

**UNIVERSIDAD DE EL SALVADOR
FACULTAD MULTIDISCIPLINARIA DE OCCIDENTE
DEPARTAMENTO DE MATEMÁTICA**



Licenciatura en Estadística

Control Estadístico del Paquete R

**”UNIDAD DOS”
PRÁCTICA 07 - ANÁLISIS ESTADÍSTICO DE DATOS
UNIVARIADOS DISCRETOS CON R.**

**Alumna:
Martha Yoana Medina Sánchez**

**Fecha de elaboración
Santa Ana - 27 de noviembre de 2015**

Práctica 07-Análisis estadístico de datos univariados discretos con R.

ANÁLISIS ESTADÍSTICO DE LOS DATOS.

1. *# Activar el directorio de trabajo*

```
getwd()  
  
## [1] "C:/Users/User/Documents/PRACTICAS_YOANA_MEDINA/Yoana/PRACTICAS DE R"  
  
setwd("C:/Users/User/Documents/PRACTICAS_YOANA_MEDINA/Yoana/PRACTICAS DE R")
```
2. *# Crear un nuevo Script y llamarle "Script07-DatosDiscretos"*
3. *# Crear el vector de datos.*

```
Hijos <- c(2, 1)  
data.entry(Hijos)  
Hijos  
  
## [1] 2 1  
  
length(Hijos)  
  
## [1] 2
```
4. *# Guardar el vector de datos en un archivo de texto.*

```
write(Hijos, "Hijos.txt")
```
5. *# Limpiar el Área de trabajo (Workspace)*

```
ls()  
  
## [1] "Hijos"  
  
rm(list=ls(all=TRUE)); ls()  
  
## character(0)
```

6. *# Leer o recuperar el vector de datos o archivo de texto*

```
X <- scan("Hijos.txt", what = integer(0), na.strings = "NA", flush=FALSE)
ls()
```

```
## [1] "X"
```

Si el vector contiene caracteres se usa:

```
what = character()
```

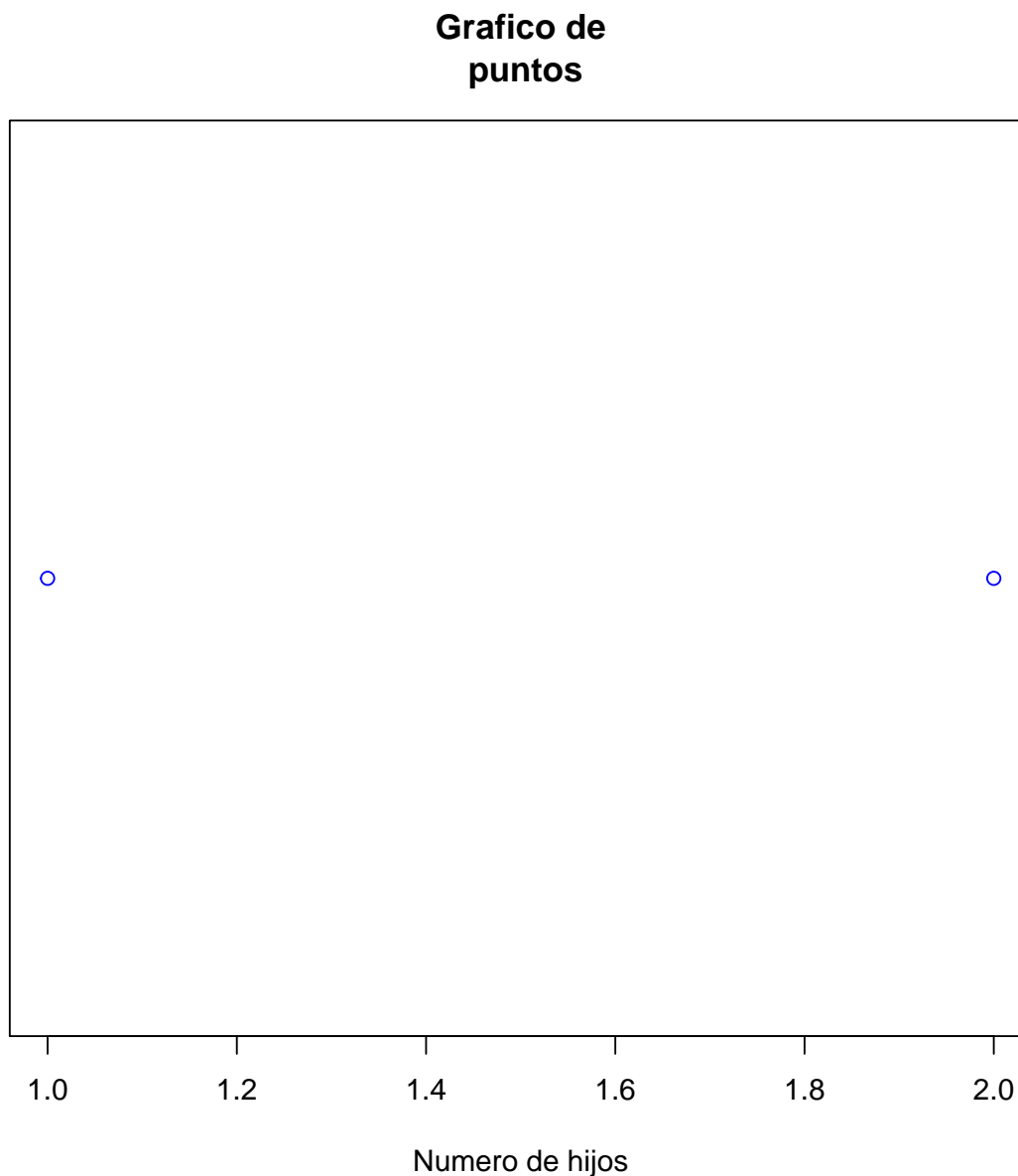
Si el vector contiene reales se ocupa:

```
what = double(0)
```

7. Elaborar el gráfico de puntos y diagrama de tallo-hojas (stem-and-leaf)

Grafico de puntos

```
stripchart(X, method="stack", vertical=FALSE, col="blue", pch=1,
main="Grafico de\n puntos", xlab="Numero de hijos")
```



Observación: method puede ser:

”overplot” (los puntos coincidentes son superpuestos)

”jitter” (los puntos se ven como alejados o inquietos)

”stack” (los puntos coincidentes son apilados, uno tras otro)

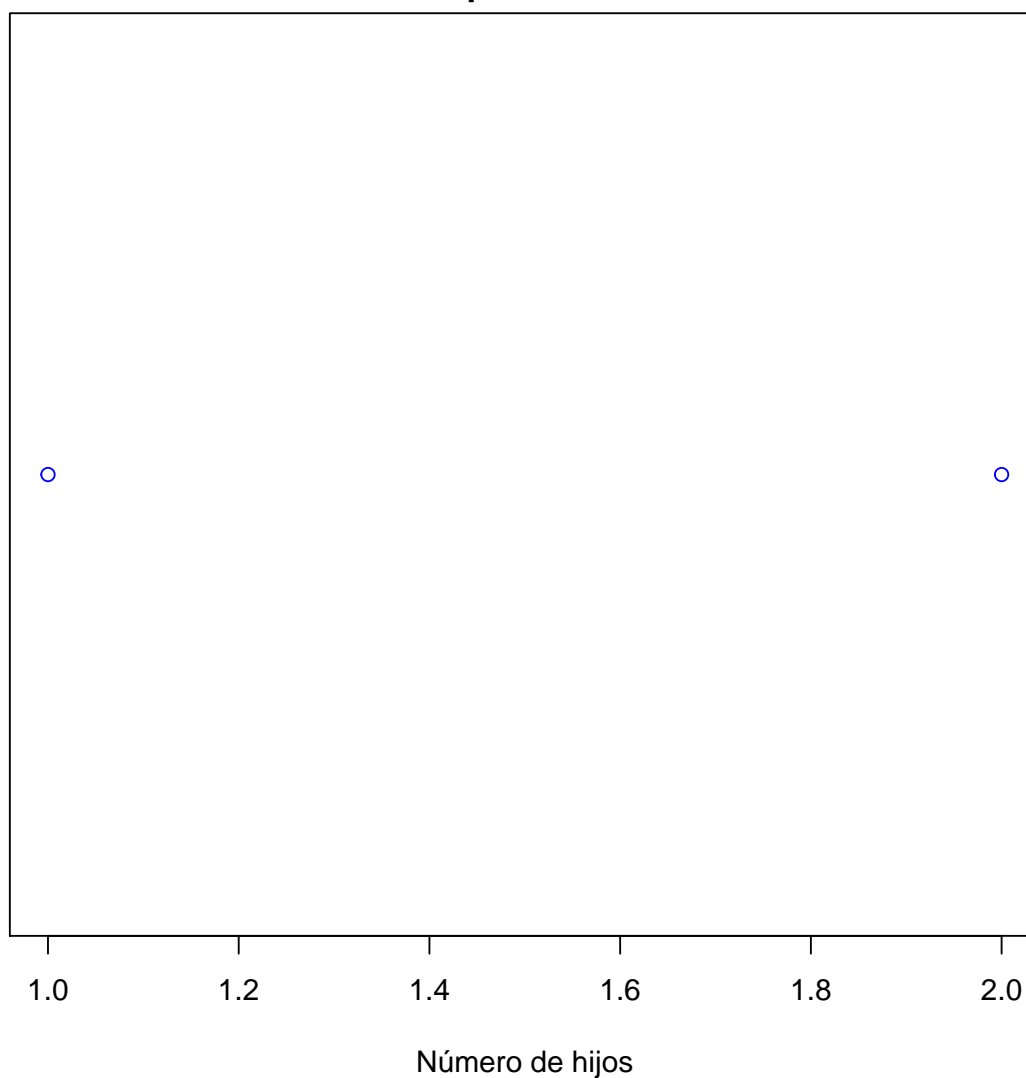
```
# Gráfico de puntos
```

```
stripchart(X, overplot="stack", vertical=FALSE, col="blue", pch=1, main="Gráfico de  
puntos", xlab="Número de hijos")
```

```
## Warning in plot.window(xlim, ylim, log, ...): "overplot" is not a graphical  
parameter
```

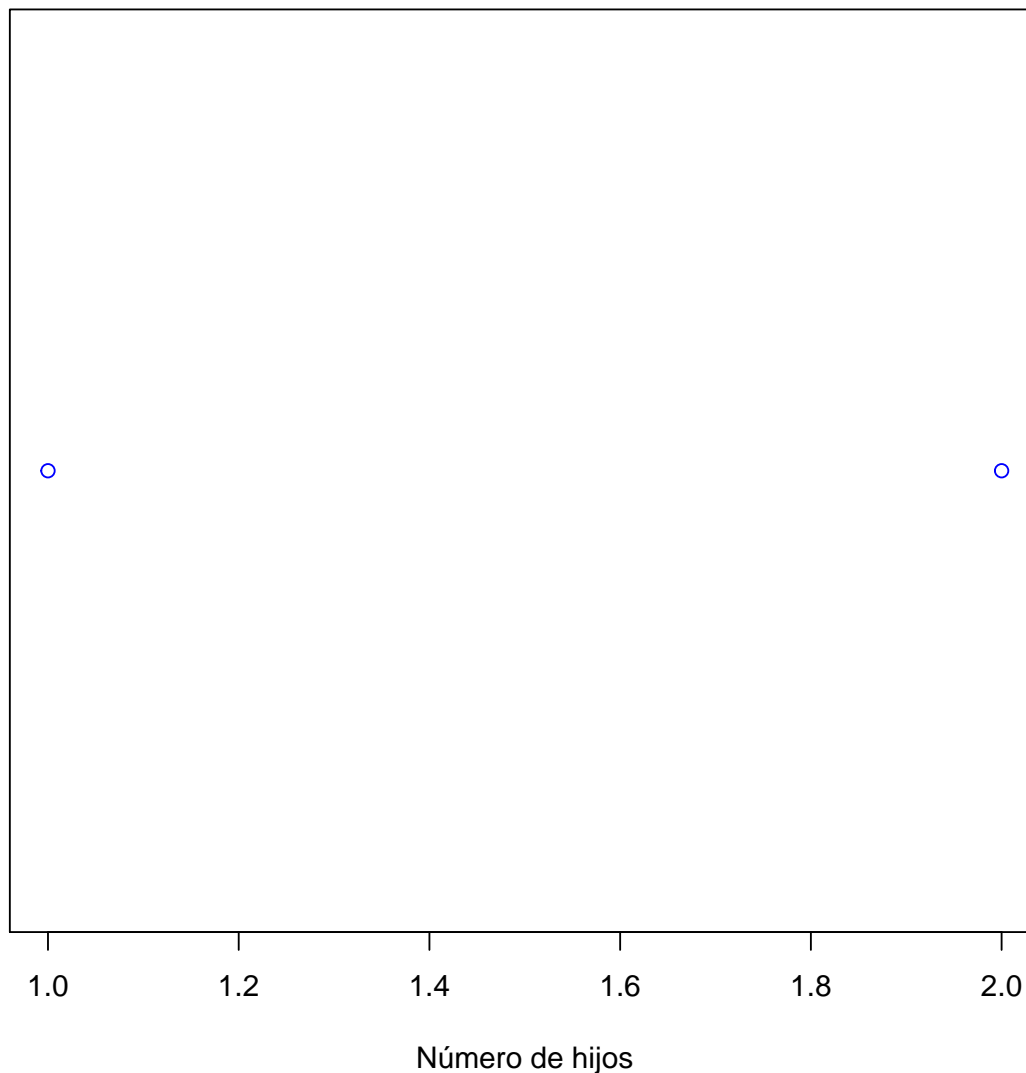
```
## Warning in axis(side = side, at = at, labels = labels, ...): "overplot" is
not a graphical parameter
## Warning in title(xlab = xlab, ylab = ylab, ...): "overplot" is not a graphical
parameter
## Warning in plot.xy(xy.coords(x, y), type = type, ...): "overplot" is not a
graphical parameter
```

Gráfico de puntos



```
# Gráfico de puntos
stripchart(X, jitter="stack", vertical=FALSE, col="blue", pch=1, main="Gráfico de\n
puntos", xlab="Número de hijos")
```

Gráfico de puntos



```
# Gráfico de puntos
```

```
stripchart(X, stack="stack", vertical=FALSE, col="blue", pch=1, main="Gráfico de\npuntos", xlab="Número de hijos")
```

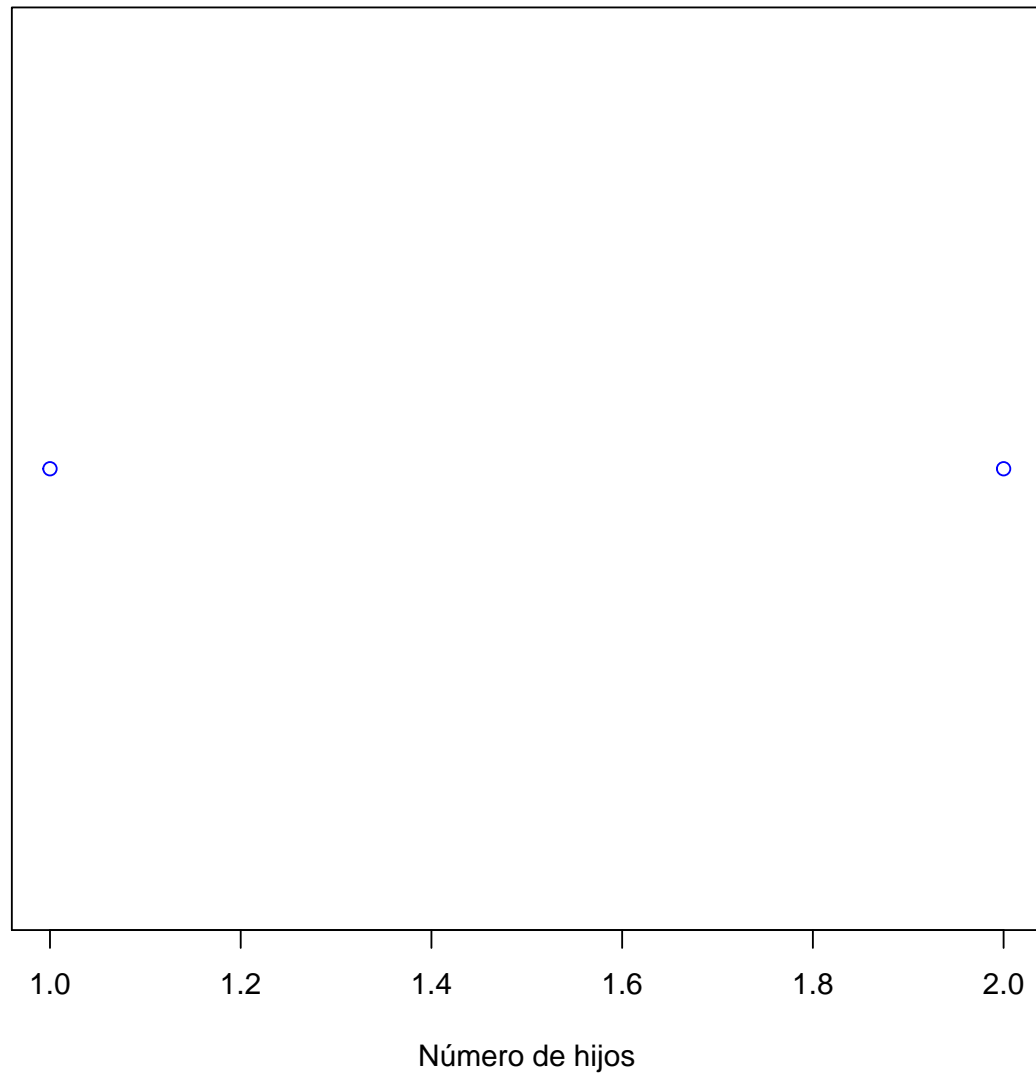
```
## Warning in plot.window(xlim, ylim, log, ...): "stack" is not a graphical parameter
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "stack" is not a graphical parameter
```

```
## Warning in title(xlab = xlab, ylab = ylab, ...): "stack" is not a graphical parameter
```

```
## Warning in plot.xy(xy.coords(x, y), type = type, ...): "stack" is not a graphical parameter
```

Gráfico de puntos



```
8. # Crear la tabla de frecuencias completa
# frecuencias individuales

fab <- table(X); fab # frecuencias absolutas

## X
## 1 2
## 1 1

fre <- fab/length(X); fre # frecuencias relativas

## X
```

```
##    1    2
## 0.5 0.5

Fac <- cumsum(fab); Fac # frecuencias acumuladas

## 1 2
## 1 2

Far <- Fac/length(X); Far # frecuencias acumuladas relativas

##    1    2
## 0.5 1.0
```

```
# tabla de frecuencias completa

options(digits=2)
tabla <- data.frame(fab=fab, fre=fre, Fac=Fac, Far=Far)
names(tabla) <- c("X", "fab", "free.X", "fre", "Fac", "Far")
tabla

##    X fab free.X fre Fac Far
## 1 1    1      1 0.5    1 0.5
## 2 2    1      2 0.5    2 1.0

tfre <- data.frame(X=tabla$X, fab=tabla$fab, fre=tabla$fre, Fac=tabla$Fac,
                   Far=tabla$Far)
tfre

##    X fab fre Fac Far
## 1 1    1 0.5    1 0.5
## 2 2    1 0.5    2 1.0
```

```
# Note que el cuadro resultante no tiene la presentacion deseada para
# presentarla en un informe. Sin embargo, si estamos utilizando LATEX podemos
# utilizar la siguiente instruccion xtable(tfre) y con esto nos genera el
# codigo correspondiente para incorporarlo en nuestro archivo.
```

9. *# Calcular los estadísticos descriptivos de la variable*

```
# Estadísticos de tendencia central de los datos
```

```
media <- mean(X, na.rm = FALSE); media
```



```
## [1] 1.5
```

```
# na.rm = FALSE, le indica a R que los datos faltantes son omitidos en el  
# calculo de la media.
```

```
for(i in 1:length(X)) if (fab[i] == max(fab)) break()  
moda <- names(fab[i]); moda
```

```
## [1] "1"
```

```
# R no tiene incorporada una funcion para la moda
```

```
mediana <- median(X); mediana
```

```
## [1] 1.5
```

```
# Estadísticos de dispersión o variabilidad de los datos
```

```
range(X)
```

```
## [1] 1 2
```

```
# Devuelve el valor minimo y maximo del conjunto de datos.
```

```
cuasivar <- var(X); cuasivar
```

```
## [1] 0.5
```

```
s <- sd(X); s
```

```
## [1] 0.71
```

```
# Devuelve la cuasivarianza y la cuasivarianza muestral
```

```
quantile(X,c(0.25, 0.5, 0.75))
```

```
## 25% 50% 75%
```

```
## 1.2 1.5 1.8
```

```
# Calculo de Q1, Q2, Q3
```

```
quantile(X, 0.6)

## 60%
## 1.6

# En general se pueden encontrar cualquier percentil
```

```
# Conocer un resumen de los datos

resumen <- summary(X); resumen

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.0     1.2     1.5     1.5     1.8     2.0

# Min, Q1, Median, Mean, Q3, Max
```

```
fivenum(X)

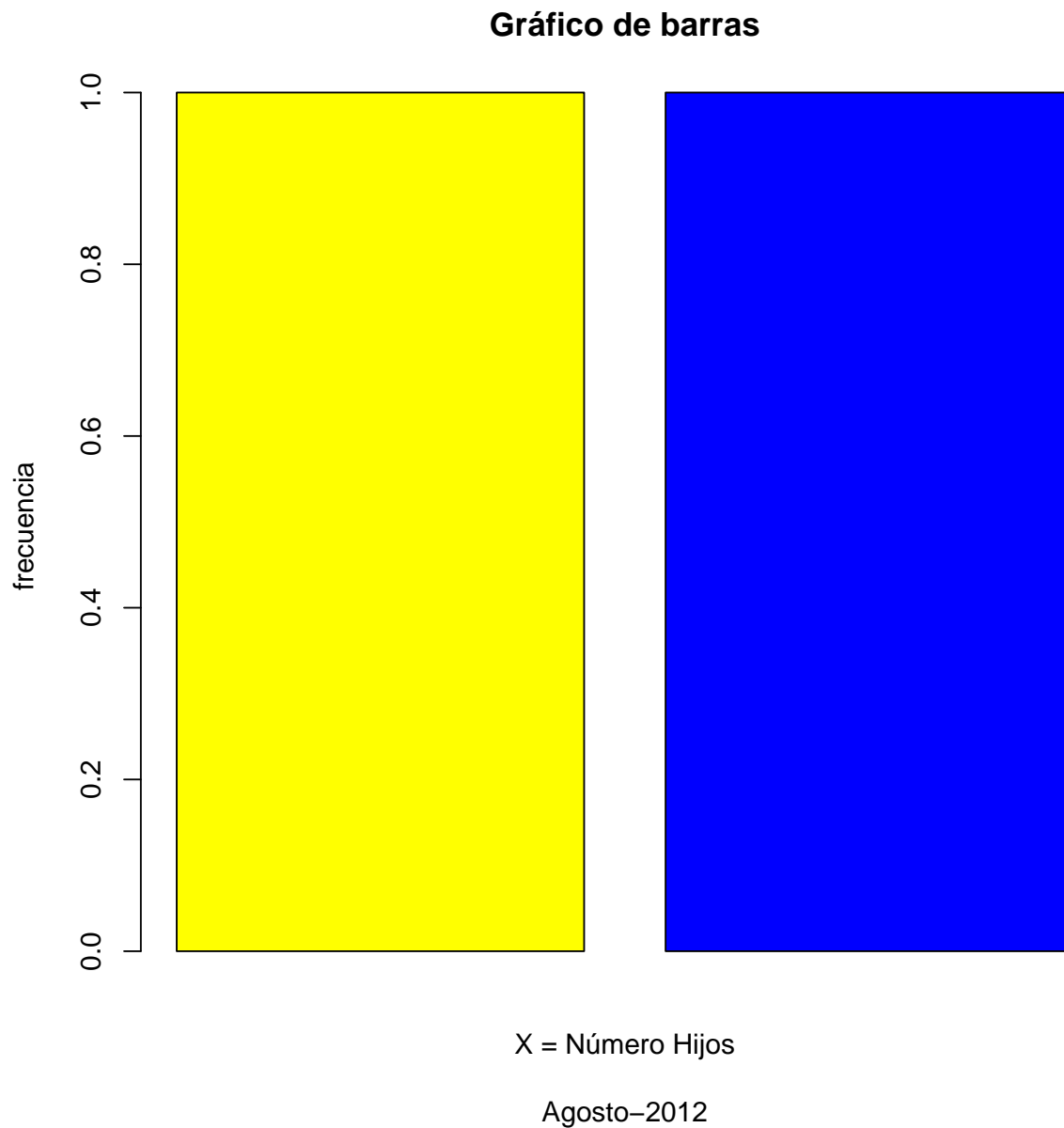
## [1] 1.0 1.0 1.5 2.0 2.0

# min, cuartil menor, mediana, cuartil mayor, max
```

10. Elaborar los gráficos que se le pueden aplicar a la variable discreta

```
# Grafico de barras (por ser pocos valores)

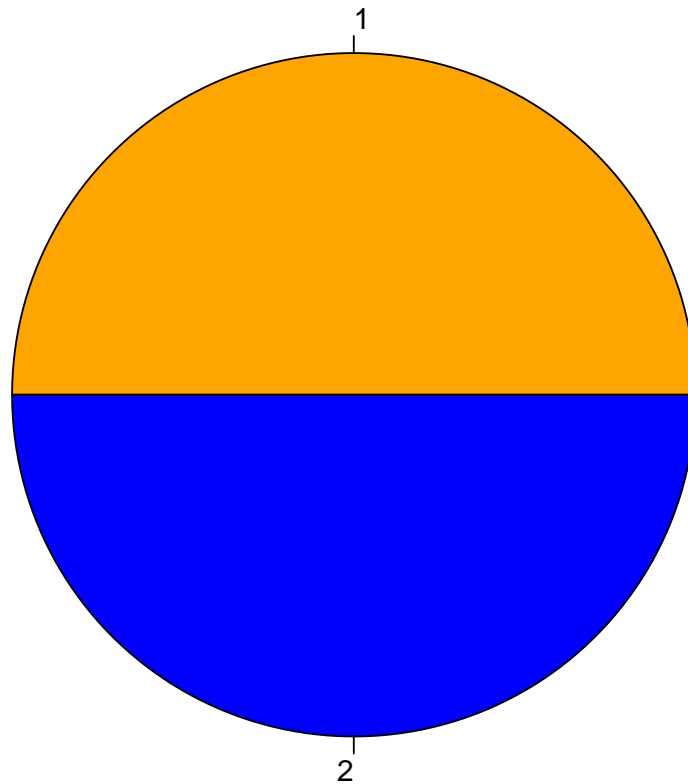
barplot(tfre[[2]], main="Gráfico de barras", xlab="X = Número Hijos\n",
        ylab="frecuencia", col=c("yellow", "blue", "white", "orange",
        "cyan", "red"), sub="Agosto-2012")
```



```
# Grafico de pastel (por ser pocos valores)

pie(tfre[[2]], main="Gráfico de pastel", xlab="Número Hijos \n",
    col=c("orange", "blue", "white", "yellow", "cyan", "blue"),
    sub="Agosto-2012")
```

Gráfico de pastel



Número Hijos

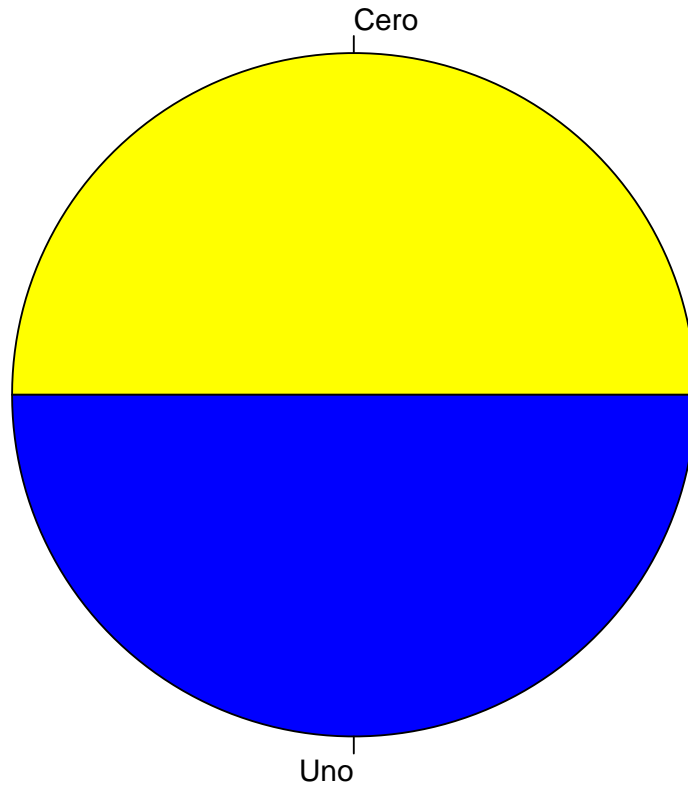
Agosto-2012

```
# Se puede especificar nombres para las categorías

names(fab) = c("Cero", "Uno")

pie(fab, main="Gráfico de pastel", xlab="X = Número Hijos\n",
    col=c("yellow", "blue", "white", "orange", "cyan", "red"),
    sub="Agosto-2012")
```

Gráfico de pastel

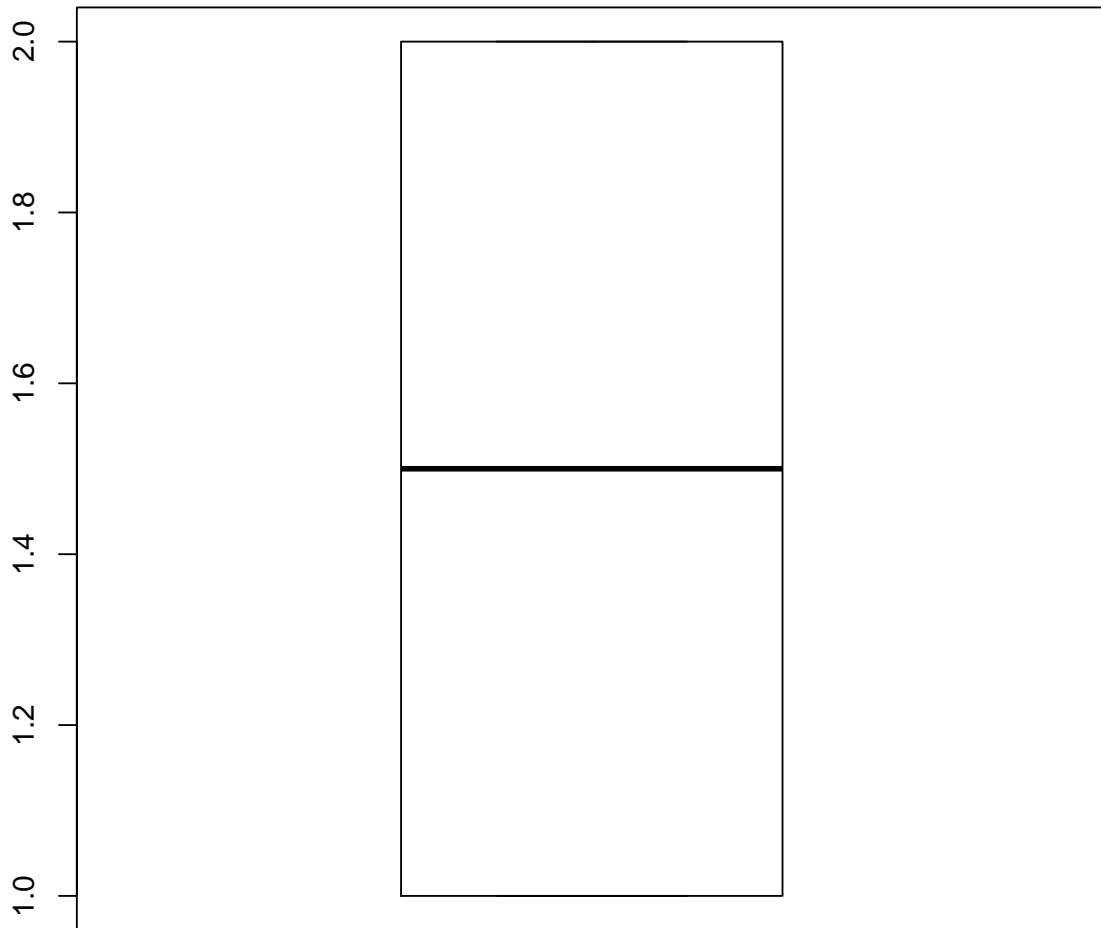


X = Número Hijos

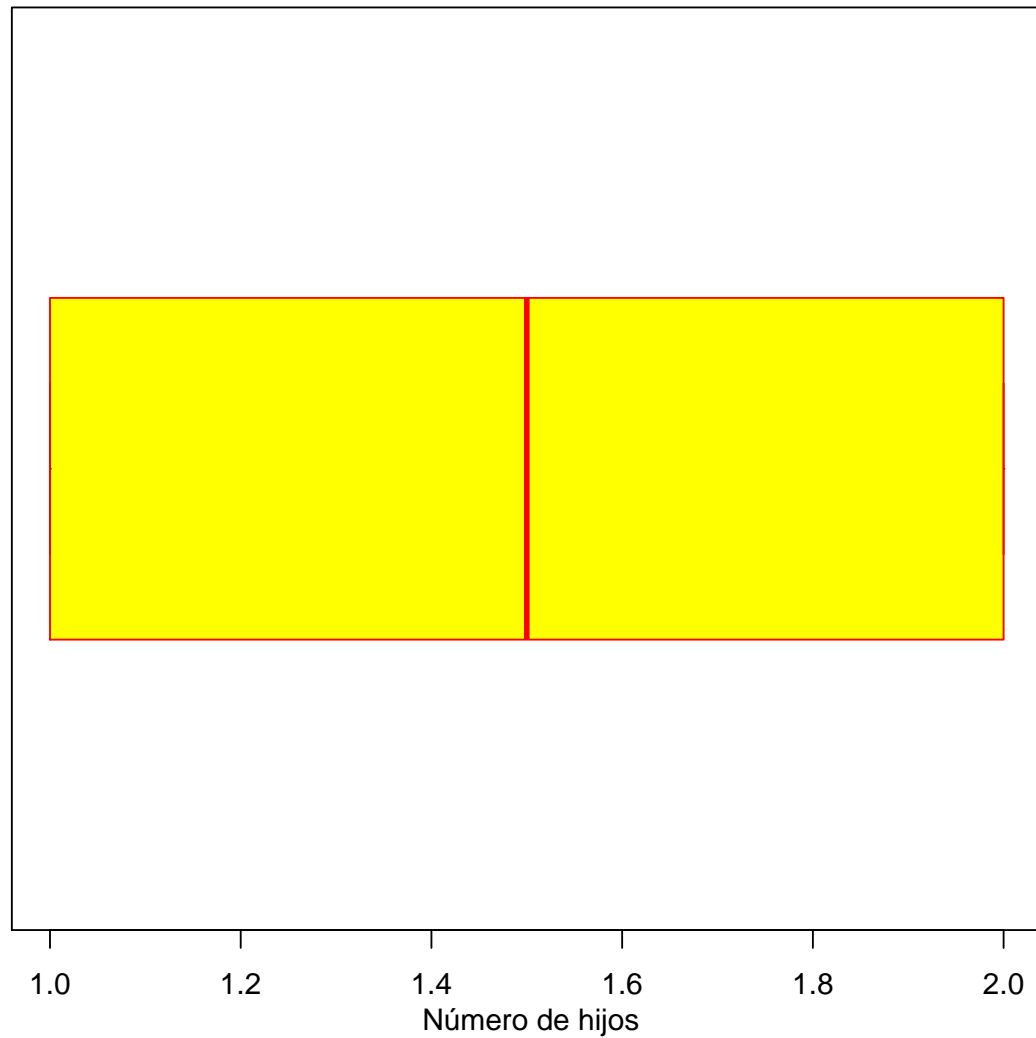
Agosto-2012

```
# Gráfico de cajas (box-plot) es la representación gráfica de los cinco  
# números Horizontal
```

```
boxplot(X, main="Gráfico de caja", ylab="Número de hijos\n")
```

Gráfico de caja

```
# Vertical  
boxplot(X, main="Gráfico de caja", xlab=" Número de hijos\n", plot=TRUE,  
        border="red", col="yellow", horizontal=TRUE)
```

Gráfico de caja

```
# NOTE QUE TODOS LOS GRAFICOS DE BARRAS Y DE PASTEL SON REALIZADOS  
# APARTIR DE UNA TABLA DE FRECUENCIA, LA CUAL SE INDICA EN tfre[[2]].  
# TAMBIEN SE PUDO UTILIZAR tabla[[2]].
```