

Tarea10

Yimmy Eman

2022-07-01

Instala y carga el paquete MASS en R. En él hay una tabla de datos llamada birthwt sobre factores que pueden incidir en el peso de los niños al nacer.

```
library(MASS)
df <- birthwt
```

Pregunta 1

Utiliza str() y head() para explorar la estructura, y con help(), mirar el significado de cada variable.

```
str(df)
```

```
## 'data.frame':    189 obs. of  10 variables:
## $ low  : int  0 0 0 0 0 0 0 0 0 0 ...
## $ age  : int  19 33 20 21 18 21 22 17 29 26 ...
## $ lwt  : int  182 155 105 108 107 124 118 103 123 113 ...
## $ race : int  2 3 1 1 1 3 1 3 1 1 ...
## $ smoke: int  0 0 1 1 1 0 0 0 1 1 ...
## $ ptl  : int  0 0 0 0 0 0 0 0 0 0 ...
## $ ht   : int  0 0 0 0 0 0 0 0 0 0 ...
## $ ui   : int  1 0 0 1 1 0 0 0 0 0 ...
## $ ftv  : int  0 3 1 2 0 0 1 1 1 0 ...
## $ bwt  : int  2523 2551 2557 2594 2600 2622 2637 2637 2663 2665 ...
```

```
head(df, 5)
```

```
##    low age lwt race smoke ptl ht ui ftv  bwt
## 85   0  19 182   2     0   0 0 1   0 2523
## 86   0  33 155   3     0   0 0 0   3 2551
## 87   0  20 105   1     1   0 0 0   1 2557
## 88   0  21 108   1     1   0 0 1   2 2594
## 89   0  18 107   1     1   0 0 1   0 2600
```

Significado de las variable:

- low: Indicador de peso al nacer inferior a 2,5 kg.
- race: raza de la madre (1 = blanca, 2 = negra, 3 = otra).

Pregunta 2

Calcula una tabla de frecuencias relativas marginales de los pares raza de la madre y peso inferior a 2.5 kg o no, que permita ver si la raza de la madre influye en el peso del bebé.

```
df1 <- df[,c("race", "low")]
head(df1,10)
```

```
##      race low
## 85      2   0
## 86      3   0
## 87      1   0
## 88      1   0
## 89      1   0
## 91      3   0
## 92      1   0
## 93      3   0
## 94      1   0
## 95      1   0
```

```
prop.table(table(df1), margin = 1)
```

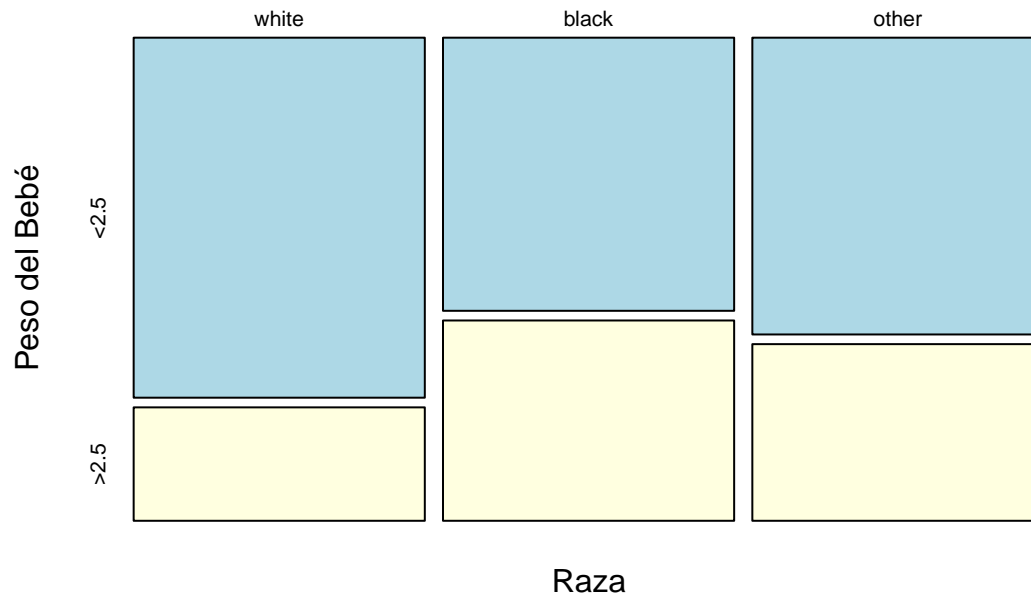
```
##      low
## race      0      1
##   1 0.7604167 0.2395833
##   2 0.5769231 0.4230769
##   3 0.6268657 0.3731343
```

Pregunta 3

Dibuja un diagrama de mosaico de la tabla del ejercicio anterior.

```
t <- table(df1)
colnames(t) <- c("<2.5", ">2.5")
rownames(t) <- c("white", "black", "other")
plot(prop.table(t, margin = 1),
     col = c("lightblue", "lightyellow"),
     main = "Gráfico de mosaico variable \"race\" y \"low\"",
     xlab = "Raza",
     ylab = "Peso del Bebé")
```

Gráfico de mosaico variable "race" y "low"

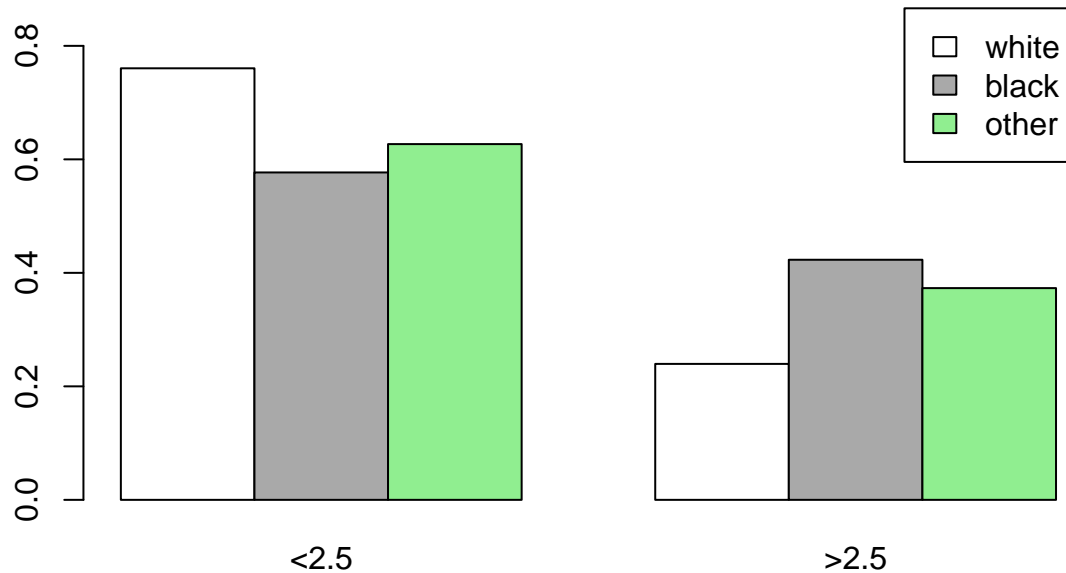


Pregunta 4

Dibuja un diagrama bidimensional de barras, con las barras organizadas en bloques, que permita visualizar la información de los ejercicios anteriores. Aprovecha para poner nombres adecuados a los bloques, colores a las barras y añadir una leyenda que explique qué representa cada barra.

```
barplot(prop.table(t, margin = 1),
        legend.text = row.names(t),
        beside = T,
        col = c("white", "darkgray", "lightgreen"),
        ylim = c(0,0.9),
        main = "Gráfico de barras variable \"race\" y \"low\"",
        )
```

Gráfico de barras variable "race" y "low"



Pregunta 5

¿Se puede obtener alguna conclusión de la tabla y el diagrama de barras anterior? Argumenta tu respuesta.

Se puede concluir que el peso de los bebés mayor a 2.5 kg la raza Negra supera a la raza Blanca por más del 50%.

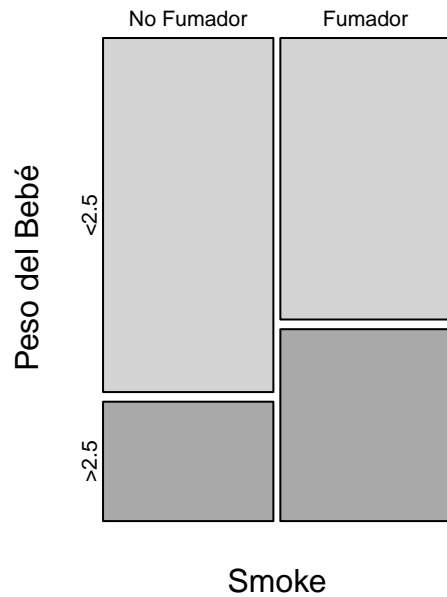
Pregunta 6

Repita los cuatro ejercicios anteriores para los pares madre fumadora o no, peso inferior a 2.5 kg o no y para los pares madre hipertensa o no y peso inferior a 2.5 kg o no.

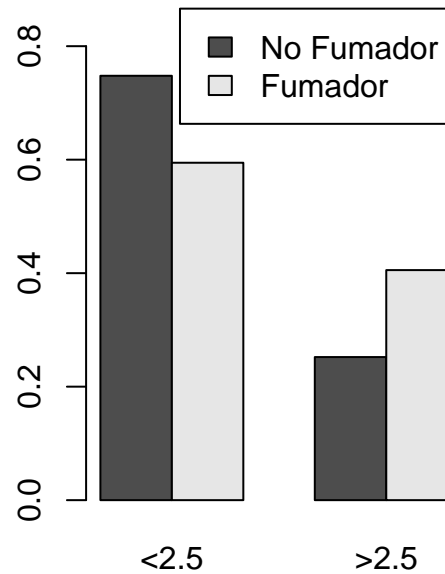
```
par(mfrow = c(1,2))
df2 <- df[,c("smoke", "low")]
t2 <- table(df2)
colnames(t2) <- c("<2.5", ">2.5")
rownames(t2) <- c("No Fumador", "Fumador")
plot(prop.table(t2, margin = 1),
     col = c("lightgray", "darkgray"),
     main = "Gráfico de mosaico variable \n\"smoke\" y \"low\"",
     xlab = "Smoke",
     ylab = "Peso del Bebé")

barplot(prop.table(t2, margin = 1),
       legend.text = row.names(t2),
       beside = T,
       ylim = c(0,0.9),
       main = "Gráfico de barras variable \n\"smoke\" y \"low\"",
       )
```

**Gráfico de mosaico variable
"smoke" y "low"**



**Gráfico de barras variable
"smoke" y "low"**



```
par(mfrow = c(1,1))
```

```
par(mfrow = c(1,2))
df3 <- df[, c("ht", "low")]
t3 <- table(df3)
colnames(t3) <- c("<2.5", ">2.5")
rownames(t3) <- c("Sin Hipertención", "Con Hipertención")
plot(prop.table(t3, margin = 1),
     col = c("gray", "lightyellow"),
     main = "Gráfico de mosaico variable \n \"ht\" y \"low\"",
     xlab = "Hipertención",
     ylab = "Peso del Bebé")

barplot(prop.table(t3, margin = 1),
        legend.text = row.names(t3),
        beside = T,
        ylim = c(0,0.9),
        col = c("gray", "lightyellow"),
        main = "Gráfico de barras variable \n \"ht\" y \"low\"",
        )
```

Gráfico de mosaico variable "ht" y "low"

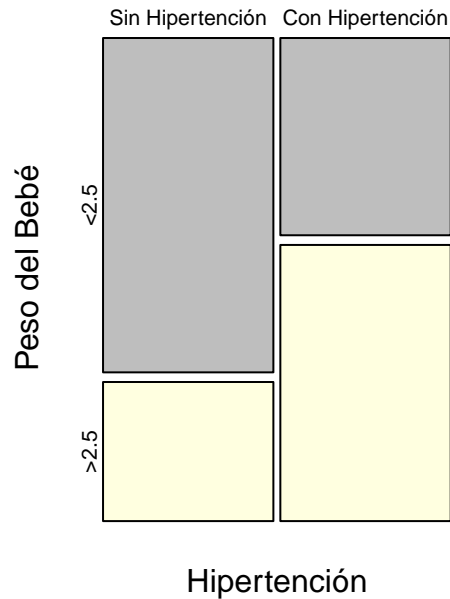
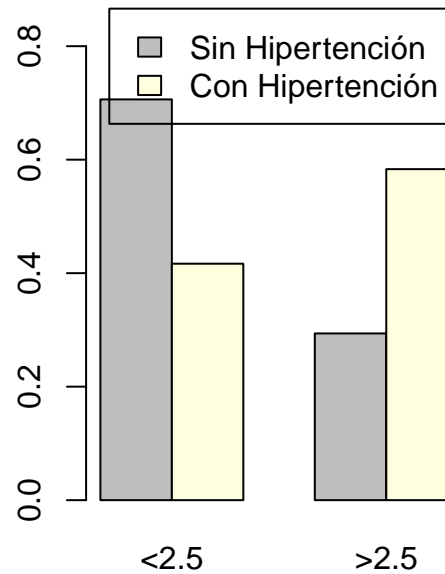


Gráfico de barras variable "ht" y "low"



```
par(mfrow = c(1,1))
```

Conclusión: Las Madres con diagnóstico de hipertensión tienden a tener bebés con peso inferior a 2.5kg.

Pregunta 7

Calcula una tabla de frecuencias relativas marginales de la terna raza de la madre, madre fumadora o no y peso inferior a 2.5 kg o no, que permita ver si la raza de la madre y su condición de fumadora o no fumadora influyen en el peso del bebé.

```
df4 <- df[,c("smoke", "low", "race")]
t4 <- table(df4)
prop.table(t4, margin = 1)
```

```
## , , race = 1
##
##      low
## smoke      0      1
##      0 0.34782609 0.03478261
##      1 0.44594595 0.25675676
##
## , , race = 2
##
##      low
## smoke      0      1
##      0 0.09565217 0.04347826
##      1 0.05405405 0.08108108
##
## , , race = 3
```

```
##
##      low
## smoke      0      1
##      0 0.30434783 0.17391304
##      1 0.09459459 0.06756757
```

```
prop.table(t4, margin = c(1,3))
```

```
## , , race = 1
##
##      low
## smoke      0      1
##      0 0.90909091 0.09090909
##      1 0.63461538 0.36538462
##
## , , race = 2
##
##      low
## smoke      0      1
##      0 0.68750000 0.31250000
##      1 0.40000000 0.60000000
##
## , , race = 3
##
##      low
## smoke      0      1
##      0 0.63636364 0.36363636
##      1 0.58333333 0.41666667
```

Pregunta 8

Dibuja un diagrama de mosaico de la tabla del ejercicio anterior.

```
plot(prop.table(table(df[c("smoke", "low", "race")])), margin=c(1, 3),
     main = "Gráfico de mosaico de las variables \"race\", \"smoke\" y \"low\"",
     col =c("lightgreen","lightblue", "pink"))
```

Gráfico de mosaico de las variables "race", "smoke" y "low"

