

Tarea 13

Imanol

26/2/2021

Lee atentamente las transparencias del tema 9 e implementa las reglas que se comentan en los ejercicios

```
crabs = read.table("../data/datacrab.txt", header = TRUE)
str(crabs)
```

```
## 'data.frame': 173 obs. of 6 variables:
## $ input : int 1 2 3 4 5 6 7 8 9 10 ...
## $ color : int 3 4 2 4 4 3 2 4 3 4 ...
## $ spine : int 3 3 1 3 3 3 1 2 1 3 ...
## $ width : num 28.3 22.5 26 24.8 26 23.8 26.5 24.7 23.7 25.6 ...
## $ satell: int 8 0 9 0 4 0 0 0 0 0 ...
## $ weight: int 3050 1550 2300 2100 2600 2100 2350 1900 1950 2150 ...
```

```
cw = crabs$width # Cogemos solo la columna de tamaños
```

Yo

Preguntas de esta tarea

1. Da el algoritmo para reproducir el proceso de generación de clases y sus marcas respectivo para el número de clases obtenido con la regla de la Scott en Python.

Esto esta en Jupyter Notebook Estadística

2. Da el algoritmo para reproducir el proceso de generación de clases y sus marcas respectivo para el número de clases obtenido con la regla de la raíz en R o en Python.

Paso 1 Numero de Intervalos

```
n = length(cw)
k1 = ceiling((sqrt(n)))
k1
```

```
## [1] 14
```

Paso 2 Amplitud

```
A = diff(range(cw)) / k1 # Diferencia del rango de intervalo dividido por el numero de intervalos
A
```

```
## [1] 0.8928571
```

```
A = 0.9 # Se redondea a una cifra decimal por ser la precision de los datos del data frame
```

Paso 3 Extremos

```
# Como nos pide tener 14 intervalos debemos hacer de L1 a L15
```

```
L1 = min(cw)-1/2*0.1 # Valor mínimo de la variable dividido la mitad de la precisión
```

```
L1
```

```
## [1] 20.95
```

```
L = L1 + A*(0:k1) # Asi calculamos todas las L desde 0 hasta 14
L
```

```
## [1] 20.95 21.85 22.75 23.65 24.55 25.45 26.35 27.25 28.15 29.05 29.95 30.85
## [13] 31.75 32.65 33.55
```

Paso 4 Marcas de clase

```
x1 = (L[1]+L[2])/2 # Es el punto medio de cada una de los intervalos
x1
```

```
## [1] 21.4
```

```
x = x1 + A*(0:k1-1) # Asi calculamos todas las x desde 0 hasta 13
x
```

```
## [1] 20.5 21.4 22.3 23.2 24.1 25.0 25.9 26.8 27.7 28.6 29.5 30.4 31.3 32.2 33.1
```

3. Da el algoritmo para reproducir el proceso de generación de clases y sus marcas respectivo para el número de clases obtenido con la regla de la Sturges en R o en Python

Paso 1 Numero de Intervalos

```
k1 = nclass.Sturges(cw)
k1
```

```
## [1] 9
```

Paso 2 Amplitud

```
A = diff(range(cw)) / k1 # Diferencia del rango de intervalo dividido por el numero de intervalos
A
```

```
## [1] 1.388889
```

```
A = 1.4 # Se redondea a una cifra decimal por ser la precision de los datos del data frame
```

Paso 3 Extremos

```
# Como nos pide tener 9 intervalos debemos hacer de L1 a L10
L1 = min(cw)-1/2*0.1 # Valor mínimo de la variable dividido la mitad de la precisión
L1
```

```
## [1] 20.95
```

```
L = L1 + A*(0:k1) # Asi calculamos todas las L desde 0 hasta 9
L
```

```
## [1] 20.95 22.35 23.75 25.15 26.55 27.95 29.35 30.75 32.15 33.55
```

Paso 4 Marcas de clase

```
x1 = (L[1]+L[2])/2 # Es el punto medio de cada una de los intervalos
x1
```

```
## [1] 21.65
```

```
x = x1 + A*(0:k1-1) # Asi calculamos todas las x desde 0 hasta 8
x
```

```
## [1] 20.25 21.65 23.05 24.45 25.85 27.25 28.65 30.05 31.45 32.85
```

4. Da el algoritmo para reproducir el proceso de generación de clases y sus marcas respectivo para el número de clases obtenido con la regla de la Freedman-Diaconis en R o en Python.

Paso 1 Numero de Intervalos

```
k1 = nclass.FD(cw)
k1
```

```
## [1] 13
```

Paso 2 Amplitud

```
A = diff(range(cw)) / k1 # Diferencia del rango de intervalo dividido por el numero de intervalos
A
```

```
## [1] 0.9615385
```

```
A = 1.0 # Se redondea a una cifra decimal por ser la precision de los datos del data frame
```

Paso 3 Extremos

```
# Como nos pide tener 13 intervalos debemos hacer de L1 a L14
L1 = min(cw)-1/2*0.1 # Valor mínimo de la variable dividido la mitad de la precisión
L1
```

```
## [1] 20.95
```

```
L = L1 + A*(0:k1) # Asi calculamos todas las L desde 0 hasta 13
L
```

```
## [1] 20.95 21.95 22.95 23.95 24.95 25.95 26.95 27.95 28.95 29.95 30.95 31.95
## [13] 32.95 33.95
```

Paso 4 Marcas de clase

```
x1 = (L[1]+L[2])/2 # Es el punto medio de cada una de los intervalos
x1
```

```
## [1] 21.45
```

```
x = x1 + A*(0:k1-1) # Asi calculamos todas las x desde 0 hasta 12
x
```

```
## [1] 20.45 21.45 22.45 23.45 24.45 25.45 26.45 27.45 28.45 29.45 30.45 31.45
## [13] 32.45 33.45
```