

GUÍA TEÓRICO PRÁCTICA 4.

Parte 1 Estimación de densidades y boxplot

La función **density** le permite obtener un estimador de densidad de un conjunto de datos. Utilice el **help**.

1- Para los datos correspondientes a las variables `faithful$waiting` y `faithful$eruptions` obtenga estimadores de densidad, con distintos tamaños de ventanas y grafíquelos. Pruebe modificar las distintas opciones del comando `density`.

2- Superponga la curva de densidad a los histogramas construidos en la práctica anterior.

La función **boxplot** le permite obtener un boxplot o “gráfico de caja” de un conjunto de datos. Utilice el **help**.

3- Realice boxplots para las variables `faithful$eruptions`, `abbey` (de la library (MASS)) `mdeaths`, y `fdeaths`.

Cuando corresponda identifique los outliers.

4- Obtenga una muestra de tamaño 100 de una distribución normal estándar.

Para la muestra obtenida grafique el histograma con la curva de densidad superpuesta y el boxplot.

5- Repita lo pedido en el ejercicio anterior para una t con 10 y 20 de grados de libertad, y para una χ^2 con 10 y 20 grados de libertad.

Parte 2 - Funciones

1- Las siguientes funciones toman como argumento un vector y devuelven un número como respuesta. Utilice el **help** y realice sus propios ejemplos:

```
max() min() var() mean() sum() prod()
```

a) Genere un vector numérico con algunos valores anómalos. Reemplace los valores anómalos por NA. ¿Qué ocurre con las funciones anteriores si se aplican a vectores numéricos que tienen uno ó más valores faltantes (NA)?

b) Muestre dos maneras distintas de aplicar las funciones a los valores no-faltantes.

Método 1) Eliminando los elementos con valores faltantes (`na.exclude`).

Método 2) Utilizando argumentos de las funciones.

c) ¿Qué ocurre si se aplica una de estas funciones a un vector de datos lógicos?

2- Escribiendo funciones

```
mifuncion <- function(argumento1,argumento2,...){  
  primera instrucción  
  segunda instrucción  
  etc.  
  objeto que devuelve la funcion }
```

a) Abra una ventana de escritura (script) en R o S-plus para escribir sus funciones

b) El **desvío estándar muestral s** de un conjunto de datos x_1, x_2, \dots, x_n , está definido como la raíz cuadrada de la **varianza muestral**:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Escriba la siguiente función que calcula el desvío estándar utilizando la función var().

```
# desvío estándar  
  
std.dev <- function(x) sqrt(var(x))
```

Ejecute la función: `std.dev(1:5)`

¿Cuántas instrucciones tiene la función `std.dev`? ¿Cuántos argumentos?

c) Genere un vector (w) de longitud 150, con valores pseudo aleatorios Normales con media 5 y desvío 2 y calcule su media y su desvío estándar. Repita varias veces y compare.

3- Escriba la función “medidas.resumen” que tiene como argumento una matriz de datos numéricos y calcula la media, la mediana y el desvío estándar de cada columna y devuelve como resultado una matriz de tres filas.

```
medidas.resumen <-function(x){  
  media<-apply(x,2,mean)  
  mediana<-apply(x,2,median)  
  desv<-sqrt(apply(x,2,var))  
  result<-rbind(media,mediana,desv)  
  dimnames(result)<-  
  list(c("media","mediana","desvíoestándar"),NULL )  
  result}  
}
```

a) Cree algunas matrices (

```
por ejemplo: A1 <- matrix(w,50,3);  
A2 <-  
cbind(rnorm(10),rnorm(10,5),rnorm(10,5,5))  
)
```

y obtenga las medidas resumen de cada columna

b) ¿Qué ocurre cuando aplica la función `medidas.resumen` si la matriz tiene algunos valores faltantes? Construya sus propias funciones `mean`, `median`, `var` de manera que calculen dichas medidas para un vector con valores faltantes. (vea ej 1 b2)). Modifique la función `medidas.resumen` y de manera que calcule las medidas resumen de las columnas inclusive cuando tienen algunos valores faltantes.