

Estadística Computacional

EJERCICIOS DE REPASO DE LOS TEMAS 2 Y 3

Ejercicio 1

Crea un vector **x** con 100 números aleatorios desde una distribución uniforme en el intervalo (0,1), fijando la semilla de generación con `set.seed(1)`. A partir de dicho vector:

- Calcula la media de sus elementos (**mx**), localiza en qué posición se encuentra el elemento del vector más cercano a dicha media e imprímelo.
- Calcula el número de elementos en **x** que están por debajo de **mx**.
- Elimina los elementos en **x** que has localizado en el apartado anterior.
- Crea una matriz (**A**) con dos columnas, la primera es el vector **x**, y la segunda el vector que contenga la distancia en valor absoluto de cada elemento de **x** a **mx**.

Ejercicio 2

El fichero *hatco.txt* en PRADO contiene datos de un estudio de mercado relativos a 100 empresas, clientes de un gran distribuidor industrial, en relación a sus percepciones sobre un producto determinado. Cada fila corresponde a una empresa y las columnas son distintas variables observadas cuyos nombres se recogen en la primera fila del fichero. A partir de dichos datos se pide realizar las siguientes tareas:

1. Lee los datos en R y almacénalos en un data frame con nombre **hatco**.
2. La última columna (**x8**) del fichero corresponde a una variable categórica indicando si se trata de una empresa pequeña (valor 0) o grande (valor 1). Convierte esta columna en el data frame que has creado en un factor asignándole los nombres correspondientes a sus niveles.
3. La primera columna es una etiqueta que identifica el cliente (la empresa). Transforma esta columna en un vector de tipo carácter.
4. Cuenta cuántas empresas hay de cada tipo (pequeña o grande) e imprime los datos de una de cada tipo elegida al azar.
5. La novena columna (**y**) es un indicador del nivel de fidelidad de la empresa. Calcula la nivel de fidelidad medio de todas las empresas. Calcula también la media para las empresas pequeñas y grandes.
6. Cuenta cuántas empresas pequeñas tienen un nivel de fidelidad por encima de su media. Haz lo mismo para las empresas grandes.
7. Crea un segundo data frame (**hatco2**) con las columnas numéricas de **hatco** tipificadas.

Ejercicio 3

La progresión aritmética es una sucesión definida recurrentemente a partir del primer elemento a_1 , y la diferencia d , mediante la relación:

$$a_{n+1} = a_1 + d * n$$

Crea una función en R que permita calcular los n primeros términos de la progresión aritmética a_n , 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 (`suma`) usando la función `sum`.
- El producto de sus elementos (`producto`) usando la función `prod`.

Crea una segunda versión de la función que incluya un argumento adicional (`explicit`) de tipo lógico con valor `FALSE` por defecto. Cuando tome el valor `TRUE` el cálculo de la suma y el producto de los elementos de la progresión se realizará usando las siguientes expresiones explícitas:

$$\sum_{i=1}^n a_i = n \frac{a_1 + a_n}{2} \quad \prod_{i=1}^n a_i = d^n \frac{\Gamma(a_1/d + n)}{\Gamma(a_1/d)}$$