Laboratorio 1, Tópicos en análisis datos 1

Joshua Isaac Cervantes Artavia

2023-08-22

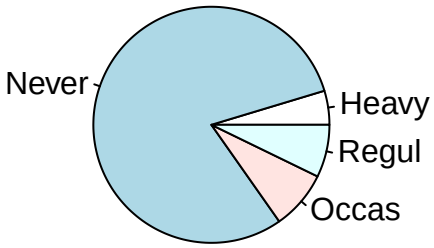
Table of contents

1 Carga de tablas de datos

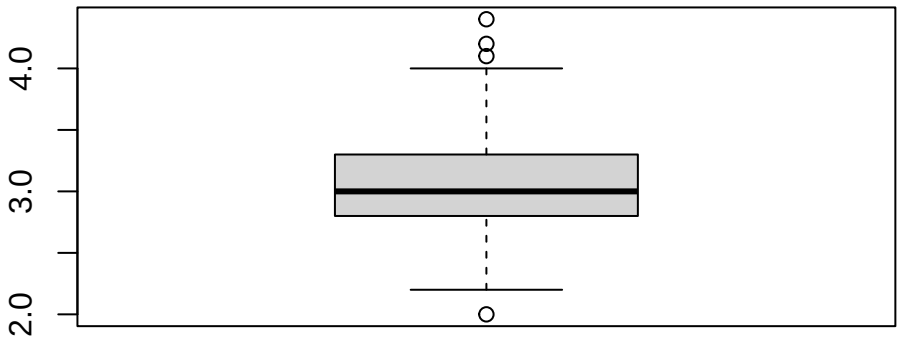
```
library(MASS)
data(survey)
names(survey)
```

```
 [1] "Sex"    "Wr.Hnd" "NW.Hnd" "W.Hnd"  "Fold"   "Pulse"  "Clap"   "Exer"
 [9] "Smoke"  "Height" "M.I"    "Age"
```

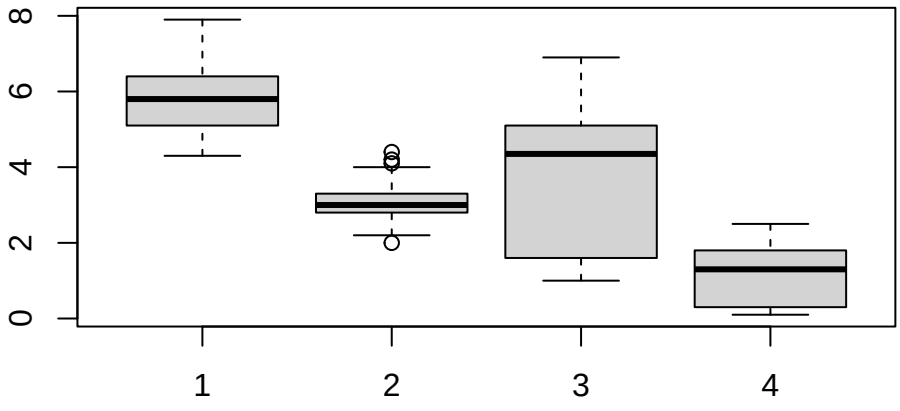
```
# Se hace un grafico
pie(table(survey$Smoke))
```




```
# Se cargan los datos de iris
data(iris)


# Se realiza un grafico de caja
boxplot(iris$Sepal.Width)
```




```
boxplot(iris$Sepal.Length, iris$Sepal.Width, iris$Petal.Length, iris$Petal.Width)
```



```r
# Se ven las instrucciones de la funcion boxplot
?boxplot
```

```
starting httpd help server ... done
```

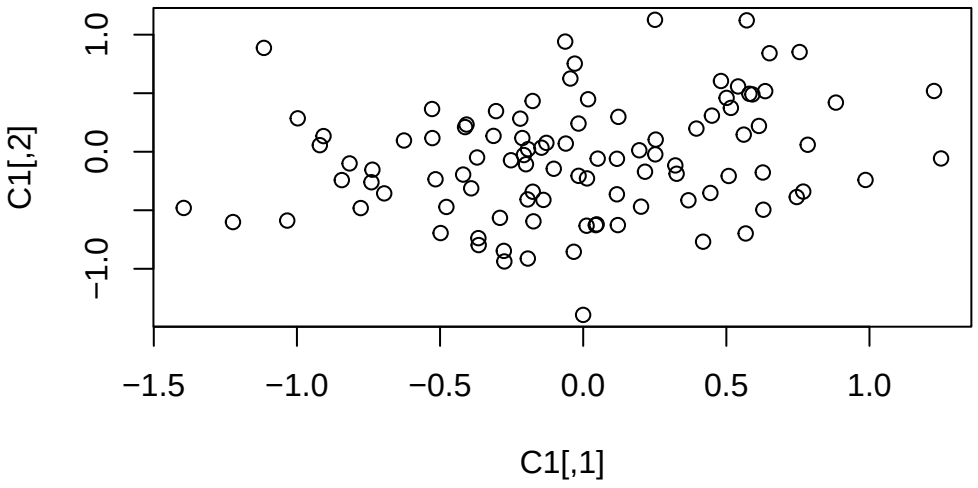
2 Gráficos de dispersión

Se generan dos series de datos normales:

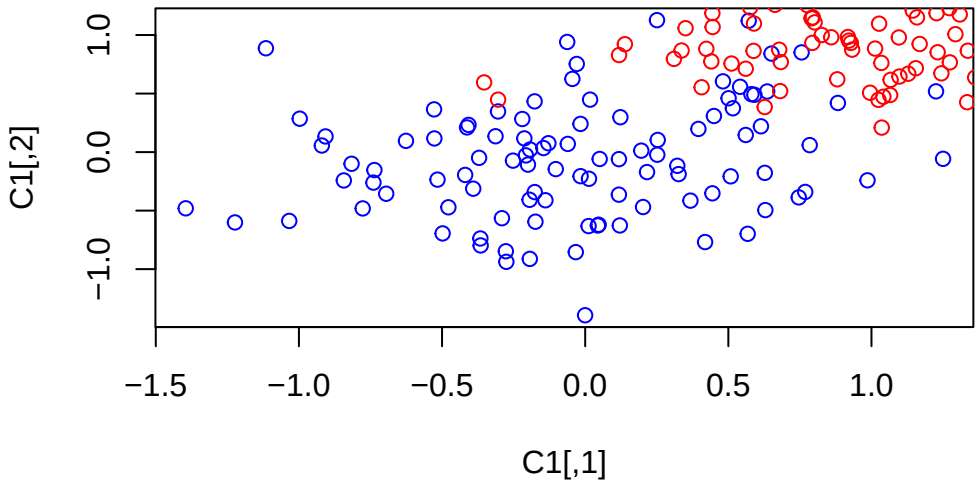
```r
C1 <- matrix(rnorm(200, sd = 0.5), ncol = 2)
C2 <- matrix(rnorm(200, mean = 1, sd = 0.5), ncol = 2)

# Se unen las matrices
mat <- rbind(C1, C2)
```

Se grafica C1

```r
plot(C1)
```

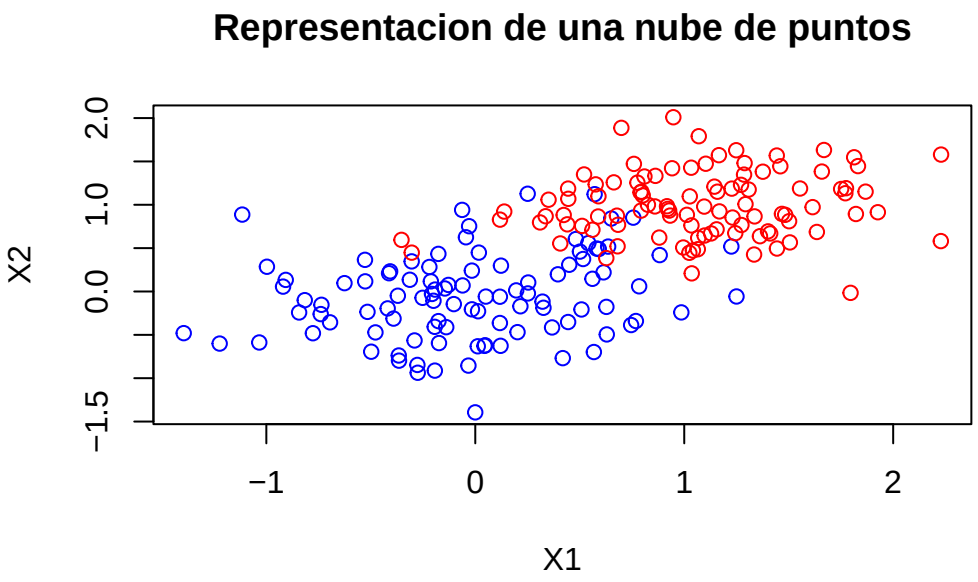


```r
plot(C1, col = "blue")
# Se añaden los puntos de C2
points(C2, col = "red")
```



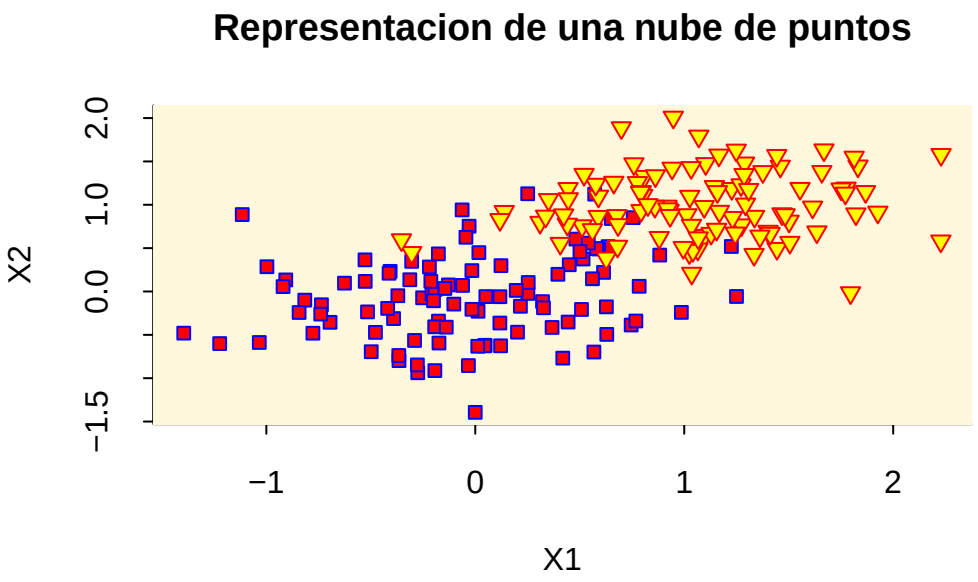
Se procede a ajustar el tamaño del gráfico

```r
plot(C1,
    col = "blue",
    xlim = range(mat[, 1]),
    ylim = range(mat[, 2]),
    main = "Representacion de una nube de puntos",
    xlab = "X1", ylab = "X2"
)
points(C2, col = "red")
```



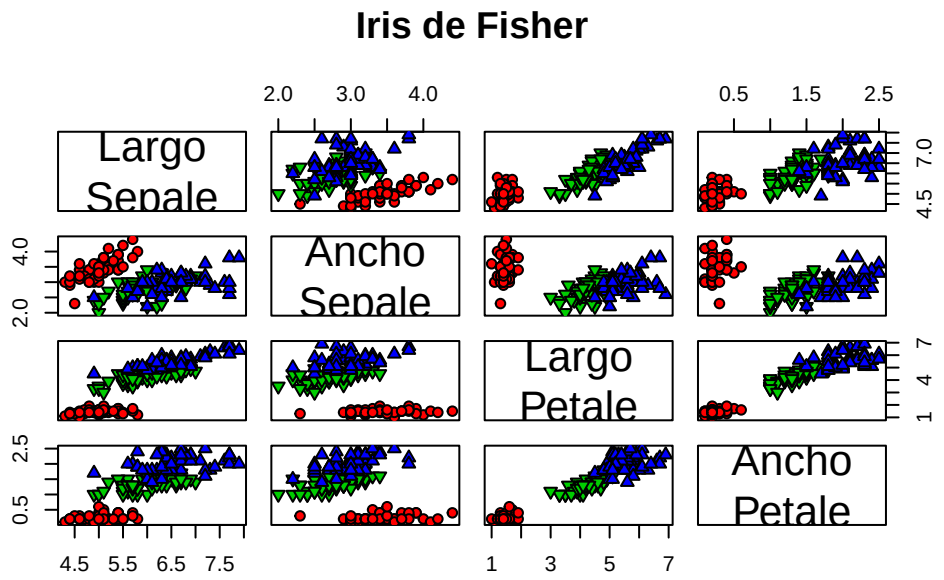
Se retoca el gráfico

```
plot(1,
    xlim = range(mat[, 1]),
    ylim = range(mat[, 2]),
    main = "Representacion de una nube de puntos",
    xlab = "X1", ylab = "X2",
    bty = "l", tcl = -.25
)
rect(-3, -3, 3, 3, col = "cornsilk")
points(C1, col = "blue", pch = 22, bg = "red")
points(C2, col = "red", pch = 25, bg = "yellow")
```



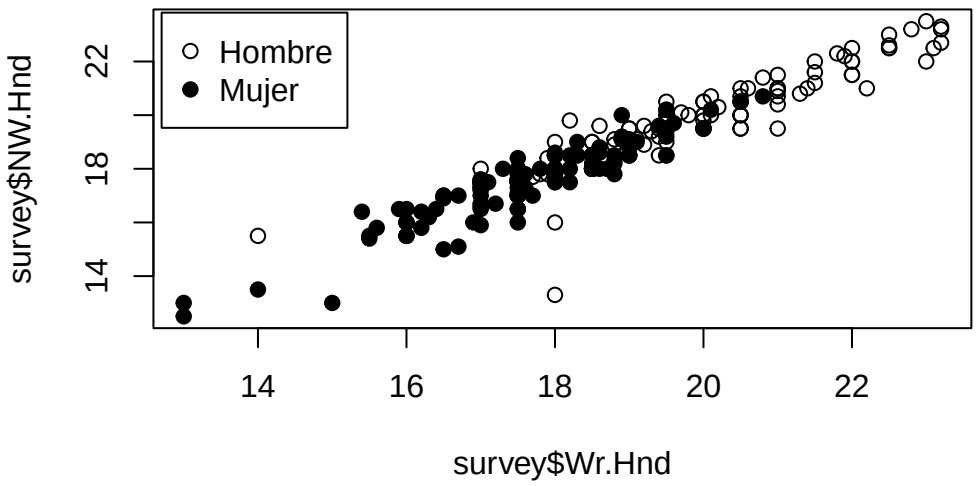
Se generan varios gráficos al mismo tiempo

```
plot(iris[, 1:4],
    bg = c("red", "green3", "blue")[iris[, 5]],
    pch = c(21, 25, 24)[iris[, 5]],
    main = "Iris de Fisher",
    labels =
        c(
            "Largo\nSepale",
            "Ancho\nSepale",
            "Largo\nPetale",
            "Ancho\nPetale"
        )
)
```



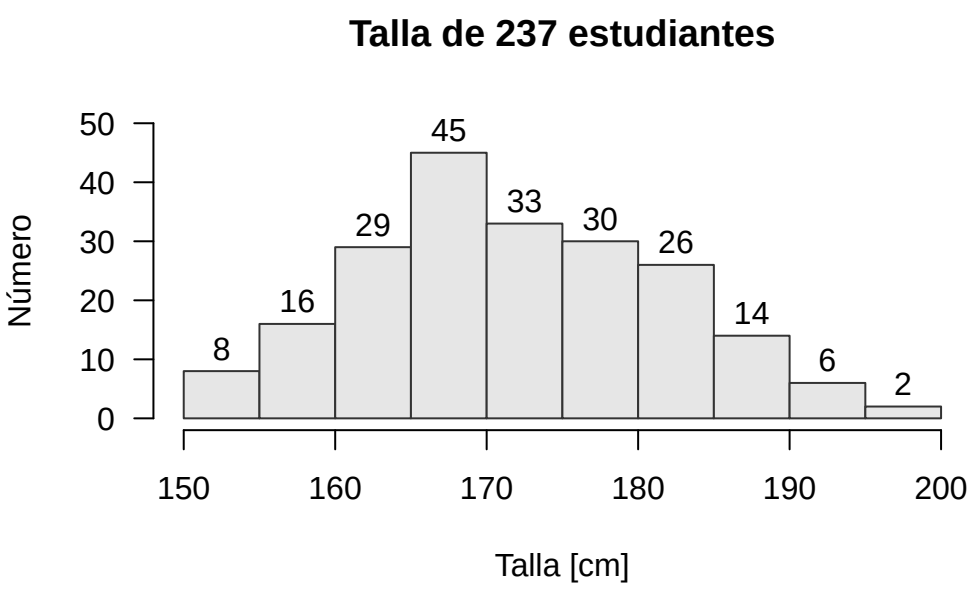
Se genera un gráfico de dispersión separando colores según sexo

```
plot(survey$Wr.Hnd, survey$NW.Hnd, pch = ifelse(survey$Sex == "Male", 1, 19))
legend("topleft", inset = 0.01, c("Hombre", "Mujer"), pch = c(1, 19))
```



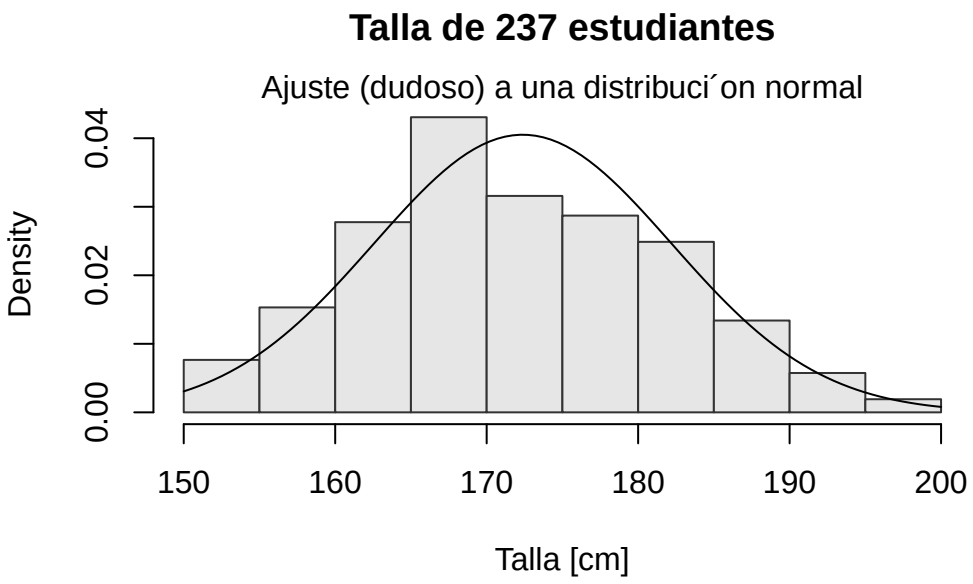
3 Creación de histogramas básicos

```r
hist(survey$Height,
    col = grey(0.9),
    border = grey(0.2),
    main = paste("Talla de", nrow(survey), "estudiantes"),
    xlab = "Talla [cm]",
    ylab = "Número",
    labels = TRUE,
    las = 1,
    ylim = c(0, 50)
)
```



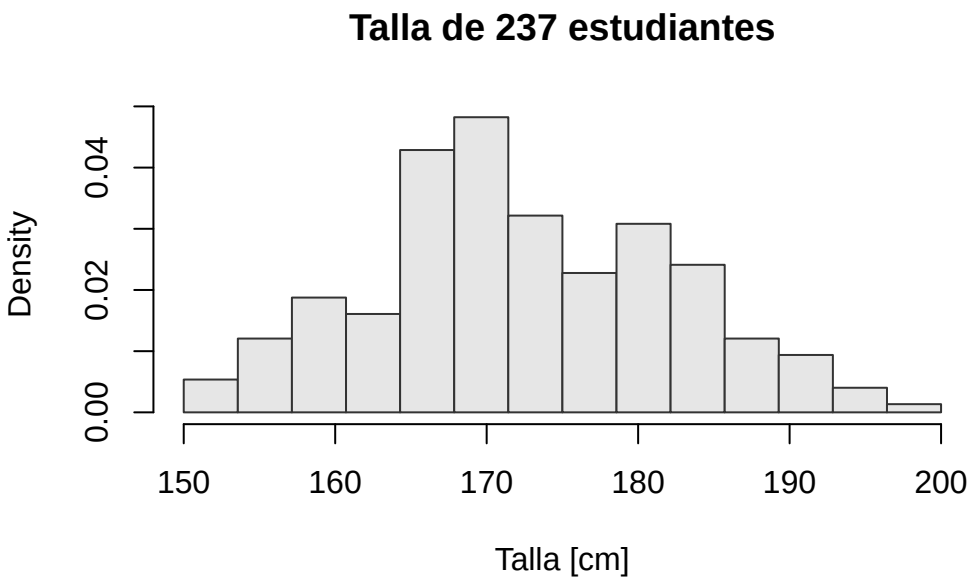
4 Frecuencias relativas, incluyendo ajuste a una curva normal

```
hist(survey$Height,
    col = grey(0.9), border = grey(0.2),
    main = paste("Talla de", nrow(survey), "estudiantes"),
    xlab = "Talla [cm]", proba = TRUE
)
x <- seq(
    from = min(survey$Height, na.rm = T),
    to = max(survey$Height, na.rm = T), length = 100
)
lines(x, dnorm(
    x, mean(survey$Height, na.rm = TRUE),
    sd(survey$Height, na.rm = TRUE)
))
mtext("Ajuste (dudoso) a una distribuci´on normal")
```



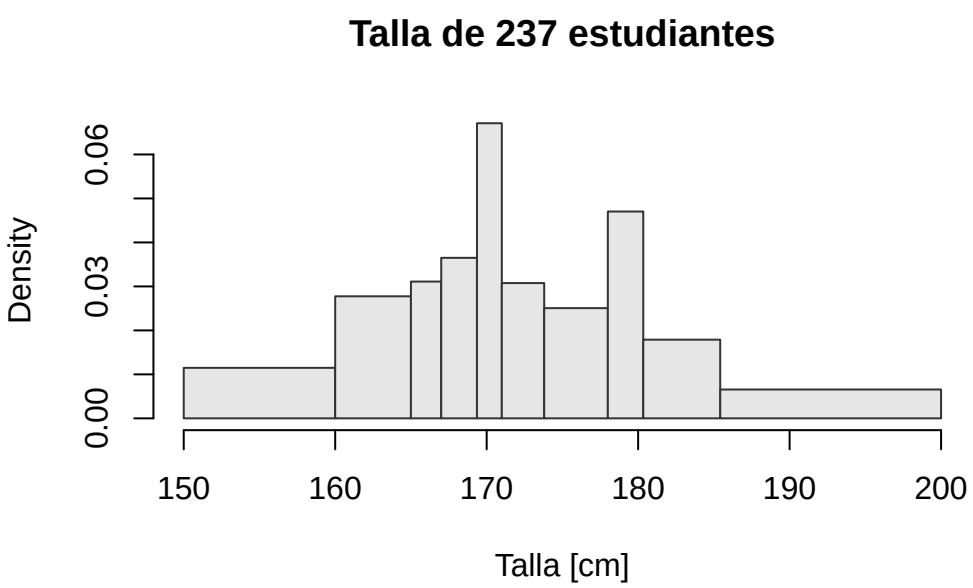
5 Control del ancho de los intervalos en los hitogramas

```r
hist(survey$Height,
    col = grey(0.9), border = grey(0.2),
    main = paste("Talla de", nrow(survey), "estudiantes"),
    xlab = "Talla [cm]",
    proba = TRUE,
    breaks = seq(from = 150, to = 200, length = 15)
)
```



6 Usando intervalosdel mismo efectivo

```
isohist <- function(x, nclass, ...) {
    breaks <- quantile(x,
        seq(from = 0,
        to = 1,
        length = nclass + 1),
        na.rm = TRUE)
    invisible(hist(x, breaks = breaks, ...))
}
isohist(survey$Height, 10,
    col = grey(0.9), border = grey(0.2),
    main = paste("Talla de", nrow(survey), "estudiantes"),
    xlab = "Talla [cm]", proba = TRUE
)
```


Talla de 237 estudiantes
Density
0.00
0.03
0.06
150
160
170
180
190
200
Talla [cm]