UNIVERSIDAD NACIONAL DE LOJA

PERIODO ACADÉMICO: OCTUBRE 2019 – MARZO 2020
PRACTICA # 7
ASIGNATURA: SIMULACIÓN
RESULTADO DE APRENDIZAJE DE LA PRÁCTICA: Entiende los métodos de las diferentes distribuciones de probabilidad usando R y aplica estos métodos para la resolución de problemas a través de la simulación
TIEMPO PLANIFICADO: 3 HORAS
NUMERO DE ESTUDIANTES: Sexto ciclo (Paralelo A)

NOMBRE: Deiby Patricio Calva. **FECHA:** 09/02/2020

LINK ARCHIVO: https://github.com/DeibyCalva/Laboratorio_en-_R_simulacion/blob/master/LABORATORIOS%20EN%20R/Practica_7DeibyCalva/Practica_7DeibyCalva.R

1. **TEMA: Aplicaciones de la simulación usando las distribuciones de probabilidad y variables aleatorias**
2. **OBJETIVOS:**

- Comprende los métodos de distribuciones de probabilidad en R.
- Aplica la simulación para la resolución de problemas prácticos.

3. **RECURSOS NECESARIOS:**

- R
- Computador de Laboratorios

4. **INSTRUCCIONES:**

- Prohibido consumo de alimentos
- Prohibido equipo de diversión, celulares etc.
- Prohibido jugar
- Prohibido mover o intercambiar los equipos de los bancos de trabajo
- Prohibido sacar los equipos del laboratorio sin autorización.
- Ubicar los equipos y accesorios en el lugar dispuesto por el responsable del laboratorio, luego de terminar las prácticas.
- Uso adecuado de equipos

5. **ACTIVIDADES POR DESARROLLAR:**

1. Supongamos que la proporción defectuosa es de 0,15 para una operación de fabricación. Simule el número de defectuosos por cada hora de un período de 24 horas, suponiendo que se producen 25 unidades por hora. Compruebe si el número de defectuosos supera alguna vez los 5. Repita, asumiendo que p = 0,2 y luego p = 0,25.

```r
## PREGUNTA 1 ##

defectuosos = rbinom(24, 25, 0.15)
defectuosos
any(defectuosos>5)
sum(defectuosos>5)

# P=0.2
defectuosos = rbinom(24, 25, 0.2)
defectuosos
any(defectuosos>5)
sum(defectuosos>5)

#P=0.25
defectuosos = rbinom(24, 25, 0.25)
defectuosos
any(defectuosos>5)
sum(defectuosos>5)

```

```
> ## PREGUNTA 1 ##
>
> defectuosos = rbinom(24, 25, 0.15)
> defectuosos
 [1] 5 3 4 1 2 3 4 2 5 3 1 4 4 2 4 3 7 2 5 1 7 7 4 3
> any(defectuosos>5)
[1] TRUE
> sum(defectuosos>5)
[1] 3
>
> # P=0.2
> defectuosos = rbinom(24, 25, 0.2)
> defectuosos
 [1] 7 3 7 8 4 6 5 7 7 2 4 5 2 5 8 6 8 5 2 4 9 3 8 6
> any(defectuosos>5)
[1] TRUE
> sum(defectuosos>5)
[1] 12
>
> #P=0.25
> defectuosos = rbinom(24, 25, 0.25)
> defectuosos
 [1]  5  9  5  5 10  5  4  8  2  9 11  6  5  5  9  9 11  6  5  4  4  6  6  6
> any(defectuosos>5)
[1] TRUE
> sum(defectuosos>5)
[1] 13
```

2. Simular 10 000 números pseudoaleatorios binomiales con los parámetros 20 y 0,3, asignándolos a un vector llamado binsim. Dejemos que X sea una variable aleatoria binomial (20, 0,3). Utiliza los números simulados para estimar lo siguiente:

a) $P(X \leq 5)$.

```
binsim=rbinom(10000,20,0.3)
length(binsim[binsim<=5])/length(binsim)
pbinom(5,20,0.3)
```

```
> length(binsim[binsim<=5])/length(binsim)
[1] 0.4232
> pbinom(5,20,0.3)
[1] 0.4163708
```

(b) $P(X = 5)$.

```
sum(binsim==5)/length(binsim)
dbinom(5,20,0.3)
```

```
> sum(binsim==5)/length(binsim)
[1] 0.1863
> dbinom(5,20,0.3)
[1] 0.1788631
```

(c) E[X].

```
mean(binsim)
20*0.3
```

```
> mean(binsim)
[1] 6.0064
> 20*0.3
[1] 6
```

(d) Var(X).

```
var(binsim)
20*0.3*0.7
```

```
> var(binsim)
[1] 4.133972
> 20*0.3*0.7
[1] 4.2
```

(e) el percentil 95 de X (puede utilizar la función quantile()).

```
#con funcion quantile y con el qbinom
quantile(binsim, prob=c(95,99,99.9999)/100)
qbinom(0.95,20,0.3)
```

```
> #con funcion quantile y con el qbinom
> quantile(binsim, prob=c(95,99,99.9999)/100)
     95%      99% 99.9999%
    9.00    11.00    14.99
> qbinom(0.95,20,0.3)
[1] 9
```

(f) el percentil 99 de X.

```
qbinom(0.99,20,0.3)
```

```
> qbinom(0.99,20,0.3)
[1] 11
>
```

(g) el cuantil 99,9999 de X. En cada caso, compare sus estimaciones con los valores reales. ¿Qué se necesita para estimar con precisión las cantidades extremas?

```
qbinom(0.999999,20,0.3)
```

```
> qbinom(0.999999,20,0.3)
[1] 16
```

3. Usar la simulación para estimar la media y la varianza de una variable aleatoria binomial con n = 18 y p = 0,76. 4. Comparar con los valores teóricos.

```
r=rbinom(n=100,size=18,p=0.76)
mean(r)
var(r)
```

```
> r=rbinom(n=100,size=18,p=0.76)
> mean(r)
[1] 13.83
> var(r)
[1] 3.395051
```

4. Considere la siguiente función que está diseñada para simular las variaciones pseudoaleatorias del binomio utilizando el llamado método de inversión:

```
ranbin <- function(n, size, prob) {
    cumpois <- pbinom(0:(size - 1), size, prob)
    singlenumber <- function() {
        x <- runif(1)
        N <- sum(x > cumpois)
        N
    }
    replicate(n, singlenumber())
}
```

a. Estudie esta función con cuidado y escriba la documentación correspondiente. Note, en particular, para qué sirven las operaciones de la función de *singlenumbe*r()-.

La función de replicate() nos permite llamar repetidamente singlenumber (), asignando n resultados a un vector.

b. Usar ranbin() para simular vectores de longitud 1000, 10000 y 100000 de la distribución binomial con el parámetro de tamaño 10 y el parámetro de probabilidad 0.4. Utilice la función system.time () para comparar los tiempos de ejecución de estas simulaciones con los tiempos de ejecución correspondientes cuando se usa rbinom().

```
ranbin <- function(n, size, prob){
  cumpois <- pbinom(0:(size - 1), size, prob)
  singlenumber <- function(){
    x <- runif(1)
    N <- sum(x > cumpois)
    N
  }
  replicate(n, singlenumber())
}
system.time(gcFirst = TRUE, ranbin(n = 1000, size = 10, prob = 0.4))
system.time(gcFirst = TRUE, ranbin(n = 10000, size = 10, prob = 0.4))
system.time(gcFirst = TRUE, ranbin(n = 100000, size = 10, prob = 0.4))
```

```
> ranbin <- function(n, size, prob){
+   cumpois <- pbinom(0:(size - 1), size, prob)
+   singlenumber <- function(){
+     x <- runif(1)
+     N <- sum(x > cumpois)
+     N
+   }
+   replicate(n, singlenumber())
+ }
> system.time(gcFirst = TRUE, ranbin(n = 1000, size = 10, prob = 0.4))
   user  system elapsed 
      0       0       0 
> system.time(gcFirst = TRUE, ranbin(n = 10000, size = 10, prob = 0.4))
   user  system elapsed 
   0.05    0.00    0.05 
> system.time(gcFirst = TRUE, ranbin(n = 100000, size = 10, prob = 0.4))
   user  system elapsed 
   0.36    0.01    0.37 
```

5. La siguiente función simula números pseudoaleatorios binomiales sumando las correspondientes variables aleatorias de Bernoulli independientes

```
ranbin2 <- function(n, size, prob) {
    singlenumber <- function(size, prob) {
        x <- runif(size)
        N <- sum(x < prob)
        N
    }
    replicate(n, singlenumber(size, prob))
}
```

a. Estudie esta función con cuidado y escriba la documentación correspondiente. Tenga en cuenta, en particular, cuáles son las funciones de la función *singlenumber().*

La función de replicate() nos permite llamar repetidamente singlenumber (), asignando n resultados a un vector.

b. Use ranbin2() para simular vectores de longitud 10000 de la distribución binomial con los parámetros de tamaño 10, 100 y 1000, y el parámetro de probabilidad 0.4. Use la funcion system.time()para comparar los tiempos de ejecución de estas simulaciones con los tiempos de ejecución correspondientes cuando se usa rbinom(). Compare con los tiempos de ejecución de la función ranbin() creada en el ejercicio anterior.

```r
ranbin2 <- function(n, size, prob){
  singlenumber <- function(size, prob) {
    x <- runif(size)
    N <- sum(x < prob)
    N
  }
  replicate(n, singlenumber(size, prob))
}

system.time(gcFirst = TRUE, ranbin2(n = 10000, size = 10, prob = 0.4))
system.time(gcFirst = TRUE, ranbin2(n = 10000, size = 100, prob = 0.4))
system.time(gcFirst = TRUE, ranbin2(n = 10000, size = 1000, prob = 0.4))
```

```
> ### PREGUNTA 5###
> ranbin2 <- function(n, size, prob){
+   singlenumber <- function(size, prob) {
+     x <- runif(size)
+     N <- sum(x < prob)
+     N
+   }
+   replicate(n, singlenumber(size, prob))
+ }
>
> system.time(gcFirst = TRUE, ranbin2(n = 10000, size = 10, prob = 0.4))
   user  system elapsed
   0.06    0.00    0.07
> system.time(gcFirst = TRUE, ranbin2(n = 10000, size = 100, prob = 0.4))
   user  system elapsed
   0.09    0.00    0.10
> system.time(gcFirst = TRUE, ranbin2(n = 10000, size = 1000, prob = 0.4))
   user  system elapsed
   0.47    0.00    0.47
```

6. INVESTIGACIÓN COMPLEMENTARIA

DISTRIBUCIÓN BINOMIAL

Una forma corriente de descripción de los experimentos aleatorios equiprobables con variable discreta es la distribución binomial. En este tipo de distribución se estudia la probabilidad de que se produzca un cierto resultado, que se describe por medio de dos parámetros: el número de repeticiones realizadas del experimento y la probabilidad individual del suceso aleatorio que se persigue como resultado.

Condiciones para una distribución binomial.

Una distribución se denomina binomial cuando se cumplen las condiciones siguientes:

- El experimento aleatorio de base se repite n veces, y todos los resultados obtenidos son mutuamente independientes.
- En cada prueba se tiene una misma probabilidad de éxito (suceso A), expresada por p. Asimismo, existe en cada prueba una misma probabilidad de fracaso (suceso), que es igual a q = 1 - p.
- El objetivo de la distribución binomial es conocer la probabilidad de que se produzca un cierto número de éxitos. La variable aleatoria X, que indica el número de veces que aparece el suceso A (éxito), es discreta, y su recorrido es el conjunto {0, 1, 2, 3, ..., n}

Formula de la distribución binomial

La distribución binomial se caracteriza porque su función de probabilidad viene dada por la expresión siguiente:

$$P_{(x)} = \binom{n}{x} p^x q^{n-x}$$

Donde:

n = número de ensayos/experimentos
x = número de éxitos
p = probabilidad de éxito
q = probabilidad de fracaso (1-p)

Es importante resaltar que la expresión entre corchetes no es una expresión matricial, sino que es un resultado de una combinatoria sin repetición. Este se obtiene con la siguiente formula:

$$C_{n,x} = \binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

Esperanza, varianza y desviación típica

En una distribución binomial denotada por B (n, p), donde n es el número de repeticiones del experimento y p la probabilidad de que se produzca un cierto suceso (éxito), la esperanza matemática de la variable aleatoria X viene dada por la expresión siguiente:

$$\mathrm{E}\,[\mathrm{X}] = \mathrm{n} \cdot \mathrm{p}$$

Análogamente, la varianza de la variable aleatoria X, al ser ésta de tipo discreto, se calcula como:

$$\mathrm{V}\,[\mathrm{X}] = \mathrm{n} \cdot \mathrm{p} \cdot \mathrm{q}$$

siendo q la probabilidad de no éxito (fracaso). La desviación típica es, como de costumbre, la raíz cuadrada de la varianza:

$$\sigma = \sqrt{\mathrm{n} \cdot \mathrm{p} \cdot \mathrm{q}}$$

7. DISCUSIÓN

Para el diseño de la secuencia de actividades se identificaron los diferentes parámetros que hacen parte de la definición formal del concepto de distribución binomial. A partir de esto, se diseñaron 2 actividades, donde los estudiantes definieron, calcularon y relacionaron los siguientes conceptos: Espacio muestral, combinatoria, eventos independientes, probabilidad, caracterización de la variable aleatoria. Potenciando así la noción del concepto de distribución binomial, todo esto bajo las fases de acción, formulación y validación.

8. CONCLUSIONES

- La función `dbinom` devuelve el valor de la función de densidad y la función `pbinom` devuelve el valor de la función de densidad acumulada de la distribución binomial.

- La función `qbinom` devuelve el valor de la función de densidad acumulativa inversa de la distribución binomial y la función `rbinom` genera un vector de variables aleatorias distribuidas binomiales dada una longitud de vector *n,* número de ensayos (tamaño) y probabilidad de éxito en cada ensayo.

9. RECOMENDACIONES

- Seguir utilizando el software R para tener un mejor entendimiento ya que sus funciones son muy útiles para este tipo de cálculos.

- Plantear actividades que permitan el desarrollo y construcción de las distribuciones discretas.

BIBLIOGRAFÍA:

- Hamza, K. (1995). The smallest uniform upper bound on the distance between the mean and the median of the binomial and Poisson distributions. Statist. Probab. Lett. 23 21–25.
- Mode, Elmer B. (1990). *Elementos de probabilidad y estadística*. Reverte. p. 171. ISBN 9788429150926. Consultado el 5 de diciembre de 2017.
- R. Matthews *Maximally Periodic Reciprocals* Bulletin of the Institute of Mathematics and its Applications 28 147-148 1992
-