CIMAT Unidad MTY

MAESTRIA EN CÓMPUTO ESTADÍSTICO

Inferencia estadística

TAREA 1

---*Elaborada por: Román Castillo Casanova*

EJERCICIO 1

A) Todas la tarjetas tienen probabilidad de $\frac{1}{29}$ de ser escogidas, tambien hay 16 tarjetas etiquetadas con un número mayor a 13, entonces si X es la etiqueta que se observa al escoger una tarjeta al azar :

$$P(x > 13) = \frac{16}{29}$$

Entonces la probabilidad de que Anastasio escoja una tarjeta marcada con un número mayor a 13 es de $\frac{16}{29}$

B) La función de densidad para $X \sim U(11)$ es:

$$p(x) = \begin{cases} 1/11 & \text{si } 19 \leq x \leq 29 \\ 0 & \text{de otro modo} \end{cases}$$

Los primeros tres momentos de la función:

$$E(x) = \cdot \sum_{i=19}^{29} \frac{1}{11} i = 19(\frac{1}{11}) + 20(\frac{1}{11}) + \cdots + 28(\frac{1}{11}) + 29(\frac{1}{11}) = (\frac{264}{11}) = 24$$

$$E(x^2) = \cdot \sum_{i=19}^{29} \frac{1}{11} i^2 = 19^2(\frac{1}{11}) + 20^2(\frac{1}{11}) + \cdots + 28^2(\frac{1}{11}) + 29^2(\frac{1}{11}) = (\frac{6446}{11}) = 586$$

$$E(x^3) = \cdot \sum_{i=19}^{29} \frac{1}{11} i^3 = 19^3(\frac{1}{11}) + 20^3(\frac{1}{11}) + \cdots + 28^3(\frac{1}{11}) + 29^3(\frac{1}{11}) = (\frac{159984}{11}) = 14544$$

C) Ya que todos los números tienen la misma posibilidad de ser elegidos, calculamos la varianza correspondiente a una distribución uniforme discreta

$$V(x) = \frac{20^2 - 1}{12} = \frac{399}{12} = 33.25$$

D) La esperanza para un dado balanceado se calcula para una uniforme discreta (n=6).

$$E(x) = \frac{6+1}{2} = 3.5$$

EJERCICIO 2

Tomemos a x como la demanda del producto que puede tomar tomar valores entre $\{0, N\}$, también llamamos como ϕ al stock producido con antelación, que toma valores en el mismo rango de la demanda. La función de beneficio en términos de la demanda está dada por:

$$b(x) = \begin{cases} gx - p(\phi - x) & \text{si } x \leq \phi \\ g\phi & \text{si } x > \phi \end{cases}$$

Así que el valor esperado de la ganancia es:

$$E(b(x)) = \sum_X b(x_i) \cdot p(x_i)$$

Sustituyendo:

$$E(b(x)) = \sum_0^{\phi}[gx_i - p(\phi - x_i)] \cdot (\frac{1}{N+1}) \ + \sum_{\phi+1}^{N} g\phi \cdot (\frac{1}{N+1})$$

$$E(b(x)) = \frac{1}{N+1}[g\sum_0^{\phi} x_i \ - \sum_0^{\phi} p\phi + p\sum_0^{\phi} x_i \ + \sum_{\phi+1}^{N} g\phi]$$

$$E(b(x)) = \frac{1}{N+1}[g\frac{\phi \cdot (\phi+1)}{2} - \phi(\phi+1)p + p\frac{\phi \cdot (\phi+1)}{2} + \phi(N-\phi)g]$$

Reordenando:

$$E(b(x)) = \frac{1}{N+1}[\phi \cdot (\phi+1)(\frac{g-p}{2}) + g\phi(N-\phi)]$$

Para maximizar la esperanza en términos de ϕ:

$$\frac{dE(b(x))}{d\phi} = \frac{1}{N+1}[(2\phi+1)(\frac{g-p}{2}) + gN - 2g\phi]$$

Igualamos a cero y despejamos ϕ:

$$\phi = \frac{g(2N+1) - p}{2(g+p)}$$

Tomamos como producción a $\phi*$ que es el entero superior o inferior a ϕ que tiene el mayor valor en $B(x)$

EJERCICIO 3

A) Para calcular la probabilidad de que no más de dos bits de un paquete se corrompan, calculamos la probabilidad binomial (N=12, p=0.1)

$$P(X \leq 2) = \sum_{i=0}^{2} \binom{12}{i} (0.1)^i (0.9)^{12-i}$$

Se realiza el cálculo apoyado en R:

```
round(pbinom(2,12,0.1,lower.tail = T),3)
```

```
## [1] 0.889
```

La probabilidad de que no más de dos bits se corrompan es de 0.889

B) Para determinar la probabilidad de que al menos un paquete de los 6 enviados tengan 3 o más bit corruptos, se calcula la probabilidad de un paquete tenga a lo menos 3 bits corruptos, se observa que esta probabilidad es el complemento de la calculada en el ejercicio anterior:

$$p = P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.889 = 0.111$$

Los paquetes son independientes entre sí, y sea Y el número de paquetes con 3 o más bits corrompidos ($Y \sim Bin(6, 0.111)$):

$$P(Y \geq 1) = 1 - P(Y = 0)$$

$$P(Y \geq 1) = 1 - (0.889)^6 = 1 - 0.493 = 0.507$$

La probabilidad de que al menos un paquete tenga al menos 3 bits corruptos es de 0.507.

C) Para el caso anterior se obtiene la media ($\mu = (6)(0.111) = 0.666$) y desviacion estándar ($\sigma = \sqrt{(6)(0.111)(0.889)} = 0.769$). Exceder en dos desviaciones a la media, implica que X tome valores mayores de 2.204, es decir de 3 en adelante. Esta probabilidad es:

$$P(X \geq 3) = \sum_{i=3}^{6} \binom{6}{i} (0.111)^i (0.899)^{12-i}$$

Al realizar el cálculo se obtiene:

```
## Redondeamos a 3 cifras
round(pbinom(2,6,0.11, lower.tail = F),3)
```

```
## [1] 0.021
```

Hay 0.021 de probabilidad de que los bits corrompidos excedan en 2 desviaciones a la media

EJERCICIO 4

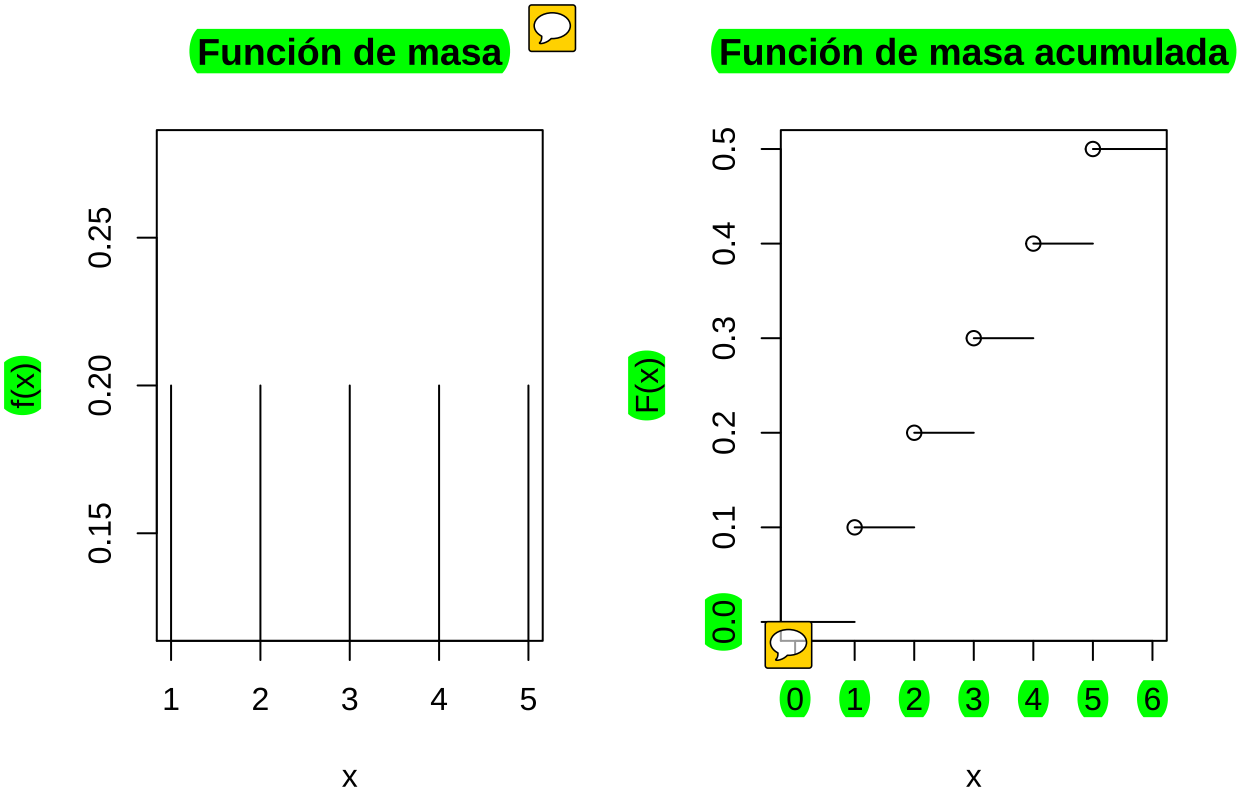
Dado que X es una variable aleatoria que se distribuye hipergeométrica ($X \sim Hiper(12, 3, 4)$), su esperanza es:

$$E(X) = 3(\frac{12}{16}) = 2.25$$

EJERCICIO 5

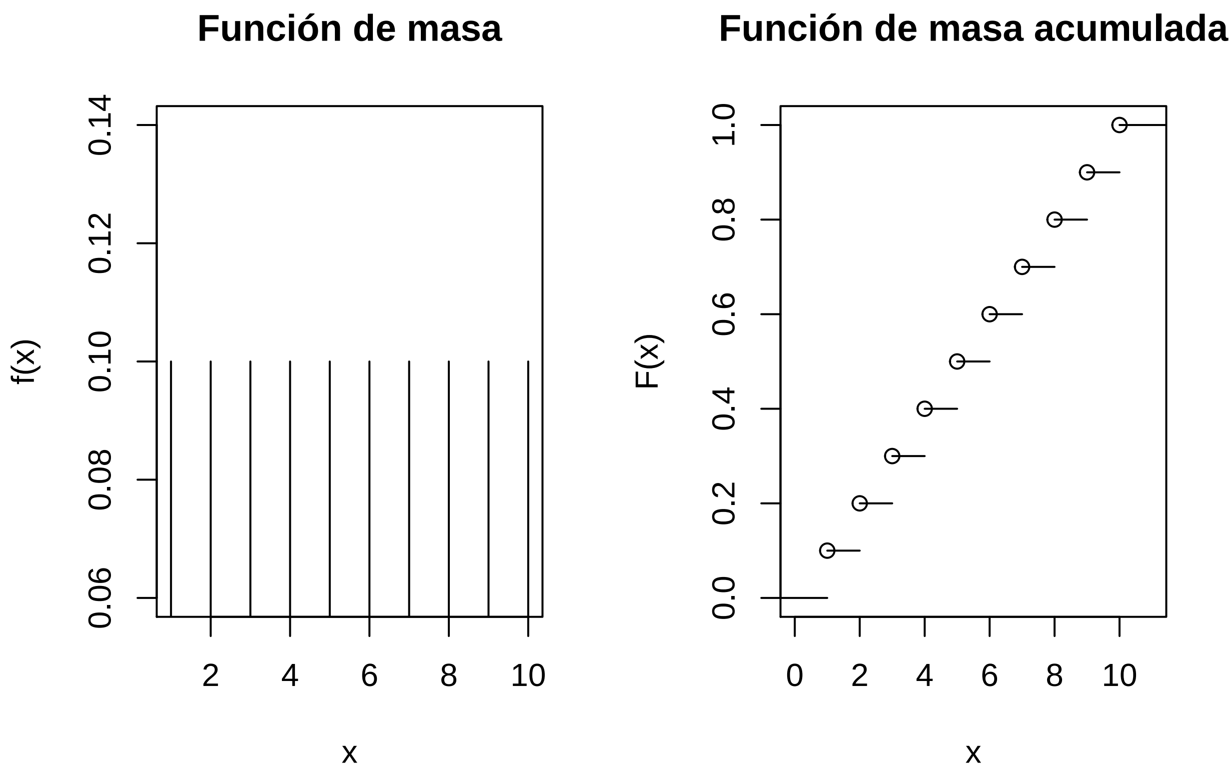
A) **Distribución uniforme para N=5**

```
#Se crea una retícula de 1,2
par(mfrow=c(1,2))
n<-5
x<-c(1:n)
y<-rep(1/n,n)
z<-0.1*c(0:n)
#Función de masa
plot(x,y,type = "h", xlab="x",ylab="f(x)", main = "Función de masa")
#Función de masa acumulada
plot(stepfun(x,z), verticals= FALSE, xlab="x",ylab="F(x)", main = "Función de masa acumulada" )
```



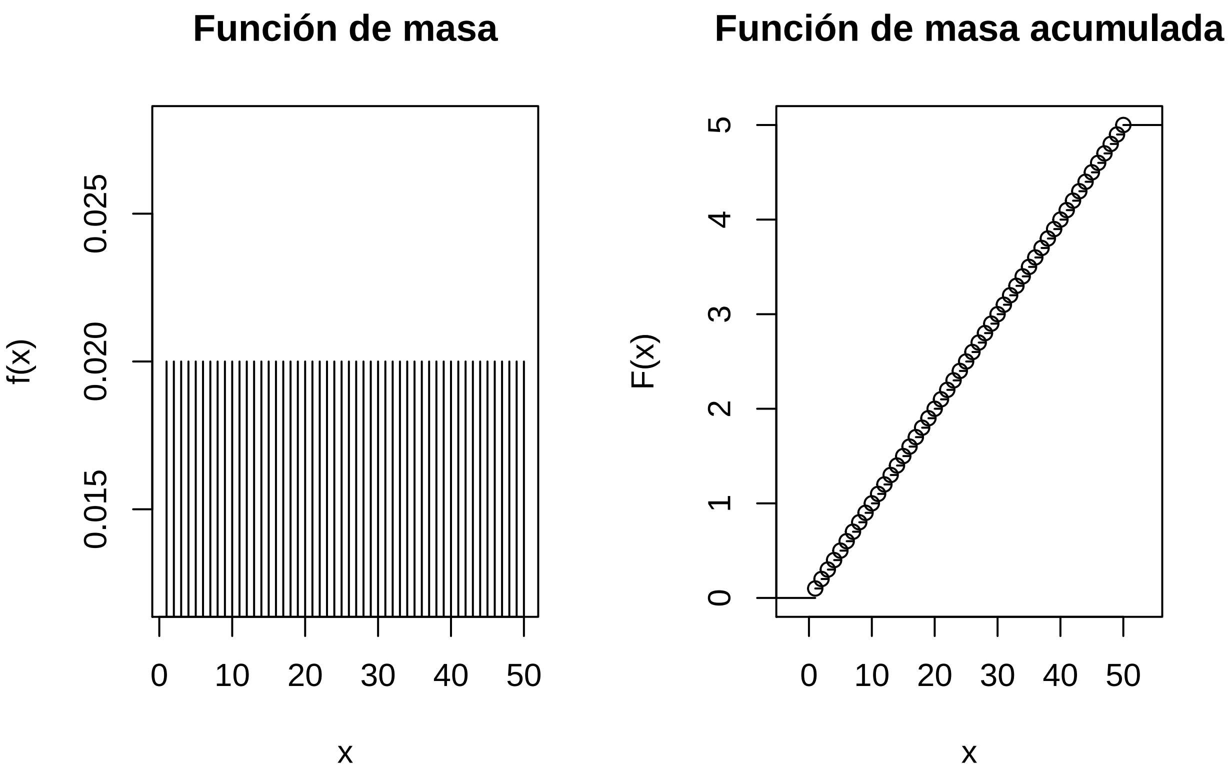
Distribución uniforme para N=10

```
par(mfrow=c(1,2))
n<-10
x<-c(1:n)
y<-rep(1/n,n)
z<-0.1*c(0:n)
#Función de masa
plot(x,y,type = "h", xlab="x",ylab="f(x)", main = "Función de masa")
#Función de masa acumulada
plot(stepfun(x,z), verticals= FALSE, xlab="x",ylab="F(x)", main = "Función de masa acumulada" )
```



Distribución uniforme para N=50

```r
par(mfrow=c(1,2))
n<-50
x<-c(1:n)
y<-rep(1/n,n)
z<-0.1*c(0:n)
#Función de masa
plot(x,y,type = "h", xlab="x",ylab="f(x)", main = "Función de masa")
#Función de masa acumulada
plot(stepfun(x,z), verticals= FALSE, xlab="x",ylab="F(x)", main = "Función de masa acumulada" )
```



B) ______________________________

C)

```r
set.seed(13)

# se guarda la muestra de la distribución uniforme
muestra<-sample(1:10,10000, replace = T, prob = NULL)
#cálculo de la media, usando dos cifras significativas
mu<-round(mean(muestra),2)
#cálculo de la varianza, usando dos cifras significativas
s<-round(var(muestra),2)
#Se muestran los resultados: tabla de frecuencias, histograma, media, varianza
print(paste("-----","TABLA DE FRECUENCIAS","-----"))
```

```
## [1] "----- TABLA DE FRECUENCIAS -----"
```

```r
table(muestra)
```

```
## muestra
##    1    2    3    4    5    6    7    8    9   10
## 1015 1001  962 1013 1020  982  991  926 1069 1021
```

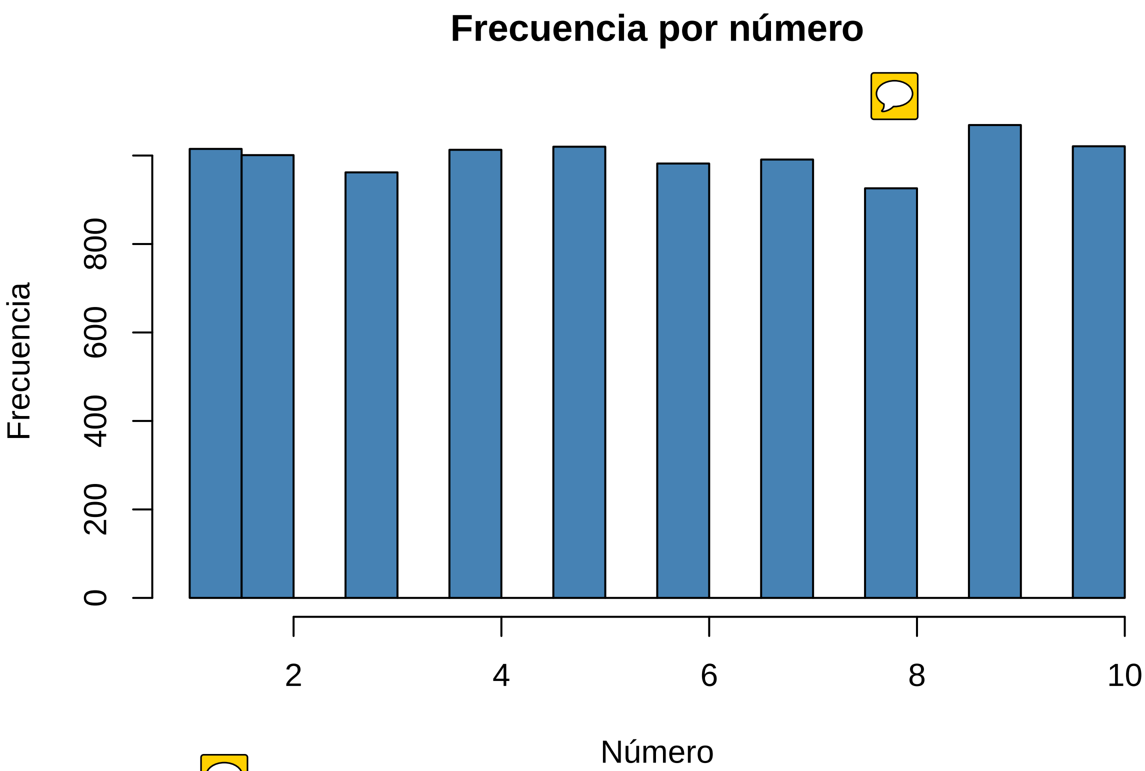
```r
paste("La media es: ",mu," y la varianza es: ", s)
```

```
## [1] "La media es:  5.51  y la varianza es:  8.34"
```

```r
print(paste("-----","HISTOGRAMA","-----"))
```

```
## [1] "----- HISTOGRAMA -----"
```

```r
hist(muestra, main = "Frecuencia por número", xlab = "Número", ylab="Frecuencia", col = "steelblue")
```



EJERCICIO 6

INCISOS A y B

```r
set.seed(13)
# Codificar como águila=1, sol=0
moneda<-c(1,0)
# Se crea un vector para guardar la cantidad de águilas por corrida de 10
aguila_freq<-rep(0,10^6)
i<-1

# Se realizan el experimento, se guardan la cantidad de aguilas
while(i<=10^6){
  aguila_freq[i]<-sum(sample(moneda,10,replace = T, prob = c(0.5,0.5)))

  if(i<2){
    print(paste("-----","RESULTADOS DE LOS PRIMEROS TRES LANZAMIENTOS","-----"))
  }
  # Se indica la impresión de los primeros 3 lanzamientos
  if(i<4){
  print(paste("El número de águilas en el",i,"° tiro es de",aguila_freq[i]))}
  i=i+1}
```

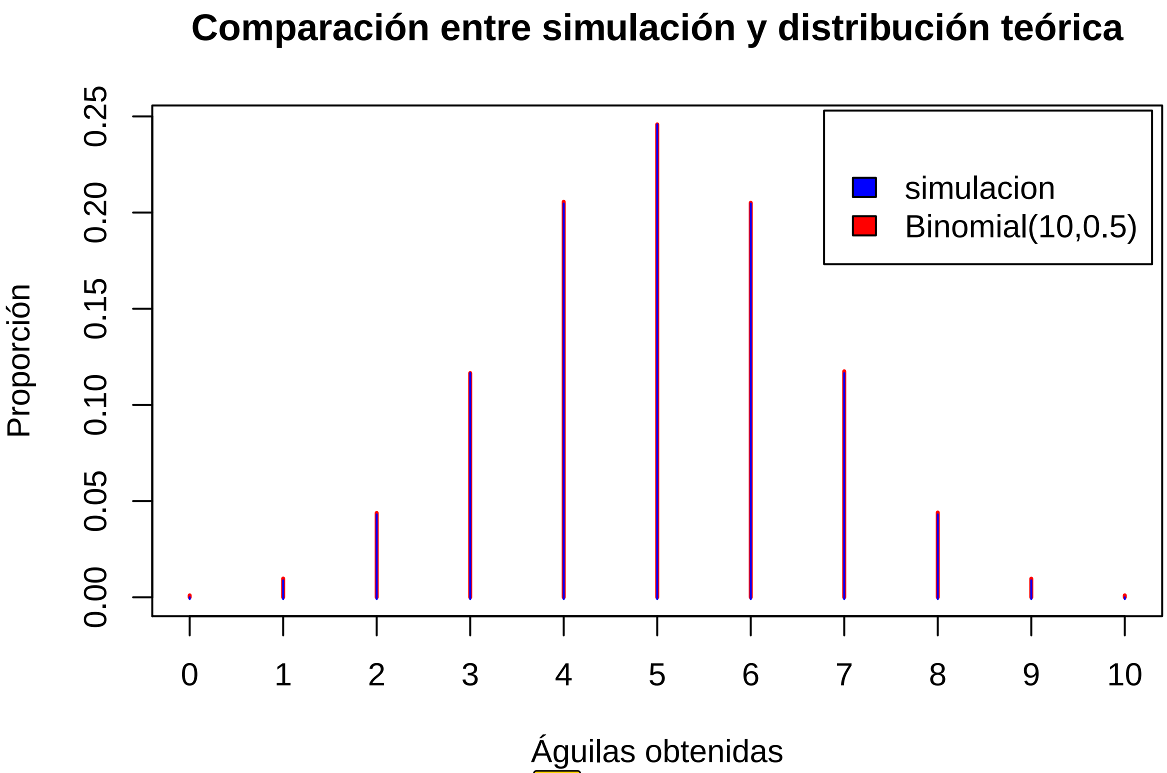
```
## [1] "----- RESULTADOS DE LOS PRIMEROS TRES LANZAMIENTOS -----"
## [1] "El número de águilas en el 1 ° tiro es de 5"
## [1] "El número de águilas en el 2 ° tiro es de 8"
## [1] "El número de águilas en el 3 ° tiro es de 5"
```

```r
# Construcción de la tabla de proporciones
prop.table(table(aguila_freq))
```

```
## aguila_freq
```

```
##        0        1        2        3        4        5        6        7
## 0.001004 0.009740 0.043841 0.116592 0.205615 0.245846 0.205099 0.117499
##        8        9       10
## 0.044071 0.009681 0.001012
```

```
x<-c(0:10)

#Grafica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(aguila_freq)),type="h",col="red", xlab = "", ylab="", main = "Comparación entre s
par(new=TRUE)
#Gráfica de la distribución teórica Bin(10,0.5)
plot(dbinom(x,10,0.5),type="h", axes= F, col="blue",xlab = "Águilas obtenidas", ylab="Proporción" )
legend("topright", inset=.01, title="",
    c("simulacion","Binomial(10,0.5)"), fill=c("blue","red"), horiz=F)
```



C) Repetición con moneda no justa $p = 0.3$

```
set.seed(13)
# Codificar como águila=1, sol=0
moneda<-c(1,0)
# Se crea un vector para guardar la cantidad de águilas por corrida de 10
aguila_freq<-rep(0,10^6)
i<-1

# Se realizan el experimento, se guardan la cantidad de aguilas
while(i<=10^6){
  aguila_freq[i]<-sum(sample(moneda,10,replace = T, prob = c(0.3,0.7)))

  if(i<2){
    print(paste("-----","RESULTADOS DE LOS PRIMEROS TRES LANZAMIENTOS","-----"))
  }
```

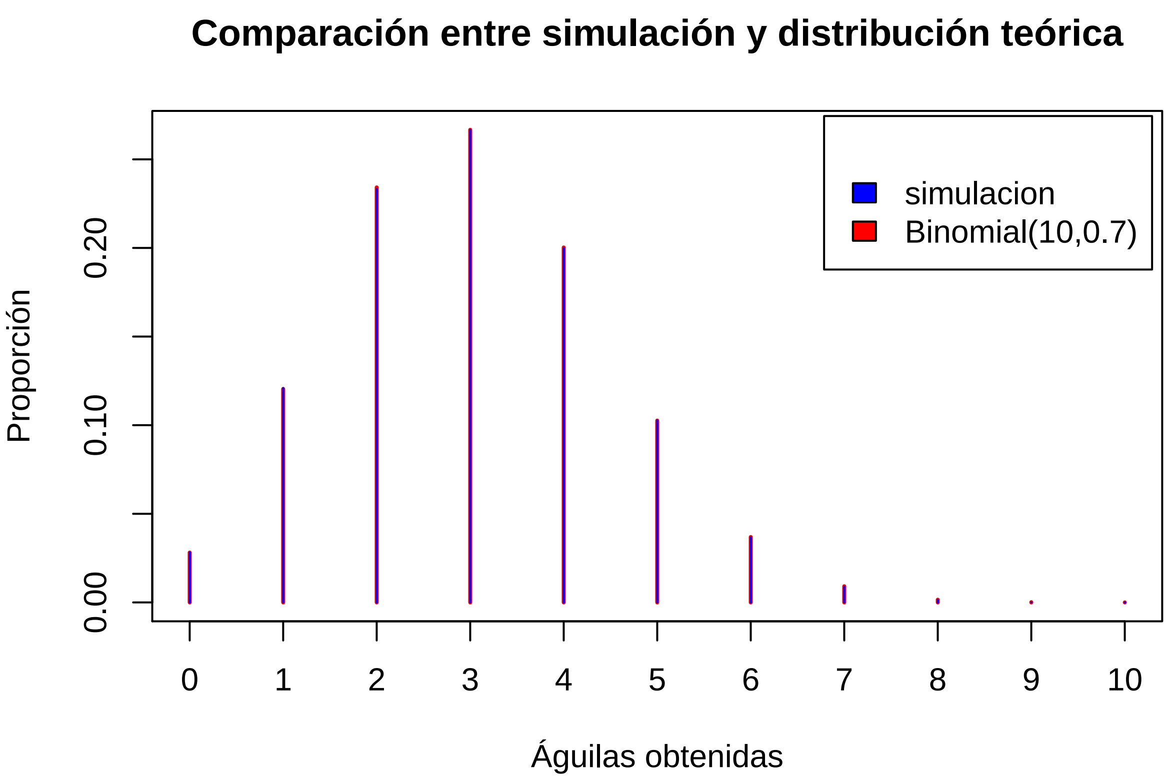
```r
  # Se indica la impresión de los primeros 3 lanzamientos
  if(i<4){
  print(paste("El número de águilas en el",i,"° tiro es de",aguila_freq[i]))}
  i=i+1}
```

```
## [1] "----- RESULTADOS DE LOS PRIMEROS TRES LANZAMIENTOS -----"
## [1] "El número de águilas en el 1 ° tiro es de 4"
## [1] "El número de águilas en el 2 ° tiro es de 3"
## [1] "El número de águilas en el 3 ° tiro es de 0"
```

```r
# Construcción de la tabla de proporciones
prop.table(table(aguila_freq))
```

```
## aguila_freq
##        0        1        2        3        4        5        6        7
## 0.028127 0.120457 0.234168 0.266659 0.200285 0.102604 0.036910 0.009121
##        8        9       10
## 0.001523 0.000141 0.000005
```

```r
x<-c(0:10)

#Gráfica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(aguila_freq)),type="h",col="red", xlab = "", ylab="", main = "Comparación entre si
par(new=TRUE)
#Gráfica de la distribución teórica Bin(10,0.5)
plot(dbinom(x,10,0.3),type="h", axes= F, col="blue",xlab = "Águilas obtenidas", ylab="Proporción" )
legend("topright", inset=.01, title="",
    c("simulacion","Binomial(10,0.7)"), fill=c("blue","red"), horiz=F)
```



Es de observar que a mayor número de repiticiones del experimento, más se aproxima a las distribución teórica.

EJERCICIO 7

INCISOS A y B

Se define una función para cálcular probabilidades para $X \sim B(123, 0.31)$

```
mibinom <-function(a,b){
  #Se calcula la probabilidad directamente la función de masa
  #Se ingresa por defecto p=0.31, N=123. Se agregan el rango de x
  #como parámetros de la función
  if(a==b){x<-choose(123,a)*(0.31^a)*(1-0.31)^(123-a)}
  if(a<b){
    x<-0
    for(i in a:b){
      y<-choose(123,i)*(0.31^i)*(1-0.31)^(123-i)
      x<-x+y
    }
  }
 return(x)
  }
```

Ahora calculamos las probabilidades propuestas, se muestra primero la salida de la función escrita(a) y posteriormente la probabilidad calculada con las funciones *pbinom* y *dbinom* (b), se observa coincidencia en los cálculos:

$P(X = 0) =$

```
mibinom(0,0)
```

```
## [1] 1.508128e-20
```

```
dbinom(0,123,0.31)
```

```
## [1] 1.508128e-20
```

$P(X = 123) =$

```
mibinom(123,123)
```

```
## [1] 2.738346e-63
```

```
dbinom(123,123,0.31)
```

```
## [1] 2.738346e-63
```

$P(X = 62) =$

```
mibinom(62,62)
```

```
## [1] 3.275387e-06
```

```
dbinom(62,123,0.31)
```

```
## [1] 3.275387e-06
```

ii) $P(0 \leq X \leq 10) =$

```
mibinom(0,10)
```

```
## [1] 9.364108e-10
```

```
pbinom(10,123,0.31)
```

```
## [1] 9.364108e-10
```

$P(0 < X \leq 10) =$

```
mibinom(1,10)
```

```
## [1] 9.364108e-10
```

```
pbinom(10,123,0.31)-dbinom(0,123,0.31)
```

```
## [1] 9.364108e-10
```

$P(0 \leq X < 10) =$

```
mibinom(0,9)
```

```
## [1] 1.783427e-10
```

```
pbinom(9,123,0.31)
```

```
## [1] 1.783427e-10
```

$P(X > 11)$

```
mibinom(12,123)
```

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

```
pbinom(11,123,0.31,lower.tail = F)
```

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

$P(X \geq 10) =$

```
mibinom(10,123)
```

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

```
pbinom(9,123,0.31,lower.tail = F)
```

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

C) Debido a que se trata de una distribución discreta, para el cálculo de cuantiles tomaremos la primera x que cumpla que $P(X \leq x) =\geq q$. Se escribe una función para estimar cuantiles para $X \sim B(123, 0.31)$, tambien se realiza el cálculo con la función *qbinom* de la base de R para confirmación de resultados:

```
ElCuantil<-function(x){
 #Se calcula la función acumulada hasta que se encuentra la primera que  supere al cuantil deseado
   y<-0
  i<-0
  while(y<x){
    y<-pbinom(i,123,0.31)
    i<-i+1}
  print(i-1)}
```

Q=0.25

```
ElCuantil(0.25)
```

```
## [1] 35
```

```
qbinom(0.25,123,.31)
```

```
## [1] 35
```

Q=0.50

```r
ElCuantil(0.5)
```

```
## [1] 38
```

```r
qbinom(0.5,123,.31)
```

```
## [1] 38
```

Q=0.75

```r
ElCuantil(0.75)
```

```
## [1] 42
```

```r
qbinom(0.75,123,.31)
```

```
## [1] 42
```

EJERCICIO 8

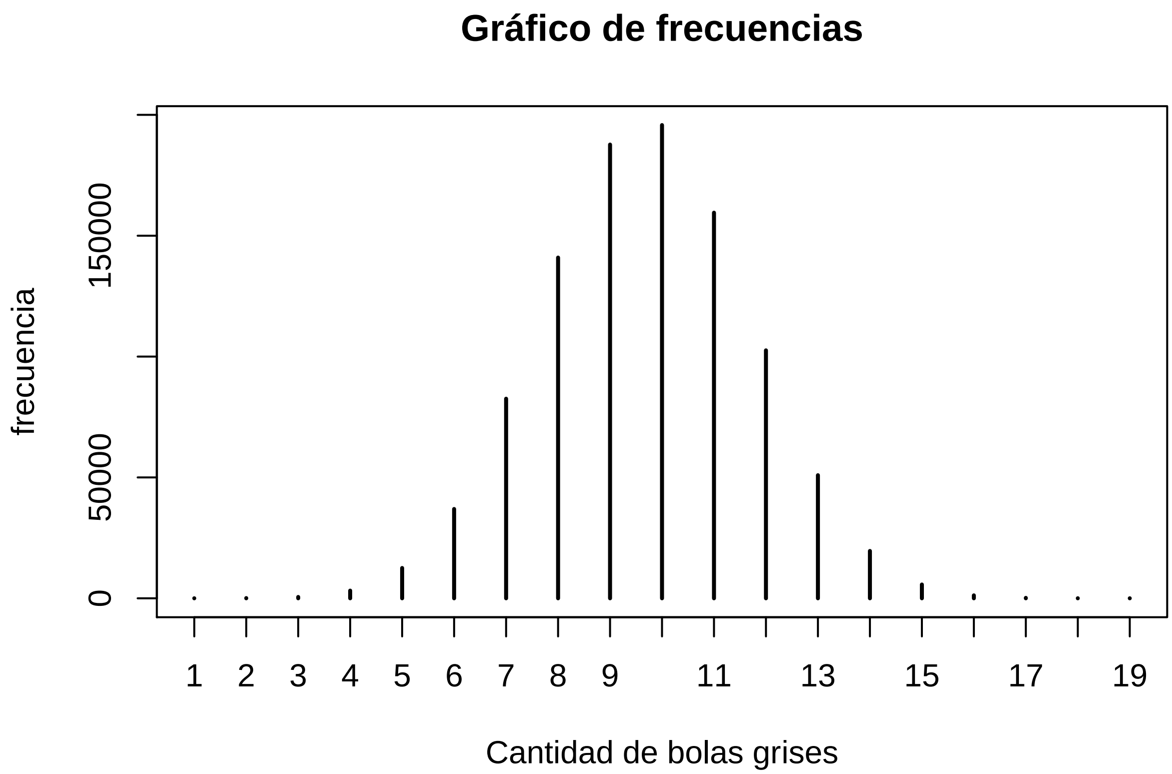
```r
set.seed(13)
# Se codifica bola gris= 1,  bola blanca=0
bolas<-c(rep(1,46),rep(0,49))
# Vector en el cuál se guardarán los conteos de grises
gris_freq<-rep(0,10^6)
i<-1
# Se realiza el millón de repeticiones
while(i<=10^6){
# Impresión de los primeros tres resultados
  gris_freq[i]<-sum(sample(bolas,20,replace = F))
  if(i<4){print(paste("En el",i,"° tiro, se obtuvo ",gris_freq[i], "bolas grises"))}

  i=i+1}
```

```
## [1] "En el 1 ° tiro, se obtuvo  7 bolas grises"
## [1] "En el 2 ° tiro, se obtuvo  11 bolas grises"
## [1] "En el 3 ° tiro, se obtuvo  11 bolas grises"
```

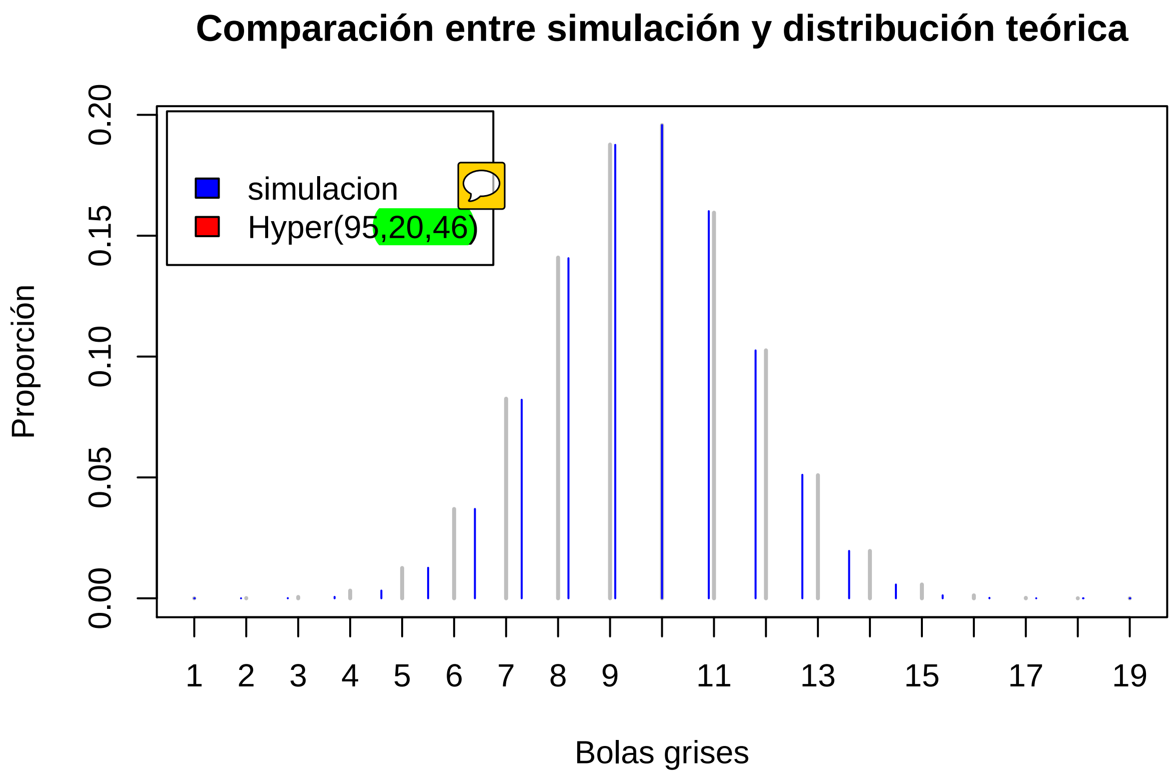
```r
table(gris_freq)
```

```
## gris_freq
##      1      2      3      4      5      6      7      8      9     10
##      7     82    550   3180  12540  36948  82573 140951 187707 195760
##     11     12     13     14     15     16     17     18     19
## 159520 102571  50923  19583   5701   1212    173     17      2
```

```r
plot(table(gris_freq), xlab = "Cantidad de bolas grises", ylab = "frecuencia", main="Gráfico de frecuen
```




```
# Comparación con la distribución teorica
x<-c(0:20)
plot((prop.table(table(gris_freq))),type="h",col="gray", xlab = "", ylab="", main = "Comparación entre s
par(new=TRUE)
plot(x,dhyper(x,46,49,20),type="h", axes= F, col="blue",xlab = "Bolas grises", ylab="Proporción")
legend("topleft", inset=.01, title="",
    c("simulacion","Hyper(95,20,46)"), fill=c("blue","red"), horiz=F)
```



Por último la probabilidad de que al extraer 20 bolas se obtengan 5 grises es:

$$P(X=5)=\frac{\binom{5}{46}\cdot\binom{15}{49}}{\binom{95}{20}}$$

Al correrse en R, se obtiene

```
round(dhyper(5,46,49,20),3)
```

```
## [1] 0.013
```

Hay una probabilidad de 0.13 de obtener 5 bolas grises en las 20 que se sacan