Tarea 2 inferencia

Román Castillo C.

EJERCICIO 1

A.

Si nombramos como X, a la cantidad de hijos antes de que tengan una niña, podemos decir que $X \sim Geom(p = 0.5)$. La probabilidad de tener más de 4 hijos antes de una niña se expresa por:

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - (1 - q^3) = q^3$$

Sustituyendo $q = 0.5$:

$$P(X \geq 4) = (0.5)^3 = 0.125$$

La probabilidad de que tengan más de cuatro hijos antes de tener una niña es de 0.125

B.

El valor esperado para la familia, sabiendo que $X \sim Geom(p = 0.5)$ es:

$$E(x) = \frac{1}{0.5} = 2$$

Es decir se espera que haya un par de hijos (un varón y una mujer).

EJERCICIO 2

A.

Si llamamos a Y como la cantidad de artículos producidos por la máquina mal ajustada hasta producir 3 defectusos, podemos notar que $Y \sim BN(r = 3, p = 0.15)$. La probabilidad de que se produzcan al menos 5 artículos por la máquina antes de ser detenida se expresa como:

$$P(X \geq 5) = 1 - P(X \leq 4)$$

Para el cálculo de esta utilizamos la función *pnbinom* en R

```
1- pnbinom(1,3,0.15)
```

```
## [1] 0.9880187
```

La probabilidad de que se produzcan al menos 5 artículos antes de detectar 3 defectuosos es de 0.98

B.

El valor esperado de artículos producidos hasta tener que detener la máquina es :

$$E(X) = \frac{3}{0.15} = 20$$

Se espera que la máquina mal ajustada tenga que detenerse después de producir 20 artículos.

EJERCICIO 3

A.

La probabilidad deseada puede obtenerse usando una binomial negativa con parámetros ($p = 0.4$, $r = 3$). Realizamos el cálculo de la probabilidad:

```
dnbinom(7,3,0.4)
```

```
## [1] 0.06449725
```

Se obtiene una probabilidad de 0.0645 de analizar 10 trabajadores para encontrar los 3 deseados

EJERCICIO 4

A.

Código para la simulación de lanzamientos:

```
simgeo<-function(N,p){

##Moneda 1:águila 0:sol
moneda<-c(1,0)

## Creación del vector donde se guardan resultados
resultados<-rep(0,N)
## Número de éxitos para terminar la simulación
r<-1

## Inicio de la simulación

for(i in 1:N){
  #contador del número de tiros
  tiros<-0
  #contador del número de éxitos
  exitos<-0
## Se tira la moneda obtener r éxitos
  while(exitos<r){
    if(sample(moneda,1,replace=T, prob = c(p,1-p))==1)
      { exitos=1}
    tiros<-tiros +1
  }
## Se guarda el número de tiros
  resultados[i]<-tiros
}
   return(resultados)
}
```

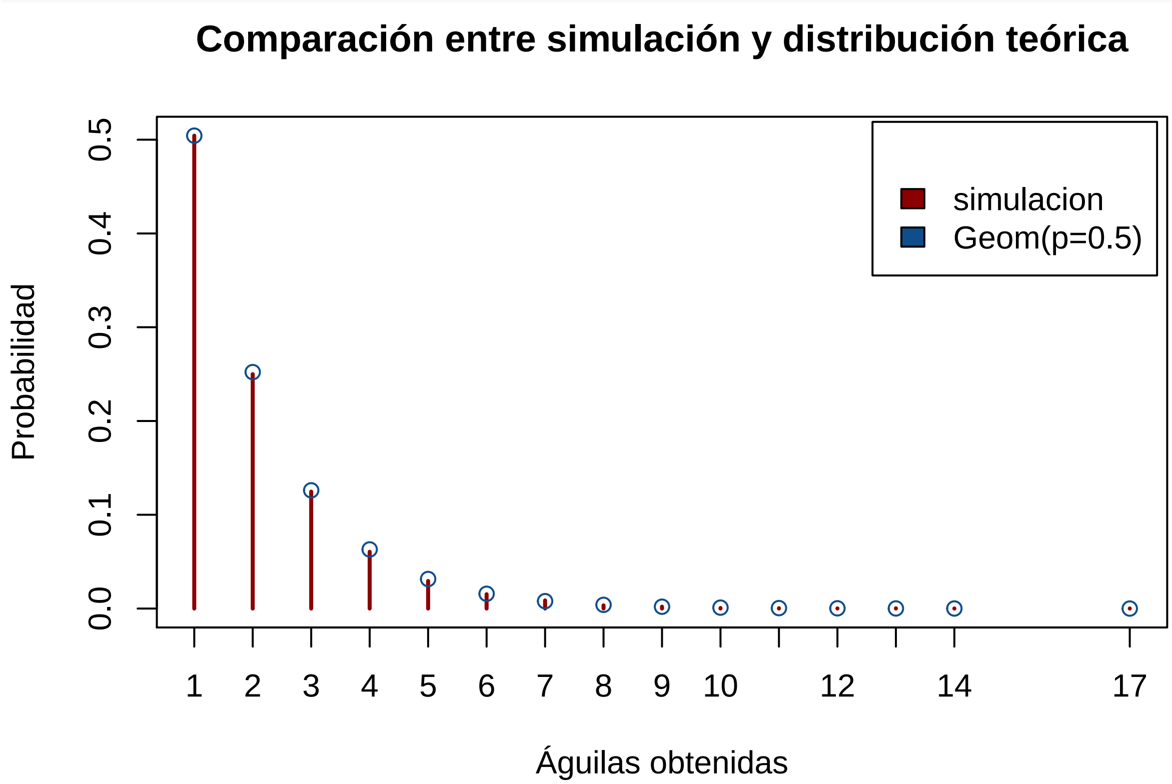
B.

Se realizan 10^4 simulaciones con $p_1 = 0.5$,$p_2 = 0.1$,$p_3 = 0.01$

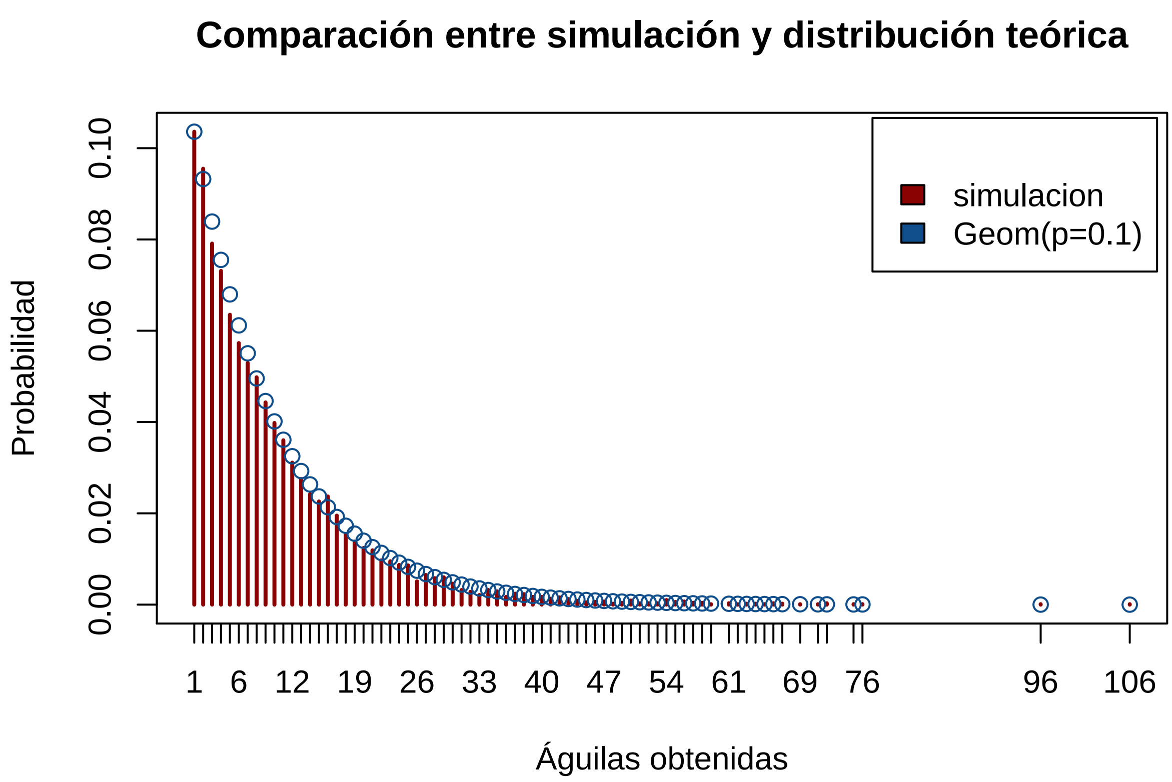
```
set.seed(15)
p_0.5<-simgeo(10^4,0.5)
p_0.1<-simgeo(10^4,0.1)
p_0.01<-simgeo(10^4,0.01)
```

Ahora se realizan los gráficos solicitados con la comparación de masa correspondiente

```
# Retícula para los gráficos
#par(mfrow=c(3,1))


##layout(matrix(c(1,2,3), 3, 1, byrow = TRUE),
    ##widths=c(1,1), heights=c(1,1))

#-------------------------------------p=0.5

#Grafica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(p_0.5)),
     type="h",
     col="darkred",
     xlab = "",
     ylab="",
     main = "Comparación entre simulación y distribución teórica" )
par(new=TRUE)
#Gráfica de la distribución teórica Geom(p=0.5)

x<-sort(unique(p_0.5))
plot((x+0), dgeom(x,0.5),
     type="p",
     axes= F,
     col="dodgerblue4",
     xlab = "Águilas obtenidas",
     ylab="Probabilidad" )
legend("topright", inset=.01, title="",
    c("simulacion","Geom(p=0.5)"), fill=c("darkred","dodgerblue4"), horiz=F)
```

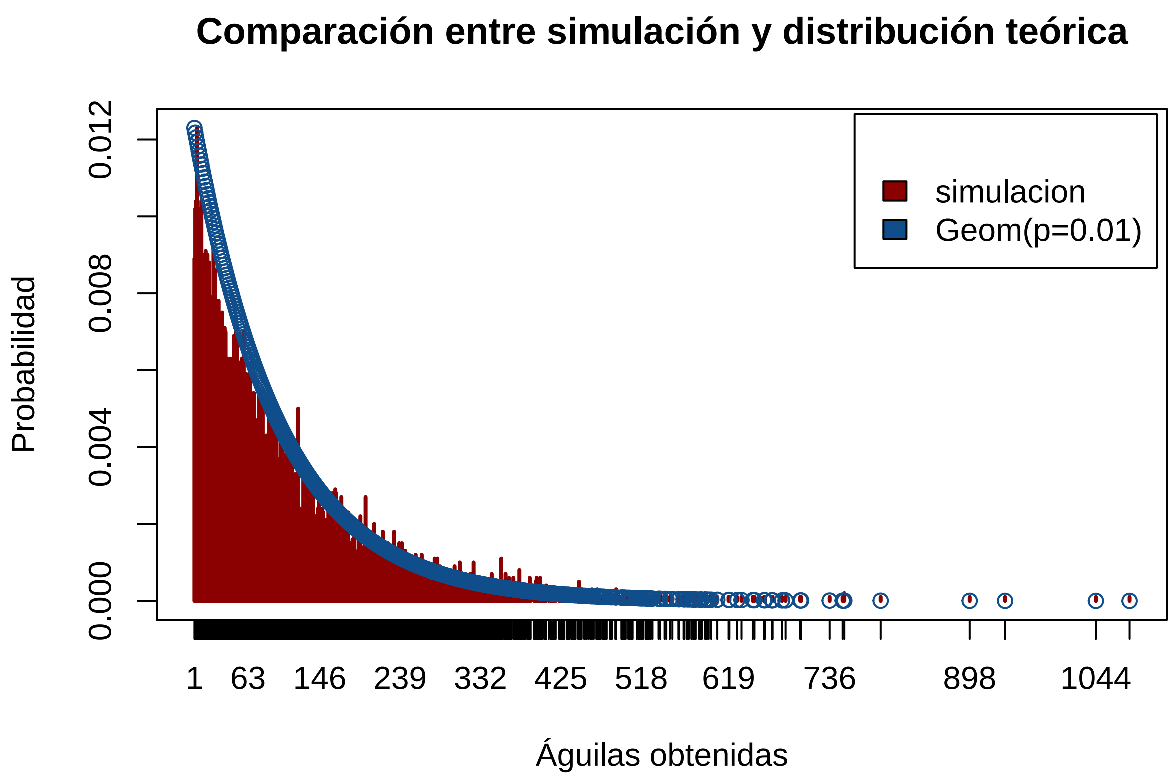


```
#------------------------------------p=0.1

#Grafica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(p_0.1)),
     type="h",
     col="darkred",
     xlab = "",
     ylab="",
     main = "Comparación entre simulación y distribución teórica" )
par(new=TRUE)
#Gráfica de la distribución teórica Geom(p=0.1)
y<-sort(unique(p_0.1))
plot((y), dgeom(y,0.1),
     type="p",
     axes= F,
     col="dodgerblue4",
     xlab = "Águilas obtenidas",
     ylab="Probabilidad" )
legend("topright", inset=.01, title="",
    c("simulacion","Geom(p=0.1)"), fill=c("darkred","dodgerblue4"), horiz=F)
```




```
#------------------------------------p=0.01

#Grafica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(p_0.01)),
     type="h",
     col="darkred",
     xlab = "",
     ylab="",
```

```
     main = "Comparación entre simulación y distribución teórica" )
par(new=TRUE)
#Gráfica de la distribución teórica Geom(p=0.1)
z<-sort(unique(p_0.01))
plot(z, dgeom(z,0.01),
     type="p",
     axes= F,
     col="dodgerblue4",
     xlab = "Águilas obtenidas",
     ylab="Probabilidad" )
legend("topright", inset=.01, title="",
    c("simulacion","Geom(p=0.01)"), fill=c("darkred","dodgerblue4"), horiz=F)
```



C.

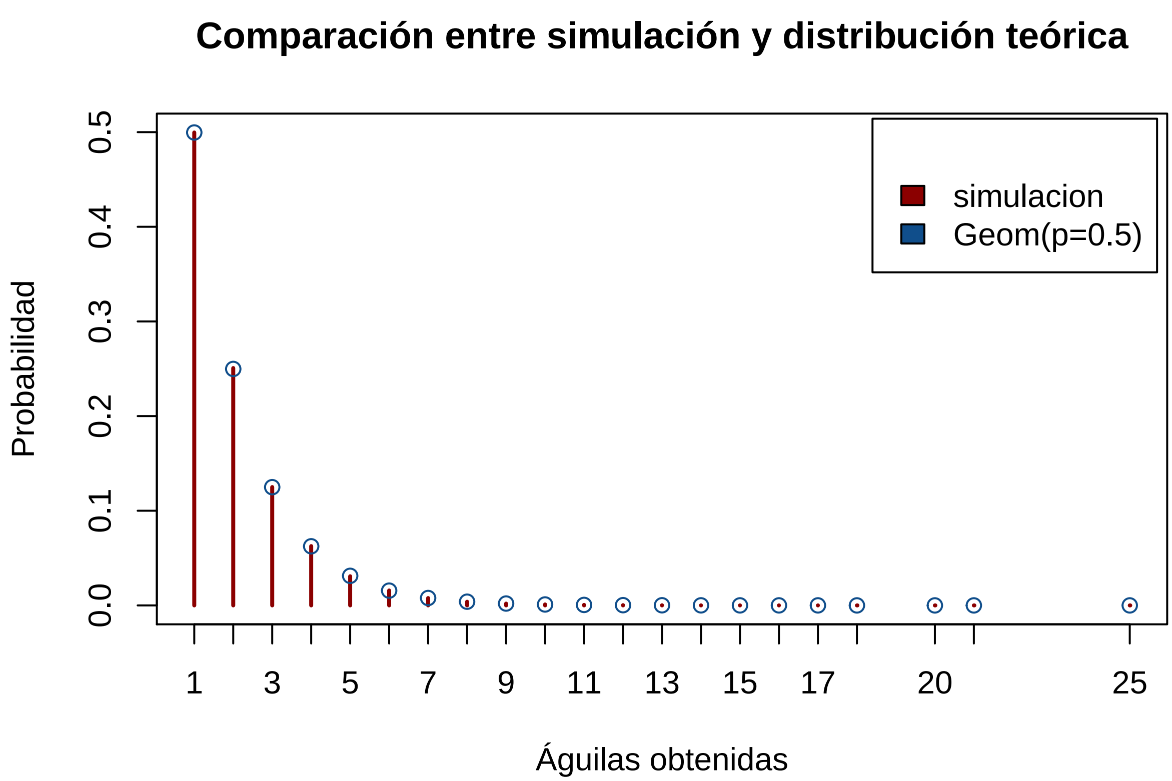
Se repite el ejercicio anterior 10^6 simulaciones

```
set.seed(15)
p2_0.5<-simgeo(10^6,0.5)
p2_0.1<-simgeo(10^6,0.1)
p2_0.01<-simgeo(10^6,0.01)
```

Ahora se realizan los gráficos solicitados con la comparación de masa correspondiente

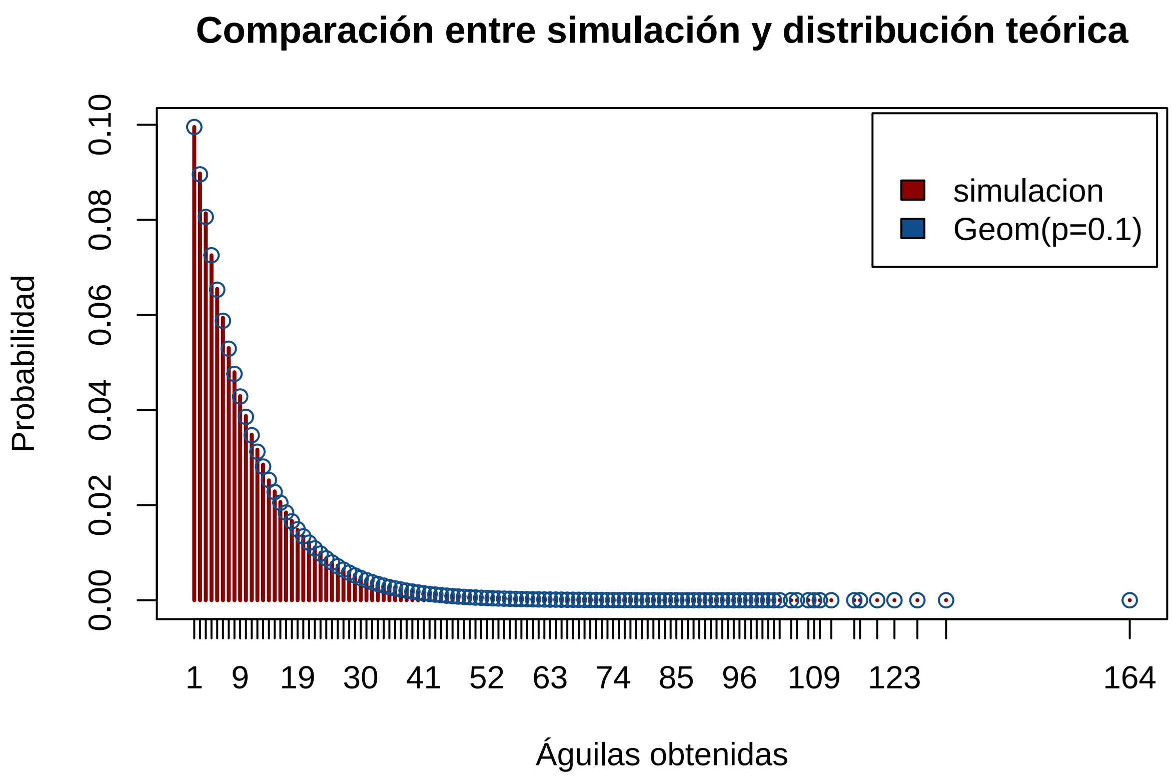
```
# Retícula para los gráficos
##par(mfrow=c(3,1))

#-------------------------------------p=0.5

#Grafica de las proporciones por la cantidad de águilas observadas
```

```r
plot(prop.table(table(p2_0.5)),
     type="h",
     col="darkred",
     xlab = "",
     ylab="",
     main = "Comparación entre simulación y distribución teórica" )
par(new=TRUE)
#Gráfica de la distribución teórica Geom(p=0.5)

x<-sort(unique(p2_0.5))
plot((x+1), dgeom(x,0.5),
     type="p",
     axes= F,
     col="dodgerblue4",
     xlab = "Águilas obtenidas",
     ylab="Probabilidad" )
legend("topright", inset=.01, title="",
    c("simulacion","Geom(p=0.5)"), fill=c("darkred","dodgerblue4"), horiz=F)
```



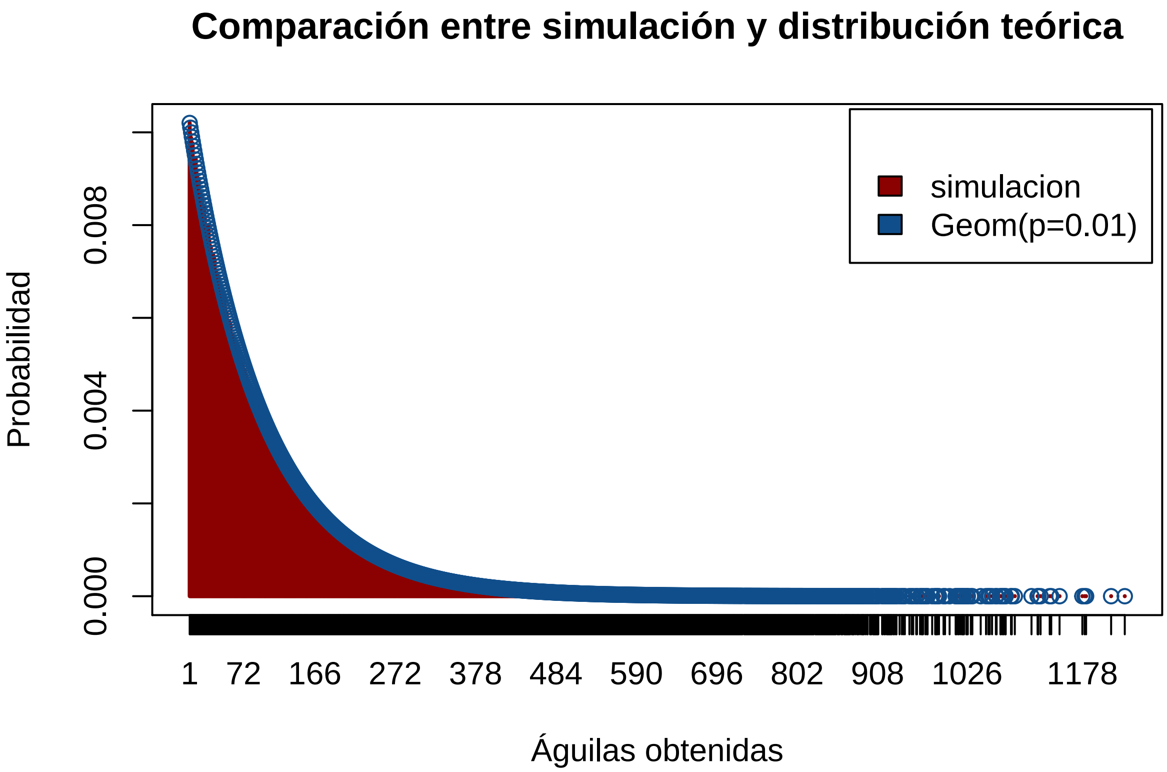

```r
#-----------------------------------p=0.1

#Grafica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(p2_0.1)),
     type="h",
     col="darkred",
     xlab = "",
     ylab="",
     main = "Comparación entre simulación y distribución teórica" )
par(new=TRUE)
```

```
#Gráfica de la distribución teórica Geom(p=0.1)
y<-sort(unique(p2_0.1))
plot((y+1), dgeom(y,0.1),
     type="p",
     axes= F,
     col="dodgerblue4",
     xlab = "Águilas obtenidas",
     ylab="Probabilidad" )
legend("topright", inset=.01, title="",
    c("simulacion","Geom(p=0.1)"), fill=c("darkred","dodgerblue4"), horiz=F)
```




```
#-----------------------------------p=0.01

#Grafica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(p2_0.01)),
     type="h",
     col="darkred",
     xlab = "",
     ylab="",
     main = "Comparación entre simulación y distribución teórica" )
par(new=TRUE)
#Gráfica de la distribución teórica Geom(p=0.1)
z<-sort(unique(p2_0.01))
plot((z+1), dgeom(z,0.01),
     type="p",
     axes= F,
     col="dodgerblue4",
     xlab = "Águilas obtenidas",
     ylab="Probabilidad" )
```

```
legend("topright", inset=.01, title="",
    c("simulacion","Geom(p=0.01)"), fill=c("darkred","dodgerblue4"), horiz=F)
```



Se solicita el cálculo de la media y varianza de las simulaciones:

```
cat(
  "para p=0.5 se obtuvo una media de: ",round(mean(p2_0.5),2),
  "y una varianza de,",round(sqrt(var(p2_0.5)),2),"\n",
  "para p=0.1 se obtuvo una media de: ",round(mean(p2_0.1),2),
  "y una varianza de,",round(sqrt(var(p2_0.1)),2),"\n",
  "para p=0.01 se obtuvo una media de: ",round(mean(p2_0.01),2),
  "y una varianza de,",round(sqrt(var(p2_0.01)),2),"\n")
```

```
## para p=0.5 se obtuvo una media de:  2 y una varianza de, 1.41
##  para p=0.1 se obtuvo una media de:  10.01 y una varianza de, 9.5
##  para p=0.01 se obtuvo una media de:  99.99 y una varianza de, 99.38
```

El valor esperado para distribución teorica en cada eso:

- Para $Geom(0.5)$ su valor esperado es 2 y su desviación estándar 2
- Para $Geom(0.1)$ su valor esperado es 10 y su desviación estándar 9.48
- Para $Geom(0.01)$ su valor esperado es 100 y su desviación estándar 99.49

Puede notarse que al aumentar las simulaciones, los valores esperados de la simulación se aproximan a los valores teóricos.

EJERCICIO 5

A.

Código para la simulación de lanzamientos hasta r éxitos:

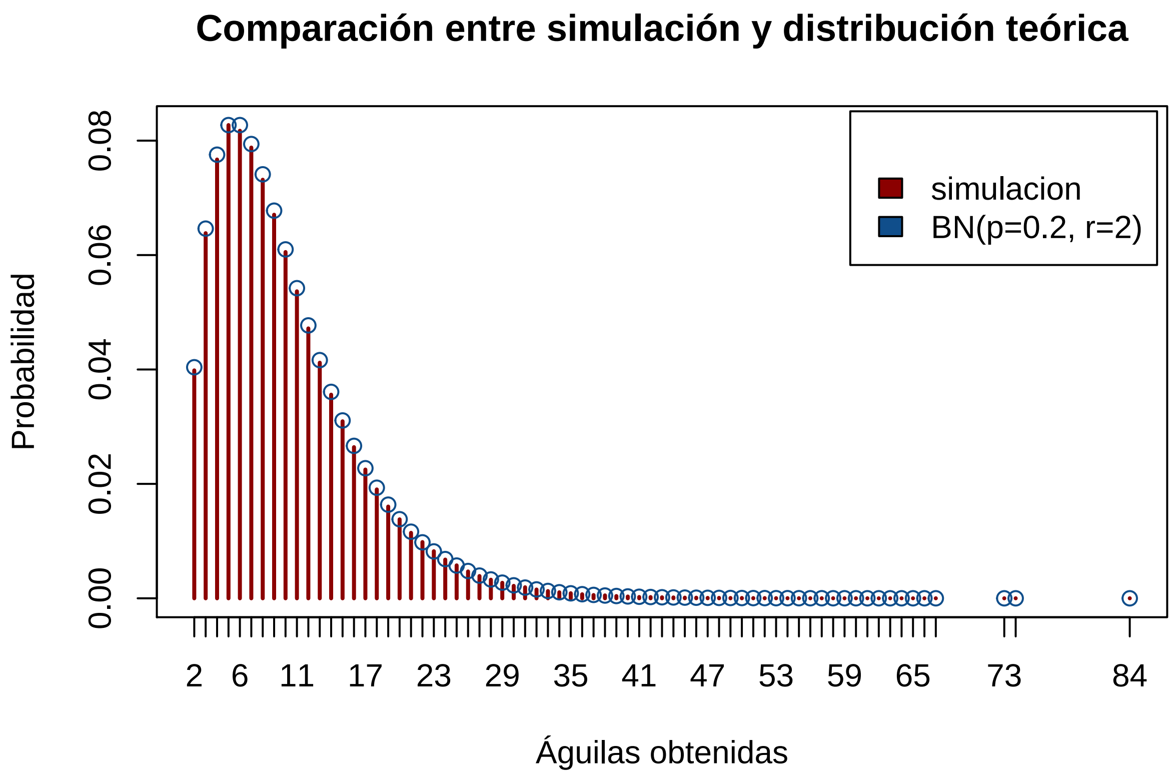
```r
simbneg<-function(N,p,r){

##Moneda 1:águila 0:sol
moneda<-c(1,0)

## Creación del vector donde se guardan resultados
resultados<-rep(0,N)

## Inicio de la simulación

for(i in 1:N){
  #contador del número de tiros
  tiros<-0
  #contador del número de éxitos
  exitos<-0
## Se tira la moneda hasta obtener r éxitos
  while(exitos<r){
    if(sample(moneda,1,replace=T, prob = c(p,1-p))==1)
      { exitos= exitos+1}
    tiros<-tiros +1
  }
## Se guarda el número de tiros
  resultados[i]<-tiros
}
   return(resultados)
}
```

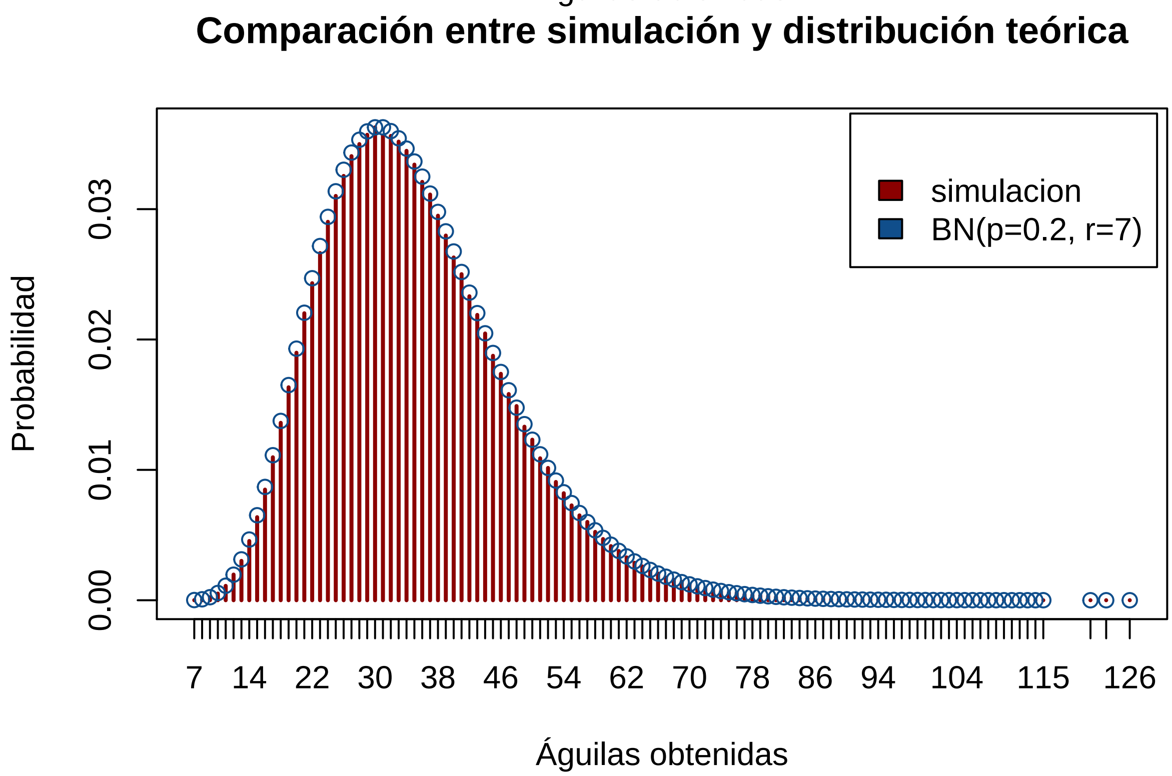
B.

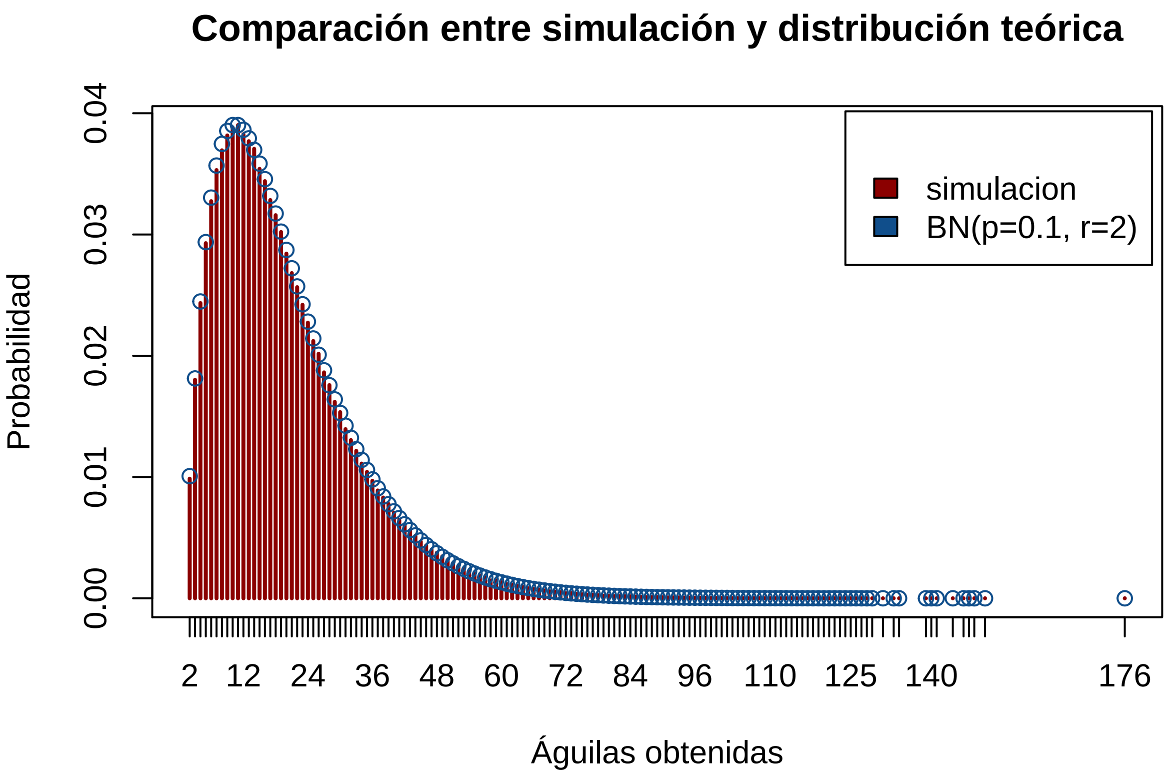
Se realizan las 4 simulaciones solicitadas * \$N=10^6,$p = 0.2$,$r = 2$ * \$N=10^6,$p = 0.2$,$r = 7$ * \$N=10^6,$p = 0.1$,$r = 2$ * \$N=10^6,$p = 0.1$,$r = 7$

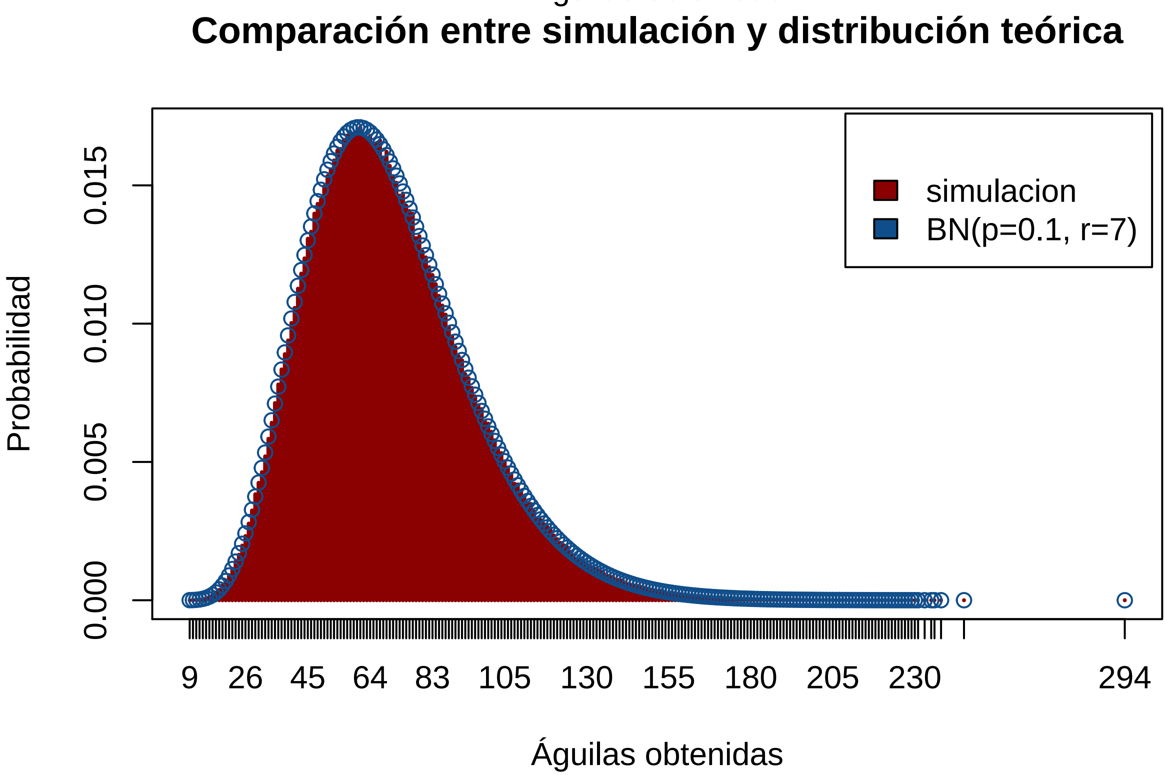
```r
set.seed(15)
bn<-simbneg(10^6,0.2,2)
bn1<-simbneg(10^6,0.2,7)
bn2<-simbneg(10^6,0.1,2)
bn3<-simbneg(10^6,0.1,7)
```

Los gráficos de comparación se presentan a continuación:


Comparación entre simulación y distribución teórica
simulacion
BN(p=0.2, r=2)
Probabilidad
0.00
0.02
0.04
0.06
0.08
2
6
11
17
23
29
35
41
47
53
59
65
73
84
Águilas obtenidas


Comparación entre simulación y distribución teórica
simulacion
BN(p=0.2, r=7)
Probabilidad
0.00
0.01
0.02
0.03
7
14
22
30
38
46
54
62
70
78
86
94
104
115
126
Águilas obtenidas





EJERCICIO 6

Y es una función de x, que es una variable aleatoria, por tanto, tambien es una variable aleatoria discreta, ya que puede tomar como valores 0 y 1. Entonces su densidad debe tener la forma:

$$f(y) = \begin{cases} a & \text{si } y = 0 \\ 1-a & \text{si } y = 1 \\ 0 & \text{de otro modo} \end{cases}$$

Y su funcion acumulada:

$$F(y) = \begin{cases} 0 & \text{si } y < 0 \\ a & \text{si } 0 \leq y < 1 \\ 1 & \text{si } y \geq 1 \end{cases}$$

En la función definimos a a como la probabilidad asignada a cero, por tanto corresponde a $p(x \notin A)$ y su complemento a $p(x \in A)$. Podemos escribir estas probabilidades en términos de la función de distribución de $x \rightarrow F_x$ Considerando que x toma valores en la recta real, y definiendo a x_a y x_b como los límites del intervalo A, se llega a la conclusión de que:

$$P(x \in A) = F_x(b) - F_x(a)$$

y por tanto:

$$P(x \notin A) = 1 - (F_x(b) - F_x(a))$$

Al sustituir en la F_y, se obtiene

$$F(y) = \begin{cases} 0 & \text{si } y < 0 \\ 1 - (F_x(b) - F_x(a)) & \text{si } 0 \leq y < 1 \\ 1 & \text{si } y \geq 1 \end{cases}$$

EJERCICIO 7

A.

La variable Y, definida como el número de meteoros vistos en un cuarto de hora, se distribuye Poisson con parámetro $\lambda = 6$. El valor esperado para una variable Poisson es su mismo parámetro, así que se desea calcular:

$$P(X \geq \lambda/2) = P(X \geq 3) = 1 - (P(X = 0) + (X = 1) + (X = 2))$$

Se realiza el cálculo usando la función ppois en R

```
ppois(2,6, lower.tail = FALSE)
```

```
## [1] 0.9380312
```

Se obtuvo que la probabilidad de que se vea al menos de la mitad de lo esperado es 0.94

EJERCICIO 8

A.

El alumno reprueba con una $y < 0.4$, usamos la densidad de probabilidad para el cálculo de la probabilidad:

$$P(Y \leq 0.4) = \int_0^{0.4} 6y(1-y)dy = 0.352$$

La probabilidad de que al alumno repruebe es de 0.352

B.

La probabilidad que dos de los seis alumnos que tomaron el examen reprueben, se calcula con el modelo binomial con $p = 0.352$, $n = 6$ y $k = 2$. Se obtiene

$$P(X = 2) = \binom{5}{2}(0.352)^2(1 - 0.352)^{5-3} = 0.33714$$

Hay una probabilidad de 0.337 de que reprueben dos de los seis estudiantes que tomen el examen

EJERCICIO 9

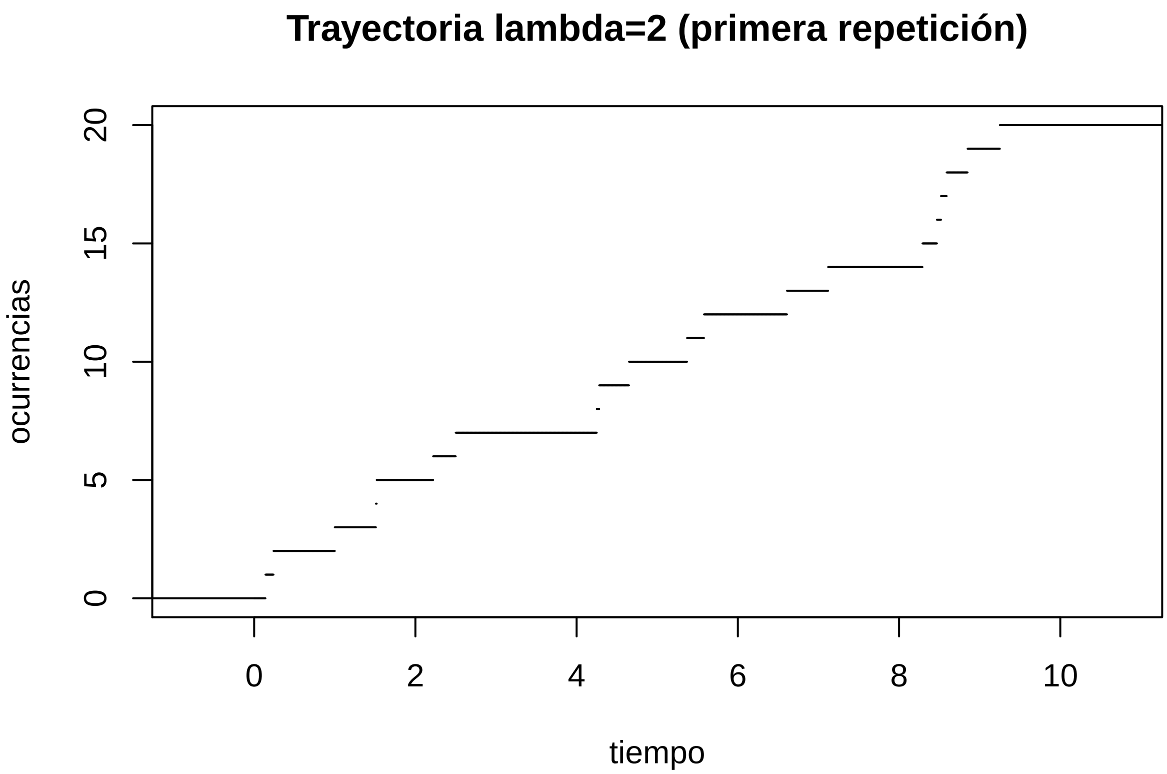
A.

Función para simular un proceso poisson en intervalo [0, T] y una lambda conocida

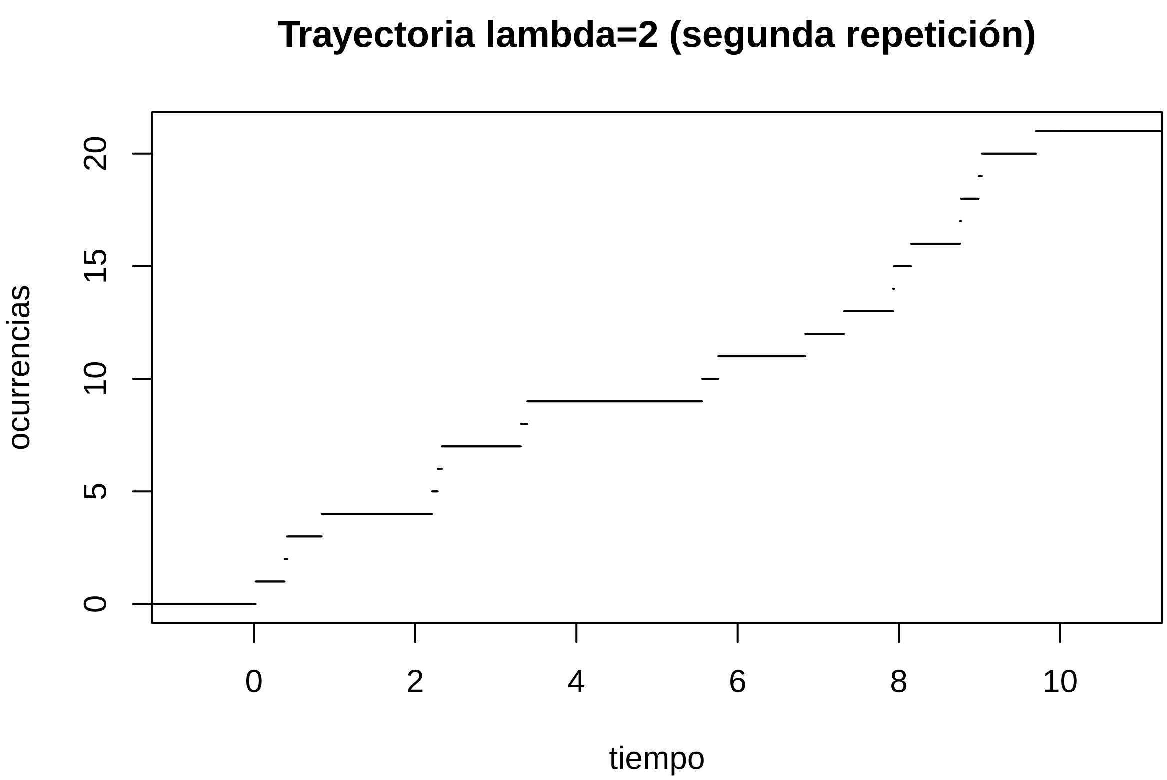
```
procesoPoisson<-function(lambda, t){

intervalos<-1000
dt<-t/intervalos
tiempos<-seq(0, t , by= dt)
resultados <- rep(0, intervalos+2)
for(i in 1:intervalos+1 ){

  resultados[i+1]<- resultados[i] + rbinom(1,1,lambda*dt)
}
 return(resultados)
}
```

Se realizan las 3 simulaciones y posteriormente se muestra la gráfica correspondiente

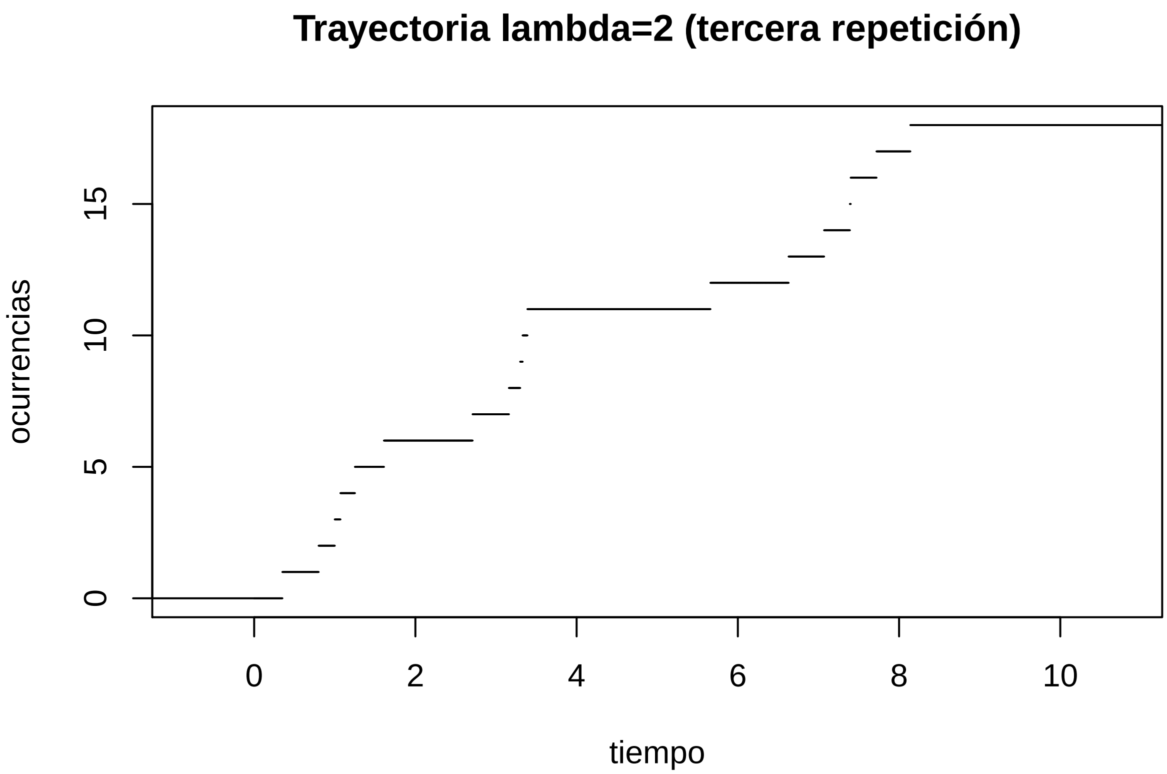
```
tiempos<-seq(0, 10 , by= 10/1000)

ppss1<-procesoPoisson(2,10)
ppss2<-procesoPoisson(2,10)
ppss3<-procesoPoisson(2,10)


plot(stepfun(tiempos,ppss1), verticals= FALSE, xlab="tiempo",ylab="ocurrencias", main = "Trayectoria lan
```




```
plot(stepfun(tiempos,ppss2), verticals= FALSE, xlab="tiempo",ylab="ocurrencias", main = "Trayectoria lam
```




```
plot(stepfun(tiempos,ppss3), verticals= FALSE, xlab="tiempo",ylab="ocurrencias", main = "Trayectoria lam
```



B.

Modificamos la función anterior para ajustarla a los objetivos del ejercicio:

```
procesoPoisson2<-function(N,lambda,t){

intervalos<-1000
dt<-t/intervalos
tiempos<-seq(0, t , by= dt)
resultados <- rep(0, intervalos+2)
for(i in 1:N){
acum<-0
for(i in 1:intervalos +1){

  acum= acum + rbinom(1,1,(lambda*dt +10^-6))
}
 resultados[i+1]=acum

}
return(resultados)
}
```

Ahora realizamos las simulación (10^4 simulaciones, tiempo (0,1), $\lambda = 0.5$)

```
simpropois <- procesoPoisson2(10^4,0.5,1)
```

Realizamos la comparativa con la distribucion Poisson($\lambda = 0.5$):

```
#Grafica de las proporciones por la cantidad de águilas observadas
plot(prop.table(table(simpropois)),
     type="h",
```

```
      col="darkred",
      xlab = "",
      ylab="",
      main = "Comparación entre simulación y distribución teórica" )
par(new=TRUE)
#Gráfica de la distribución teórica Geom(p=0.1)
z<-sort(unique(simpropois))
plot((z), dpois(z,0.5),
      type="p",
      axes= F,
      col="dodgerblue4",
      xlab = "Éxitos",
      ylab="Probabilidad" )
legend("topright", inset=.01, title="",
    c("simulacion","Poisson(lambda=1/2)"), fill=c("darkred","dodgerblue4"), horiz=F)
```