



UNIVERSIDAD
CATÓLICA DE
TEMUCO

DEPARTAMENTO DE
INGENIERÍA INFORMÁTICA
FACULTAD DE INGENIERÍA

Taller Formativo 8

Mario Castillo Sanhueza

Docente: Dr. Julio Rojas Mora

Departamento de Ingeniería Informática
Facultad de Ingeniería
Universidad Católica de Temuco

Agosto 18, 2022

1. Problemática.

Usando la base de datos disponible aquí referente a datos de predicción en fallas del corazón, construir la tabla de frecuencia usando la columna "platelets" (plaquetas).

2. Código

```
library(Hmisc)
library(TTR)
hfData <- read.csv("heart_failure_clinical_records_dataset.csv",
header = TRUE, sep = ",")
platelets <- hfData$platelets
dataFrc <- data.frame(Platelets=platelets)

# Numero de intervalos con la Regla de Sturges
#Number of intervals with Sturges Rule
n <- nrow(dataFrc)
m <- 1+ceiling(3.322*log10(n))

# Rango y amplitud usando Sturges calculando el maximo
#menos el minimo y luego dividiendo por la division de clases //
#Range and amplitude using Sturges calculating the maximum
#minus the minimum and then dividing by the division of classes
I <- max(dataFrc$Platelets)-min(dataFrc$Platelets)
C <- I/m

#Genera la secuencia de datos separados por el intervalo c
#we generate the sequence of data separated by the interval c
cuts <- seq(min(dataFrc$Platelets),max(dataFrc$Platelets), C)

#Definida PlateletsFact como variable categorica , con los datos
#y cortes calculados crear intervalos de 5 digitos
#we define PlateletsFact as a categorical variable , with
#the calculated data and cuts we create 5-digit intervals
dataFrc$PlateletsFact <- cut2(dataFrc$Platelets , cuts , digits=5)
```

```

#Creamos las marcas de clase calculando el promedio con
#un valor por la izuquierda y uno por la derecha
#We create the class marks by calculating the average with
#a value on the left and one on the right
MC <- runMean(cuts, 2)

#Descartamos el primer valor
#We discard the first value
MC<-MC[-1]

#Obtenemos los niveles en los que se dividen los datos
#We obtain the levels in which the data is divided
frcTable <- data.frame(Platelets=levels(dataFrc$PlateletsFact))

#Creamos y asignamos los valores ya calculados a la columna MC
# We create and assign the values already
#calculated to the MC column
frcTable$MC <- MC

#Calculamos la frecuencia absoluta
# Calculate the absolute frequency
frcTable$f <- table(dataFrc$PlateletsFact)

#Calculamos la frecuencia relativa
# Calculate the relative frequency
frcTable$fr <- frcTable$f/n

#Calculamos la frecuencia absoluta acumulada
# Calculate the cumulative absolute frequency
frcTable$F <- cumsum(frcTable$f)

#Calculamos la frecuencia relativa acumulada
# Calculate the cumulative relative frequency
frcTable$Fr <- cumsum(frcTable$fr)

frcTable

```

```

> frcTable
      Platelets      MC      f      fr      F      Fr
1  [ 25100,107590)  66345      9 0.030100334      9 0.03010033
2  [107590,190080) 148835     46 0.153846154     55 0.18394649
3  [190080,272570) 231325    136 0.454849498    191 0.63879599
4  [272570,355060) 313815     67 0.224080268    258 0.86287625
5  [355060,437550) 396305     27 0.090301003    285 0.95317726
6  [437550,520040) 478795      9 0.030100334    294 0.98327759
7  [520040,602530) 561285      2 0.006688963    296 0.98996656
8  [602530,685020) 643775      1 0.003344482    297 0.99331104
9  [685020,767510) 726265      1 0.003344482    298 0.99665552
10 [767510,850000] 808755      1 0.003344482    299 1.00000000
>

```