

Universidad Nacional de Rosario

FACULTAD DE CIENCIAS EXACTAS, INGENIERÍA Y AGRIMENSURA

PROBABILIDAD Y ESTADÍSTICA

Tp final

Autor:

Demagistris, Santiago Ignacio

Agosto 2020

0.1 Ejercicio 1

a) El espacio muestral sobre el cual estamos trabajando es $S=\{0,1\}$, donde cara es 1 y cruz es 0. I_n y S_n

$n = 100$ Cantidad de tiradas

```
simulacion = simular(n)
```

```
simular <- function(n){
```

```
  In <- tirarMoneda(n) (n tiradas de una moneda)
```

```
  Xn <- cumsum(In) (numero de caras al momento n - Frecuencia relativa acumulada -)
```

```
  vectorTiradas <- c(1:n)
```

```
  return (data.frame(Xn,vectorTiradas)) }
```

```
tirarMoneda<-function(n){ (Tirar moneda n veces)
```

```
  return(sample(0:1,100,replace = TRUE)) (Vector de longitud n donde cada posicion tiene 1 o 0) }
```

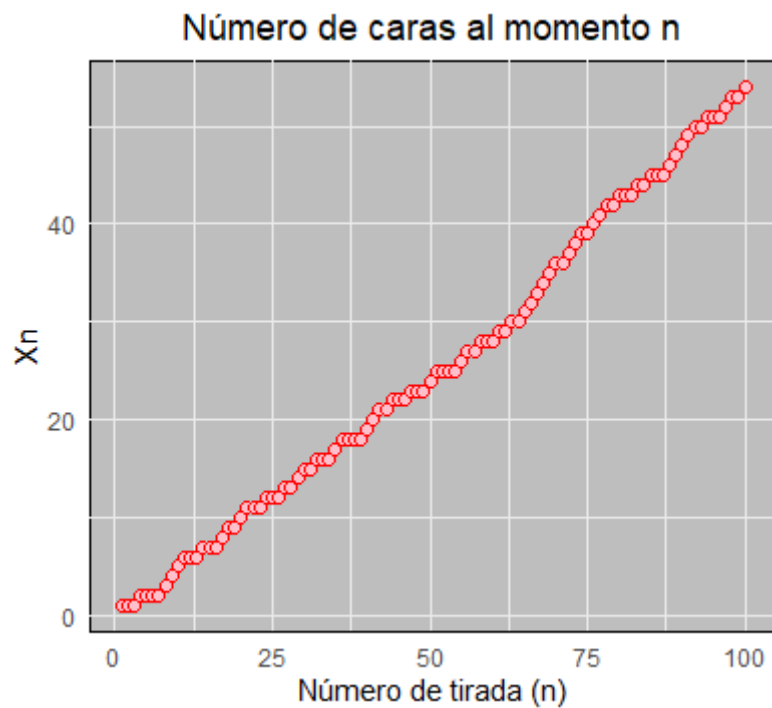


Figure 1: Tiradas de moneda

Sea $A=\{\text{"Obtener cara al lanzar la moneda"}\}$. Podemos observar que se obtuvo un total de 54 caras en este proceso, por lo tanto $P(A) \simeq \frac{54}{100} = 0.54 = p$

b)

($\mathbf{P(X=1)}$). Sabemos que $X = \{\text{"Número de veces que ocurrió el evento A"}\}$ y que $X \sim Bi(n, 0.54)$ (siendo $p = 0.54$ observado en la simulación). Por lo tanto $P(X = 1) = \binom{n}{1}(0.54)(0.46)^{n-1}$ y $E(X) = n0.54$.

c) Sea $Y = \text{Número de tiradas hasta que salgan 3 caras}$. $Y \sim Pascal$, por lo que

$$P(Y = k) = \binom{k-1}{3-1} 0.54^3 0.46^{k-3} = \binom{k-1}{2} 0.54^3 0.46^{k-3}$$

y $E(Y) = \frac{r}{p} = \frac{3}{0.54} \simeq 5.55$. Es decir que se espera que entre la 5ta y la 6ta tirada de la moneda obtengamos 3 caras.

d) La implementación que utilice para sesgar la moneda es la siguiente:

```
tirarMonedaSesgada <- function(n){
  return (sample(0:2,100, replace =TRUE)%%2) }
```

De esta manera trabajo sobre un espacio muestral $S_x\{0, 1, 2\}$, donde $A = \{1\} = \{\text{"Obtuve cara"}\}$ y $A^c = \{\text{"Obtuve Cruz"}\}$

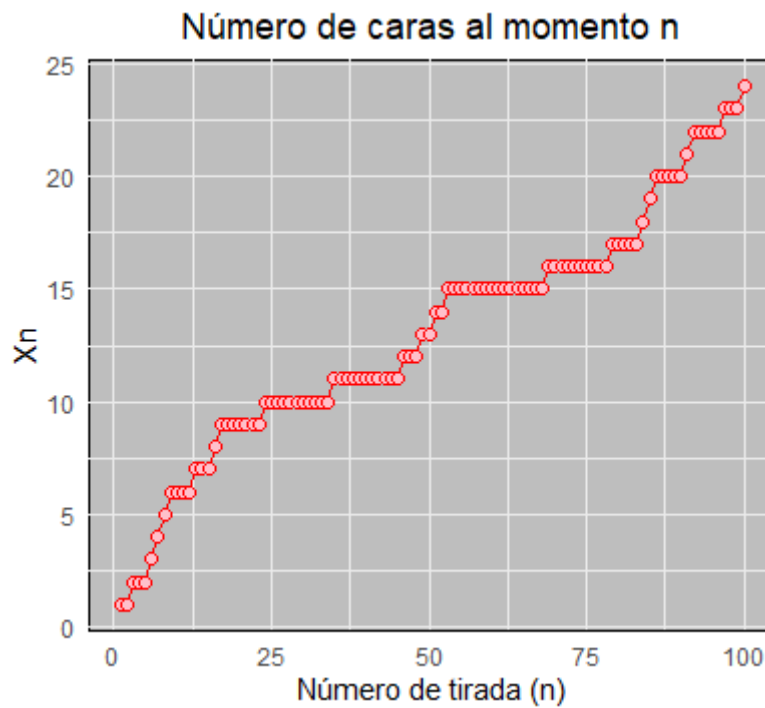


Figure 2: Tiradas de moneda sesgada

Al simular el proceso con $n=100$ obtuve que salieron en total 24 caras, por lo tanto podría aproximar $P(A) \simeq 0.24$. Por lo tanto $P(X = 1) = \binom{n}{1}(0.24)(0.76)^{n-1}$ y $E(X) = n0.24$

0.2 Ejercicio 2

Para simular la situacion implemente lo siguiente:

```
simular <- function(n){  
  
  In <- tirarMoneda(n) - Tiro n veces la moneda  
  
  Tiradas <- c(1:n) - Genero el eje X  
  
  Dn = variableAleatoria(In) - Obtengo la variable aleatoria Dn apartir de In  
  
  Xn = cumsum(Dn) - Obtengo el desplazamiento como la frecuencia relativa acumulada de Dn  
  
  proceso1 <- data.frame(In,Tiradas)  
  
  proceso2 <- data.frame(Dn,Tiradas)  
  
  midf <- data.frame( "Proceso1" = In, "Proceso2" = Xn )  
  
  return(midf)  
  
}  
  
tirarMoneda <- function(n){ - Tirar moneda n veces  
  
  return(sample(0:1,n,replace = TRUE))  
  
}  
  
variableAleatoria <- function(In){ - Obtengo Dn  
  
  Dn <- c()  
  
  for (i in In){  
  
    Dn <- c(Dn,(2 * i) - 1)  
  
  }  
  
  return (Dn)  
  
}
```

Moneda vs Posicion vs Desplazamiento

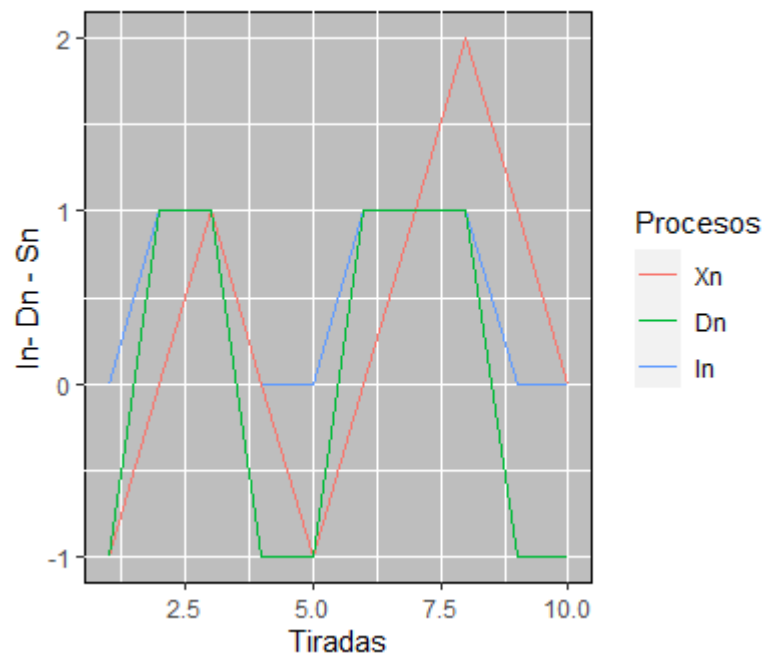


Tabla de frecuencias

In	Dn	Sn
0	-1	-1
1	1	0
1	1	1
0	-1	0
0	-1	-1
1	1	0
1	1	1
1	1	2
0	-1	1
0	-1	0

La trayectoria descrita por X_n es causa del lanzamiento de la moneda (10 veces en este caso). Cada vez que salga cara en la moneda el sujeto se desplazará 1 unidad en el espacio y cuando salga cruz un 0. (Cara representado con 1 y cruz representado con 0). La decisión de desplazamiento esta dada por D_n , mientras que X_n solo describe la trayectoria, es decir la frecuencia acumulada de los resultados obtenidos.

0.3 Ejercicio 3

a) En R implemente lo siguiente para simular el proceso:

```
proceso <- function(plataInicial,S,p){

  limiteInferior = 0

  limiteSuperior = S

  plataActual = plataInicial

  ganar = p

  perder = 1-ganar

  vectorProbabilidad = c(ganar,perder)

  trayectoria = c(plataInicial)

  termino = FALSE

  while (!termino) {

    if(tirada(vectorProbabilidad) == ganar){

      plataActual = plataActual + 1 }

    else{

      plataActual = plataActual - 1 }

    trayectoria = c(trayectoria,plataActual)

    if(plataActual == limiteSuperior){termino = TRUE}

    if(plataActual == limiteInferior){termino = TRUE}

  }

  return (trayectoria[length(trayectoria)] == S)

}
```

```

tirada<-function(probabilidades){

probabilidades = sort(probabilidades)

intervaloAleatorio = 1000000 - Probabilidades de hasta 6 digitos despues del .

tirada = sample(1:intervaloAleatorio,1)

posicion = 1

for (i in probabilidades){

if (posicion == length(probabilidades)){

return (i) } if(tirada <= i*1000000){

return (i) Devuelvo la probabilidad que ocurrio }

posicion = posicion +1

} }

```

- b) Para realizar la simulacion usare $k=50$, $p=0.55$, $S=100$, $\text{perder}=0$. El resultado obtenido fue el siguiente:

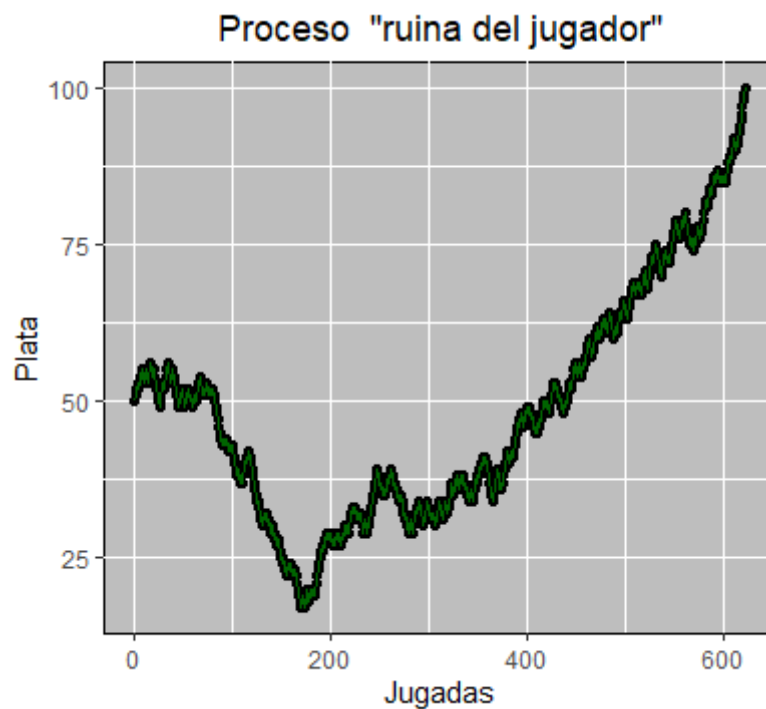


Figure 3: Trayectoria. Ruina del jugador

c) Para simular esto, implemente en R: `probabilidadRuina <- function(plataInicial,S,p,n){`

```
frecRuina = 0
```

```
for (i in c(1:n)){
```

```
  if(!proceso(plataInicial,S,p)){
```

```
    frecRuina = frecRuina+1
```

```
  } } return (frecRuina/n)
```

```
}
```

Con $k=20$, $S=60$, $p=0.5001$ y 1000 trayectorias obtuve una aproximacion a la probabilidad de caer en la ruina de 0.678

0.4 Ejercicio 4

a) Para simular la situacion implemente en R lo siguiente:

```
decision <- function(x){  
  if (x==izquierda){  
    return (caminoIzquierdo(salidaRetorno()))  
  } else{  
    return (3 + decision(elegirCamino())) }  
}  
  
caminoIzquierdo <- function(x){  
  if (x==izquierda){  
    return (2)  
  } else{  
    return (5 + decision(elegirCamino()))  
  } }  
  
elegirCamino <- function(){ return (sample(0:1,1)) }  
  
salidaRetorno <- function(){  
  return (sample(0:2,1)) }  
  
esperanza<-function(){  
  acumulador=0  
  for(i in (1:n)){  
    acumulador = acumulador + decision(elegirCamino())  
  } return (acumulador / n)  
}
```

Al realizar la simulacion observe 50000 repeticiones del experimento y obtuve un promedio de 21.08484 minutos.

- b) Sea $X = \{\text{"Tiempo en que tarda el raton en salir del laberinto"}\}$, donde X va a depender de las decisiones del raton, estas pueden ser "ii", "id", "d1". Las elecciones se desarrollan en dos etapas, en la primer etapa nos encontramos con que

$$P(\{ \text{"d1"} \}) = 1/2$$

En el caso de elegir el camino hacia la izquierda (con una probabilidad de $1/2$) nos encontramos en la segunda etapa en la cual:

$$P(\{ \text{"i2"} \} | \{ \text{"i1"} \}) = 1/3$$

y

$$P(\{ \text{"d2"} \} | \{ \text{"i1"} \}) = 2/3$$

De donde obtenemos que

$$P(\{ \text{"ii"} \}) = P(\{ \text{"i2"} \} \cap \{ \text{"i1"} \}) = P(\{ \text{"i2"} \} | \{ \text{"i1"} \}) P(\{ \text{"i1"} \}) = 1/3 \cdot 1/2 = 1/6$$

y

$$P(\{ \text{"id"} \}) = P(\{ \text{"d"} \} \cap \{ \text{"i1"} \}) = P(\{ \text{"d"} \} | \{ \text{"i1"} \}) P(\{ \text{"i1"} \}) = 2/3 \cdot 1/2 = 2/6 = 1/3$$

Podemos observar que las posibles elecciones del raton son "ii", "id" y "d". Por lo cual $E(X) = P(\{ \text{"ii"} \})E(X | \text{"ii"}) + P(\{ \text{"id"} \})E(X | \text{"id"}) + P(\{ \text{"d1"} \})E(X | \text{"d1"})$ donde

$$\begin{cases} E(X | \text{"ii"}) &= 2 \\ E(X | \text{"id"}) &= E(X) + 5 \\ E(X | \text{"d1"}) &= E(X) + 3 \end{cases}$$

Por lo tanto

$$E(X) = 1/6 \cdot 2 + 2/6 (E(X) + 5) + 1/2 (E(X) + 3)$$

$$= 5/6 E(X) + 7/2$$

$$\Rightarrow E(X) = 21$$

Este resultado concuerda con el de la simulacion.

0.5 Ejercicio 5

- a) La matriz de transición P es no-regular. Esta cadena de Markov presenta 4 clases de comunicación (una por cada estado). El estado M es absorbente, por lo cual la cadena de Markov es absorbente, los demás estados son de transición.
- b) Para $n=10$ pasos obtenemos la matriz P^n , donde P_{ij}^n representa la probabilidad de llegar a j desde i en n pasos.

$$P^{10} \simeq \begin{pmatrix} 0.99004 & 0.00969 & 0.000214 & 0.00005069 \\ 0 & 0.94159 & 0.04078 & 0.0176248 \\ 0 & 0 & 0.425037 & 0.5749629 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

Partiendo desde un estado luego de 10 unidades de observación en la evolución de esta persona podemos observar:

”Suceptible”. **Suceptible.** Partiendo desde suceptible hay una probabilidad de que el sujeto siga en su estado de suceptibilidad de 0.99004, luego de 10 unidades de observación.

VIH. Partiendo desde suceptible hay una probabilidad de que el sujeto contraiga VIH de 0.00969, luego de 10 unidades de observación.

SIDA. Partiendo desde suceptible hay una probabilidad de que el sujeto contraiga SIDA de 0.000214, luego de 10 unidades de observación.

Muerto. Partiendo desde suceptible hay una probabilidad de que el sujeto muera a causa de la enfermedad de 0.00005069, luego de 10 unidades de observación.

Por lo tanto si el sujeto es suceptible, en 10 unidades de observación, es altamente probable que se mantenga en esta condición.

”VIH”. **Suceptible.** Partiendo desde VIH no hay una probabilidad de que el sujeto vuelva al estado ”suceptible”, luego de 10 unidades de observación.

VIH. Partiendo desde VIH hay una probabilidad de que el sujeto se mantenga en su condición de 0.94159, luego de 10 unidades de observación.

SIDA. Partiendo desde VIH hay una probabilidad de que el sujeto contraiga SIDA de 0.04078, luego de 10 unidades de observación.

Muerto. Partiendo desde VIH hay una probabilidad de que el sujeto muera a causa de la enfermedad de 0.0176248, luego de 10 unidades de observación.

Por lo tanto si el sujeto está diagnosticado con VIH, tiene altas probabilidades de mantener su condición y no puede librarse de su enfermedad.

”SIDA”. **Suceptible.** Partiendo desde SIDA no hay una probabilidad de que el sujeto vuelva al estado de ”suceptibilidad”, luego de 10 unidades de observación.

Suceptible. Partiendo desde SIDA no hay una probabilidad de que el sujeto vuelva al estado de ”VIH”, luego de 10 unidades de observación.

SIDA. Partiendo desde SIDA hay una probabilidad de que el sujeto se mantenga en su condición de 0.425037, luego de 10 unidades de observación.

Muerto. Partiendo desde SIDA hay una probabilidad de que el sujeto muera a causa de la enfermedad de 0.5749629, luego de 10 unidades de observación.

Por lo tanto si el sujeto está diagnosticado con SIDA, tiene un 57% de morir luego de 10 unidades de observación. No puede volver a la condición de VIH, ni a la de suceptible, mientras que se puede mantener con SIDA con una probabilidad de 0.425037%.

Por otro lado si está muerto... bueno, digamos que es el estado absorbente por excelencia.

c) Sabemos que

$$Q = \begin{pmatrix} 0.999 & 0.001 & 0 \\ 0 & 0.994 & 0.006 \\ 0 & 0 & 0.918 \end{pmatrix} \quad (2)$$

Y por lo tanto,

$$F = (I - Q)^{-1} = \begin{pmatrix} 1000 & \frac{500}{3} & \frac{500}{41} \\ 0 & \frac{500}{3} & \frac{41}{500} \\ 0 & 0 & \frac{41}{500} \end{pmatrix} \quad (3)$$

Donde F es la matriz fundamental, en la cual F_{ij} es el numero esperado de visitas a j dado que la cadena comienza en i. Por lo tanto, si iniciamos en el estado

"Suceptible" el numero esperado de pasos para llegar al estado de muerte a causa de la enfermedad es de 1178.862

"VIH" el numero esperado de pasos para llegar al estado de muerte a causa de la enfermedad es de 178.862

"SIDA" el numero esperado de pasos para llegar al estado de muerte a causa de la enfermedad es de 12.19

Cada uno de los resultados obtenidos de acuerdo al estado desde el que se comienza a contabilizar, es el promedio de unidades de observacion para que la persona llegue al estado de muerte.

d) Para la busqueda de la distribucion estacionaria utilizare la tecnica vista en clase. Busco un vector $x=(x_1,x_2,x_3,1)$ tal que

$$xP = x$$

Por lo cual obtenemos el siguiente sistema de ecuaciones lineales:

$$(1) \quad x_1 0.999 + x_2 0 + x_3 0 + 0 = x_1$$

$$(2) \quad x_1 0.001 + x_2 0.994 + x_3 0 + 0 = x_2$$

$$(3) \quad x_1 0 + x_2 0.006 + x_3 0.918 + 0 = x_3$$

$$(4) \quad x_1 0 + x_2 0 + x_3 0.082 + 1 = 1$$

$$(4) \Rightarrow X_3 = 0$$

$$(4)(3) \Rightarrow X_2 = 0$$

$$(4)(3)(2) \Rightarrow X_1 = 0$$

Se puede observar que las componentes en x suman 1 por lo tanto $x = \pi$ es la distribucion estacionaria de esta cadena de Markov.

e) Al ser una cadena absorbente tenemos que:

$$R = \begin{pmatrix} 0 \\ 0 \\ 0.082 \end{pmatrix} \quad F = (I - Q)^{-1} = \begin{pmatrix} 1000 & \frac{500}{3} & \frac{500}{41} \\ 0 & \frac{500}{3} & \frac{41}{500} \\ 0 & 0 & \frac{41}{500} \end{pmatrix} \quad FR = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\lim_{n \rightarrow \infty} P^n = \left(\begin{array}{c|c} 0 & (I-Q)^{-1}R \\ \hline 0 & I \end{array} \right) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Por lo tanto esta cadena tiene distribucion limite.

0.6 Ejercicio 6

a) El modelado de este PageRank como cadena de Markov produce la siguiente matriz de transición:

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \end{pmatrix} \quad (4)$$

b) Para simular esta situación en R, implemente lo siguiente:

```
P = matrix(nrow=7,ncol=7)

Xa = c(1/2,5,6)

Xb = c(1/3, 1, 3,6)

Xc = c(1/2,4,6)

Xd = c(1,6)

Xe = c(1/4,1,4,6,7)

Xf = c(1/2,1,2)

Xg = c(1/7,1,2,3,4,5,6,7)

P = cargarMatriz(P)

contadorPaginas = c(1:7)

probabilidad= simular(100)

simular<- function(n){

  contadorPaginas= crearVector(contadorPaginas)

  paginaInicial = sample(1:7,1) Inicio de manera aleatoria

  print("La pagina inicial es")

  print(paginaInicial)

  paginaActual = paginaInicial

  for (i in (1:n)){
```

```

paginaActual = surfear(paginaActual)

contadorPaginas[paginaActual] = contadorPaginas[paginaActual] + 1

print("La pagina actual es")

print(paginaActual)

print("Contador")

print(contadorPaginas) } return (contadorPaginas/n)

}

```

```

surfear <- function (x){

paginaActual = obtenerFila(x)

proximaPagina = paginaActual[elegirProximaPagina(paginaActual)]

return (proximaPagina) }

```

```

elegirProximaPagina <-function(x){

```

```

unoEntre = length(x)-1

return (sample(1:unoEntre,1)+1) }

```

```

cargarMatriz <- function(P){

```

```

P = crearMatriz(P)

```

```

indiceFilas = 0

```

```

P = cargarFila(P,Xa,1)

```

```

P = cargarFila(P,Xb,2)

```

```

P = cargarFila(P,Xc,3)

```

```

P = cargarFila(P,Xd,4)

```

```

P = cargarFila(P,Xe,5)

```

```
P = cargarFila(P,Xf,6)
```

```
P = cargarFila(P,Xg,7)
```

```
return (P) }
```

```
cerarVector <- function(v){
```

```
for (i in (1:7)){
```

```
v[i] = 0 }
```

```
return (v) }
```

```
cerarMatriz <- function(P){
```

```
for (i in (1:7)){
```

```
for (j in (1:7)){
```

```
P[i,j] =0 } }
```

```
return (P) }
```

```
obtenerFila <- function(x){
```

```
if (x == 1){
```

```
return (Xa) }
```

```
if (x == 2){
```

```
return (Xb) }
```

```
if (x == 3){
```

```
return (Xc) }
```

```
if (x == 4){
```

```
return (Xd) }
```

```

if (x == 5){
  return (Xe) }

if (x == 6){
  return (Xf) }

if (x == 7){
  return (Xg) }

}

cargarFila <- function(P,x,fila){

  indiceFila = 0

  for (i in x){

    indiceFila = indiceFila + 1

    if (indiceFila != 1){

      P[fila,i] = x[1]

    } }

  return (P)

}

```

- c) Obtuve el siguiente vector de frecuencias relativas : [0.25, 0.16, 0.04, 0.06, 0.10, 0.34, 0.05]. Por lo cual con $n=100$ suponemos que el rango de las paginas a,b,c...,g corresponden con las posiciones desde 1 a 7 del vector obtenido por medio de la simulacion.

0.7 Ejercicio 7

- a) Sabemos que Juan recibe 10 mensajes por hora, por lo tanto $\lambda = 10$ y que nos situamos dentro del modelo de proceso de Poisson. Sea N_t el numero de mensajes que Juan recibe al tiempo t , como el analisis comienza desde las 10am $t=0$ equivale a las 10 am. Deseamos hallar la probabilidad $P(N_2 = 18, N_7 = 70) = P(N_2 = 18, N_7 - N_2 = 52)$, ya que $12 - 10 = 2$, $17 - 10 = 7$ y $70 - 18 = 52$. Como estos intervalos de tiempo son disjuntos $((0,2], (2,5])$ entonces N_2 y N_5 son variables aleatorias independientes. Por lo tanto, $P(N_2 = 18, N_7 - N_2 = 52) = P(N_2 = 18)P(N_7 - N_2 = 52) = P(N_2 = 18)P(N_5 = 52) = \frac{e^{-20}(20)^{18}}{18!} \frac{e^{-50}(50)^{52}}{52!} = 0.004481022$

0.8 Ejercicio 8

- a) Dada la distribucion inicial π_0 , y la matriz de transicion P, luego de dos semanas nos encontramos con la siguiente probabilidad de ataques a los puertos:

$$\pi_0 = (0 \quad 0 \quad 0 \quad 0 \quad 1)$$

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & \frac{8}{13} & \frac{3}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{1}{16} & \frac{3}{16} & \frac{8}{16} & \frac{1}{4} & \frac{1}{8} \\ 0 & \frac{1}{11} & \frac{4}{11} & \frac{5}{11} & \frac{1}{11} \\ 0 & \frac{1}{8} & \frac{1}{2} & \frac{1}{8} & \frac{1}{4} \end{pmatrix}$$

$$\pi_0 P^2 = (0 \quad \frac{1}{8} \quad \frac{1}{2} \quad \frac{1}{8} \quad \frac{1}{4})$$

Los puertos con menos probabilidad de ser atacados son el 135 y 445, mientras que el que tiene mas probabilidad de ser atacado es el 139. Se puede observar que el puerto 80 no tiene posibilidad de ser atacado.

- b) Esta cadena presenta una matriz de transicion regular ya que:

$$P^3 = \begin{pmatrix} \frac{1}{4387} & \frac{61}{11964081} & \frac{885}{5400719} & \frac{1019}{3203451} & \frac{167}{2113687} \\ \frac{237952}{3087} & \frac{34027136}{5128733} & \frac{17013568}{3700267} & \frac{17013568}{2425947} & \frac{17013568}{1558587} \\ \frac{146432}{579} & \frac{20939776}{1591501} & \frac{10469888}{633335} & \frac{10469888}{1898595} & \frac{10469888}{504509} \\ \frac{25168}{885} & \frac{7198048}{1268269} & \frac{1799512}{911497} & \frac{7198048}{152189} & \frac{3599024}{399807} \\ \frac{36608}{5234944} & & \frac{2617472}{654368} & & \frac{2617472}{654368} \end{pmatrix}$$

presenta todas sus entradas mayores a 0. Por lo visto en la teoria si una matriz de transicion de una cadena de Markov es regular, entonces la cadena tiene una distribucion limite, la cual es la unica distribucion estacionaria de la cadena. Por lo tanto si buscamos c y x de forma tal que $cxP = cx$ y $\sum_{i \in cx} cx_i = 1$, entonces obtendremos $\pi = cx$ (la distribucion estacionaria). Gracias a la ayuda de un software puedo concluir que:

$$\pi =$$

$$\pi = (0, 02146667, 0, 2669333, 0, 3434667, 0, 2273333, 0, 1408)$$

Es la distribucion limite de la matriz P.