## Parcial1

#### Guillermo Alvarez Morales Cc. 1152715109

### 29/4/2022

1) Se realizará un experimento con la prueba exacta de Fisher, similar al experimento del famoso ejemplo de la dama que declara que conoce si en una taza de té con leche fue colocado primero el té o la leche.

El experimento a desarrollar consiste en que una persona va a determinar si en la preparación de una tasa de choco krispis con leche, se aplicó primero el cereal que la leche o fue primero la leche que el cereal, para este experimento se tomara una muestra de tamaño 10 como se hizo en el experimento famoso de la dama y la tasa de té.

 $H_0$ : La mujer hacierta si en la preparacion de una tasa de choco krispis con leche, se aplicó primero el cereal o la leche.

 $H_1$ : La mujer en realidad no puede hacertar en saber que fue lo primero que se aplico en una tasa de choco krispis.

```
data.t = matrix(c(4,1,2,3),ncol = 2,byrow = T)
colnames(data.t)=c("Choco","Leche")
rownames(data.t)=c("choco","Leche")
data.t
##
         Choco Leche
## choco
            4
                   3
## Leche
fisher.test(data.t, y = NULL, workspace = 200000, hybrid = FALSE,
or = 1, alternative = "two.sided", conf.level = 0.95)
##
##
   Fisher's Exact Test for Count Data
##
## data: data.t
## p-value = 0.5238
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
      0.218046 390.562917
## sample estimates:
## odds ratio
    4.918388
```

Con un P-value: 0.5238 > a un  $\alpha = 0.05$  hay evidencia suficiente para no rechazar una hipotesis nula, por tanto la mujer si puede determinar que se agrega primero a la hora de prepara una tasa de chocho krispis en leche.

2) La Encuesta de calidad de vida 2007 presenta información sobre si en un hogar tienen vehículo particular (0:No y 1: Si), si tienen motocicleta y si tienen algún negocio en la casa además del estrato. La información se presenta en la siguiente tabla. Realice un análisis descriptivo.

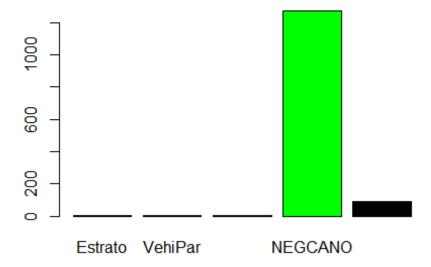
```
punto.2 <- read.csv("C:/Users/Guillermo/Desktop/DATOS CATEGORICO/punto 2.
txt", sep="")
View(punto.2)</pre>
```

Realizaremos a continuación el analisis descriptivo de la información sobre la calidad de vida del 2007.

- Media de las variables del informe de calidad de vida:

```
a = sapply(punto.2, mean, na.rm = TRUE)
а
##
                   VehiPar Motocicleta
       Estrato
                                             NEGCANO
                                                         NEGCASI
       3.50000
##
                   0.50000
                                0.50000
                                           635.41667
                                                        45.45833
m = sapply(punto.2, sd, na.rm = TRUE)
m
##
                     VehiPar Motocicleta
        Estrato
                                                 NEGCANO
                                                               NEGCASI
##
      1.7445568
                   0.5222330
                                 0.5107539 1271.0139710
                                                           92.6714529
barplot(m, col=c("orange", "blue", "red", "green", "black"), main = "Desviacion"
estandar de las variables")
```

### Desviacion estandar de las variables

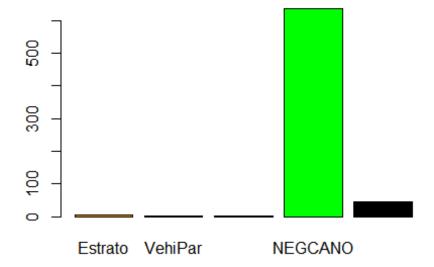


Podemos apreciar gráficamente que la desviación estándar de las variables estrato, vehipar, motocicleta presentan una misma desviación, en cambio en las variables negcano y negcasi presentan desviaciones diferentes y podemos apreciar que es mayor la desviación de las personas que no tienen un negocio en casa que las personas que si lo tienen.

Valores mínimos, máximos y media de cada una de las variables a trabajar del informe de calidad de vida.

<pre>summary(punto.2)</pre>										
## I	Estr	ato	Veh	iPar	Motoci	icleta	NEGO	CANO	NEGC	:AS
## 0.0	Min. 0	:1.0	Min.	:0.0	Min.	:0.0	Min.	: 0.00	Min.	:
## 2.0	_	:2.0	1st Qu	:0.0	1st Qu.	:0.0	1st Qu	.: 30.75	1st Qu.	:
## 5.0	Median 0	:3.5	Median	:0.5	Median	:0.5	Median	: 116.00	Median	:
## 45.4	Mean 46	:3.5	Mean	:0.5	Mean	:0.5	Mean	: 635.42	Mean	:
## 47.	_	:5.0	3rd Qu	:1.0	3rd Qu.	:1.0	3rd Qu	.: 451.00	3rd Qu.	:
## 85.	Max. 00	:6.0	Max.	:1.0	Max.	:1.0	Max.	:5109.00	Max.	:3
##			NA's							
<pre>barplot(a,col=c("orange","blue","red","green","black"),main = "Medias de las variables")</pre>										

# Medias de las variables

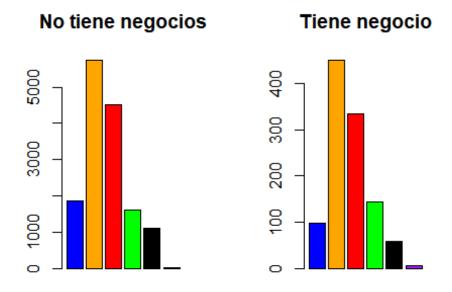


Podemos apreciar gráficamente que la variable estrato, vehipar, motocicleta presentan una misma media, en cambio en las variables negcano y negcasi presentan medias diferentes y podemos apreciar que es mayor la media de las personas que no tienen un negocio en casa que las personas que si lo tienen.

```
tabla.2 <- read.csv("C:/Users/Guillermo/Desktop/DATOS CATEGORICO/tabla 2.
txt", sep="")

par(mfrow=c(1,2))
barplot(tabla.2$NEGCANO, col = c("blue","orange","red","green","black","p
urple"), main = "No tiene negocios")

barplot(tabla.2$NEGCASI, col = c("blue","orange","red","green","black","p
urple"),main = "Tiene negocio")</pre>
```



Estrati1, Estrato2, Estrato3, Estrato4, Estrato5

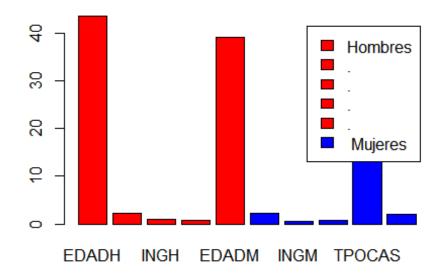
En las gráficas anteriores podemos notar que para ambos casos de si tienen negocios en casa o no tienen negocios en casa, el estrato que presenta un mayor comportamiento es el estrato 2, también podemos notar que el número de personas que no tienen negocios en casa por estrato son mayores que las personas que si los tienen, además de esto en ambas variables apreciamos que los estratos 2 y 3 hay menos personas con un negocio en casa que las personas que no lo tienen, también notamos que para el estrato 6 parece que se presenta el mismo comportamiento sin importar si tienen un negocio en casa o no lo tiene.

Pero en conclusión mas general en sin importar los estratos notamos que son menos las personas que tienen negocios en casa que las personas que si los tienen.

3) Realizar un analisis descriptivo de la base de datos PAREJAS

```
PAREJAS <- read.csv("C:/Users/Guillermo/Desktop/DATOS CATEGORICO/PAREJAS.
txt", sep="")
base0 = PAREJAS[,c(3,4,5,6,8,9,10,11,12,13)]
b = sapply(base0, mean)
b
##
                              INGH
                                       CASARH
        EDADH
                  ESCOLH
                                                   EDADM
                                                             ESCOLM
INGM
## 43.6510067
              2.2751678
                         0.9463087 0.8322148 39.1610738
                                                         2.2751678 0.6
040268
##
       CASARM
                 TPOCAS
                            NHIJOS
    0.7449664 15.3288591 2.0604027
##
barplot(b,col=c("red","red","red","red","blue","blue","blue","blue"
"blue", "blue", "blue"), main = "Medias de las variables", legend.text=c("Ho
mbres",".","."," Mujeres"))
```

### Medias de las variables



Del grafico anterior pordemos apreciar que la media de los hombres es mayor que la de la mujer en la edad, tambien notamos que son mas los hombres que estan casados que las mujeres, igualmente notamos que los hombres hablan ingles con mayor fluides que las mujeres.

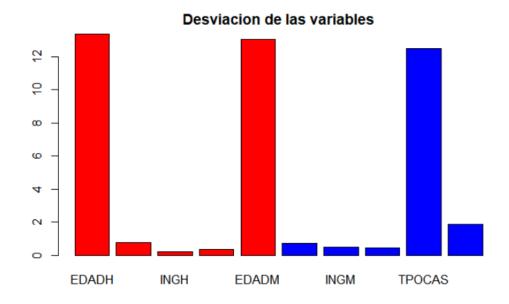
Desviacion estandar

```
c = sapply(base0, sd)
c

## EDADH ESCOLH INGH CASARH EDADM ESCOLM
INGM
## 13.3876336 0.7610365 0.2261677 0.3749358 13.0777649 0.7057587 0.4
907081
```

```
## CASARM TPOCAS NHIJOS
## 0.4373502 12.5231909 1.8680391

barplot(c,col=c("red","red","red","red","blue","blue","blue","blue","blue","blue","blue","blue","blue","blue"),main = "Desviacion de las variables")
```

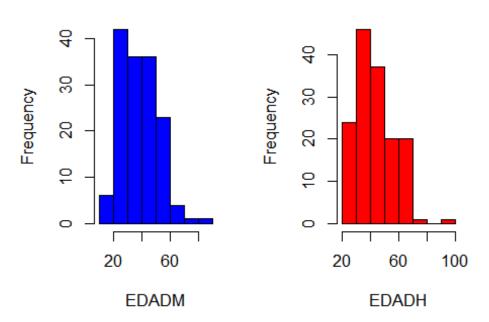


**HOMBRES**, MUJERES

Podemos apreciar que los hombres presentan una mayor desviación estándar que las mujeres y que el tiempo de casados presenta una desviación estándar notoria.

```
summary(base0)
                                           INGH
                                                            CASARH
##
        EDADH
                         ESCOLH
    Min.
           :20.00
                     Min.
                            :1.000
                                      Min.
                                             :0.0000
                                                        Min.
                                                               :0.0000
##
    1st Qu.:33.00
                     1st Qu.:2.000
                                      1st Qu.:1.0000
                                                        1st Qu.:1.0000
##
##
    Median :41.00
                     Median :2.000
                                     Median :1.0000
                                                        Median :1.0000
##
    Mean
           :43.65
                     Mean
                            :2.275
                                      Mean
                                             :0.9463
                                                        Mean
                                                               :0.8322
##
    3rd Qu.:53.00
                     3rd Qu.:3.000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.0000
                                                               :1.0000
##
    Max.
           :93.00
                     Max.
                            :3.000
                                      Max.
                                             :1.0000
                                                        Max.
##
        EDADM
                         ESCOLM
                                           INGM
                                                           CASARM
##
                            :1.000
    Min.
           :18.00
                     Min.
                                      Min.
                                             :0.000
                                                      Min.
                                                              :0.000
##
    1st Qu.:28.00
                     1st Qu.:2.000
                                      1st Qu.:0.000
                                                       1st Qu.:0.000
##
    Median :38.00
                     Median :2.000
                                      Median :1.000
                                                      Median :1.000
                                             :0.604
##
    Mean
           :39.16
                            :2.275
                                                              :0.745
                     Mean
                                      Mean
                                                      Mean
##
    3rd Qu.:49.00
                     3rd Qu.:3.000
                                      3rd Qu.:1.000
                                                       3rd Qu.:1.000
           :89.00
                            :3.000
                                             :1.000
##
    Max.
                     Max.
                                      Max.
                                                      Max.
                                                              :1.000
##
        TPOCAS
                         NHIJOS
           : 0.00
                          : 0.00
##
    Min.
                     Min.
##
    1st Ou.: 4.00
                     1st Qu.: 1.00
    Median :12.00
                     Median: 2.00
##
    Mean :15.33
                     Mean : 2.06
##
```

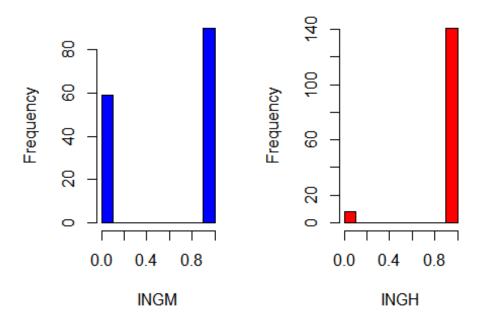
# Edad de los mujeres Edad de los hombres



Notamos que en esta base de datos se presentan más información de hombres que mujeres entre los 20 y 40 años, pero que entre los 40 y 70 años hay más información de las mujeres.

```
par(mfrow=c(1,2))
hist(base0$INGM, xlab = "INGM", col= "blue", main = "Ingles de los mujere
s")
hist(base0$INGH, xlab = "INGH", col= "red", main = "Ingles de los hombres
")
```

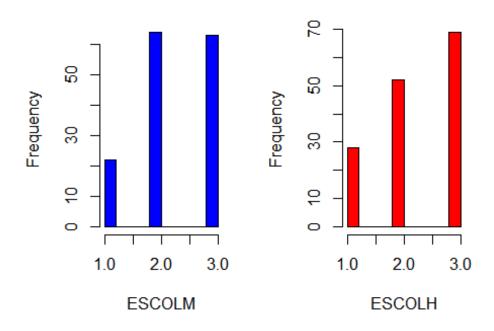
# Ingles de los mujeres Ingles de los hombres



Se aprecia que son mas los hombres que hablan ingles con fluides que las mujeres que lo hacen.

```
par(mfrow=c(1,2))
hist(base0$ESCOLM, xlab = "ESCOLM", col= "blue", main = " Grado de escola
ridad mujeres")
hist(base0$ESCOLH, xlab = "ESCOLH", col= "red", main = "Grado de escolari
dad hombres")
```

# Grado de escolaridad mujGrado de escolaridad hom



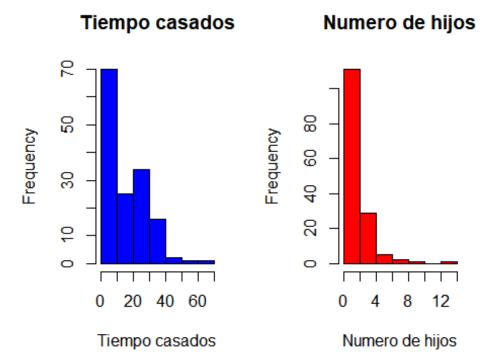
Para ambos géneros se presenta una alta frecuencia en el grado de escolaridad 3, aunque en los hombres se ve una frecuencia más alta que en las mujeres, fuera de esto el grado de escolaridad 2 es más frecuente en las mujeres que en los hombres.

```
par(mfrow=c(1,2))
hist(base0$CASARM, xlab = "CASARM", col= "blue", main = "Mujeres casada")
hist(base0$CASARH, xlab = "CASARH", col= "red", main = "Hombres casado")
```

# Mujeres casada Hombres casado 8 8 Frequency Frequency 80 8 2 0 0.0 0.4 8.0 0.0 0.4 8.0 CASARM CASARH

Como notamos son mas los hombres que se encuentran casados que las mujeres, algo podiamos suponer desde antes de observar esta grafica por la media que se analiso anteriormente.

```
par(mfrow=c(1,2))
hist(base0$TPOCAS, xlab = "Tiempo casados", col= "blue", main = "Tiempo c
asados")
hist(base0$NHIJOS, xlab = "Numero de hijos", col= "red", main = "Numero de hijos")
```



Notamos que lo más normal es que se tenga un número de hijos entre 0 y 2, que si lo pesamos es lo más usual que se pueda presentar en una familia, también podemos apreciar que el grafico de número de hijos presenta una caída a medida que aumenta la cantidad, ya que no es muy normal que se presenten familias en el 2007 con un número de hijos muy amplio.

También notamos que las personas no duran mucho tiempo casadas pues a medida que este tiempo de casados incrementa la frecuencia disminuye.

4) En este ejercicio se conseguirán foto de 50 personas donde 25 serán hombres y 25 mujeres, luego de esto se seleccionarán otras 10 personas donde 5 serán hombres y 5 serán mujeres las cuales se encargarán de catalogar si la persona fue o no fue atractiva, con el fin de mirar quienes son más atractivos si los hombres o las mujeres, tomaremos 1 para atractivo y 0 para no atractivo.

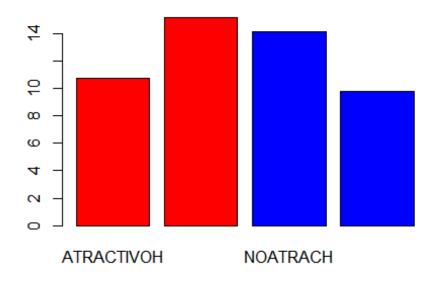
```
Datosmi <- read.csv("C:/Users/Guillermo/Desktop/DATOS CATEGORICO/Datosmi.
txt", sep="")
View(Datosmi)
bas = Datosmi[,3:6]
summary(bas)
##
      ATRACTIVOH
                     ATRACTIVOM
                                       NOATRACH
                                                      NOATRACM
##
   Min.
          : 7.0
                   Min.
                           :12.00
                                    Min.
                                           : 9.0
                                                   Min.
                                                          : 5.00
                   1st Qu.:13.25
##
   1st Ou.:10.0
                                    1st Ou.:14.0
                                                   1st Qu.: 8.25
##
   Median :11.0
                   Median :14.50
                                    Median :14.0
                                                   Median :10.50
                           :15.20
##
   Mean
           :10.8
                   Mean
                                    Mean
                                           :14.2
                                                   Mean
                                                          : 9.80
                                                   3rd Qu.:11.75
    3rd Qu.:11.0
                   3rd Qu.:16.75
                                    3rd Qu.:15.0
##
   Max. :16.0
                   Max. :20.00
                                           :18.0
                                                   Max. :13.00
##
                                    Max.
```

```
e = sapply(bas, mean)
e

## ATRACTIVOH ATRACTIVOM NOATRACH NOATRACM
## 10.8 15.2 14.2 9.8

barplot(e, col=c("red","red","blue","blue"), main = "Media atractivo y no atractivo por genero")
```

# Media atractivo y no atractivo por genero



AtractivoH, AtractivoM, NOatracH, NOatracM

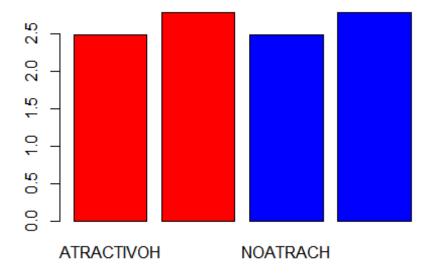
Podemos observar que la media de los hombres que son considerados atractivos es la menor que la media de las mujeres que son consideradas atractivas, esto mismo ocurre con la media de los no atractivos son considerados que los hombres son menos atractivos que las mujeres al poseer una media más alta.

```
f = sapply(bas, sd)
f

## ATRACTIVOH ATRACTIVOM NOATRACH NOATRACM
## 2.485514 2.780887 2.485514 2.780887

barplot(f, col=c("red","red","blue","blue"), main = "Desviacion de atract
ivo y no atractivo para hombres y mujeres")

GRAFICO DE DESVIACION ESTANDAR DE LAS VARIABLES A TRATAR
```



AtractivoH, AtractivoM, NOatracH, NOatracM

Apreciamos que las mujeres poseen una desviación estándar mayor que la de los hombres tanto en el campo de atractivo y no atractivo.

```
ber = Datosmi[1:5,c(3,4,5,6)]
bis = Datosmi[6:10,c(3,4,5,6)]
```

- Medias de las decisiones tomada por los hombres sobre atractivo para ambos géneros de las fotos

```
mh = sapply(ber, mean)
mh

## ATRACTIVOH ATRACTIVOM NOATRACH NOATRACM
## 11.0 16.2 14.0 8.8
```

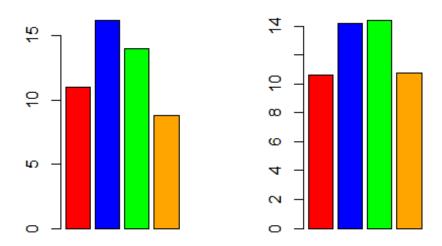
- Medias de las decisiones tomada por las mujeres sobre atractivo para ambos géneros de las fotos

```
mm = sapply(bis, mean)
mm

## ATRACTIVOH ATRACTIVOM NOATRACH NOATRACM
## 10.6 14.2 14.4 10.8

par(mfrow=c(1,2))
barplot(mh,col = c("red","blue","green","orange"), main = "Media por deci sion hombres" )
barplot(mm,col = c("red","blue","green","orange"), main = "Media por deci sion mujeres" )
```

# Media por decision homb Media por decision mujer



AtractivoH, AtractivoM, NOatracH, NOatracM

Se observa que tanto para los hombres como las mujeres las personas menos atractivas del estudio son los hombres, además se observa que para ambos sexos hay más hombres no atractivos que mujeres no atractivas.

Por otro lado, podemos notar que, para las mujeres, las mujeres atractivas son proporcionales a las no atractivas, además para los hombres las mujeres que son no atractivas son muy pocas, en cambio que para las mujeres no es de esta manera, dado que, para las mujeres, las mujeres no atractivas tienen una media proporcional a los hombres atractivos.

En las tablas a continuación veremos los resultados obtenidos por evaluador

### Evaluador 1 Hombre

```
data1 = matrix(c(10,15,14,11),ncol = 2,byrow = T)
colnames(data1)=c("Atractivo","NO.atractivo")
rownames(data1)=c("Hombre","Mujer")
data1

## Atractivo NO.atractivo
## Hombre 10 15
## Mujer 14 11
```

#### Evaluador 2 Hombre

```
data2 = matrix(c(16,9,14,11),ncol = 2,byrow = T)
colnames(data2)=c("Atractivo","NO.atractivo")
rownames(data2)=c("Hombre","Mujer")
data2
```

```
## Atractivo NO.atractivo
## Hombre 16 9
## Mujer 14 11
```

#### Evaluador 3 Hombre

#### Evaluador 4 Hombre

#### Evaluador 5 Hombre

```
data5 = matrix(c(7,18,20,5),ncol = 2,byrow = T)
colnames(data5)=c("Atractivo","NO.atractivo")
rownames(data5)=c("Hombre","Mujer")
data5

## Atractivo NO.atractivo
## Hombre 7 18
## Mujer 20 5
```

### Evaluador 6 Mujer

```
data6 = matrix(c(8,17,19,6),ncol = 2,byrow = T)
colnames(data6)=c("Atractivo","NO.atractivo")
rownames(data6)=c("Hombre","Mujer")
data6

## Atractivo NO.atractivo
## Hombre 8 17
## Mujer 19 6
```

#### Evaluador 7 Mujer

```
data7 = matrix(c(11,14,12,13),ncol = 2,byrow = T)
colnames(data7)=c("Atractivo","NO.atractivo")
```

```
rownames(data7)=c("Hombre","Mujer")
data7

## Atractivo NO.atractivo
## Hombre 11 14
## Mujer 12 13
```

#### Evaluador 8 Mujer

```
data8 = matrix(c(10,15,13,12),ncol = 2,byrow = T)
colnames(data8)=c("Atractivo","NO.atractivo")
rownames(data8)=c("Hombre","Mujer")
data8

## Atractivo NO.atractivo
## Hombre 10 15
## Mujer 13 12
```

### Evaluador 9 Mujer

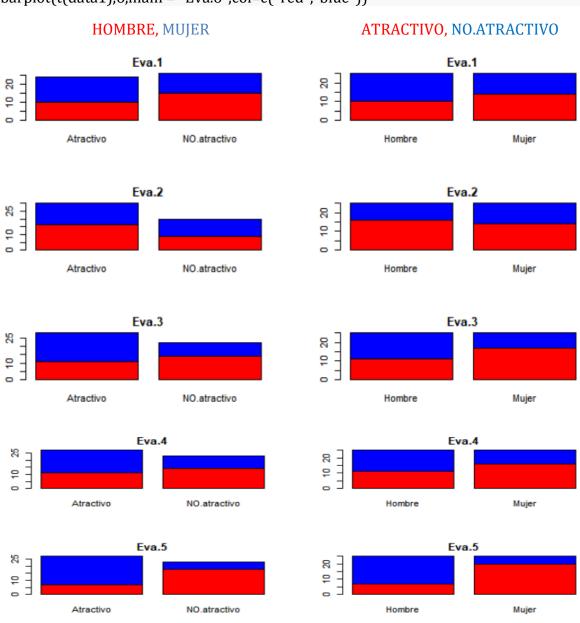
### Evaluador 10 Mujer

Le haremos un analicis a cada evaluador:

Primero haremos los graficos de las respuestas de los hombres:

```
par(mfrow=c(3,2))
barplot(data1,1, main = "Eva.1", col=c("red","blue"))
barplot(t(data1),2,main = "Eva.1",col=c("red","blue"))
barplot(data2,3, main = "Eva.2", col=c("red","blue"))
barplot(t(data2),4,main = "Eva.2",col=c("red","blue"))
```

barplot(data3,5, main = "Eva.3", col=c("red","blue"))
barplot(t(data3),6,main = "Eva.3",col=c("red","blue"))
barplot(data4,1, main = "Eva.4", col=c("red","blue"))
barplot(t(data4),2,main = "Eva.4",col=c("red","blue"))
barplot(data5,3, main = "Eva.5", col=c("red","blue"))
barplot(t(data5),4,main = "Eva.5",col=c("red","blue"))
barplot(data1,5, main = "Eva.6", col=c("red","blue"))
barplot(t(data1),6,main = "Eva.6",col=c("red","blue"))



Como podemos aprciar en los histogramas, notamos un comportamiento considerable en el campo de no atractivo ya que si observamos bien, apreciamos que dependiendo del evaluador hombre esta variable puede incrementar como no hacerlo para el caso del hombre no atractivo, pero en el caso de la mujer se aprecia una tendencia a disminuir la cantidad de mujeres no atractivas, lo cual nos indica que los hombres son menos atractivos que las mujeres, esta situación se presemta similarmente en el caso de atractivos, en donde los hombres suelen tener una amplia variabilidad la cual en las mujeres no lo presentan, se podria incluso suponer que para los hombres la mayoria de las mujeres en el experimento eran atractivas.

```
par(mfrow=c(3,2))
pie(data1,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador1")
pie(data2,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador2")
pie(data3,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador3")
pie(data4,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador4")
pie(data5,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador5")
```



NO.atractivo y Hombre NO.atractivo y Mujer

Evaluador1

#### Evaluador2

Atractivo y mujer

NO.atractivo y Hombre

NO.atractivo y Mujer

#### Evaluador3

Atractivo y mujer Atractivo y Hombre
NO.atractivo y Mujer

#### Evaluador4

Atractivo y mujer Atractivo y Hombre
NO.atractivo y Hombre

#### Evaluador5

Atractivo y mujer

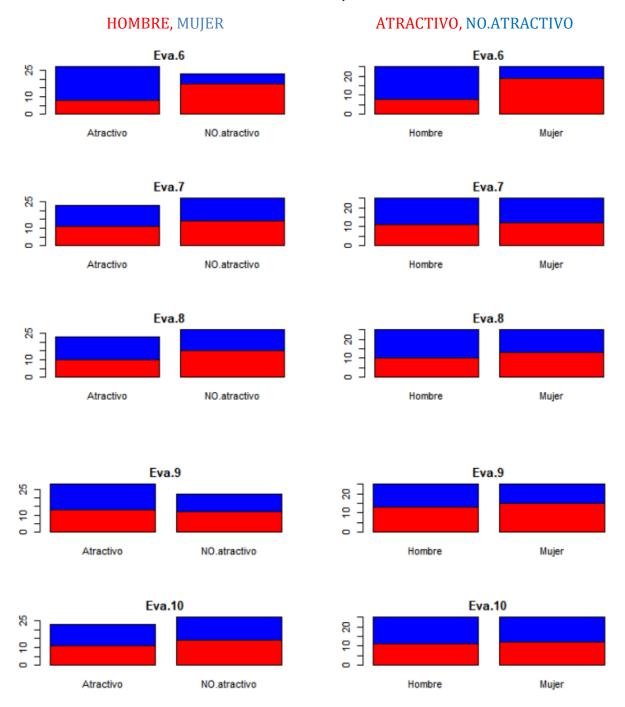
Atractivo y Hombre

NO.atractivo y Mujer

En los diagramas observamos que atractivo y mujer presentan una amplia proporción de la torta indispensablemente de quien sea el evaluador, y también notamos que la porción más pequeña siempre suele ser la mujer no atractiva, lo cual nos indica que el

análisis que hicimos con anterioridad tiene sentido, fuera de esto las demás variables como el No.atractivo y hombre se presenta en casi todas las gráficas con una proporción similar, dándonos a entender que el hombre no atractivo se presentó demasiado en los resultados dados por los evaluadores .

Ahora haremos el analisis de las evaluadoras mujeres:



Se observa que para la mayoría de las evaluadoras mujeres estas determinaron que el sexo más atractivo es el femenino, además notamos también que los hombres son quienes fueron considerados menos atractivos de la investigación para todas las evaluadoras, lo cual nos permite concluir que para las evaluadoras mujeres de las 50 fotos, donde 25 eran hombres y 25 mujeres, el género menos atractivo en definitiva eran los hombres.

```
par(mfrow=c(3,2))
pie(data6,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador6")
pie(data7,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador7")
pie(data8,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador8")
pie(data9,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador9")
pie(data10,labels=c('Atractivo y Hombre','Atractivo y mujer',
'NO.atractivo y Hombre','NO.atractivo y Mujer'),main = "Evaluador10")
```



# Evaluador10

Atractivo y mujer Atractivo y Hombre
NO.atractivo y Hombre

Del grafico anterior podemos observar que para casi todas las evaluadoras que la mujer es considerada la mas atractiva, y que la proporción mas grande del grafico la suele ocupar no atractivo y hombre dándonos a entender que para las evaluadoras la persona menos atractiva en definitiva son los hombres en el estudio, algo que ya habíamos notado en los gráficos anteriores.

Como conclusión general podemos decir que tanto para evaluadores hombres como para mujeres, el género más atractivo es el femenino y el menos atractivo es el masculino.