

Laboratorio 2

Entrega: 4 de setiembre de 2023; 14:00

Instrucciones: trabaje las siguientes instrucciones en R, creando uno o varios scripts. Entregue los scripts y genere un archivo Word o L^AT_EX con las respuestas.

1. Datos europeos.

Cargar tabla de datos europa.csv con:

```
> europa <- read.table("europa.csv", dec=".", sep=";", header = T)
```

Verifiquemos los datos:

```
> names(europa)
> head(europa)
> str(europa)
> summary(europa)
> range(europa$Duracion.horas.)
> sd(europa$Duracion.horas.)
```

Algunos gráficos:

```
> plot(europa)
> boxplot(europa$Duracion.horas., ylab="Duración (horas)")
> points(1, mean(europa$Duracion.horas.), pch=2)
```

pch es una opción gráfica que define el símbolo que representa a las observaciones.

2. Considere la tabla de los Iris de Fisher en la librería MASS.

- (a) Dé las dimensiones de la tabla, los nombres de las variables y el tamaño de la muestra.
- (b) Para cada variable cuantitativa, calcule la media, la mediana, la desviación estándar, los cuartiles, el máximo, el mínimo, el rango.
- (c) Haga un solo gráfico con las 3 cajas de dispersión (boxplot) de la longitud del sépalos por cada especie de iris.
- (d) Haga un gráfico de contornos entre la longitud y el ancho del pétalo, con:

```
densidad <- kde2d(iris$Petal.Length, iris$Petal.Width)
filled.contour(densidad, color = topo.colors, xlab = "Longitud del pétalo",
+ ylab = "Ancho del pétalo")
```

3. Manipulación de matrices:

- (a) Creación de matrices:

```
# Creación de matrices
> m1 <- matrix(1, nr = 2, nc = 2)
> m2 <- matrix(2, nr = 2, nc = 2)
# Concatenación de matrices por filas
> m <- rbind(m1, m2)
> m
# Concatenación de matrices por columnas
> cbind(m1, m2)
```

(b) Usando muestreo con repetición:

```
> a=sample(1:10,30,rep=T)
> a=matrix(a,nrow=6)
# Anulación de elementos menores que 3
> a[a<3]=0
# Extracción de la diagonal
> diag(a)
# Creación de la identidad 5x5
> diag(5)
```

(c) Suma de elementos de la matriz:

```
> sum(m1)
#Suma por columnas
> apply(m,2,sum)
#Suma por filas
> apply(m,1,sum)
```

(d) Operaciones matriciales:

```
# Suma
> m1+m2
# Resta
> m1-m2
# Producto
> m1%%m2
# Producto componente por componente
> m1*m2

# Transpuesta
> a
> t(a)
```

(e) Construya la matriz Z siguiente:

$$Z = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$$

Haga lo siguiente:

- El elemento de Z contenido en la primera fila y tercera columna.
- La primera fila de Z .
- La tercera columna de Z .
- La submatriz después de haber quitado la primera fila y la primera columna de Z .

- Calcule la transpuesta de Z .
- Determine la submatriz al quitar la segunda columna de Z , calcule la inversa, si ésta existe. Verifique que es efectivamente la inversa.

4. Instalar el paquete: `Losck5Data` y cargarlo:

```
> install.packages("Losck5Data")
> library(Losck5Data)
```

Cargar la tabla de datos `AllCountries`:

```
> data(AllCountries)
```

- Escriba una función R que tome una muestra de algunos de los países en la tabla de datos, y que calcule la media de la superficie (`LandArea`) en Km cuadrados.
 - Use esa función para tomar una muestra de 10 países y reporte el resultado.
 - Haga la caja de dispersión (`boxplot`) de la variable de población.
 - Haga el gráfico de dispersión entre las variables superficie y población.
- Considere el ejercicio 6 del Capítulo 1 del libro. Construya la tabla de datos con dos variables (sexo, profesión). A partir de ella construya, con instrucciones de R, la tabla disyuntiva completa.
 - Considere la tabla de contingencia `HairEyeColor` que cruza el color del cabello con el color de los ojos de un grupo de personas. Calcule el χ^2 de asociación y el T^2 de Chuprov entre las dos variables.
 - Considere la tabla de los Iris de Fisher. Esta tabla contiene 4 variables cuantitativas y 1 variable cualitativa. Calcule el cociente de correlación η entre cada una de las variables cuantitativas (`Sepal.Length` `Sepal.Width` `Petal.Length` `Petal.Width`) con la variable cualitativa (`Species`).