Práctica 2 MEE - Muestreo y diseño de experimentos

1 - Utilizando estos datos, realizar un análisis descriptivo que permita comparar el nivel de colesterol en la muestra de hombres y en la de mujeres. Redactar una breve síntesis con las conclusiones de este análisis aportando los datos que consideréis relevantes.

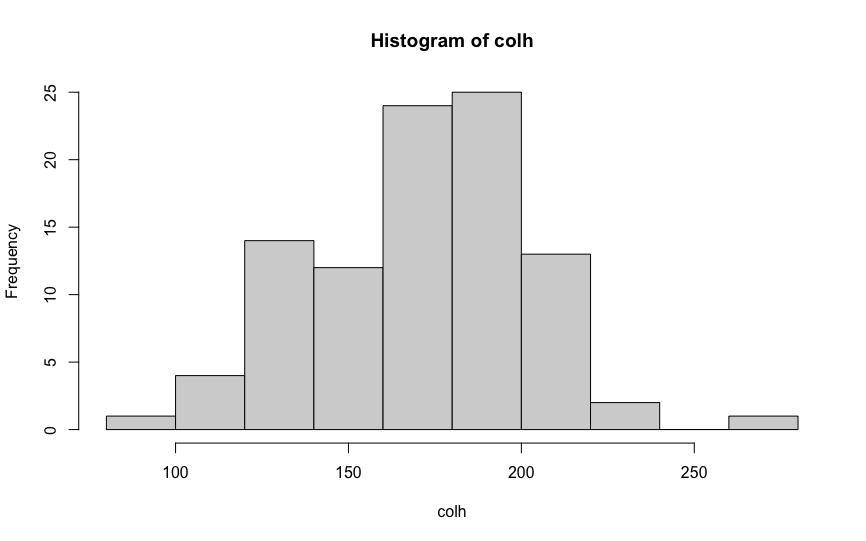
Objetivos: comparar el nivel de colesterol de hombres y mujeres en un mismo rango de edad

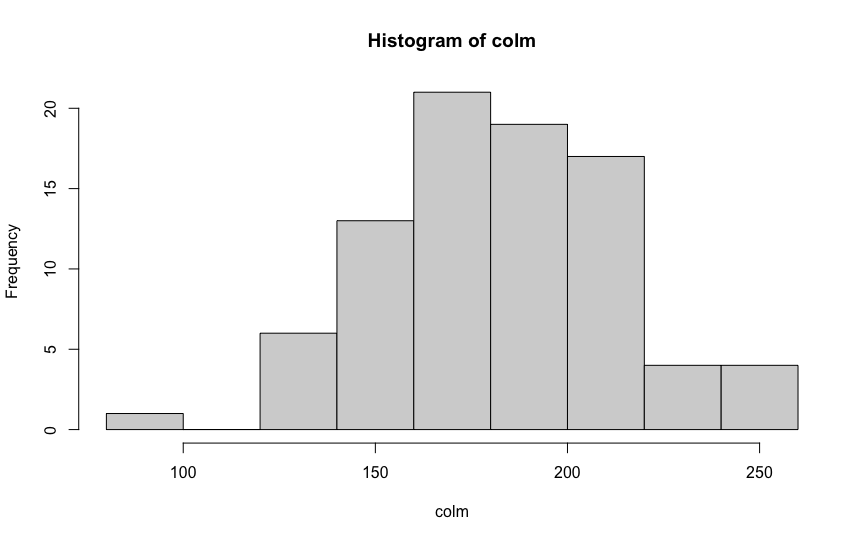
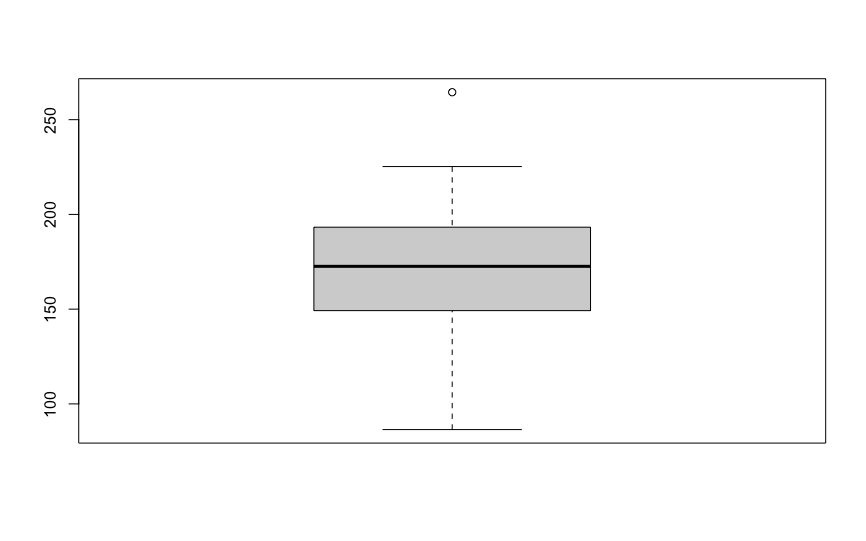
|  |  |
| --- | --- |
| summary(colmujeres)  colm  Min. : 95.98  1st Qu.:160.13  Median :182.32  Mean :181.09  3rd Qu.:202.80  Max. :255.91 | summary(colhombres)  colh  Min. : 86.44  1st Qu.:149.56  Median :172.60  Mean :170.81  3rd Qu.:193.09  Max. :264.49 |

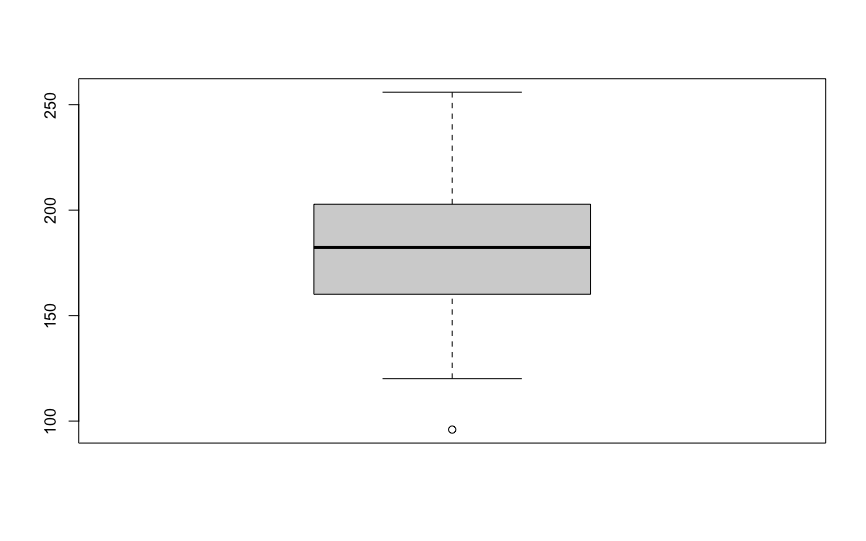
|  |  |
| --- | --- |
| var(colhombres)  colh  colh 933.5336 | var(colmujeres)  colm  colm 948.4498 |

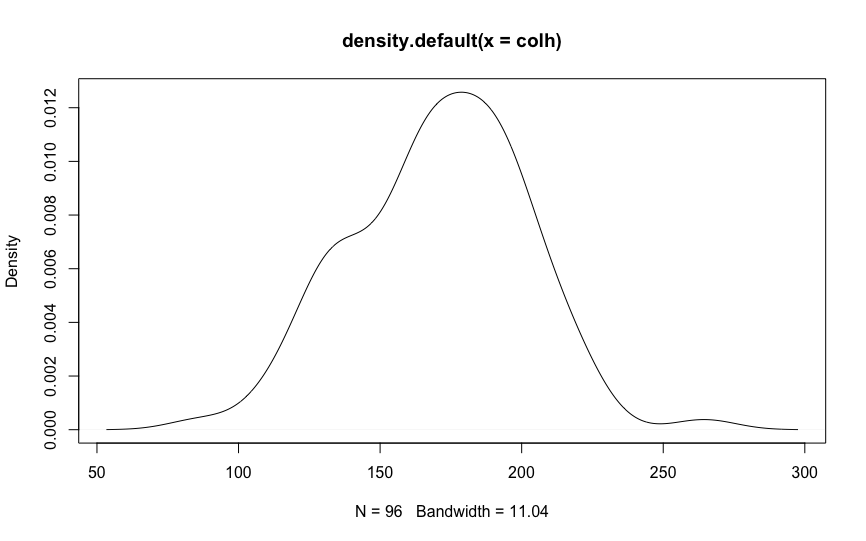
DESVIACIÓN TÍPICA:

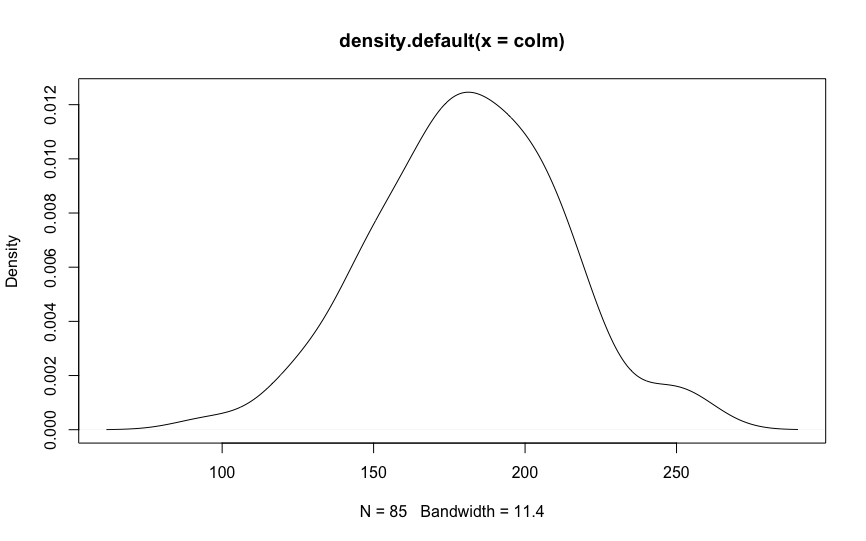
|  |  |
| --- | --- |
| sd(colh)  [1] 30.55378 | sd(colm)  [1] 30.79691 |

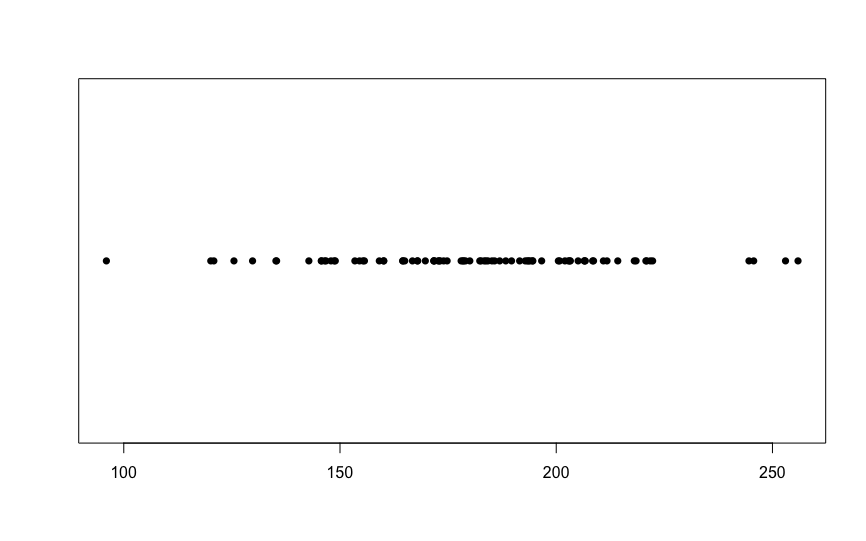
HOMBRES:

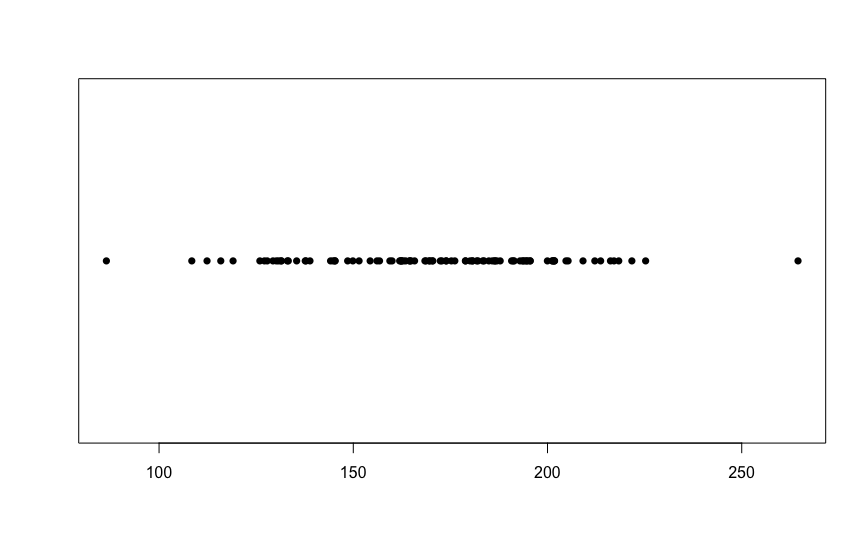
MUJERES:HOMBRES:

MUJERES:DENSIDAD EMPÍRICA:

HOMBRES:

MUJERES:GRÁFICA DE PUNTOS:

MUJERES:

HOMBRES:

2. Con estas premisas:

a) Identificar las poblaciones implicadas en el estudio:

Una población la constituyen aquellos hombres y mujeres de un determinado rango de edad y otra población la constituyen aquellas mujeres de un determinado rango de edad.

b) Muestras utilizadas:

Se ha tomado una muestra (de tamaño n=96) de hombres de entre 20 y 29 años de edad y otra muestra (de tamaño n=85) de mujeres de entre 20 y 29 años de edad.

c) Variables aleatorias:

nivel de colesterol total en sangre de cada persona

d) Parámetro más relevante de la población sobre el que se desea información y estadístico muestral utilizado para estimarlo:

Nivel MEDIO de colesterol en sangre en cada población. Se trata de una muestra de voluntarios, no probabilístico

e) El método de muestreo utilizado para seleccionar la muestra de individuos ¿es adecuado? ¿por qué?

No es adecuado, porque al ser una muestra de voluntarios, no es representativo, pues presenta cierto sesgo que puede favorecer que solo aquellos individuos que crean tener problemas de colesterol se presenten voluntarios.

3. distribución Normal con µ = 171 mg/dl y σ = 30 mg/dl:

a) Obtener el porcentaje de varones en la población dentro de este rango de edad, que tendrá un nivel de colesterol total en sangre superior a 240 mg/dl. (Utilizar R para obtener la probabilidad, y comprobar que el resultado es similar al obtenido mediante tablas. Marcar el área correspondiente en un gráfico y utilizar una notación matemática coherente).

pnorm(240,mean=171,sd=30,lower.tail=F)

[1] 0.01072411

b) Completar las siguientes frases, justificando la respuesta:

En la población, un 15 % de los varones con edades comprendidas entre 20 y 29 años presenta un nivel de colesterol total en sangre inferior a 139.907 mg/dl.

qnorm(0.15,mean=171,sd=30)

[1] 139.907

En la muestra, el 15 % de varones con edades comprendidas entre 20 y 29 años presenta un nivel de colesterol total en sangre inferior 139.907 a mg/dl.

4. Este estudio, ¿es observacional o un experimento? Justifica tu respuesta.

Se trata de un experimento, puesto que se ha dividido la muestra en dos grupos y se ha provado sobre cada individuo un tratamiento para observar su respuesta, Se trata de determinar la causa de un determinado efecto; en este caso la aplicación de música antes de la operación, y la aplicación de una droga antes de la operación.

¿Cuántos factores se han estudiado? 1, tipo de ayuda que reciben ¿Cuántos tratamientos? 2, aplicación de música y aplicación de droga ansiolítica

Identifica la variable respuesta. ¿Es una variable categórica o cuantitativa?

Puntuación obtenida en el test STAI que mide el nivel de ansiedad de cada paciente. La variable es numérica, variable cuantitativa

Realiza el esquema del diseño que se ha utilizado. El esquema debe mostrar los diferentes tratamientos, el número de sujetos en cada tratamiento y el modo en que éstos fueron asignados a los mismos.

(1) Tomar midazolam | (2) Escuchar música

:----------------:|:----------------:

186 pacientes | 186 pacientes

Si a los pacientes se les asigna un número entre 1 y 372, ¿qué comando de R utilizarías para asignar totalmente al azar estos pacientes a los tratamientos de forma que el diseño sea equilibrado?

Cojo 186 individuos con sample de la muestra(1:372) y los asigno a música:

Música <- sample(1:372,186,replace=F) y el resto a droga

O bien:

tratamientos=rep(1:2, each=186)

sample(tratamientos)

¿Cuántos pacientes no completaron el protocolo de este estudio? 9 en el grupo que escuchó música relajante y 36 en el grupo que tomó midazolam.

¿Piensas que la música puede ayudar a relajarte? Prueba con esta canción especialmente compuesta para ello: Weightless del grupo Marconi Union.

5. Utiliza los comandos apropiados de R para obtener las siguientes probabilidades y percentiles (redondea todos los resultados a 3 cifras decimales):

a) P(N(8, 3) < 6) = pnorm(6,mean=8,sd=3)

[1] 0.2524925

0.252

b) P(N(8, 3) > 6) = pnorm(6,mean=8,sd=3,lower.tail=F)

[1] 0.7475075

0.748

c) P(−1 < N(0, 1) < 2) = round((pnorm(2,mean=0,sd=1,lower.tail = TRUE)-pnorm(-1,mean=0,sd=1,lower.tail = TRUE)),3)

[1] 0.819

d) P(5 < N(8, 2) < 6) = round((pnorm(6,mean=8,sd=2,lower.tail = TRUE)-pnorm(5,mean=8,sd=2,lower.tail = TRUE)),3)

[1] 0.092

e) Valor z/P(N(0, 1) < z) = 0,95 z = qnorm(0.95,mean=0,sd=1)

[1] 1.644854

f ) Valor z/P(N(0, 1) > z) = 0,95 z = qnorm(0.95,mean=0,sd=1,lower.tail=F)

[1] -1.644854

g) Valor x/P(N(100, 9) > x) = 0,25 x = qnorm(0.25,mean=100,sd=9,lower.tail=F)

[1] 106.0704

h) Si X ∼ N(µ, σ = 5) y P(X > 20) = 0,85, el valor de la media será (Nota: este valor no puede obtenerse directamente con ningún comando de R. Realiza los cálculos que consideres para obtener el resultado.)

µ=25.182165