Tema-09: Densidades y funciones de distribución

Alonso Pizarro Lagunas

24/11/2021

Densidades y funciones de distribución

- Usando la función density.
- Una función para calcular histogramas de frecuencias más completos

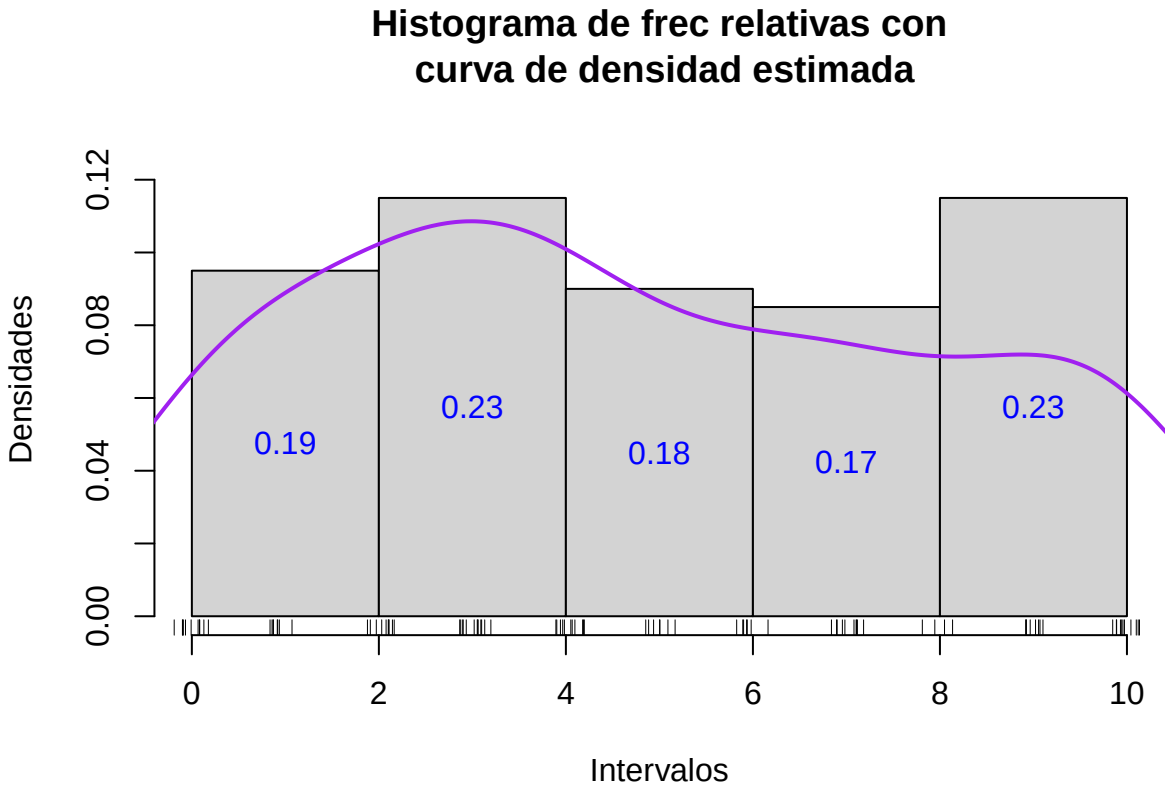
```
histRel = function(x,L){
  h = hist(x, breaks = L, right = FALSE, plot = FALSE)
  t = round(1.1*max(max(density(x)[[2]]),h$density),2)
  plot(h, freq = F, col = "lightgray",
       main = "Histograma de frec relativas con\n curva de densidad estimada",
       xaxt = "n", ylim=c(0,t), xlab = "Intervalos", ylab ="Densidades")
  axis(1, at = L)
  text(h$mids, h$density/2, labels= round(h$counts/length(x),2), col = "blue")
  lines(density(x), col = "purple", lwd = 2)
  }
```

- Vamos a probar la función histRel

```
set.seed(4)
notas = sample(0:10,100, replace = T)
set.seed(NULL)
notas
```

```
##   [1]  7 10  2  2  6  2  5  4  9  2  7  5  1  7  0  3 10  2 10  4  1  4  5  4  0
##  [26]  5 10  4  3  0  7  5 10  3  4  8  1  9  3  7  9  1  9 10  5 10 10  9  5  0
##  [51]  3  1  3  2  0  6  6  4  7  4  7  3  9  0  7  0  3  0  3  3  1  4 10  9  1
##  [76]  4  0  6 10  0 10  1  0  2  6  4  8  2  3  7  7  3  3  8  2  6  6  2  8  9
```

```
L =  c(0,2,4,6,8,10)
histRel(notas, L)
rug(jitter(notas))
```

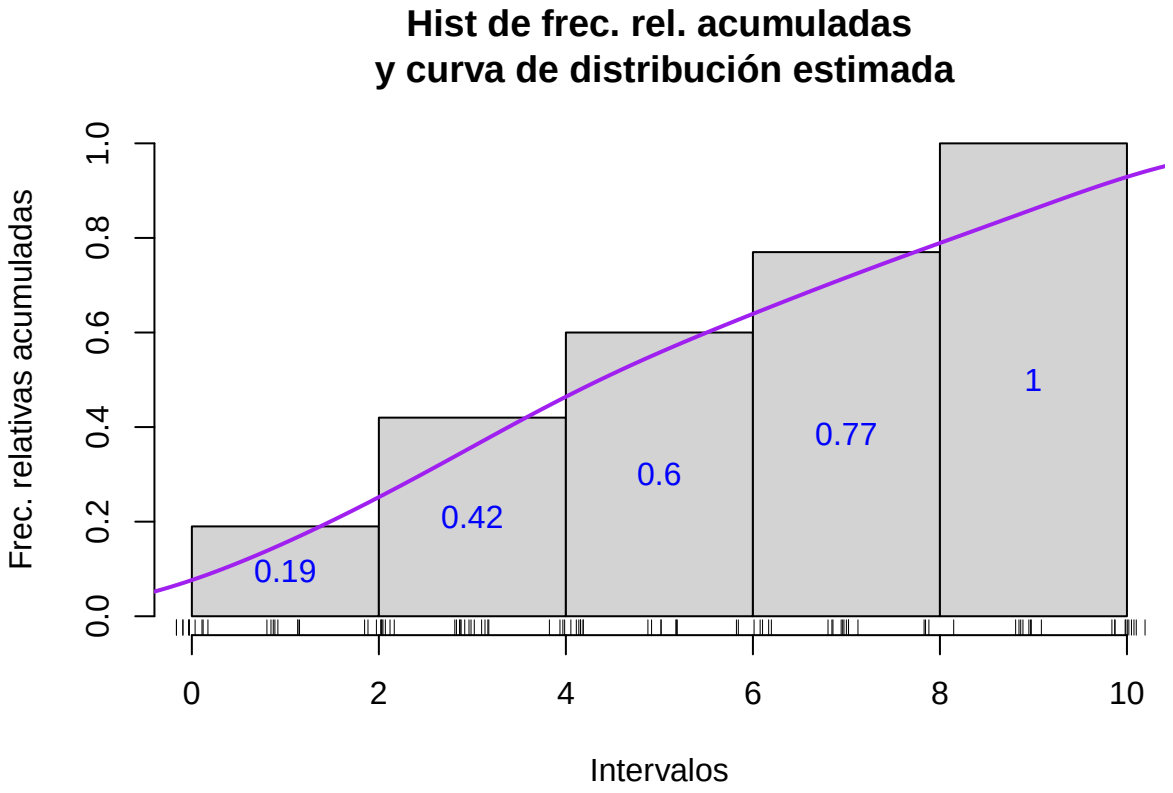


- Ahora crearemos un histograma de frecuencias relativas acumuladas

```
histRelCum=function(x,L){
  h = hist(x, breaks = L, right = F, plot = F)
  h$density = cumsum(h$counts)/length(x)
  plot(h, freq = FALSE,
  main = "Hist de frec. rel. acumuladas\n y curva de distribución estimada",
  xaxt = "n", col = "lightgray", xlab = "Intervalos",
  ylab ="Frec. relativas acumuladas")
axis(1,at = L)
text(h$mids, h$density/2, labels = round(h$density,2), col="blue")
dens.x = density(x)
dens.x$y = cumsum(dens.x$y)*(dens.x$x[2]-dens.x$x[1])
lines(dens.x, col = "purple", lwd = 2)
}
```

- Aplicando la función anterior a un df

```
L =  c(0,2,4,6,8,10)
histRelCum(notas, L)
rug(jitter(notas))
```



Ejemplo 2: usando datos de cangrejos

Histogramas básico con función `hist`

```
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
L1 = min(cw)-1/2*0.1
A = 1.3
L = L1 + A*(0:10)
L
```

```
##  [1] 20.95 22.25 23.55 24.85 26.15 27.45 28.75 30.05 31.35 32.65 33.95
```

```
hist(cw,breaks = L, right = F, plot = F)
```

```
## $breaks
##  [1] 20.95 22.25 23.55 24.85 26.15 27.45 28.75 30.05 31.35 32.65 33.95
##
```

```
## $counts
##  [1]  2 14 27 44 34 31 15  3  2  1
##
## $density
##  [1] 0.008892841 0.062249889 0.120053357 0.195642508 0.151178301 0.137839040
##  [7] 0.066696309 0.013339262 0.008892841 0.004446421
##
## $mids
##  [1] 21.6 22.9 24.2 25.5 26.8 28.1 29.4 30.7 32.0 33.3
##
## $xname
## [1] "cw"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
```

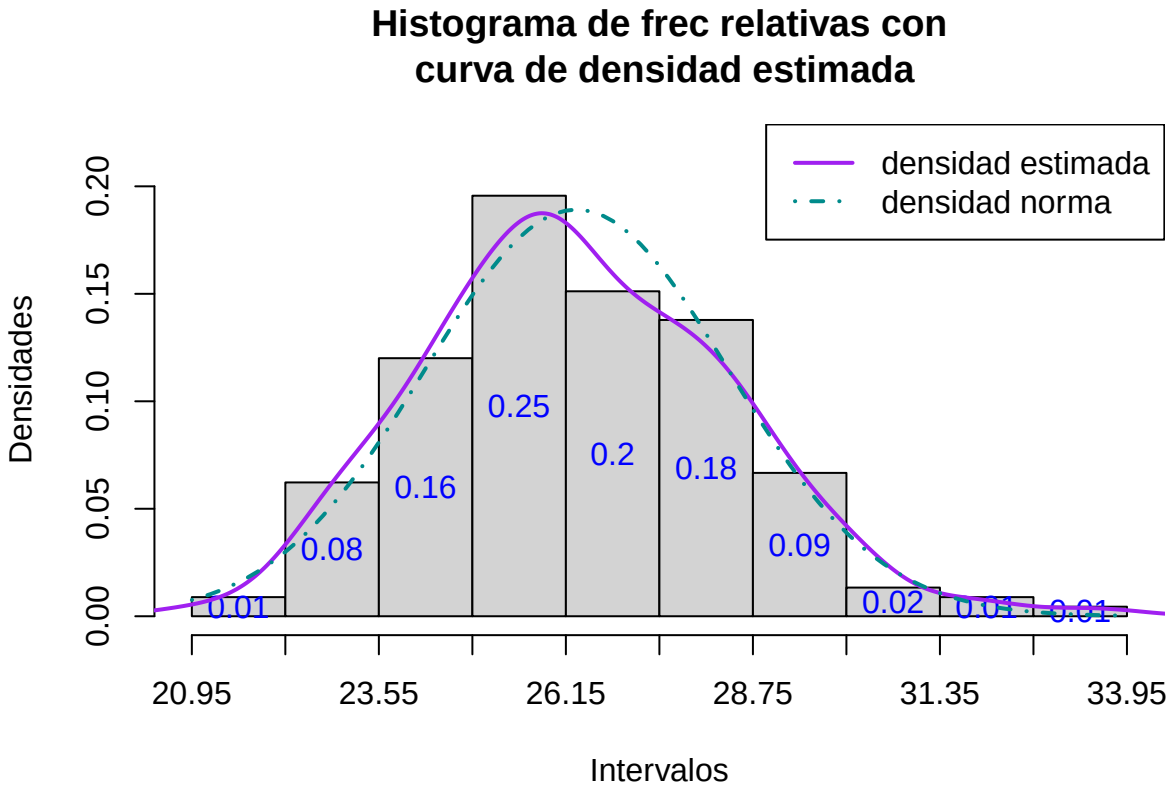
```
str(density(cw))
```

```
## List of 7
##  $ x        : num [1:512] 19 19 19.1 19.1 19.1 ...
##  $ y        : num [1:512] 3.90e-05 4.50e-05 5.17e-05 5.94e-05 6.82e-05 ...
##  $ bw       : num 0.671
##  $ n        : int 173
##  $ call     : language density.default(x = cw)
##  $ data.name: chr "cw"
##  $ has.na   : logi FALSE
##  - attr(*, "class")= chr "density"
```

```
histRel(cw,L)

# agregamos una curva normal


histRel(cw,L)
curve(dnorm(x, mean(cw),sd(cw)), col =c("cyan4"), lty = 4, lwd =2,
      add =TRUE)
legend("topright", lwd=c(2,2), lty = c(1,4), col = c("purple","cyan4"),
       legend = c("densidad estimada", "densidad norma"))
```


Histograma de frec relativas con
curva de densidad estimada
densidad estimada
densidad norma
Densidades
0.00
0.05
0.10
0.15
0.20
0.01
0.08
0.16
0.25
0.2
0.18
0.09
0.02
0.01
0.01
20.95
23.55
26.15
28.75
31.35
33.95
Intervalos