# UNIDAD 3: Práctica 15 - Distribuciones de probabilidad continuas.

Caterine Melissa Guerrero España

2022-09-09

## DISTRIBUCIONES CONTINUAS.

## CÁLCULO DE PROBABILIDADES

- **Ejemplo 1:** Una persona informal hace esperar a su pareja aleatoriamente entre 0 y 90 minutos. Harto de esta situación, la persona que sufre la espera se plantea un ultimátum; sí al día siguiente su pareja tarda menos de 15 minutos mantiene la relación, sí la espera está entre 15 y 55 minutos, decide en la siguiente cita con los mismos criterios, mientras que si tarda más de 55 minutos la relación termina en ese momento.

a) Calcule la probabilidad de que la relación continúe hasta la siguiente cita.

```
x<-55
a=0
b<-90
#usando la función propia de R
punif(x, min=a, max=b, lower.tail=TRUE)
```

```
## [1] 0.6111111
```

b) Calcule la probabilidad de que la relación termine en la segunda cita.

Suponiendo que el tiempo de espera en una cita es independiente respecto de otras citas, se calcula la probabilidad $P(15 < X < 55) = P(X < 55) - p(X \leq 15) = 0.4445$

```
F55=punif(55, min=a, max=b, lower.tail=TRUE)
F15=punif(15, min=a, max=b, lower.tail=TRUE)
F55-F15
```

```
## [1] 0.4444444
```

que es la probabilidad de que aplace la decisión para la segunda cita y, en la segunda cita, la probabilidad de que lo deje definitivamente es P(X>55) = 0.3888

```
F55=punif(55, min=a, max=b, lower.tail=TRUE)
F55
```

```
## [1] 0.6111111
```

luego multiplicando ambas probabilidades se obtiene el valor pedido 0.1728.

```
(1-F55)*( F55-F15)
```

```
## [1] 0.1728395
```

- **Ejemplo 2:** Una empresa está buscando personal para su departamento de mercadeo. El perfil solicitado es el de sujetos extrovertidos y creativos. Se han presentado 50 candidatos y la empresa ha establecido como criterio de selección que los candidatos superen el percentil 80 en creatividad y extroversión. Sabiendo que la variable extroversión (X) se distribuye según una Normal de media 5 y desviación típica 1, que la variable creatividad (Y) sigue una t-Student de 10 grados de libertad y que las puntuaciones de creatividad y extroversión son independientes:

a) ¿Cuántos candidatos serán seleccionados? Al ser X e Y independientes, la probabilidad

Como se han presentado 50 aspirantes, serán seleccionadas (50) ( 0.04) = 2 personas

b) ¿Qué puntuaciones debe superar un aspirante en creatividad y extroversión para ser admitido?

Según el criterio de selección se debe superar el percentil 80, en ambas variables, para ser admitido. Se calculará pues el percentil 80 de la variable X e Y, utilizando los cuantiles-normales para la variable X:

y los cuantiles-normales para la variable X:

```
p<-c(0.80)
media=5
d.t=1
qnorm(p, mean=media, sd=d.t, lower.tail=TRUE)
```

```
## [1] 5.841621
```

y los cuantiles-t para la variable Y:

```
p<-c(0.80)
g.l<-10
qt(p, df=g.l, lower.tail=TRUE)
```

```
## [1] 0.8790578
```

c) Si se extraen al azar 16 candidatos, ¿cuál es la probabilidad de que su media aritmética en extroversión sea mayor que 4.5?

Se sabe que al extraer una muestra de una población normal de tamaño n, la media muestral, sigue otra distribución normal de media igual que la poblacional y desviación típica $\frac{\sigma}{\sqrt{n}}$.

Como se desea calcular $P(\overline{x} \geq 4.5)$:

```
n<-16
x<-4.5
mu=5
sigma=1
d.t=sigma/sqrt(n)
pnorm(x, mean=mu, sd=d.t, lower.tail=FALSE)
```

```
## [1] 0.9772499
```

- **Ejemplo 3:** La duración media de un modelo de marcapasos es de 7 años.

a) ¿Cuál es la probabilidad de que dure al menos 5 años? ¿y menos de 3 años? Suponiendo que la variable X="tiempo de funcionamiento del marcapasos" sigue una distribución exponencial con parámetro $\lambda = \frac{1}{\theta} = \frac{1}{7}$ con $\theta = E[X]$ tiempo promedio

La probabilidad $P(X \geq 5)$ se obtiene así

```
x<-5
teta=7
pexp(x, rate=1/teta, lower.tail=FALSE)
```

```
## [1] 0.4895417
```

y de igual forma P(X<3):

```
x<-3
teta=7
pexp(x, rate=1/teta, lower.tail=TRUE)
```

```
## [1] 0.3485609
```

b) Si han transcurrido ya 4 años desde su implantación, ¿cuál es la probabilidad de que dure otros 4? Nos piden $P(X \geq 8/X \geq 4)$

Teniendo en cuenta que la función de distribución es la única distribución continua no tiene memoria resulta que $P(X \geq 8/X \geq 4) = P(X \geq 4) = 0.5647182$

```
pexp(4, rate=1/teta, lower.tail=FALSE)
```

```
## [1] 0.5647181
```

c) ¿Cuánto tiempo debería funcionar un marcapasos para estar entre el 10% de los que más duran?

Hay que calcular el percentil 90:

```
p<-0.9
teta<-7
qexp(p, rate=1/teta, lower.tail=TRUE)
```

```
## [1] 16.1181
```

```
#resultando 16.12 años.
```

d) Calcular el valor que deben tener a y b para que P(X<a)=0.5 y P(X>b)=0.32 De forma análoga al apartado anterior, en el primer caso habría que calcular la mediana (percentil 50), a=4.852,

```
qexp(0.5, rate=1/teta, lower.tail=TRUE)
```

```
## [1] 4.85203
```

y en el segundo caso, el percentil 68, b = 7.97

```
qexp(0.68, rate=1/teta, lower.tail=TRUE)
```

```
## [1] 7.97604
```

o de esta otra manera

```
qexp(0.32, rate=1/teta, lower.tail=FALSE)
```

```
## [1] 7.97604
```

# GENERACIÓN DE MUESTRAS ALEATORIAS DE LAS DISTRIBUCIONES

- **Ejemplo 1:** Generar 100 números aleatorios de una distribución Uniforme en [-2, 4]

Definir los parámetros apropiados

```
min<- -2
max<-4
```

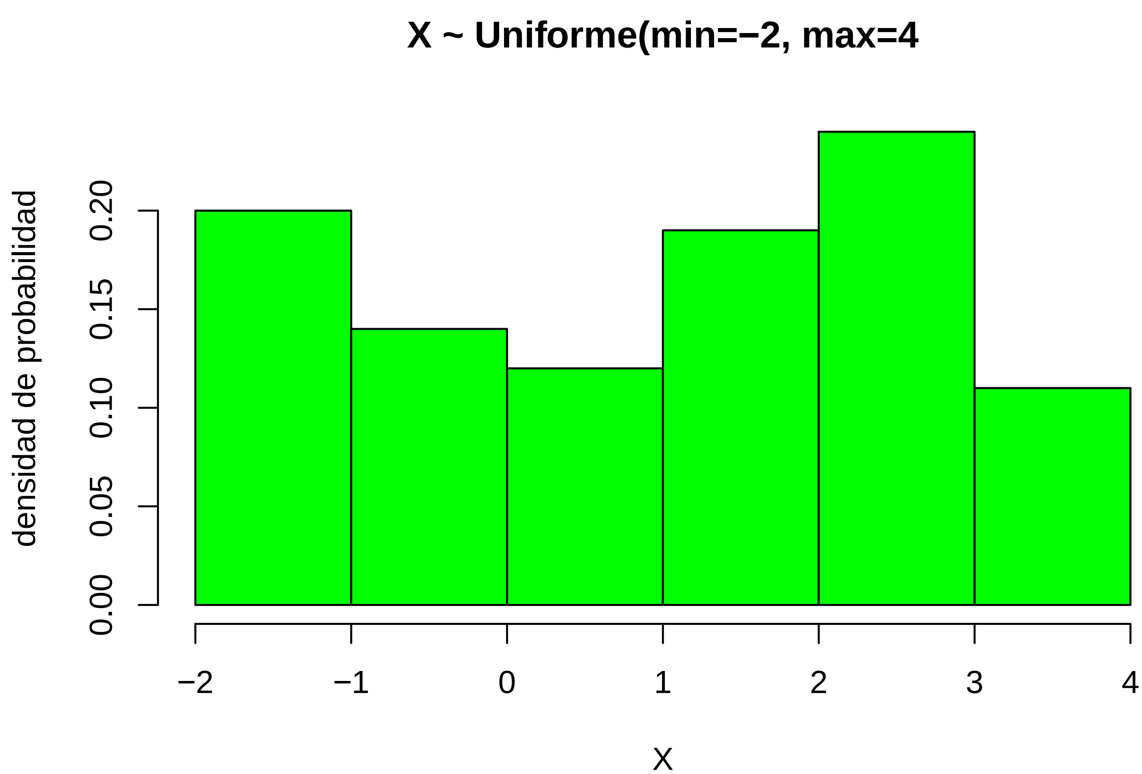
generar 100 números aleatorios de la distribución

```
x=runif(100, min, max)
x
```

```
##   [1]  2.41802236  2.94859544  3.10164053  2.55898250  3.40283227 -0.80106513
##   [7]  2.38156517  2.39957476  2.62660836  1.73575303  3.03311581  0.64714012
##  [13] -0.65820465 -0.48039754  2.13224225  2.88282900  3.24304616  1.30158276
##  [19]  2.70582460 -1.21657837  2.48762307  2.71747570  0.91506333 -1.40776284
##  [25]  2.33523300  0.88303224  1.94892765  0.65930188 -0.04984915  2.68505801
##  [31] -1.83430813  2.34596345  2.42724002  3.14156356  1.59263392  2.46499887
##  [37] -1.90274285  0.36570776  2.68791024 -1.33673388 -1.75308059  1.24000330
##  [43]  1.83479894 -1.33008668  3.74854232  3.57590088  0.35051186  0.22893836
##  [49]  0.65418243 -1.31116404  0.77637603 -0.51135167  1.41529375  1.35877214
##  [55]  1.41223492  1.18587028  1.35346659 -1.02774205 -1.59670553  1.56483030
##  [61]  1.69947987  3.10129173  0.71016036 -1.90389707  1.83352535  0.82003090
##  [67]  1.94945756  2.28303290  1.21645906 -1.35968903 -1.98496417 -1.58641506
##  [73]  2.52157041  2.73440067  1.29353076  2.35849926 -1.75405995 -0.42946944
##  [79]  3.21972144 -0.74183693 -1.28524981  1.16776448 -0.31278653  3.40204173
##  [85] -1.88122613 -1.61165648 -1.87216411  1.56470574 -0.80096270  0.63436756
##  [91]  2.77848677 -0.96431678 -0.47197390  3.77062772  2.34507784 -0.43610903
##  [97]  2.40492803 -0.56449176 -1.52371320 -0.16195717
```

Histograma para la nuestra aleatoria de tamaño 100

```r
hist(x, main="X ~ Uniforme(min=-2, max=4", xlab="X", ylab="densidad de probabilidad", probability=TRUE,
```




Graficar la función de densidad, use la función curve() para variable continua

```r
#curve(dunif(x, min, max), col="blue", add=TRUE)
```

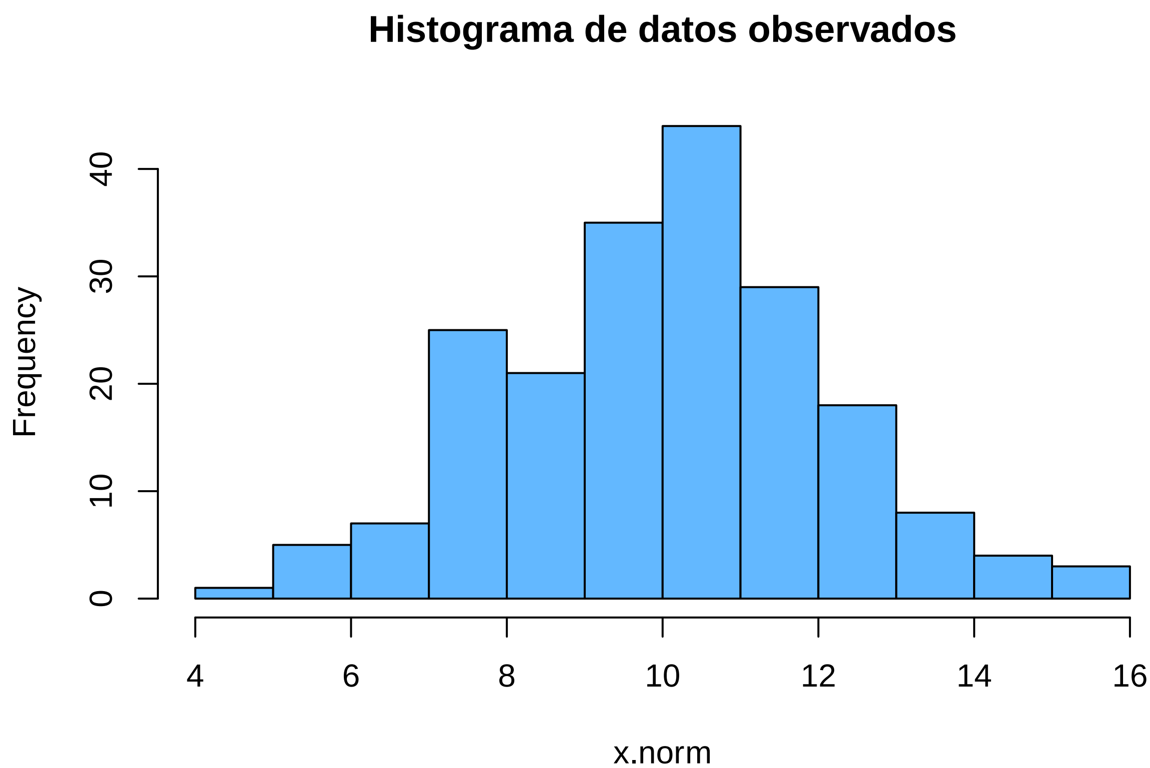
- **Ejemplo 2:** Supongamos que tenemos una muestra de tamaño n=200 perteneciente a una población normal N(10,2) con $\mu = 10$ y $\sigma = 2$:

genera los valores aleatorios de la distribución

```
x.norm<-rnorm(n=200,mean=10, sd=2)
```

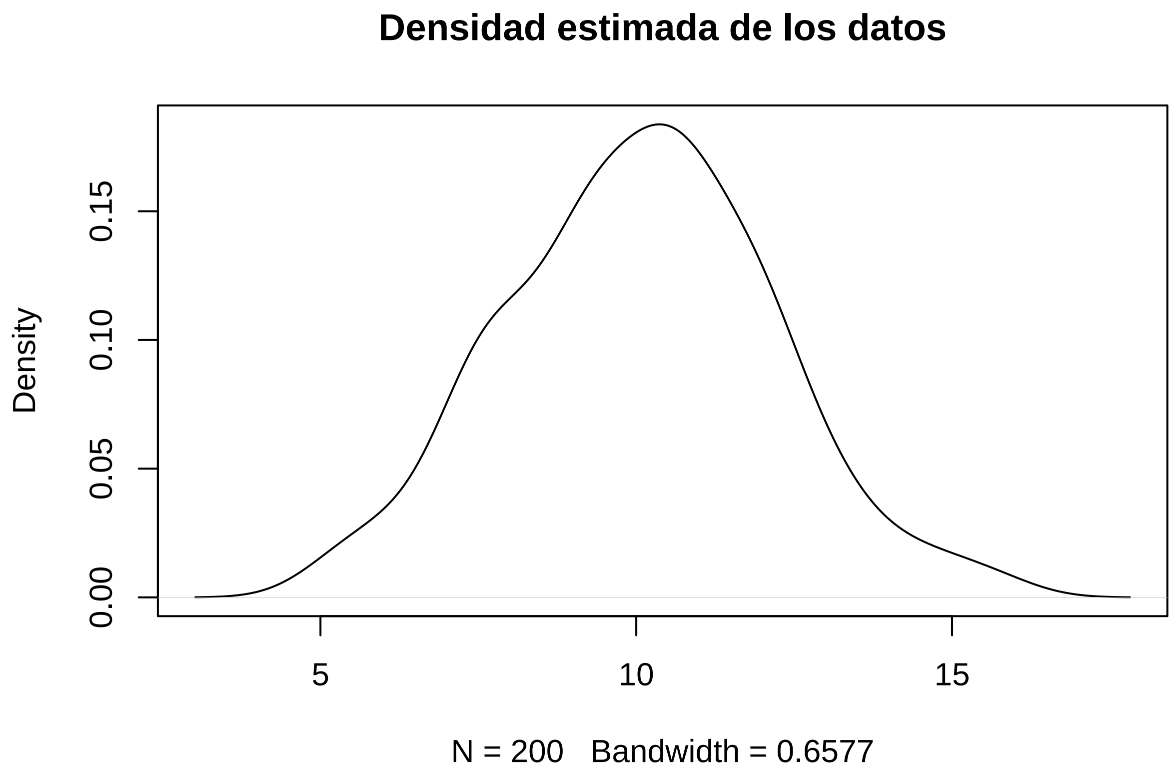
Podemos obtener un histograma usando la función hist()

```
hist(x.norm, breaks = "Sturges", freq = TRUE, probability = FALSE, include.lowest = TRUE, right= TRUE,
```



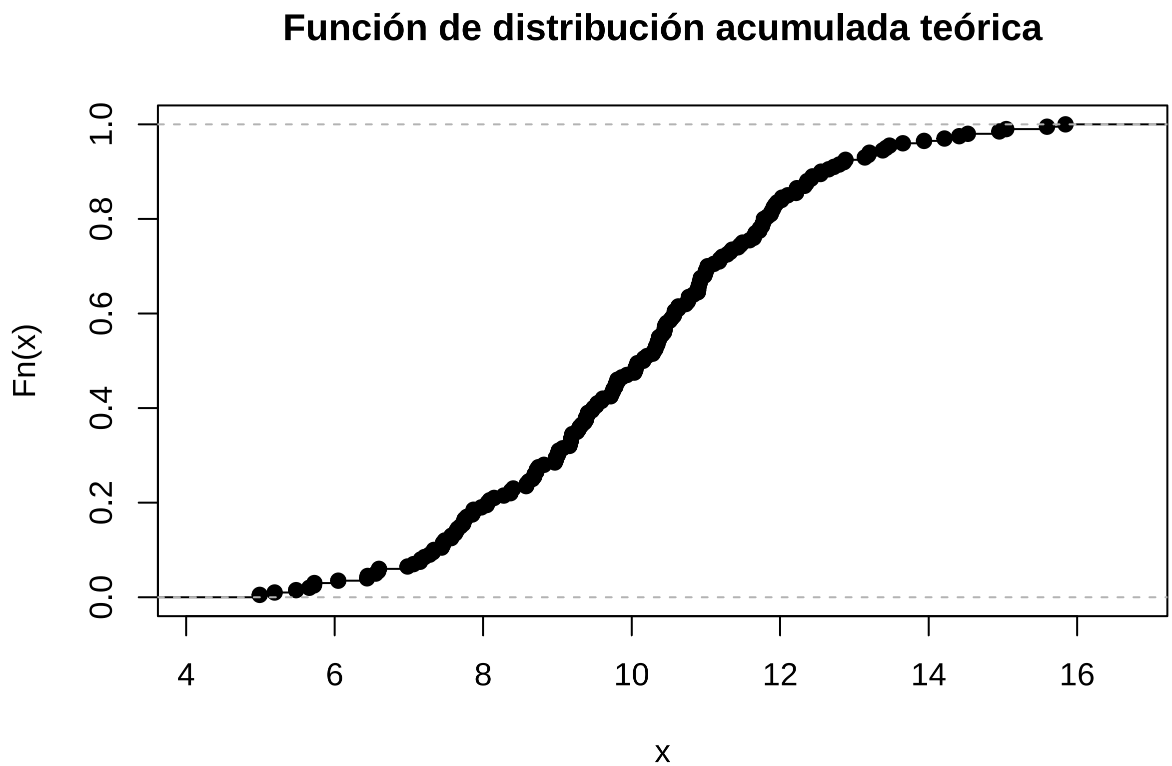

Podemos estimar la densidad de frecuencia usando la función density() y plot() para dibujar su gráfica

```
plot(density(x.norm), main="Densidad estimada de los datos")
```




R permite calcular la función de distribución acumulada teórica con ecdf()

```
plot(ecdf(x.norm),main="Función de distribución acumulada teórica")
```




- **Ejemplo 3:** Generar 100 números aleatorios de una distribución Normal con media 4.5 y desviación estándar 0.75

Definir los parámetros apropiados

```
media<-4.5
desviacion<-0.75
```

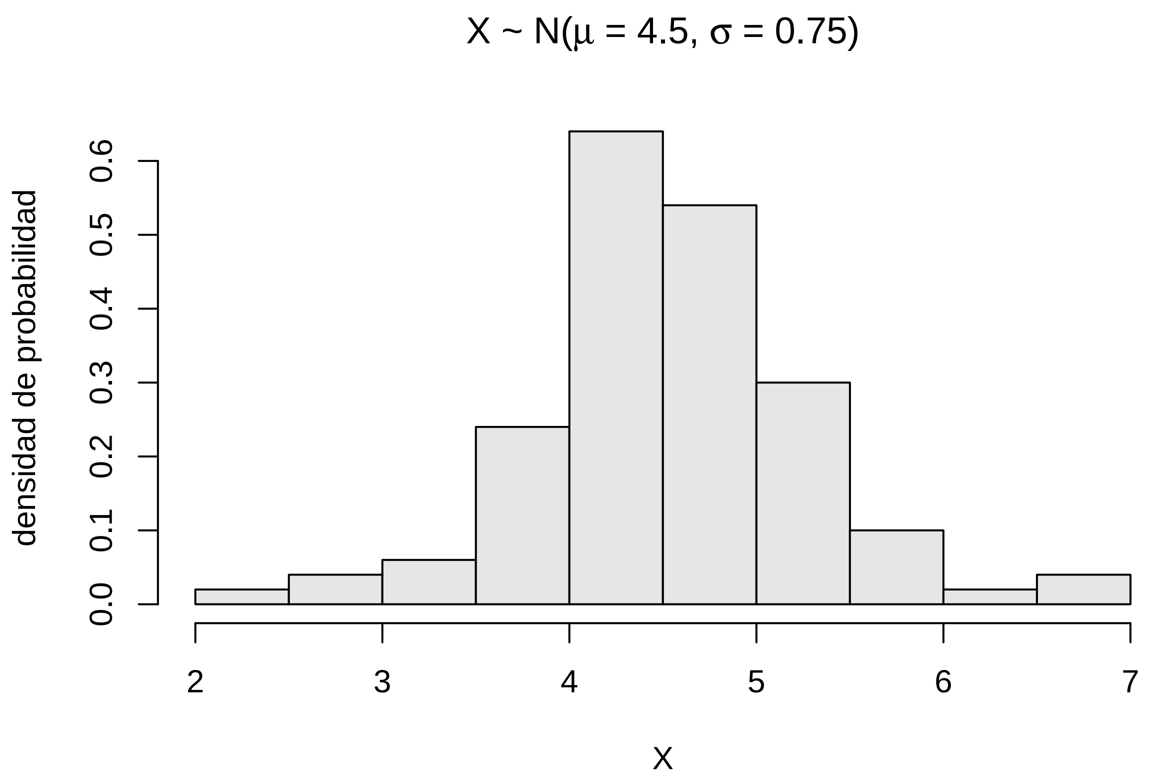
generar 100 números aleatorios de la distribución

```
x=rnorm(100, media, desviacion)
x
```

```
##   [1] 3.537795 5.546762 5.296478 4.933714 4.711293 4.851977 3.756819 4.084655
##   [9] 3.522807 4.856845 6.079890 5.705423 4.481350 3.037650 4.906168 4.658226
##  [17] 4.282941 3.959695 5.767411 5.234269 3.469770 4.940719 4.580466 4.314350
##  [25] 4.690446 3.711894 4.064927 3.772731 4.295040 5.343734 4.662040 4.259375
##  [33] 4.565537 4.443803 5.017033 5.067673 5.086332 3.619817 4.298816 4.501889
##  [41] 4.218637 5.036773 4.098609 4.290715 5.040159 4.641580 4.679107 4.247394
##  [49] 3.720178 4.196308 4.327183 5.016447 4.492368 2.645170 5.448838 4.159169
##  [57] 5.717168 3.567749 4.565497 3.870356 4.102712 4.874747 2.497272 4.746846
##  [65] 5.155140 3.476230 4.168634 4.068238 2.677986 4.397240 4.855089 6.622194
##  [73] 4.296698 4.641414 4.489380 4.261066 4.335666 4.780646 4.118977 4.964300
##  [81] 4.015162 5.368147 4.967600 3.856876 4.788898 4.753032 4.295922 4.783853
##  [89] 3.905351 4.355774 5.022423 4.175843 4.295245 4.897184 6.509485 4.094256
##  [97] 5.166767 4.590483 5.274572 5.511094
```

Histograma para la nuestra aleatoria de tamaño 100

```
hist(x,main=expression(paste("X ~ N(", mu, " = 4.5, ", sigma, " = 0.75)")), xlab="X", ylab="densidad de
```




Graficar la función de densidad teórica, usando la función curve()

```
#curve(dnorm(x, media, desviacion), col="red", lwd=2, add=TRUE)
```

- **Ejemplo 4:** Generar números aleatorios de una distribución exponencial. Por ejemplo, si la vida media de un bulbo de luz es 2500 horas, uno puede pensar que el tiempo de vida es aleatorio con una distribución exponencial que tiene media 2500. El único parámetro es la razón = 1/media.

Definir el parámetro apropiado

```
media<-2500
razon<-1/media
n=100
```

generar 100 números aleatorios de la distribución

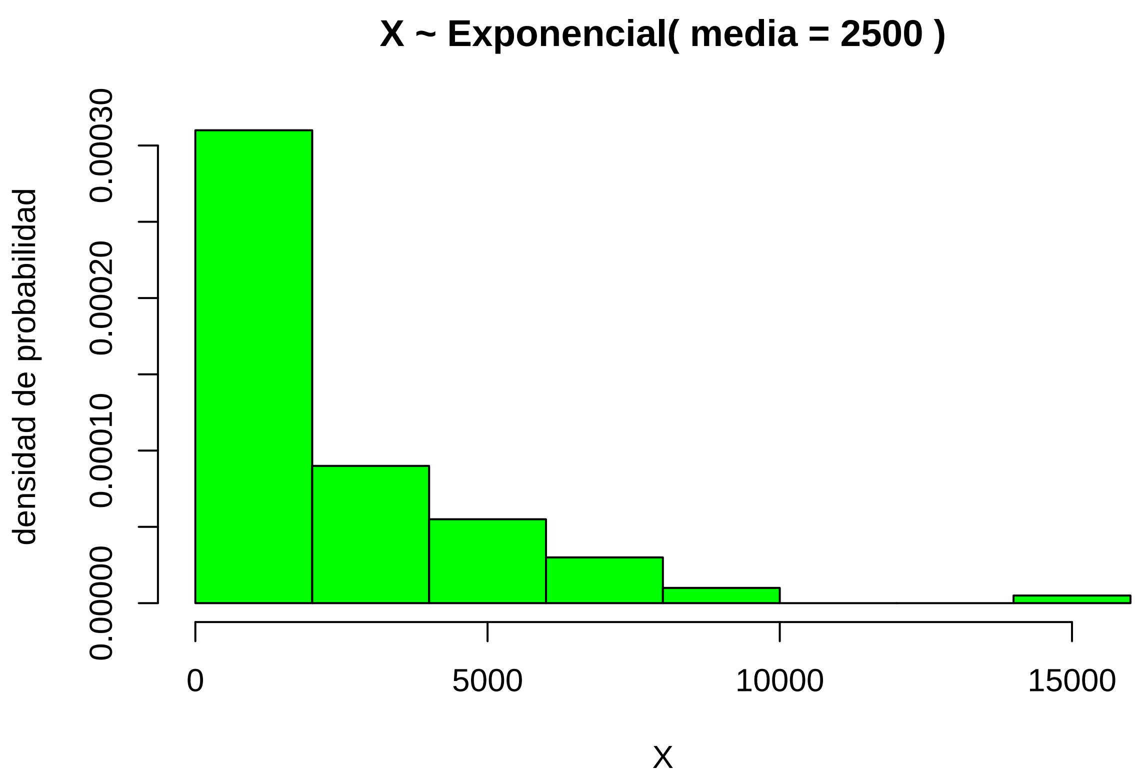
```
x=rexp(n, razon)
x
```

```
##    [1]   207.701401  7697.413582   225.303969   367.363883  1171.906986
##    [6]  8168.894350   161.741747   279.220281 14021.795056  4553.607018
##   [11]  1027.865363   220.584550  5093.642231  1527.776722  2236.798394
##   [16]   884.265590  2133.180676  1579.745526   711.820938  5135.402493
##   [21]  1554.283295  2084.237190   531.675805   104.584993   451.257088
##   [26]  2111.772469  6702.176989  1469.852922   470.638026  3571.082653
##   [31]  5985.653898  2505.703574   423.303756    82.476464  3721.769814
##   [36]   327.593352  1847.404843  1089.030628  2505.871116  1444.369563
```

```
## [41] 1502.015018  420.756391 1598.903649 4795.774645   28.254278
## [46]  166.529545 2202.905519  410.767933 6907.917023 1169.110733
## [51]  685.234156  359.395121 1098.017516 7513.055950 2764.431827
## [56] 1191.622665 3768.099965  612.342086  235.967932 2224.919977
## [61] 9717.624499 5315.338499 5554.266548 2032.733632 1161.306809
## [66] 1764.289132 2408.939987  603.553257  305.302903 7370.888609
## [71] 2269.897542 1446.786487 1045.630452 1622.014294 1584.728145
## [76]  148.590909 3798.967212 1078.522548 1322.915222  653.898645
## [81] 1743.921877    5.865317 5772.185905 4381.908495 1757.158583
## [86] 1265.357387 4761.705141   47.192129  922.078713 1979.713979
## [91] 6032.419695  760.590290   94.389468 3888.695352 1594.310188
## [96] 5430.481848  309.951755 2219.479754  544.643416  654.754934
```

Histograma para la nuestra aleatoria de tamaño 100

```
hist(x, main="X ~ Exponencial( media = 2500 )", xlab="X", ylab="densidad de probabilidad", probability=
```




Graficar la función de densidad, usando la función curve()

```
#curve(dexp(x, razon), col="blue", lwd=2, add=TRUE)
```

# FUNCIONES DE DISTRIBUCIÓN Y SU INVERSA (LOS CUANTILES)

En R, las funciones a las que se les antepone una “p” permiten contestar cuál es la probabilidad de que una variable aleatoria X sea menor o igual que x, esto es $F(x) = P[X \leq x]$ F(x). Las funciones a las que se les antepone una “q” son lo inverso de esto, ellas permiten conocer qué valor de una variable aleatoria X corresponde a una probabilidad p dada. Esto es el cuantil q X o punto en el que los datos son partidos, q $P[X \leq x_q] = p$

- **Ejemplo 1:** Para una Variable aleatoria X con distribución normal de media 1 y desviación estándar 1, ¿cuál es la probabilidad de que sea menor que 0.7?

```
x<-0.7
p<-pnorm(x, mean=1, sd=1, lower.tail = TRUE)
p
```

```
## [1] 0.3820886
```

- **Ejemplo 2:** Para una variable aleatoria con distribución normal estándar, encontrar $P[Z \leq 0.7]$ y $P[Z > 0.7]$

```
z<-0.7
p1<-pnorm(z, mean=0, sd=1)
p1
```

```
## [1] 0.7580363
```

```
p2<-pnorm(z, mean=0, sd=1, lower.tail=FALSE)
p2
```

```
## [1] 0.2419637
```

Observación: ya que $P[Z > 0.7] = 1 - P[Z \leq 0.7]$ obtenemos el mismo resultado con

```
p3 <- 1-pnorm(z, mean=0, sd=1);p3
```

```
## [1] 0.2419637
```

- **Ejemplo 3:** ¿Qué valor de una variable aleatoria con distribución normal estándar, tiene 75% del área a la izquierda?.

```
p<- 0.75
z<-qnorm(p, mean=0, sd=1, lower.tail = TRUE)
z
```

```
## [1] 0.6744898
```

Observación: note que el valor de z que resuelve $P[Z \leq z] = 0.75$ es el tercer cuartil (Q3), esto es z=0.6744898

- **Ejemplo 4:** ¿Cuál es la probabilidad a la derecha de 18.55 para una Variable aleatoria X con distribución Chi-cuadrado de 12 grados de libertad?

```
x<-18.55
gl<-12
p<-pchisq(x, gl, lower.tail = FALSE); p
```

```
## [1] 0.09998251
```