ \text{y} \\ \text{E(Puntaje; Trat = Trat1, Sexo = Hombre)} = \text{E(Puntaje; Trat = Trat1, Sexo = Mujer)} \\ \text{y} \\ \text{E(Puntaje; Trat = Trat2, Sexo = Hombre)} = \text{E(Puntaje; Trat = Trat2, Sexo = Mujer)} \\ \textbf{VS} \end{gathered}$$

$$H_a : \text{E(Puntaje; Trat = Control, Sexo = Hombre)} \neq \text{E(Puntaje; Trat = Control, Sexo = Mujer)}$$

ó

$$\text{E(Puntaje; Trat = Trat1, Sexo = Hombre)} \neq \text{E(Puntaje; Trat = Trat1, Sexo = Mujer)}$$

ó

$$\text{E(Puntaje; Trat = Trat2, Sexo = Hombre)} \neq \text{E(Puntaje; Trat = Trat2, Sexo = Mujer)}$$

Sustituyendo con los varoles de las esperanzas obtenidos en el inciso 2

$$H_0 : \beta_0 = \beta_0 + \beta_3 \;\text{ y }\; \beta_0 + \beta_1 = \beta_0 + \beta_1 + \beta_3 + \beta_4 \;\text{ y }\; \beta_0 + \beta_2 = \beta_0 + \beta_2 + \beta_3 + \beta_5$$

VS

$$H_a : \beta_0 \neq \beta_0 + \beta_3 \;\text{ ó }\; \beta_0 + \beta_1 \neq \beta_0 + \beta_1 + \beta_3 + \beta_4 \;\text{ ó }\; \beta_0 + \beta_2 \neq \beta_0 + \beta_2 + \beta_3 + \beta_5$$

Simplificando tenemos

$$H_0 : \beta_3 = 0 \;\text{ y }\; \beta_3 + \beta_4 = 0 \;\text{ y }\; \beta_3 + \beta_5 = 0$$

VS

$$H_a : \beta_3 \neq 0 \;\text{ ó }\; \beta_3 + \beta_4 \neq 0 \;\text{ ó }\; \beta_3 + \beta_5 \neq 0$$

Como nos están pidiendo en especifico ver si el sexo tienen un efecto en el puntuaje, entonces el segundo y tercer caso ($\beta_3 + \beta_4 = 0$ y $\beta_3 + \beta_5 = 0$) son redundantes, por lo que se reduce a $\beta_4 = 0$ y $\beta_5 = 0$. Entonces, nuestra prueba de hipótesis queda así

$$H_0 : \beta_3 = 0 \;\text{ y }\; \beta_4 = 0 \;\text{ y }\; \beta_5 = 0 \quad \textbf{VS} \quad H_a : \beta_3 \neq 0 \;\text{ ó }\; \beta_4 \neq 0 \;\text{ ó }\; \beta_5 \neq 0$$

Para ver si H_0 se rechaza con un nivel de significancia de 0.025 nos apoyaremos de *multcomp* para hacer la prueba F

```r
K=matrix(c(0,0,0,1,0,0,
           0,0,0,0,1,0,
           0,0,0,0,0,1), ncol=6, nrow=3, byrow=TRUE)
m=c(0,0,0)
summary(glht(fit, linfct=K, rhs=m))
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Fit: lm(formula = Puntaje ~ Trat * Sexo, data = Datos4)
##
## Linear Hypotheses:
##        Estimate Std. Error t value Pr(>|t|)
## 1 == 0  -0.5231     0.5584  -0.937    0.623
## 2 == 0   1.3758     0.7897   1.742    0.182
## 3 == 0   3.5914     0.7897   4.548   <0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Dado que $1.827792e{-}06 = p{-}value < 0.025$ se rechaza H_0. Esto nos indica que al menos para un tratamiento existe un efecto diferenciado en el puntaje derivado del sexo de los individos De acuerdo a las indicaciones, tendremos que realizar una prueba de hipotesis simultenea para ver en tratamiento (en especifico) se puede considerar que el sexo tiene un efecto, Recordemos que habíamos quitado los casos redundantes, para esta nueva prueba, necesitaremos dejar la prueba de hipótesis con estos casos redundantes. Entonces

$$H_0 : \beta_3 = 0 \ \text{ y } \ \beta_3 + \beta_4 = 0 \ \text{ y } \ \beta_3 + \beta_5 = 0$$

VS

$$H_a : \beta_3 \neq 0 \ \text{ ó } \ \beta_3 + \beta_4 \neq 0 \ \text{ ó } \ \beta_3 + \beta_5 \neq 0$$

Nuevamente, nos apoyamos de *multcomp* para realizar esta prueba de hipotesis

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Fit: lm(formula = Puntaje ~ Trat * Sexo, data = Datos4)
##
## Linear Hypotheses:
##        Estimate Std. Error t value Pr(>|t|)
## 1 == 0  -0.5231     0.5584  -0.937    0.724
## 2 == 0   0.8527     0.5584   1.527    0.339
## 3 == 0   3.0683     0.5584   5.495   <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Notamos que, con el mismo nivel de significancia (0.025),NO se rechazan las siguientas pruebas de hipótesis

$$H_0 : \beta_0 = 0 \quad vs \quad H_a : \beta_0 \neq 0$$

Entonces es pausible considerar el modelo sin este valor (corresponde a la variable *Mujer*)

$$H_0 : \beta_3 + \beta_4 = 0 \quad vs \quad H_a : \beta_3 + \beta_4 \neq 0$$

Del mismo modo, es pausible considerar el modelo sin este valor (corresponde a la iteracción $Trat_1 \cdot Mujer$)

Dado que la 3era prueba de hipotesis se rechazó (correspondiente a la iteracción $Trat_2 \cdot Mujer$), podemos reducir nuestro modelo, quedando de la siguiente manera:

$$y_i = \beta_0 + \beta_1 x_{1\text{Trat 1}} + \beta_2 x_{1\text{Trat 2}} + \beta_3 x_{1\text{Trat 2}} \cdot x_{2\text{Mujer}}$$

Inciso V (Modelo reducido).

Nuestro modelo ajustado quedará del siguiente modo:

$$\widehat{\mathbb{E}}(Puntaje; Trat, Sexo) = \widehat{\beta_0} + \widehat{\beta_1} x_{1\text{Trat 1}} + \widehat{\beta_2} x_{1\text{Trat 2}} + \widehat{\beta_3} x_{1\text{Trat 2}} \cdot x_{2\text{Mujer}}$$

Como al inicio lo hicimos, ajustaremos el nuevo modelo de regresión con la función *lm()*

```
fit2=lm(Puntaje~Trat + I((Trat == 'Trat2')*(Sexo == 'Mujer')), data=Datos4)
summary(fit2)
```

```
##
## Call:
## lm(formula = Puntaje ~ Trat + I((Trat == "Trat2") * (Sexo ==
##     "Mujer")), data = Datos4)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6400 -1.0701  0.0033  1.0982  4.1249
##
## Coefficients:
##                                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                               10.0217     0.2806  35.709  < 2e-16
## TratTrat1                                 -2.1979     0.3969  -5.538 1.94e-07
## TratTrat2                                 -3.8087     0.4861  -7.835 2.47e-12
## I((Trat == "Trat2") * (Sexo == "Mujer"))   3.0683     0.5613   5.466 2.67e-07
##
## (Intercept)                              ***
## TratTrat1                                ***
## TratTrat2                                ***
## I((Trat == "Trat2") * (Sexo == "Mujer")) ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.775 on 116 degrees of freedom
## Multiple R-squared:  0.3838, Adjusted R-squared:  0.3678
## F-statistic: 24.08 on 3 and 116 DF,  p-value: 3.461e-12
```

Nuevamente, para analizar el modelo hay que escribir todas las expresiones que corresponde a la esperanza de la variable dependiente *Puntaje* dado el valor de una de las variables categóricas, dejando los demás valores fijos, entonces, al igual que el modelo inicial, tendremos las siguientes expresiones pero ahora con valores diferentes:

- $\mathrm{E}(\text{Puntaje};\ \text{Trat} = \text{Control}, \text{Sexo} = \text{Hombre}) = \beta_0$
- $\mathrm{E}(\text{Puntaje};\ \text{Trat} = \text{Trat 1}, \text{Sexo} = \text{Hombre}) = \beta_0 + \beta_1$
- $\mathrm{E}(\text{Puntaje};\ \text{Trat} = \text{Trat 2}, \text{Sexo} = \text{Hombre}) = \beta_0 + \beta_2$
- $\mathrm{E}(\text{Puntaje};\ \text{Trat} = \text{Control}, \text{Sexo} = \text{Mujer}) = \beta_0$
- $\mathrm{E}(\text{Puntaje};\ \text{Trat} = \text{Trat 1}, \text{Sexo} = \text{Mujer}) = \beta_0 + \beta_1$
- $\mathrm{E}(\text{Puntaje};\ \text{Trat} = \text{Trat 2}, \text{Sexo} = \text{Mujer}) = \beta_0 + \beta_2 + \beta_3$

Los valores correspondientes a cada β_i con $i = 0, ..., 3$ son

- $\widehat{\beta_0} = 10.0217$
- $\widehat{\beta_1} = -2.1979$
- $\widehat{\beta_2} = -3.8087$
- $\widehat{\beta_3} = 3.0683$

Otras estimaciones puntales serán:

```r
vcov(fit2)
```

```
##                                           (Intercept)     TratTrat1
## (Intercept)                              7.876336e-02 -7.876336e-02
## TratTrat1                               -7.876336e-02  1.575267e-01
## TratTrat2                               -7.876336e-02  7.876336e-02
## I((Trat == "Trat2") * (Sexo == "Mujer")) -6.121142e-17  9.618938e-17
##                                           TratTrat2
## (Intercept)                              -0.07876336
## TratTrat1                                 0.07876336
## TratTrat2                                 0.23629007
## I((Trat == "Trat2") * (Sexo == "Mujer")) -0.15752671
##                                          I((Trat == "Trat2") * (Sexo == "Mujer"))
## (Intercept)                                                         -6.121142e-17
## TratTrat1                                                            9.618938e-17
## TratTrat2                                                           -1.575267e-01
## I((Trat == "Trat2") * (Sexo == "Mujer"))                             3.150534e-01
```

```r
sigma(fit2)
```

```
## [1] 1.774974
```

Inciso VI (El nuevo tratamiento tiene el mejor desempeño?).

Dado las instrucciones del ejercicio, cuando se habla del "nuevo tratamiento" nos referimos al Tratamiento 2. El problema nos dice que "el nuevo tratamiento tiene el mejor desempeño". Al decir que el Tratamiento 2 es mejor, estamos diciendo que es más efectivo (por lo que disminuirá el Puntaje *i.e.,* va a disminuir el nivel de ansiedad) que el Tratamiento 1 y el no aplicar ningún tratamiento (*Control*). Entonces

1. El **Tratamiento 2** en los *hombres* tiene mejor desempeño que el **Tratamiento 1** en los *hombres*.
2. El **Tratamiento 2** en las *mujeres* tiene mejor desempeño que el **Tratamiento 1** en las *mujeres*.
3. El **Tratamiento 2** en los *hombres* tiene mejor desempeño que no aplicar ningún tratramiento (**Control**) en los *hombres*.
4. El **Tratamiento 2** en las *mujeres* tiene mejor desempeño que no aplicar ningún tratramiento (**Control**) en los *mujeres*.

Así

1. E(Puntaje; Trat = Trat2, Sexo = Hombre) < E(Puntaje; Trat = Trat1, Sexo = Hombre)
2. E(Puntaje; Trat = Trat2, Sexo = Mujer) < E(Puntaje; Trat = Trat1, Sexo = Mujer)
3. E(Puntaje; Trat = Trat2, Sexo = Hombre) < E(Puntaje; Trat = Control, Sexo = Hombre)
4. E(Puntaje; Trat = Trat2, Sexo = Mujer) < E(Puntaje; Trat = Control, Sexo = Mujer)

Llevando todo a termino de $\beta's$

1. $\beta_0 + \beta_2 < \beta_0 + \beta_1$
2. $\beta_0 + \beta_2 + \beta_3 < \beta_0 + \beta_1$
3. $\beta_0 + \beta_2 < \beta_0$
4. $\beta_0 + \beta_2 + \beta_3 < \beta_0$

Despejando

1. $0 < \beta_1 - \beta_2$
2. $0 < \beta_1 - \beta_2 - \beta_3$
3. $0 < -\beta_2$
4. $0 < -\beta_2 - \beta_3$

Por lo que nuestra prueba de hipótesis simultanea quedará del siguiente modo:

$$H_0 : \beta_1 - \beta_2 \leq 0 \quad \text{y} \quad \beta_1 - \beta_2 - \beta_3 \leq 0 \quad \text{y} \quad -\beta_2 \leq 0 \quad \text{y} \quad -\beta_2 - \beta_3 \leq 0$$
$$\textbf{vs}$$
$$H_a : \beta_1 - \beta_2 > 0 \quad \text{ó} \quad \beta_1 - \beta_2 - \beta_3 > 0 \quad \text{ó} \quad -\beta_2 > 0 \quad \text{ó} \quad -\beta_2 - \beta_3 > 0$$

Con ayuda de *multcomp* veremos si se rechaza H_0

```
K=matrix(c(0,1,-1,0,
           0,1,-1,-1,
           0,0,-1,0,
           0,0,-1,-1), ncol=4, nrow=4, byrow=TRUE)
m=c(0,0,0,0)
summary(glht(fit2, linfct=K, rhs=m, alternative = 'greater'))
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Fit: lm(formula = Puntaje ~ Trat + I((Trat == "Trat2") * (Sexo ==
##     "Mujer")), data = Datos4)
##
## Linear Hypotheses:
##        Estimate Std. Error t value  Pr(>t)
## 1 <= 0   1.6109     0.4861   3.314 0.00238 **
## 2 <= 0  -1.4574     0.4861  -2.998 1.00000
## 3 <= 0   3.8087     0.4861   7.835 < 0.001 ***
## 4 <= 0   0.7405     0.4861   1.523 0.18736
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Notemos que se rechazó H_0, es decir, no podemos decir que el nuevo tratamiento, en general, tiene un mejor desempeño. Pues, notemos que NO se rechazan las H_{0i} con $i = 2, 4$ que corresponden a la comparación del Tratamiento 2 con el Tratamiento 1 y el no aplicar ningún tratamiento en las mujeres. Por lo tanto, no hay suficiente evidencia para decir que el Tratamiento 2 es el mejor en las mujeres.Sin embargo, en los hombres se rechaza H_{0i} con $i = 1, 3$, entonces existe evidencia suficiente para decir que el Tratamiento 2 es mejor que el Tratamiento 1 y el no aplicar ningún tratamiento **solo** en los hombres.

Inciso VII (El nuevo tratamiento tiene el mejor desempeño en hombres, aunque el tratamiento actual lo tiene en mujeres?).

Para saber si la hipótesis de que *el nuevo tratamiento tiene el mejor desempeño en hombres, aunque el tratamiento actual lo tiene en mujeres* es cierta, necesitamos hacer una hipótesis simultenea parecida a la anterior. Entonces, queremos que los siguientes casos se cumplan:

1. El **Tratamiento 2** en los *hombres* tiene mejor desempeño que el **Tratamiento 1** en los *hombres*
2. El **Tratamiento 2** en los *hombres* tiene mejor desempeño que no aplicar ningún tratamiento (**Control**) en los *hombres*
3. El **Tratamiento 1** en las *mujeres* tiene mejor desempeño que el **Tratamiento 2** en las *mujeres*
4. El **Tratamiento 1** en las *mujeres* tiene mejor desempeño que no aplicar ningún tratamiento (**Control**) en las *mujeres*

Así

1. E(Puntaje; Trat = Trat2, Sexo = Hombre) < E(Puntaje; Trat = Trat1, Sexo = Hombre)
2. E(Puntaje; Trat = Trat2, Sexo = Hombre) < E(Puntaje; Trat = Control, Sexo = Hombre)
3. E(Puntaje; Trat = Trat1, Sexo = Mujer) < E(Puntaje; Trat = Trat2, Sexo = Mujer)
4. E(Puntaje; Trat = Trat1, Sexo = Mujer) < E(Puntaje; Trat = Control, Sexo = Mujer)

Llevando todo a termino de $\beta's$ tenemos

1. $\beta_0 + \beta_2 < \beta_0 + \beta_1$
2. $\beta_0 + \beta_2 < \beta_0$
3. $\beta_0 + \beta_1 < \beta_0 + \beta_2 + \beta_3$
4. $\beta_0 + \beta_1 < \beta_0$

Despejando

1. $0 < \beta_1 - \beta_2$
2. $0 < -\beta_2$
3. $0 < \beta_2 + \beta_3 - \beta_1$
4. $0 < -\beta_1$

Por lo que nuestra prueba de hipótesis simultanea quedará del siguiente modo:

$$H_0 : \beta_1 - \beta_2 \leq 0 \quad \text{y} \quad -\beta_2 \leq 0 \quad \text{y} \quad \beta_2 + \beta_3 - \beta_1 \leq 0 \quad \text{y} \quad -\beta_1 \leq 0$$
$$\textbf{vs}$$
$$H_a : \beta_1 - \beta_2 > 0 \quad \text{ó} \quad -\beta_2 > 0 \quad \text{ó} \quad \beta_2 + \beta_3 - \beta_1 > 0 \quad \text{ó} \quad -\beta_1 > 0$$

Con ayuda de *multcomp* veremos si se rechaza H0

```
K=matrix(c(0,1,-1,0,
           0,0,-1,0,
           0,-1,1,1,
           0,-1,0,0), ncol=4, nrow=4, byrow=TRUE)
m=c(0,0,0,0)
summary(glht(fit2, linfct=K, rhs=m, alternative = 'greater'))
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Fit: lm(formula = Puntaje ~ Trat + I((Trat == "Trat2") * (Sexo ==
##     "Mujer")), data = Datos4)
##
## Linear Hypotheses:
```

```
##        Estimate Std. Error t value  Pr(>t)
## 1 <= 0   1.6109     0.4861   3.314 0.00230 **
## 2 <= 0   3.8087     0.4861   7.835 < 1e-04 ***
## 3 <= 0   1.4574     0.4861   2.998 0.00621 **
## 4 <= 0   2.1979     0.3969   5.538 < 1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Notemos que hay evidencia suficiente para rechazar H_0 con un nivel de significancia del 0.05 pues, todas las H_{0i} con $i = 1, ..., 4$ se rechazaron. Esto nos quiere decir que es cierta la hipótesis que tenían los investigadores, de este modo, podemos concluir que el nuevo tratamiento tiene el mejor desempeño en hombres, aunque el tratamiento actual lo tiene en mujeres.