

Probabilidad avanzada

3 de Septiembre de 2021

Primer examen parcial

Instrucciones: lea cuidadosamente cada ejercicio y responda de manera clara y precisa cada pregunta. No se tomarán en cuenta resultados sin desarrollos lógicos autoexplicativos. Acompañe sus respuestas con gráficas y dibujos. Recuerde que la evaluación es individual y busca medir el aprendizaje, conocimiento, y dominio obtenidos durante la primera etapa de la unidad de aprendizaje de Procesos estocásticos.

Ejercicio 1

Considere el proceso estocástico $\{X(t), t > 0\}$ definido por

$$X(t) = e^{-Yt}$$

donde Y es una variable aleatoria exponencial con parámetro λ .

- (a) Cuál es el espacio de estados del proceso $X(t)$?
- (b) Dibuje en el papel, de manera cualitativa pero clara, un conjunto de trayectorias del proceso.
- (c) Obtenga la función de distribución acumulada de primer orden del proceso.
- (d) Calcule la densidad de probabilidad de primer orden del proceso.
- (e) Determine el valor esperado del proceso.
- (f) Obtenga la función de autocovarianza del proceso.
- (g) Calcule la varianza del proceso.
- (h) Determine el coeficiente de autocorrelación del proceso.

Solución.

- (a) El proceso estocástico es del tipo continuo-continuo.

Los valores del espacio de estados son los números reales del intervalo $(0, 1)$.

El valor mínimo $X(t) = 0$, para cualquier valor fijo de t , se alcanza en el límite cuando $Y \rightarrow \infty$: $\lim_{Y \rightarrow \infty} (e^{-Yt}) = 0$, mientras que el valor máximo $X(t) = 1$ se obtiene cuando $Y = 0$: $X(t) = e^{-0 \cdot t} = 1$.

(b) Cinco trayectorias muestrales del proceso se observan en la Fig. 1.

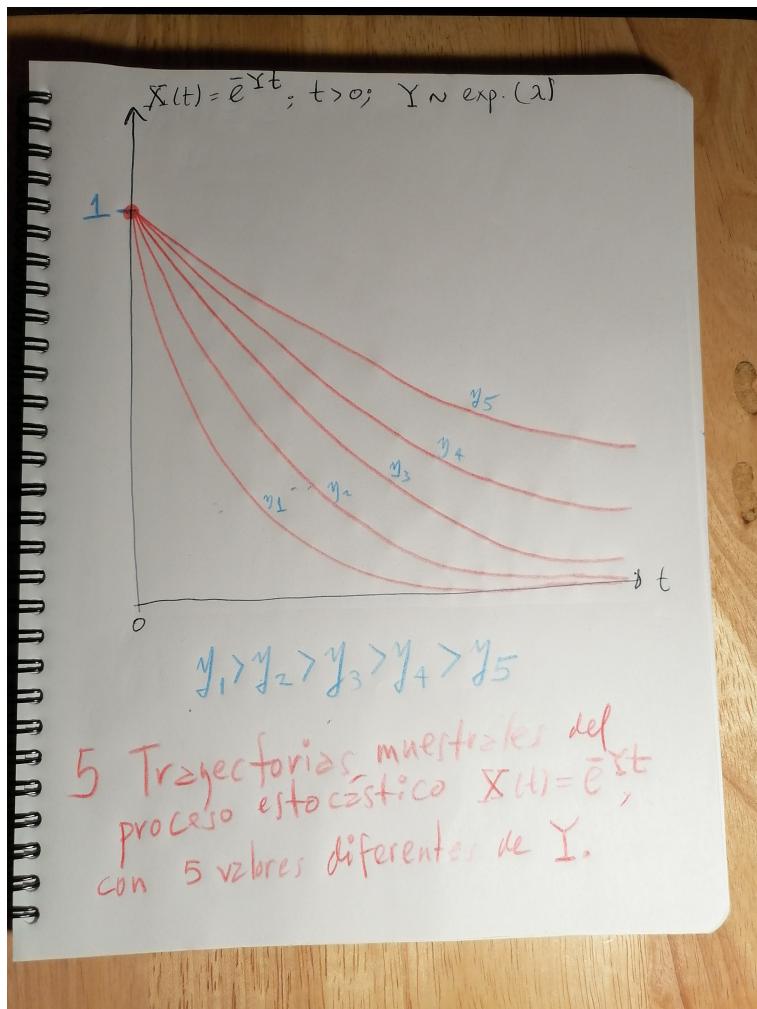


Figura 1: Una muestra de 5 trayectorias del proceso estocástico $X(t) = e^{-Yt}$ para 5 valores, sucesivamente crecientes, de la variable y .

(c) La función de distribución acumulada de primer orden del proceso es dada por

$$\begin{aligned}
 F_X(x; t) &= P(X(t) \leq x) = P(e^{-Yt} \leq x) = P[-Yt \leq \log(x)] \\
 &= P\left[Y \geq -\frac{\log(x)}{t}\right] = 1 - P\left[Y < -\frac{\log(x)}{t}\right] \\
 &= 1 - \lambda \int_0^{-\log(x)/t} e^{-\lambda y} dy = 1 - \lambda \left(\frac{e^{-\lambda y}}{-\lambda}\right) \Big|_0^{-\log(x)/t} \\
 &= 1 + \left[e^{-\lambda(-\log(x)/t)} - e^{-\lambda(0)}\right] = 1 + e^{\frac{\lambda}{t} \log(x)} - 1 \\
 &= e^{\log(x^{\lambda/t})} = x^{\lambda/t}
 \end{aligned}$$

$$F_X(x; t) = x^{\lambda/t}, \quad \text{con } 0 < x < 1$$

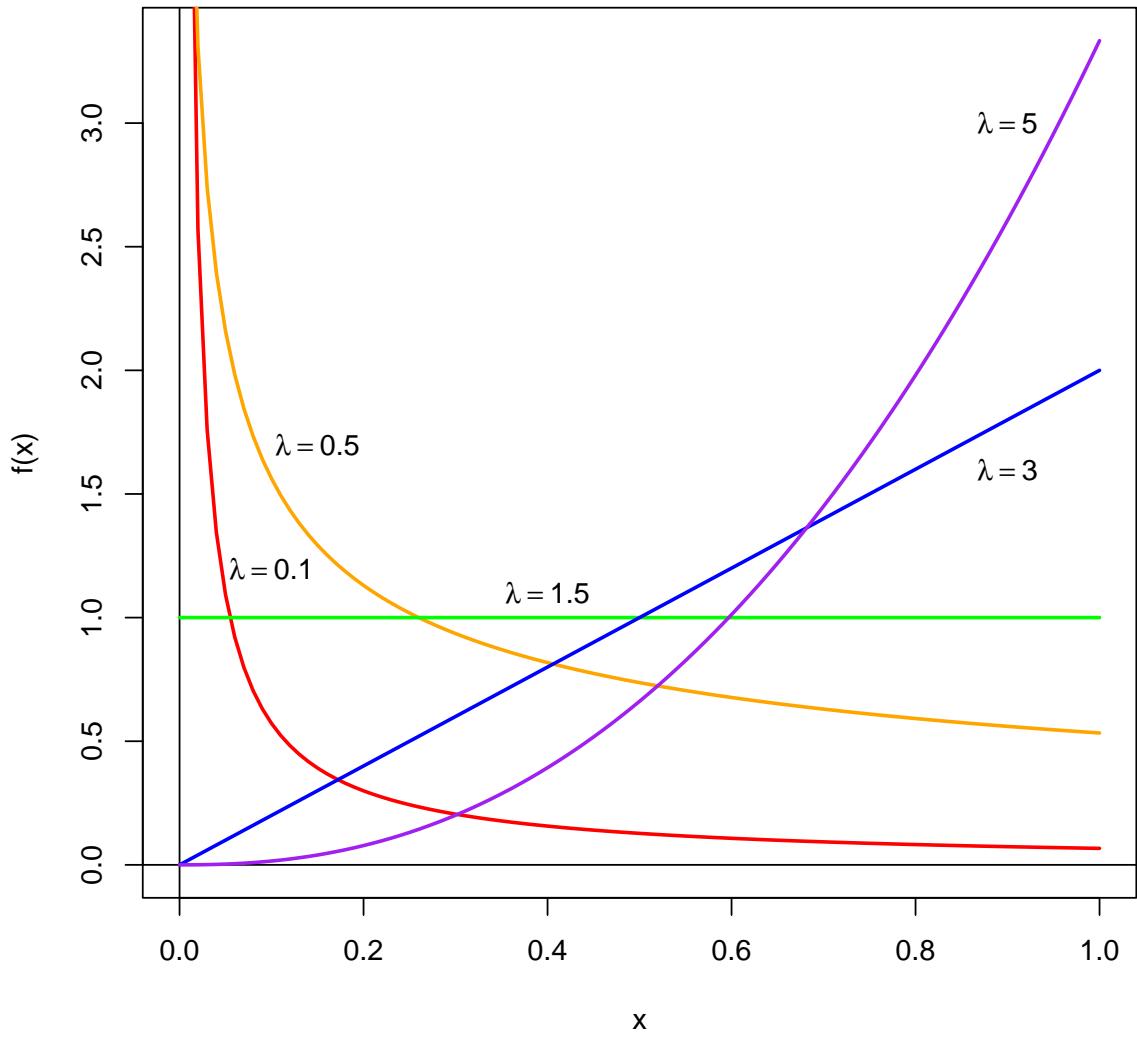
- (d) La función de densidad de probabilidad de primer orden $f_X(x; t)$ es igual a la derivada con respecto a x de la función de distribución acumulada de primer orden $F_X(x; t)$ obtenida en el inciso (c):

$$f_X(x; t) = \frac{d}{dx} F_X(x; t) = \frac{d}{dx} [x^{\lambda/t}] = \frac{\lambda}{t} x^{\lambda/t-1}$$

$$f_X(x; t) = \frac{\lambda}{t} x^{\lambda/t-1}, \quad \text{con } 0 < x < 1$$

La gráfica de la la densidad de probabilidad de primer orden para 5 valores diferentes del parámetro $\lambda = 0.1, 0.7, 1.5, 3, 5$ se muestra en la salida del siguiente programa de R:

```
rm(list=ls())#limpia la sesi'on
f <- function(x,t,L){
  (L/t)*x^((L/t)-1)
}
x <- seq(0, 1, by = 0.01) #espacio de estados del proceso
t <- 1.5 #un valor fijo del tiempo
L <- c(0.1, 0.8, 1.5, 3, 5) #valores del par'ametro lambda
C <- c("red","orange","green","blue","purple")
plot(0, 0, type="n",xlab="x",ylab="f(x)",xlim=c(0,1),ylim=c(0,max(L)/t))
abline(h=0); abline(v=0)
for(k in 1:5){
  y <- f(x,t,L[k]) #los valores de la densidad de probabilidad
  lines(x,y,col=C[k],lwd=2)
}
text(0.1,1.2,expression(lambda==0.1))
text(0.15,1.7,expression(lambda==0.5))
text(0.4,1.1,expression(lambda==1.5))
text(0.9,1.6,expression(lambda==3.0))
text(0.9,3.0,expression(lambda==5.0))
```



(e) Utilizando la densidad de Y , el valor esperado del proceso es dado por

$$\begin{aligned}
 E[X(t)] &= E[e^{-Yt}] = \int_0^\infty e^{-yt} f_Y(y) dy = \lambda \int_0^\infty e^{-yt} e^{-\lambda y} dy = \lambda \int_0^\infty e^{-(t+\lambda)y} dy \\
 &= \lambda \left[\frac{e^{-(t+\lambda)y}}{-(t+\lambda)} \right]_0^\infty = -\frac{\lambda}{t+\lambda} (0 - 1) = \frac{\lambda}{t+\lambda}
 \end{aligned}$$

$$E[X(t)] = \frac{\lambda}{t+\lambda}$$

(f) Utilizando el resultado para el valor esperado del inciso (e), la función de autocovarianza del proceso estocástico en los tiempos t_1 y t_2 es

$$\begin{aligned}
Cov(X(t_1), X(t_2)) &= E[X(t_1)X(t_2)] - E[X(t_1)]E[X(t_2)] \\
&= E[e^{-Yt_1}e^{-Yt_2}] - \frac{\lambda}{t_1 + \lambda} \frac{\lambda}{t_2 + \lambda} = E[e^{-Y(t_1+t_2)}] - \frac{\lambda^2}{(t_1 + \lambda)(t_2 + \lambda)} \\
&= E[X(t_1 + t_2)] - \frac{\lambda^2}{(t_1 + \lambda)(t_2 + \lambda)} = \frac{\lambda}{t_1 + t_2 + \lambda} - \frac{\lambda^2}{(t_1 + \lambda)(t_2 + \lambda)} \\
&= \frac{\lambda(t_1 + \lambda)(t_2 + \lambda) - \lambda^2(t_1 + t_2 + \lambda)}{(t_1 + \lambda)(t_2 + \lambda)(t_1 + t_2 + \lambda)}
\end{aligned}$$

$$Cov(X(t_1), X(t_2)) = \frac{\lambda t_1 t_2}{(t_1 + \lambda)(t_2 + \lambda)(t_1 + t_2 + \lambda)}$$

- (g) La varianza del proceso estocástico es igual a la función de autocovarianza cuando los dos tiempos t_1 y t_2 son iguales, es decir, cuando $t_1 = t_2 = t$:

$$Var(X(t)) = Cov(X(t), X(t)) = \frac{\lambda t \cdot t}{(t + \lambda)(t + \lambda)(t + t + \lambda)}$$

$$Var(X(t)) = \frac{\lambda t^2}{(t + \lambda)^2(2t + \lambda)}$$

- (h) El coeficiente de autocorrelación del proceso estocástico se obtiene al dividir la función de autocovarianza entre la raíz cuadrada del producto de las varianzas:

$$\begin{aligned}
\rho(X(t_1), X(t_2)) &= \frac{Cov(X(t_1), X(t_2))}{\sqrt{Var(X(t_1))Var(X(t_2))}} = \frac{\frac{\lambda t_1 t_2}{(t_1 + \lambda)(t_2 + \lambda)(t_1 + t_2 + \lambda)}}{\sqrt{\frac{\lambda t_1^2}{(t_1 + \lambda)^2(2t_1 + \lambda)} \frac{\lambda t_2^2}{(t_2 + \lambda)^2(2t_2 + \lambda)}}} \\
&= \frac{\frac{\lambda t_1 t_2}{(t_1 + \lambda)(t_2 + \lambda)(t_1 + t_2 + \lambda)}}{\frac{\lambda t_1 t_2}{(t_1 + \lambda)(t_2 + \lambda)\sqrt{(2t_1 + \lambda)(2t_2 + \lambda)}}}
\end{aligned}$$

$$\rho(X(t_1), X(t_2)) = \frac{\sqrt{(2t_1 + \lambda)(2t_2 + \lambda)}}{t_1 + t_2 + \lambda}$$

Ejercicio 2

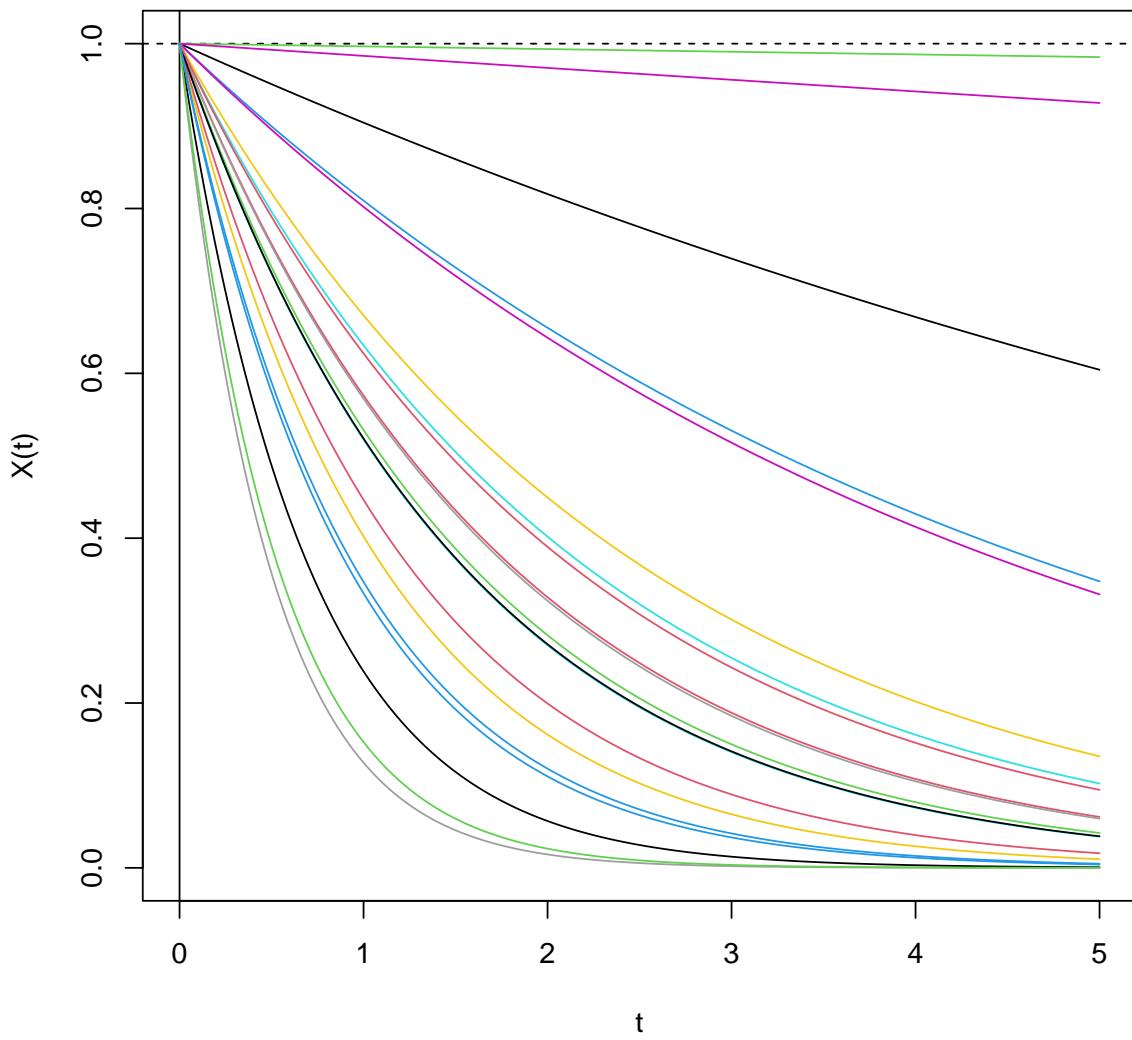
- (a) Escriba un programa en R con el cual simular un "enjambre" de realizaciones del proceso estocástico del ejercicio 1.
- (b) Utilizando la información de las realizaciones del inciso (a), genere un historgama de la variable aleatoria $X(t)$ para un valor de t fijo y muestre que el perfil empírico del histograma concide aproximadamente con la fórmula de la densidad del proceso obtenida en el inciso (d) del ejercicio 1. El programa debe dibujar la densidad de probabilidad sobre el perfil del histograma.
- (c) Utilizando las trayectorias generadas, demuestre que el valor esperado y la varianza del proceso para un valor de t fijo coincide aproximadamente con las fórmulas obtenidas

en los incisos (e) y (g) del ejercicio 1.

- (d) Con el enjambre de trayectorias de la simulación, demuestre que el coeficiente de correlación empírico para dos valores fijos del tiempo es aproximadamente igual al valor teórico del coeficiente de autocorrelación del inciso (h) del ejercicio 1.

Solución

```
(a) rm(list=ls())#limpia la sesion
t <- seq(0, 5, by = 0.01) #valores del tiempo
N <- 20#numero de realizaciones
l <- 1#valor del parametro lambda
plot(0, 0, type="n",xlim=c(0, 5), ylim=c(0, 1), xlab="t",ylab="X(t)")
abline(h=1, col = "black", lty=2)
abline(v=0)
for(i in 1:N){
  Y <- rexp(1, rate=l) #un numero aleatorio exponencial
  X <- exp(-Y*t) #una trayectoria muestral del proceso
  lines(t, X, type = "l", col = i) #grafica de cada trayectoria muestral
}
```



- (b) Se elige el valor del tiempo $t = 1.5$. Este valor corresponde a la componente 151 del vector de tiempo: $t[151] = 1.5$.

```

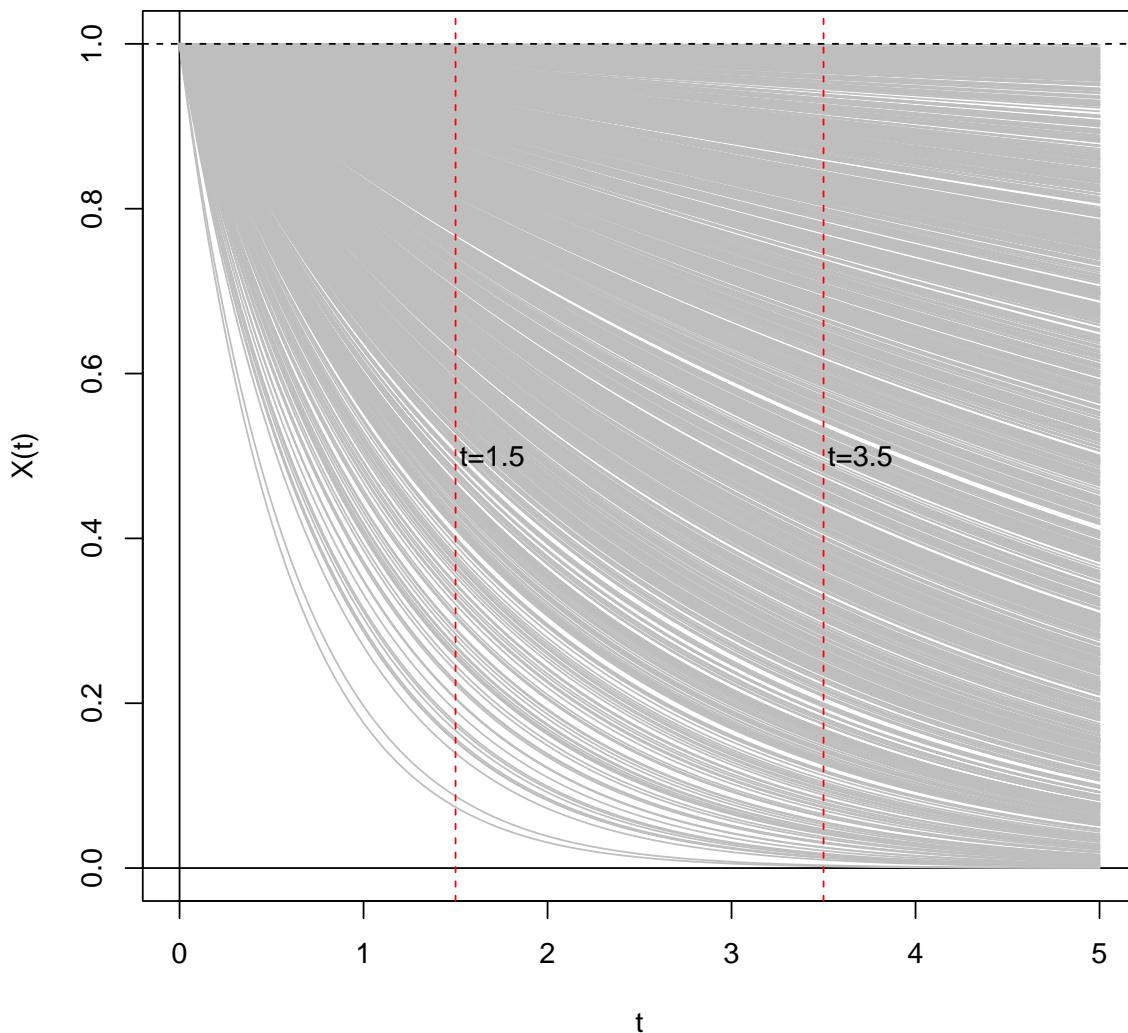
rm(list=ls())#limpia la sesion
t <- seq(0, 5, by = 0.01) #valores del tiempo
N <- 1000#numero de realizaciones
L <- 5#valor del parametro lambda
U <- vector(); V <- vector()
plot(0, 0, type="n", xlim=c(0, 5), ylim=c(0, 1), xlab="t", ylab="X(t)")
abline(v=0)
abline(h=0)
for(j in 1:N){
  Y <- rexp(1, rate=L) #un numero aleatorio exponencial
  X <- exp(-Y*t) #una trayectoria muestral del proceso
  lines(t, X, type = "l", col = "gray")#trayectoria muestral
  U[j] <- X[151]#valor del proceso cuando t=1.5 para cada trayectoria
  V[j] <- X[351]#valor del proceso cuando t=3.5 para cada trayectoria
}

```

```

}
abline(v=t[151], col = "red", lty=2)
abline(v=t[351], col = "red", lty=2)
abline(h=1, col = "black", lty=2)
text(1.7, 0.5, "t=1.5")
text(3.7, 0.5, "t=3.5")

```

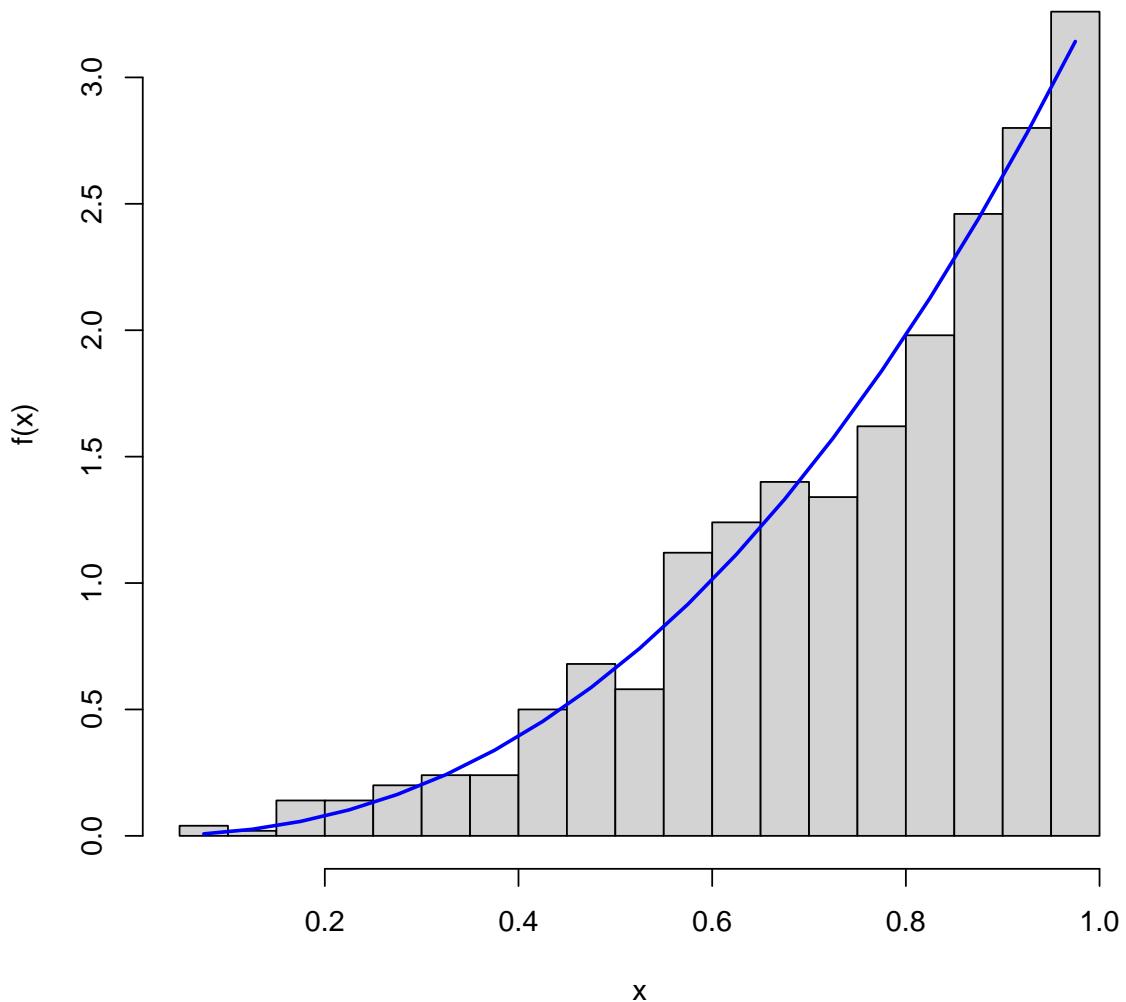


```

s <- t[151] #el valor del tiempo t=1.5
g <- function(x){#se define la densidad de probabilidad te'orica
  (L/s)*x^((L/s)-1)
}

#histograma con los valores del proceso al tiempo t=1.5
H <- hist(U,freq=FALSE,main="",breaks=30,xlab="x", ylab="f(x)")
x <- H$mid
W <- g(x)#la densidad te'orica del proceso X(t) en t=1.5
lines(x,W, col="blue",lwd=2)#grafica de la densidad te'orica

```



```
#(c) Los valores te'oricos de la media ( $m$ ) y varianza ( $v$ )
#del proceso en  $t=1.5$  y con
#lambda = 5 son, de acuerdo con los resultados de los
#incisos (g) y (h) del ejericicio 1 son, respectivamente,
m <- 5/(1.5+5)
v <- (5*(1.5)^2)/(((1.5+5)^2)*(2*1.5+5))
print(c(m, v))
```

```
## [1] 0.76923077 0.03328402
```

#Por otro lado, los valores aproximados de la media y
#la varianza del proceso en $t=1.5$ y con $\lambda=5$, son:
print(c(mean(U), var(U)))

```
## [1] 0.76426436 0.03674589
```

#(d) El valor del coeficiente de autocorrelaci'on te'orico,

```

#en los tiempo t1=1.5 y t2= 3.5, con lambda = 5,
#de acuerdo con la formula del ejercicio 1, inciso (h), es:
t1 <- 1.5; t2 <- 3.5; L <- 5
r <- sqrt((2*t1+L)*(2*t2+L))/(t1+t2+L)
print(r)

## [1] 0.9797959

#Por otro lado, el valor emp'irico (aproximado) del coeficiente de
#autocorrelaci'on es:
print(cor(U,V))

## [1] 0.9764347

```

Ejercicio 3

El flujo de un río puede ser cualquiera de los siguientes 3 estados:

- 0: flujo bajo
- 1: flujo normal
- 2: flujo alto

Suponga que el proceso estocástico $\{X_n, n = 0, 1, 2, \dots\}$, donde X_n representa el estado del flujo del río en el n -ésimo día, es una cadena de Markov. Además se estima que la probabilidad de que el flujo del río se mueva del estado i al estado j en un día es dada por la fórmula

$$P_{ij} \equiv P(X_{n+1} = j | X_n = i) = \frac{1}{2} - |i - j|a_i$$

donde $0 < a_i < 1$, para $i, j = 0, 1, 2$.

- (a) Determine los valores de las constantes a_i y escriba de manera explícita la matriz de transición de un día para el flujo del río.
- (b) Dibuje el grafo de la cadena para el flujo del río.
- (c) Mediante la cuación de Champan-Kolmogorov (sin utilizar software numérico) obtenga la probabilidad de que el flujo del río se mueva del estado 0 al estado 2 en 3 días.
- (d) Sin utilizar software numérico, determine el valor esperado del estado del flujo del río después de 2 días.

Solución

- (a) Al darle valores a los índices i y j , en la fórmula de las probabilidades de transición, la matriz de transición de un paso (un ía) adopta la forma

$$\mathbb{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} & \begin{pmatrix} 1/2 & 1/2 - a_0 & 1/2 - 2a_0 \\ 1/2 - a_1 & 1/2 & 1/2 - a_1 \\ 1/2 - 2a_2 & 1/2 - a_1 & 1/2 \end{pmatrix} \end{matrix}$$

Cada fila de la matriz de transición debe sumar uno. Al imponer esta restricción en cada fila, se determinan los valores de las constantes a_0 , a_1 y a_2 :

$$1/2 + 1/2 - a_0 + 1/2 - 2a_0 \rightarrow a_0 = 1/6$$

$$1/2 - a_1 + 1/2 + 1/2 - a_1 \rightarrow a_1 = 1/4$$

$$1/2 - 2a_2 + 1/2 - a_1 + 1/2 \rightarrow a_1 = 1/6$$

Al reemplazar la constantes obtenidas, la forma explícita de la matriz de transición de la cadena de Markov para el flujo del río adopta la forma

$$\mathbb{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} & \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \end{pmatrix} \end{matrix}$$

- (b) Una representación del grafo de la cadena de Markov para el flujo del río se muestra en la Fig. 2.

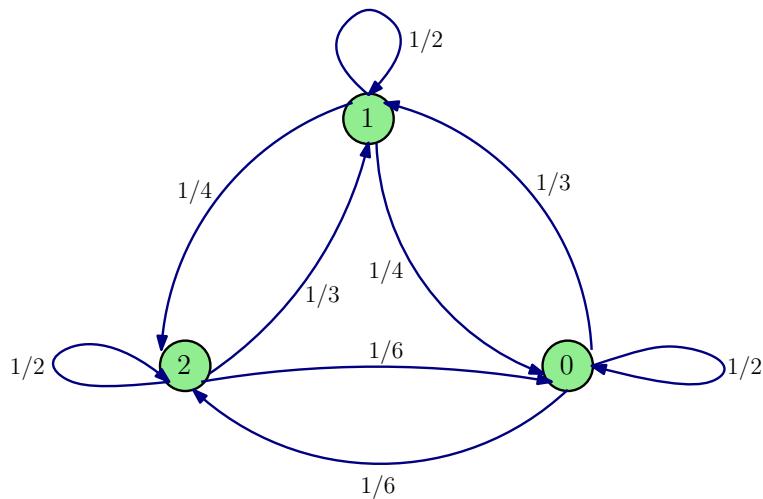


Figura 2: Grafo de la cadena de Markov para el flujo del río.

- (c) De acuerdo con la ecuación de Champan-Kolmogorov, la probabilidad de transición de ir del estado 0 al estado 2 en 3 pasos se puede escribir como

$$\begin{aligned}
P_{02}^{(3)} &= \sum_{i=0}^2 P_{0i} P_{i2}^{(2)} = \sum_{i=0}^2 \sum_{j=0}^2 P_{0i} P_{ij} P_{j2} \\
&= P_{00} P_{00} P_{02} + P_{00} P_{01} P_{12} + P_{00} P_{02} P_{22} + P_{01} P_{10} P_{02} + P_{01} P_{11} P_{12} + P_{01} P_{12} P_{22} \\
&\quad + P_{02} P_{20} P_{02} + P_{02} P_{21} P_{12} + P_{02} P_{22} P_{22} \\
&= (1/2)(1/2)(1/6) + (1/2)(1/3)(1/4) + (1/2)(1/6)(1/2) \\
&\quad + (1/3)(1/4)(1/6) + (1/3)(1/2)(1/4) + (1/3)(1/4)(1/2) \\
&\quad + (1/6)(1/6)(1/6) + (1/6)(1/3)(1/4) + (1/6)(1/2)(1/2) \\
&= \frac{1}{24} + \frac{1}{24} + \frac{1}{24} + \frac{1}{72} + \frac{1}{24} + \frac{1}{24} + \frac{1}{216} + \frac{1}{72} + \frac{1}{24} = \frac{6}{24} + \frac{2}{72} + \frac{1}{216} \\
&= \frac{1}{4} + \frac{1}{36} + \frac{1}{216} = \frac{54}{216} + \frac{6}{216} + \frac{1}{216} \\
P_{02}^{(3)} &= \frac{61}{216} \approx 0.28
\end{aligned}$$

(d) La matriz de transición \mathbb{P} se puede reescribir como

$$\mathbb{P} = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 6 & 4 & 2 \\ 3 & 6 & 3 \\ 2 & 4 & 6 \end{pmatrix}$$

Luego, la matriz de transición de 2 días (2 pasos) es

$$\mathbb{P}^2 = \frac{1}{12} \begin{pmatrix} 6 & 4 & 2 \\ 3 & 6 & 3 \\ 2 & 4 & 6 \end{pmatrix} \frac{1}{12} \begin{pmatrix} 6 & 4 & 2 \\ 3 & 6 & 3 \\ 2 & 4 & 6 \end{pmatrix} = \frac{1}{144} \begin{pmatrix} 52 & 56 & 36 \\ 42 & 60 & 42 \\ 36 & 56 & 52 \end{pmatrix}$$

Utilizando la información contenida en la matriz de transición de 2 pasos y suponiendo una distribución inicial uniforme $\boldsymbol{\alpha} = (1/3, 1/3, 1/3)$, la distribución de probabilidad del flujo del río después de 2 días es

$$\begin{aligned}
P^{(2)} &= [P(X_2 = 0), P(X_2 = 1), P(X_2 = 2)] = \boldsymbol{\alpha} \mathbb{P}^2 \\
&= (1/3, 1/3, 1/3) \frac{1}{144} \begin{pmatrix} 52 & 56 & 36 \\ 42 & 60 & 42 \\ 36 & 56 & 52 \end{pmatrix} = \frac{1}{432} (1, 1, 1) \begin{pmatrix} 52 & 56 & 36 \\ 42 & 60 & 42 \\ 36 & 56 & 52 \end{pmatrix} \\
P^{(2)} &= \frac{1}{432} (128, 172, 128)
\end{aligned}$$

$$\begin{aligned}
E[X_2] &= \sum_{k=0}^2 k P(X_2 = k) = 0P(X_2 = 0) + 1P(X_2 = 1) + 2P(X_2 = 2) \\
&= (0, 1, 2)[P(X_2 = 0), P(X_2 = 1), P(X_2 = 2)]^T \\
&= \mathbf{V}[P^{(2)}]^T = (0, 1, 2) \frac{1}{432} \begin{pmatrix} 128 \\ 172 \\ 128 \end{pmatrix} = \frac{1}{432} (0 \cdot 128 + 1 \cdot 172 + 2 \cdot 128) \\
E[X_2] &= \frac{172 + 256}{432} = \frac{428}{432} \approx 0.99
\end{aligned}$$

Ejercicio 4

Implemente un algoritmo en R para simular la cadena de Markov del ejercicio 3. Mediante la simulación, estime los valores numéricos de los incisos (c) y (d) del ejercicio 3.

Solución. Una simulación de la cadena de Markov para el flujo del río se muestra en el siguiente código R. Se calcula de manera numérica la probabilidad de que después de 3 días el estado del río pase del estado 0 al estado 2.

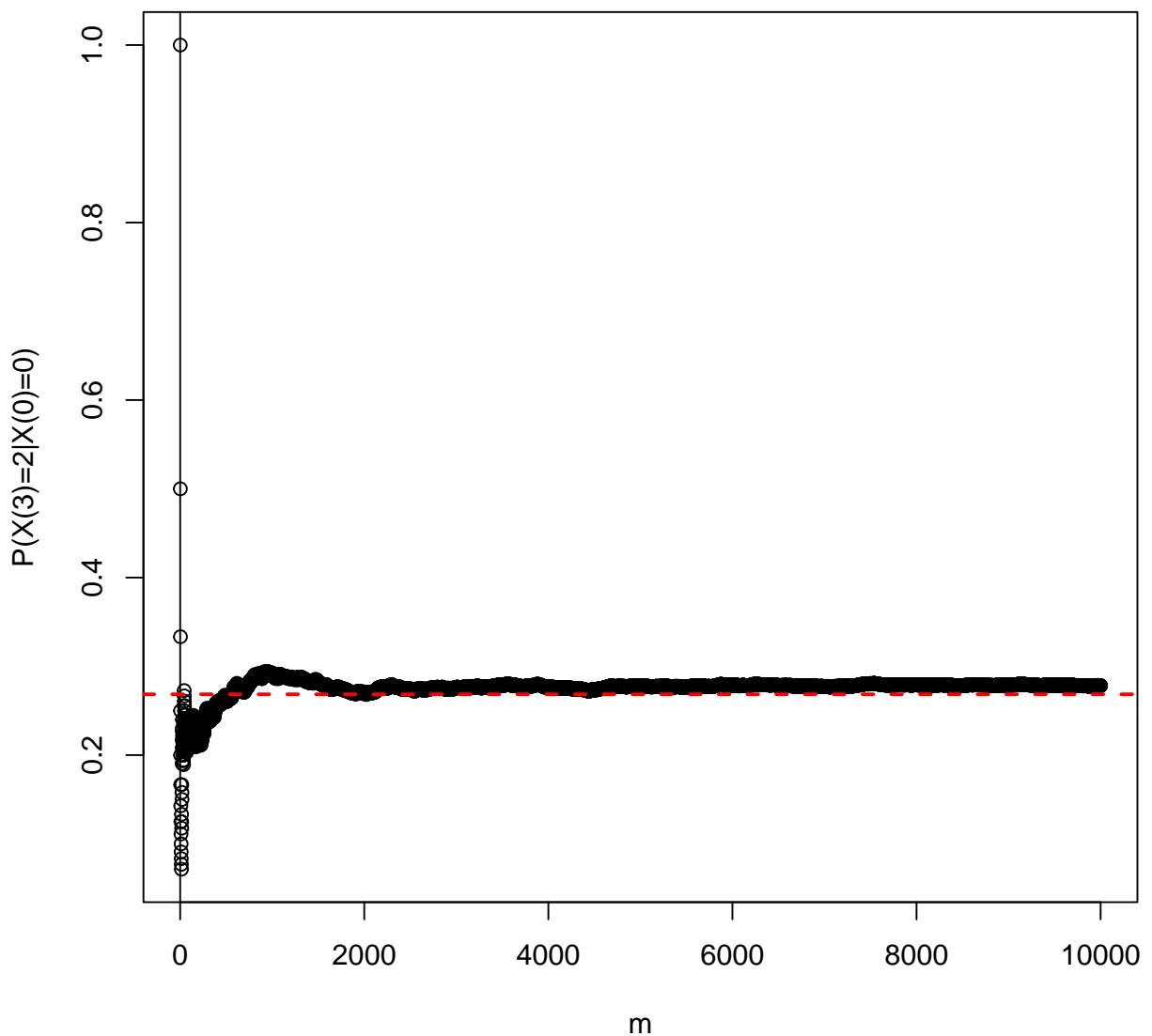
```
rm(list=ls())#para limpiar la sesi'on
library(MASS)
library(matrixcalc)
#Cadena de Markov para el flujo de un r'io
#Estados de la cadena:
estados <- c(0, 1, 2)
#Distribuci'on de probabilidad inicial:
a <- c(1, 0, 0)#inicialmente el r'io est'a en flujo bajo (estado 0)
#Matriz de transici'on:
w <- c(1/2,1/3,1/6,1/4,1/2,1/4,1/6,1/3,1/2)
M <- fractions(matrix(w,nrow=3,ncol=3,byrow=TRUE))
print(M)

##      [,1] [,2] [,3]
## [1,] 1/2  1/3  1/6
## [2,] 1/4  1/2  1/4
## [3,] 1/6  1/3  1/2

n <- 3#n'umero de pasos del proceso
m <- 10000 #n'umero de replicas o repeticiones del proceso
X <- vector(); I <- vector()
for(j in 1:m){
  X[1] <- 0
  q <- 0
  for(i in 1:n){
    X[i+1] <- sample(estados,1,prob=M[q+1,],replace=TRUE)
    q <- X[i+1]
  }
  if(X[4]==2){
    I[j] <- 1
  }else{I[j] <- 0}
}

#El valor te'orica de la probabilidad de transici'on
#de que el flujo del r'io cambie del estado 0 al estado 2
#en 3 d'ias es  $P(X(3)=2|X(0)=0)=29/108=0.27$  (Ejercicio 3, inciso (c))
#El valor emp'irico de esta probabilidad, obtenido con la simulaci'on
#de la cadena de Markov, se muestra en la siguiente figura:

plot(1:m, cumsum(I)/(1:m),xlab="m", ylab="P(X(3)=2|X(0)=0)")
abline(h=0); abline(v=0)
abline(h=29/108, col="red",lty=2,lwd=2)
```



Una simulación de la cadena de Markov para el flujo del río se muestra en el siguiente código R. Se calcula de manera numérica el valor esperado para el flujo del río después de 2 días.

```
rm(list=ls())#para limpiar la sesi'on
library(MASS)
library(matrixcalc)
#Cadena de Markov para el flujo de un r'io
#Estados de la cadena:
estados <- c(0, 1, 2)
#Distribuci'on de probabilidad inicial:
a <- c(1/3, 1/3, 1/3)
#Matriz de transici'on:
w <- c(1/2,1/3,1/6,1/4,1/2,1/4,1/6,1/3,1/2)
M <- fractions(matrix(w,nrow=3,ncol=3,byrow=TRUE))
print(M)
```

```

##      [,1] [,2] [,3]
## [1,] 1/2  1/3  1/6
## [2,] 1/4  1/2  1/4
## [3,] 1/6  1/3  1/2

n <- 2 #número de pasos del proceso
m <- 10000 #número de replicas o repeticiones del proceso
X <- vector(); Y <- vector(); I <- vector()
for(j in 1:m){
  X[1] <- sample(estados,1,prob=a,replace=TRUE)
  q <- X[1]
  for(i in 1:n){
    X[i+1] <- sample(estados,1,prob=M[q+1,],replace=TRUE)
    q <- X[i+1]
  }
  Y[j] <- X[3] #valores del flujo del río a los 2 días
}

#El valor esperado (teórico) del estado del flujo del río
#después de 2 días es  $E[X(2)] = 428/432 = 0.99$  (Ejercicio 3, inciso (d))
#El valor empírico de este valor esperado, obtenido con la simulación
#de la cadena de Markov, se muestra en la siguiente figura:

plot(1:m,cumsum(Y)/(1:m),xlab="m", ylab="E[X(2)]")
abline(h=0); abline(v=0)
abline(h=428/432,col="red",lty=2,lwd=2)

```

