

Estadística Computacional

PRIMERA PRUEBA DE CONOCIMIENTOS

EVALUACIÓN CONTINUA, 21 DE ABRIL DE 2022

Apellidos, Nombre: _____

Cuestiones teóricas [Total: 2 puntos]

1. Desde el punto de vista del lenguaje en qué se diferencia un objeto de tipo matriz de un vector atómico. Proporciona un ejemplo sencillo de cómo pasar de un tipo a otro.
2. Explica brevemente la diferencia que hay entre los operadores [, [[y \$ cuando se aplican a objetos de tipo lista.
3. Observa el siguiente código y su ejecución:

```
> y <- 10
> f1 <- function(x = y) {
+   y<-1
+   return(c(x,y))
+ }
> f1()

[1] 1 1

> f2 <- function(x = y) {
+   x
+   y<-1
+   return(c(x,y))
+ }
> f2()

[1] 10 1
```

Razona por qué se obtiene distinto resultado al evaluar f1 y f2.

Ejercicios prácticos [Total: 8 puntos]

Ejercicio 1

[2 puntos]

Crea un vector **x** con los primeros 50 números impares. A partir de dicho vector:

- Calcula la media de sus elementos (**mx**) y la cuasidesviación típica (**sx**). [Utiliza las funciones **mean** y **sd** para ello.]
- Calcula el número de elementos en **x** que disten de **mx** más de **sx** unidades.
- Sustituye los elementos en **x** que has localizado en el apartado anterior por valores perdidos.
- Cuenta cuántos múltiplos de 3 hay en el vector **x** resultante del apartado anterior.

Ejercicio 2

[3 puntos]

El conjunto de datos `airquality` del paquete `datasets` contiene datos diarios relativos a la calidad del aire en Nueva York tomados entre mayo y septiembre de 1973. Se trata de un data frame que contiene 6 columnas: concentración de ozono (`Ozone`), radiación solar (`Solar.R`), velocidad del viento (`Wind`), temperatura (`Temp`), y el mes (`Month`) y día (`Day`) donde se tomaron las mediciones. Considerando estos datos escribe código apropiado (lo más simple posible) para realizar las siguientes tareas:

1. Crea una copia del conjunto de datos en el espacio de trabajo con nombre `aire`, e imprime la estructura del objeto de datos.
2. Cuenta el número de datos perdidos que hay en cada columna del objeto `aire`.
3. Elimina todas las filas del objeto `aire` que contienen algún dato perdido. ¿Cuántas filas has eliminado?
4. Convierte la columna `Month` en un factor y asigna como nombre a sus niveles el del mes correspondiente.
5. Calcula las medianas de las columnas `Wind` y `Ozone` para cada uno de los meses considerados, esto es, para cada uno de los niveles del factor creado en el apartado anterior. [Para calcular una mediana puedes usar la función `median`.]
6. Crea un data frame (`aire.mayo`) con los datos de `aire` correspondientes al mes de mayo.

Ejercicio 3

[3 puntos]

Crea una función en R que permita calcular los n primeros términos de la progresión geométrica $a_n = a_1 r^{n-1}$, para valores n , a_1 y r dados como argumentos, y devolver un objeto de tipo lista con los siguientes resultados:

- El vector `v` con los n términos calculados.
- El resultado de la suma de sus elementos (`suma1`) usando la función `sum`.
- El resultado de la suma de sus elementos (`suma2`) usando la expresión explícita: $a_1(1 - r^n)/(1 - r)$.
- El producto de sus elementos (`producto1`) usando la función `prod`.
- El producto de sus elementos (`producto2`) usando la expresión explícita: $(\sqrt[n]{a_1^2 r^{n-1}})^n$ en el caso de $a_1, r > 0$, y `NA` en otro caso.

Evalúa la función primero en $n = 20$, $a_1 = 2$ y $r = -0.5$ y luego en $n = 20$, $a_1 = 2$ y $r = 0.5$ (imprime el resultado en ambos casos).

Notas:

- Se valorará que la función contenga filtros que controlen que el usuario proporciona valores adecuados para los argumentos, generando mensajes de error o advertencia en caso contrario.
- También se valorará que la programación de la función sea lo más simple y eficiente posible. Puedes incluir algún comentario breve justificando la estrategia de programación seguida.