

# HW\_04\_JAIRO\_LEAL.R

jairo

2023-02-24

```
# ASIGNACIÓN 4: Boxplot e Histogramas
```

```
# Maestría en Ciencias Forestales UANL
```

```
# Alumno: Jairo Alberto Leal Gómez
```

```
# Matricula: 1723093
```

```
# Histogramas -----
```

```
# Problema 1 -----
```

```
# Considere el siguiente conjunto de datos x2 con 1000 números al azar.
```

```
set.seed(9875)
```

```
size <- 1000
```

```
x2 <- round(runif(n = size, min = 0, max = 10), 2)
```

```
head(x2)
```

```
## [1] 0.45 8.62 9.59 7.64 5.93 0.34
```

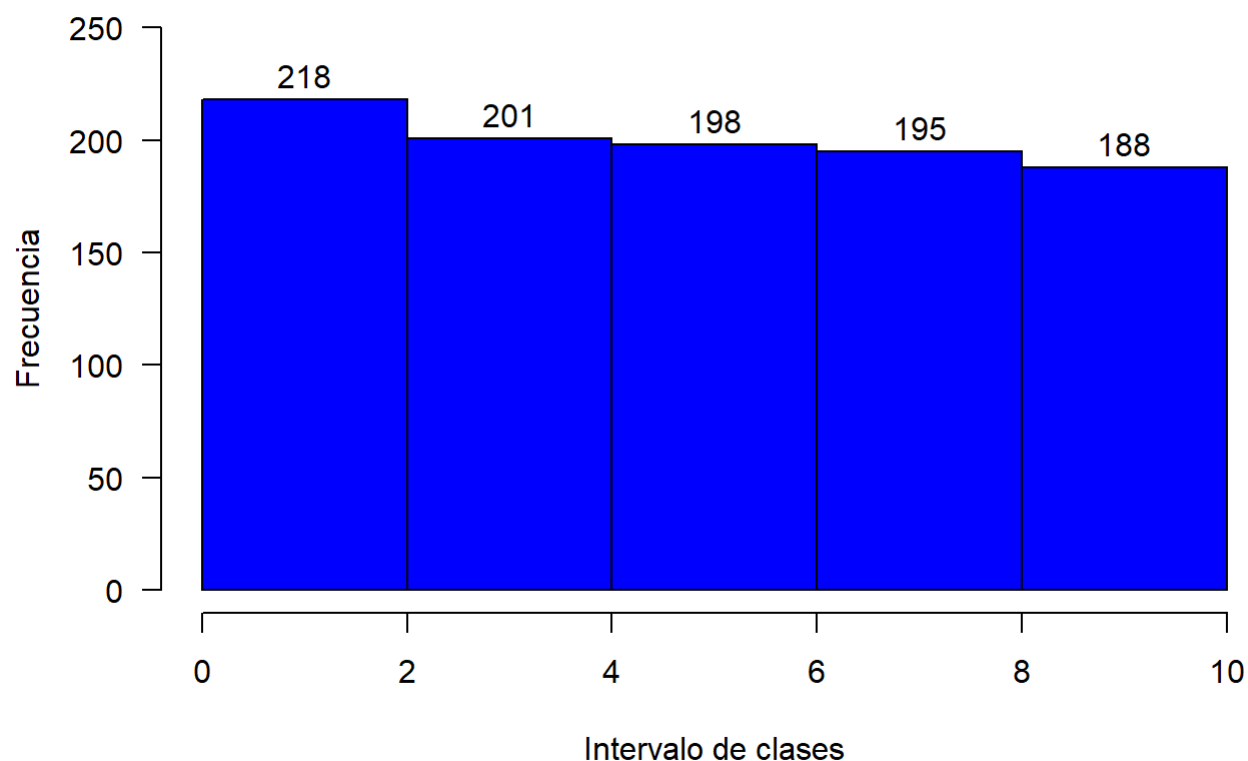
```
# Realiza dos histogramas con las siguientes características:
```

```
# 1. Incluir los siguientes intervalos de clase y personalizar con los puntos medios de cada intervalo de clase mids:
```

```
x <- hist(x2, breaks = c(0, 2, 4, 6, 8, 10), main = "Histograma", col = "blue",  
          xlab = "Intervalo de clases", ylab = "Frecuencia", las = 1, ylim = c(0, 250))
```

```
text(x$mids, x$counts, labels = x$counts, adj = c(0.5, -0.5))
```

## Histograma



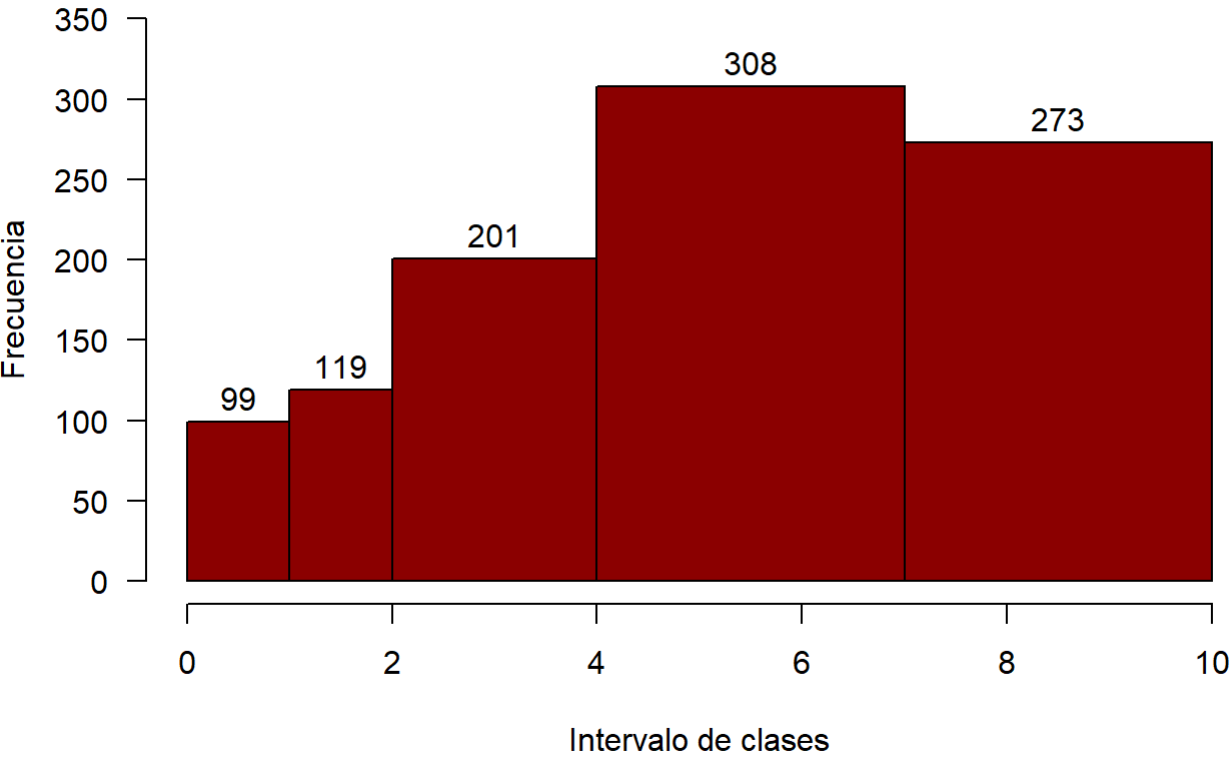
*# 2. Incluir los siguientes intervalos de clase y personalizar con los puntos medios de cada intervalo de clase mids.*

```
c <- hist(x2, breaks= c(0, 1, 2, 4, 7, 10), freq = TRUE, main = "Histograma", col = "red4",
  xlab = "Intervalo de clases", ylab = "Frecuencia", las =1, ylim = c(0,350))
```

```
## Warning in plot.histogram(r, freq = freq1, col = col, border = border, angle =
## angle, : the AREAS in the plot are wrong -- rather use 'freq = FALSE'
```

```
text(c$mids, c$counts, labels = c$counts, adj= c(0.5, -0.5))
```

Histograma



```
# Problema 2 -----
```

```
# La siguiente figura 1 presenta cuatro histogramas (A, B, C, D).
```

```
# a. ¿Cuál distribución parece estar sesgada a la derecha?
```

```
# Gráfica A
```

```
# b. ¿Cuál distribución parece estar sesgada a la izquierda?
```

```
# Gráfica D
```

```
# c. ¿Cuál distribución parece ser simétrica o en forma de “campana”?
```

```
# Gráfica C
```

```
# d. ¿Cuál distribución parece ser bimodal?
```

```
# Gráfica B
```

```
# e. ¿Cuál distribución parece mostrar una falta de intervalos?
```

```
# Gráfica D
```

```
# Problema 3 -----
```

```
# En R puede importar los datos ya pre cargados en el Software siguiendo el comando data(quakes).
# Los datos proporcionan la localidad, estaciones que detectaron el terremoto, profundidad en km
y la escala de Richter (mag) para 1000 terremotos registrados en la isla de Fiji desde el año de 1964.
```

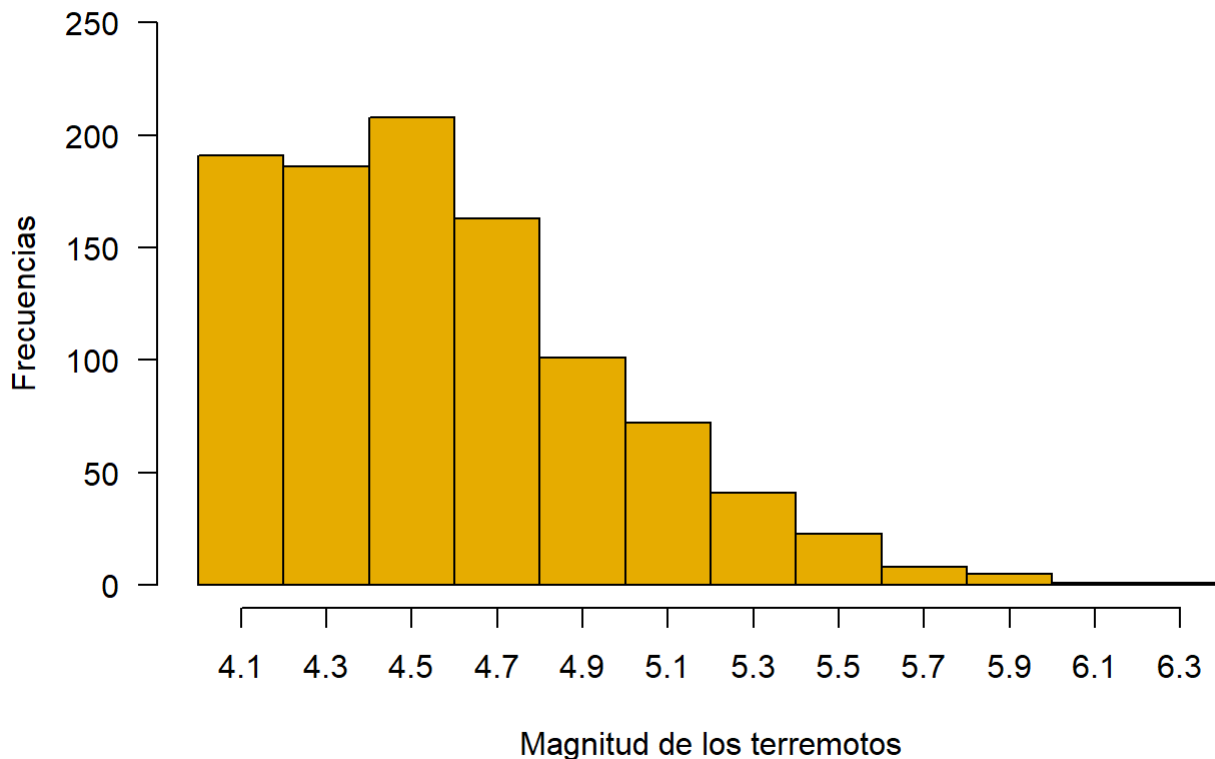
```
data(quakes)
head(quakes)
```

```
##      lat   long depth mag stations
## 1 -20.42 181.62   562 4.8         41
## 2 -20.62 181.03   650 4.2         15
## 3 -26.00 184.10    42 5.4         43
## 4 -17.97 181.66   626 4.1         19
## 5 -20.42 181.96   649 4.0         11
## 6 -19.68 184.31   195 4.0         12
```

```
mags <- hist(quakes$mag, xaxt = "n",
             col = "#e6ac00",
             xlab="Magnitud de los terremotos",
             ylab= "Frecuencias",
             main = "Terremotos en la isla Fiji 1964-actualidad",
             las = 1,
             ylim = c(0,250))

axis(1, mags$mids)
```

## Terremotos en la isla Fiji 1964-actualidad



# Puede tomar en cuenta los siguientes datos que se almacenaron en el objeto mags. para determinar las siguientes preguntas:

# a. ¿Cómo describiría la forma de esta distribución de las magnitudes de los terremotos?

# La forma de la distribución está sesgada hacia la derecha, lo cual indica que han ocurrido más terremotos de menores magnitudes que fuertes.

# b. Mencione un intervalo donde ocurren típicamente las magnitudes.

# El intervalo de 4.0 a 4.9

```
fivenum(quakes$mag)
```

```
## [1] 4.0 4.3 4.6 4.9 6.4
```

```
# c. Determine el rango de las magnitudes (Range = Max - Min).
```

```
range(quakes$mag)
```

```
## [1] 4.0 6.4
```

```
# d. ¿Qué porcentaje de los terremotos ocurren con magnitud en la clase 5.3 (5.1 : 5.4)?
```

```
clase1 <- subset(quakes, quakes$mag >= 5.1)
clase2 <- subset(clase1, clase1$mag <= 5.4)
head(clase2) # 113 observaciones
```

```
##      lat    long depth mag stations
## 3  -26.00 184.10    42 5.4        43
## 25 -19.66 180.28   431 5.4        57
## 28 -16.46 180.79   498 5.2        79
## 50 -18.97 185.25   129 5.1        73
## 74 -23.74 179.99   506 5.2        75
## 80 -28.98 181.11   304 5.3        60
```

```
(113*100)/1000
```

```
## [1] 11.3
```

```
# 11.3 %
```

```
# e. ¿Qué porcentaje de los terremotos tiene una magnitud igual o mayor a 5.0?
```

```
clase3 <- subset(quakes, quakes$mag >=5.0)
head(clase3) # 198 observaciones
```

```
##      lat    long depth mag stations
## 3  -26.00 184.10    42 5.4        43
## 15 -20.70 169.92   139 6.1        94
## 17 -13.64 165.96    50 6.0        83
## 25 -19.66 180.28   431 5.4        57
## 28 -16.46 180.79   498 5.2        79
## 50 -18.97 185.25   129 5.1        73
```

```
(198*100)/1000
```

```
## [1] 19.8
```

```
# 19.8 %
```

```
# f. ¿Qué porcentaje de los terremotos tienen una magnitud menor o igual a 4.6?
```

```
clase4 <- subset(quakes, quakes$mag <= 4.6)
head(clase4) # 585 observaciones
```

```
##      lat   long depth mag stations
## 2  -20.62 181.03   650 4.2        15
## 4  -17.97 181.66   626 4.1        19
## 5  -20.42 181.96   649 4.0        11
## 6  -19.68 184.31   195 4.0        12
## 8  -28.11 181.93   194 4.4        15
## 10 -17.47 179.59   622 4.3        19
```

```
(585*100)/1000
```

```
## [1] 58.5
```

# 58.5 %

# Boxplots -----

# Problema 4 -----

# ¿Qué porcentaje de las observaciones en una distribución se encuentran entre el primer y el tercer cuartil?

# b. 50%

# Problema 5 -----

# La siguiente figura presenta tres gráficas para los diámetros de tres especies diferentes (C, F y H).

# a. ¿Cuál especie tiene el diámetro más pequeño?

# La especie C

# b. ¿Cuál especie tiene el diámetro más grande?

# La especie F

# c. ¿Cuál especie tiene el diámetro mínimo más alto?

# La especie F

# d. ¿Cuál especie tiene la mediana de diámetro más pequeña?

# La especie C

# e. ¿Cuál especie tiene la mediana de diámetro más grande?

# Especie H

# f. ¿Cuál especie tiene el menor rango de diámetro?

# Especie F

# g. ¿Cuál especie tiene el rango intercuartil ( $Q3-Q1$ ) más grande?

# Especie C

# h. ¿Cuál especie tiene el rango intercuartil ( $Q3-Q1$ ) más pequeño?

# Especie F

# i. ¿Cuál especie tiene una distribución simétrica?

# Especie H



# j. ¿Cuál especie tiene el sesgo positivo (ver Fig. 2) más marcado ?

# Especie F

# Problema 6 -----

# Los siguientes datos muestran el número de incendios forestales ocurridos en cada semana en nuestros bosques de México.

# Los datos son del 01. de enero al 04 de marzo del 2021 de acuerdo con el reporte de CONAFOR.

```
fires <- c(78, 44, 47, 105, 126, 181, 277, 210, 155)
head(fires)
```

```
## [1] 78 44 47 105 126 181
```

#Determine lo siguiente

# 1. Valor mínimo  
`min(fires)`

```
## [1] 44
```

# 2. Valor máximo  
`max(fires)`

```
## [1] 277
```

# 3. Rango  
`range(fires)`

```
## [1] 44 277
```

# 4. Q1 (25 %)  
`quantile(fires, c(.25))`

```
## 25%
## 78
```

# 5. Q2 (50 %)  
`quantile(fires, c(.50))`

```
## 50%  
## 126
```

```
# 6. Q3 (75 %)  
quantile(fires, c(.75))
```

```
## 75%  
## 181
```

```
# 7. Media  
mean(fires)
```

```
## [1] 135.8889
```

```
# 8. Varianza  
var(fires)
```

```
## [1] 6069.111
```

```
# 9. Desviación estándar  
sd(fires)
```

```
## [1] 77.9045
```

```
# 10. Realice un boxplot personalizado con los datos de los incendios.  
boxplot(fires, main = "Incendios forestales en México  
01 de enero a 04 de marzo de 2021",  
xlab = "Número de incendios", las = 1, col = "firebrick", horizontal = TRUE)
```

## Incendios forestales en México 01 de enero a 04 de marzo de 2021

