

Introducción a los
PROCESOS ESTOCÁSTICOS

Luis Rincón
Departamento de Matemáticas
Facultad de Ciencias UNAM
Circuito Exterior de CU
04510 México DF

Versión: Enero 2011

Una versión actualizada del presente texto se encuentra disponible en formato electrónico en la dirección <http://www.matematicas.unam.mx/lars>

Prólogo

El presente texto contiene material básico en temas de procesos estocásticos a nivel licenciatura. Está dirigido a estudiantes de las carreras de matemáticas, actuaría, matemáticas aplicadas y otras carreras afines. Este trabajo es una versión ampliada de las notas de clase del curso semestral de procesos estocástico que he impartido en la Facultad de Ciencias de la UNAM a alumnos de las carreras de actuaría y matemáticas.

Los temas tratados en este texto corresponden a ciertos modelos clásicos de la teoría de los procesos estocásticos, y a ciertas preguntas o problemas matemáticos que uno puede plantearse al estudiar tales modelos. El texto inicia con una explicación del concepto de proceso estocástico y con algunas definiciones y ejemplos generales de este tipo de modelos. A manera de un acercamiento inicial se presenta primero una introducción breve al tema de caminatas aleatorias en una dimensión, y particularmente se analiza el problema de la ruina del jugador. En el segundo capítulo se estudian cadenas de Markov a tiempo discreto y después se estudia el mismo tipo de modelo para tiempos continuos, para ello se presenta primero el proceso de Poisson. Se estudian después los procesos de renovación, y brevemente también la teoría de la confiabilidad. Se presenta después una introducción a la teoría de martingalas a tiempo discreto, en donde se estudian solo algunos de sus muchos resultados y aplicaciones. Se incluye además una introducción al movimiento Browniano. El texto finaliza con una exposición compacta y ligera del cálculo estocástico de Itô. El material completo excede lo que regularmente es impartido en un curso semestral, y una selección adecuada de temas será necesaria. Cada capítulo contiene material que puede considerarse como introductorio al tema. Al final de cada uno de ellos se proporcionan algunas referencias para que el lector pueda precisar o profundizar lo que aquí se presenta.

El texto contiene una colección de ejercicios que se han colocado al final de cada capítulo, y se han numerado de manera consecutiva a lo largo del libro. Se incluyen también sugerencias o soluciones de algunos de estos ejercicios. A lo largo de la exposición el lector encontrará también algunos otros ejemplos y ejercicios, los cuales son particularmente útiles para el estudiante autodidacta, o para presentar en clase si se adopta alguna parte de este texto como material de estudio en algún curso.

Este material fue escrito en \LaTeX , las gráficas fueron elaboradas usando el excelente

paquete PSTRICKS, y las fotos fueron tomadas del archivo MacTutor (<http://www-history.mcs.st-and.ac.uk/>) de la universidad de St. Andrews en Escocia. La mayor parte de este trabajo fue elaborado mientras realizaba una estancia sabática en la universidad de Nottingham durante el año 2007, y agradezco sinceramente al Prof. Belavkin su amable hospitalidad para llevar a cabo este proyecto, y a la DGAPA-UNAM por el generoso apoyo económico recibido durante esa agradable estancia.

Agradezco sinceramente todos los comentarios, correcciones y sugerencias que he recibido por parte de alumnos y profesores para mejorar este material. Toda comunicación puede enviarse a la cuenta de correo que aparece abajo.

Luis Rincón
Enero 2011
Ciudad Universitaria UNAM
`lars@fciencias.unam.mx`

Contenido

1. Introducción	1
2. Caminatas aleatorias	5
2.1. Caminatas aleatorias	5
2.2. El problema del jugador	14
Notas y referencias	20
Ejercicios	21
3. Cadenas de Markov	23
3.1. Propiedad de Markov	23
3.2. Ejemplos	27
3.3. Ecuación de Chapman-Kolmogorov	34
3.4. Comunicación	37
3.5. Periodo	40
3.6. Primeras visitas	43
3.7. Recurrencia y transitoriedad	45
3.8. Tiempo medio de recurrencia	51
3.9. Clases cerradas	52
3.10. Número de visitas	53
3.11. Recurrencia positiva y nula	59
3.12. Evolución de distribuciones	63
3.13. Distribuciones estacionarias	65
3.14. Distribuciones límite	71
3.15. Cadenas regulares	76
3.16. Cadenas reversibles	78
Resumen de la notación	81
Notas y referencias	82
Ejercicios	84

4. El proceso de Poisson	97
4.1. Proceso de Poisson	97
4.2. Definiciones alternativas	105
4.3. Proceso de Poisson no homogéneo	109
4.4. Proceso de Poisson compuesto	112
Notas y referencias	113
Ejercicios	115
5. Cadenas de Markov a tiempo continuo	123
5.1. Procesos de saltos	124
5.2. Propiedades generales	127
5.3. Procesos de nacimiento y muerte	133
5.4. Proceso de nacimiento puro	137
Notas y referencias	140
Ejercicios	141
6. Procesos de renovación y confiabilidad	143
6.1. Procesos de renovación	143
6.2. Función y ecuación de renovación	146
6.3. Tiempos de vida	148
6.4. Teoremas de renovación	153
6.5. Confiabilidad	157
Notas y referencias	162
Ejercicios	163
7. Martingalas	167
7.1. Filtraciones	167
7.2. Tiempos de paro	169
7.3. Martingalas	171
7.4. Ejemplos	172
7.5. Procesos detenidos	175
7.6. Una aplicación: Estrategias de juego	176
7.7. Teorema de paro opcional y aplicaciones	179
7.8. Algunas desigualdades	183
7.9. Convergencia de martingalas	185
7.10. Representación de martingalas	191
Notas y referencias	195
Ejercicios	198
8. Movimiento Browniano	205

8.1. Definición	206
8.2. Función de probabilidad de transición	209
8.3. Propiedades básicas	210
8.4. Propiedades de las trayectorias	214
8.5. Movimiento Browniano multidimensional	218
8.6. El principio de reflexión	221
8.7. Recurrencia y transitoriedad	222
Notas y referencias	229
Ejercicios	232
9. Cálculo estocástico	237
9.1. Integración estocástica	237
9.2. Fórmula de Itô	250
9.3. Ecuaciones diferenciales estocásticas	252
9.4. Simulación	257
9.5. Algunos modelos particulares	258
Notas y referencias	268
Ejercicios	269
A. Conceptos y resultados varios	271

CAPÍTULO 1

Introducción

Considere un sistema que puede caracterizarse por estar en cualquiera de un conjunto de estados previamente especificado. Suponga que el sistema evoluciona o cambia de un estado a otro a lo largo del tiempo de acuerdo a una cierta ley de movimiento, y sea X_t el estado del sistema al tiempo t . Si se considera que la forma en la que el sistema evoluciona no es determinista, sino provocada por algún mecanismo azaroso, entonces puede considerarse que X_t es una variable aleatoria para cada valor del índice t . Esta colección de variables aleatorias es la definición de proceso estocástico, y sirve como modelo para representar la evolución aleatoria de un sistema a lo largo del tiempo. En general, las variables aleatorias que conforman un proceso no son independientes entre sí, sino que están relacionadas unas con otras de alguna manera particular. Más precisamente, la definición de proceso estocástico toma como base un espacio de probabilidad (Ω, \mathcal{F}, P) y puede enunciarse de la siguiente forma.

Definición. Un *proceso estocástico* es una colección de variables aleatorias $\{X_t : t \in T\}$ parametrizada por un conjunto T , llamado *espacio parametral*, y con valores en un conjunto S llamado *espacio de estados*.

En los casos más sencillos, y que son los que consideraremos en este texto, se toma como espacio parametral el conjunto discreto $T = \{0, 1, 2, \dots\}$, o bien el conjunto continuo $T = [0, \infty)$, y estos números se interpretan como tiempos. En el primer caso se dice que el proceso es a *tiempo discreto*, y en general este tipo de procesos se denotará por $\{X_n : n = 0, 1, \dots\}$, mientras que en el segundo caso el proceso es a *tiempo continuo*, y se denotará por $\{X_t : t \geq 0\}$. Es decir, seguiremos la convención

de que si el subíndice es n , entonces los tiempos son discretos, y si el subíndice es t , el tiempo se mide de manera continua. Los posibles espacios de estados que consideraremos son subconjuntos de \mathbb{Z} , y un poco más generalmente tomaremos como espacio de estados el conjunto de números reales \mathbb{R} , aunque en algunos pocos casos también consideraremos a \mathbb{Z}^n o \mathbb{R}^n . Naturalmente, espacios más generales son posibles, tanto para el espacio parametral como para el espacio de estados. En particular, para poder hablar de variables aleatorias con valores en el espacio de estados S , es necesario asociar a este conjunto una σ -álgebra. Considerando que S es un subconjunto de \mathbb{R} , puede tomarse la σ -álgebra de Borel de \mathbb{R} restringida a S , es decir, $S \cap \mathcal{B}(\mathbb{R})$.

Un proceso estocástico puede considerarse como una función de dos variables $X : T \times \Omega \rightarrow S$ tal que a la pareja (t, ω) se le asocia el estado $X(t, \omega)$, lo cual también puede escribirse como $X_t(\omega)$. Para cada valor de t en T , el mapeo $\omega \mapsto X_t(\omega)$ es una variable aleatoria, mientras que para cada ω en Ω fijo, la función $t \mapsto X_t(\omega)$ es llamada una *trayectoria* o *realización* del proceso. Es decir, a cada ω del espacio muestral le corresponde una trayectoria del proceso. Es por ello que a veces se define un proceso estocástico como una *función aleatoria*. Una de tales trayectorias típicas que además cuenta con la propiedad de ser continua se muestra en la Figura 1.1, y corresponde a una trayectoria de un movimiento Browniano, proceso que definiremos y estudiaremos más adelante.

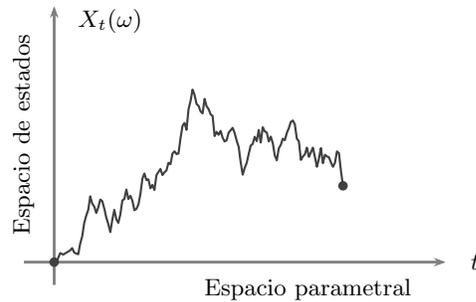


Figura 1.1:

Si A es un conjunto de estados, el evento $(X_t \in A)$ corresponde a la situación en donde al tiempo t el proceso toma algún valor dentro del conjunto A . En particular, $(X_t = x)$ es el evento en donde al tiempo t el proceso se encuentra en el estado x .

Los diferentes tipos de procesos estocásticos se obtienen al considerar las distintas posibilidades para: el espacio parametral, el espacio de estados, las características de las trayectorias, y principalmente las relaciones de dependencia entre las variables aleatorias que conforman el proceso. Los siguientes son algunos ejemplos generales de procesos estocásticos. Estos son procesos que cumplen una cierta

propiedad particular, no necesariamente excluyentes unas de otras. A lo largo del texto estudiaremos y especificaremos con mayor detalle algunos de estos tipos de procesos.

Proceso de ensayos independientes. El proceso a tiempo discreto $\{X_n : n = 0, 1, \dots\}$ puede estar constituido por variables aleatorias independientes. Este modelo representa una sucesión de ensayos independientes de un mismo experimento aleatorio, por ejemplo, lanzar un dado o una moneda repetidas veces. El resultado u observación del proceso en un momento cualquiera es, por lo tanto, independiente de cualquier otra observación pasada o futura del proceso.

Procesos de Markov. Estos tipos de procesos son importantes y son modelos en donde, suponiendo conocido el estado presente del sistema, los estados anteriores no tienen influencia en los estados futuros del sistema. Esta condición se llama *propiedad de Markov* y puede expresarse de la siguiente forma: Para cualesquiera estados x_0, x_1, \dots, x_{n-1} (pasado), x_n (presente), x_{n+1} (futuro), se cumple la igualdad

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n).$$

De esta forma la probabilidad del evento futuro ($X_n = x_n$) sólo depende el evento ($X_{n-1} = x_{n-1}$), mientras que la información correspondiente al evento pasado ($X_0 = x_0, \dots, X_{n-2} = x_{n-2}$) es irrelevante. Los procesos de Markov han sido estudiados extensamente y existe un gran número de sistemas que surgen en muy diversas disciplinas del conocimiento para los cuales el modelo de proceso estocástico y la propiedad de Markov son razonables. En particular, los sistemas dinámicos deterministas dados por una ecuación diferencial pueden considerarse procesos de Markov pues su evolución futura queda determinada por la posición inicial del sistema y una ley de movimiento especificada.

Procesos con incrementos independientes. Se dice que un proceso $\{X_t : t \geq 0\}$ tiene incrementos independientes si para cualesquiera tiempos $0 \leq t_1 < t_2 < \dots < t_n$, las variables $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ son independientes.

Procesos estacionarios. Se dice que un proceso $\{X_t : t \geq 0\}$ es estacionario (en el sentido estricto) si para cualesquiera tiempos t_1, \dots, t_n , la distribución del vector $(X_{t_1}, \dots, X_{t_n})$ es la misma que la del vector $(X_{t_1+h}, \dots, X_{t_n+h})$ para cualquier valor de $h > 0$. En particular, la distribución de X_t es la misma que la de X_{t+h} para cualquier $h > 0$, y entonces esta distribución es la misma para cualquier valor de t .

Procesos con incrementos estacionarios. Se dice que un proceso $\{X_t : t \geq 0\}$

tiene incrementos estacionarios si para cualesquiera tiempos $s < t$, y para cualquier $h > 0$, las variables $X_{t+h} - X_{s+h}$ y $X_t - X_s$ tienen la misma distribución de probabilidad. Es decir, el incremento que sufre el proceso entre los tiempos s y t sólo depende de estos tiempos a través de la diferencia $t - s$, y no de los valores específicos de s y t .

Martingalas. Una martingala a tiempo discreto es, en términos generales, un proceso $\{X_n : n = 0, 1, \dots\}$ que cumple la condición

$$E(X_{n+1} | X_0 = x_0, \dots, X_n = x_n) = x_n. \quad (1.1)$$

En palabras, esta igualdad significa que el estado promedio del proceso al tiempo futuro $n + 1$ es el valor del proceso en su último momento observado, es decir, x_n . Esto es, se trata de una ley de movimiento aleatorio que es equilibrada pues en promedio el sistema no se mueve del último momento observado. A estos procesos también se les conoce como procesos de juegos justos pues si se considera una sucesión infinita de apuestas sucesivas y si X_n denota el capital de uno de los jugadores al tiempo n , entonces la propiedad de martingala (1.1) establece que el juego es justo pues en promedio el jugador no pierde ni gana en cada apuesta.

Procesos de Lèvy. Se dice que un proceso a tiempo continuo $\{X_t : t \geq 0\}$ es un proceso de Lèvy si sus incrementos son independientes y estacionarios. Más adelante veremos que tanto el proceso de Poisson como el movimiento Browniano son ejemplos de este tipo de procesos.

Procesos Gaussianos. Se dice que un proceso a tiempo continuo $\{X_t : t \geq 0\}$ es un proceso Gaussiano si para cualesquiera colección finita de tiempos t_1, \dots, t_n , el vector $(X_{t_1}, \dots, X_{t_n})$ tiene distribución normal o Gaussiana. Nuevamente, el movimiento Browniano es un ejemplo de este tipo de procesos.

El objetivo del presente texto es el de proporcionar una introducción a algunos resultados elementales sobre ciertos tipos de procesos estocásticos.

■ Ejercicio. Demuestre que todo proceso a tiempo discreto $\{X_n : n = 0, 1, \dots\}$ con incrementos independientes es un proceso de Markov. Este resultado no es válido para procesos a tiempo continuo. ■

CAPÍTULO 2

Caminatas aleatorias

En este capítulo se presenta una introducción breve al tema de caminatas aleatorias en una dimensión. Encontraremos la distribución de probabilidad de la posición de una partícula que efectúa una caminata aleatoria en \mathbb{Z} , y la probabilidad de retorno a la posición de origen. Se plantea y resuelve después el problema de la ruina del jugador, y se analizan algunos aspectos de la solución.

2.1. Caminatas aleatorias

Una caminata aleatoria simple sobre el conjunto de números enteros \mathbb{Z} es un proceso estocástico a tiempo discreto $\{X_n : n = 0, 1, \dots\}$ que evoluciona como se muestra en la Figura 2.1.

Es decir, iniciando en el estado 0, al siguiente tiempo el proceso puede pasar al estado +1 con probabilidad p , o al estado -1 con probabilidad q , en donde $p+q = 1$. Se usa la misma regla para los siguientes tiempos, es decir, pasa al estado de la derecha con probabilidad p , o al estado de la izquierda con probabilidad q . El valor de X_n es el estado del proceso al tiempo n . Este proceso cambia de un estado a otro en dos tiempos consecutivos de acuerdo a las *probabilidades de transición* que se muestran en la Figura 2.1, válidas para

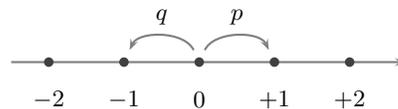


Figura 2.1:

cualquier $n \geq 0$, y para cualesquiera enteros i y j . Estas probabilidades se pueden escribir de la forma siguiente

$$P(X_{n+1} = j | X_n = i) = \begin{cases} p & \text{si } j = i + 1, \\ q & \text{si } j = i - 1, \\ 0 & \text{otro caso.} \end{cases}$$

Dado que estas probabilidades no dependen de n , se dice que son *homogéneas en el tiempo*, es decir, son las mismas para cualquier valor de n . A partir de estas consideraciones, es intuitivamente claro que este proceso cumple la propiedad de Markov, es decir, el estado futuro del proceso depende únicamente del estado presente y no de los estados previamente visitados. Una posible trayectoria de este proceso se muestra en la Figura 2.2.

Una caminata aleatoria puede también definirse de la forma siguiente: Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas. Por la idéntica distribución denotaremos a cualquiera de ellas mediante la letra ξ sin subíndice. Supondremos que $P(\xi = +1) = p$ y $P(\xi = -1) = q$, en donde, como antes, $p + q = 1$. Entonces para $n \geq 1$ se define

$$X_n = X_0 + \xi_1 + \dots + \xi_n.$$

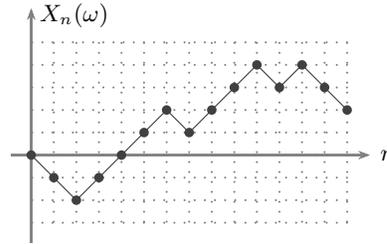


Figura 2.2:

Sin pérdida de generalidad supondremos que $X_0 = 0$. Nos interesa encontrar algunas propiedades de la variable X_n , y de su comportamiento como función de n . Por ejemplo, a partir de la expresión anterior, es inmediato encontrar su esperanza y varianza.

Proposición. Para cualquier entero $n \geq 0$,

1. $E(X_n) = n(p - q)$.
2. $\text{Var}(X_n) = 4npq$.

Demostración. Para la esperanza tenemos que $E(X_n) = \sum_{i=1}^n E(\xi_i) = n E(\xi) =$

$n(p - q)$. Por otro lado, como $E(\xi^2) = p + q = 1$ y $E(\xi) = p - q$, se tiene que $\text{Var}(\xi) = 1 - (p - q)^2 = 4pq$. Por lo tanto $\text{Var}(X_n) = \sum_{i=1}^n \text{Var}(\xi_i) = n \text{Var}(\xi) = 4npq$. \square

Analicemos estas dos primeras fórmulas. Si $p > q$, es decir, si la caminata toma pasos a la derecha con mayor probabilidad, entonces el estado promedio después de n pasos es un número positivo, es decir, su comportamiento promedio es tender hacia la derecha, lo cual es intuitivamente claro. Análogamente, si $p < q$, entonces el estado final promedio de la caminata después de n pasos es un número negativo, es decir, en promedio la caminata tiende a moverse hacia la izquierda. En ambos casos la varianza crece conforme el número de pasos n crece, eso indica que mientras mayor es el número de pasos que se dan, mayor es la incertidumbre acerca de la posición final del proceso. Cuando $p \neq q$ se dice que la caminata es *asimétrica*. Cuando $p = q = 1/2$ se dice que la caminata es *simétrica*, y en promedio el proceso se queda en su estado inicial pues $E(X_n) = 0$, sin embargo para tal valor de p la varianza es $\text{Var}(X_n) = n$, y es sencillo demostrar que ese valor es el máximo de la expresión $4npq$, para $p \in (0, 1)$.

■ Ejercicio. Demuestre que la función generadora de momentos de la variable X_n es $M(t) = E(e^{tX_n}) = (pe^t + qe^{-t})^n$. A partir de esta expresión encuentre nuevamente la esperanza y varianza de X_n . \bullet

Probabilidades de transición. Como hemos supuesto que la caminata inicia en cero, es intuitivamente claro que después de efectuar un número par de pasos la cadena sólo puede terminar en una posición par, y si se efectúan un número impar de pasos la posición final sólo puede ser un número impar. Además, es claro que después de efectuar n pasos, la caminata sólo puede llegar a una distancia máxima de n unidades, a la izquierda o a la derecha. Teniendo esto en mente, en el siguiente resultado se presenta la distribución de probabilidad de la variable X_n .

Proposición. Para cualesquiera números enteros x y n tales que $-n \leq x \leq n$, y para el caso cuando x y n son ambos pares o ambos impares,

$$P(X_n = x | X_0 = 0) = \binom{n}{\frac{1}{2}(n+x)} p^{\frac{1}{2}(n+x)} q^{\frac{1}{2}(n-x)}. \quad (2.1)$$

Para valores de x y n que no cumplen las condiciones indicadas la probabilidad en cuestión vale cero.

Demostración. Suponga que se observa la posición de la caminata después de efectuar n pasos. Sean R_n y L_n el número de pasos realizados hacia la derecha y hacia la izquierda, respectivamente. Entonces $X_n = R_n - L_n$, y además $n = R_n + L_n$. Sumando estas dos ecuaciones y substituyendo la expresión $X_n = \sum_{i=1}^n \xi_i$ se obtiene

$$R_n = \frac{1}{2}(n + X_n) = \sum_{i=1}^n \frac{1}{2}(1 + \xi_i).$$

Esta ecuación es la identidad clave para obtener el resultado buscado. Observe que esta fórmula arroja un valor entero para R_n cuando n y X_n son ambos pares o ambos impares. Como las variables independientes ξ_i toman los valores $+1$ y -1 con probabilidades p y q respectivamente, entonces las variables independientes $\frac{1}{2}(1 + \xi_i)$ toman los valores 1 y 0 con probabilidades p y q . Esto lleva a la conclusión de que la variable R_n tiene distribución binomial(n, p). Por lo tanto, para cualquier valor de x que cumpla las condiciones enunciadas se tiene que

$$\begin{aligned} P(X_n = x \mid X_0 = 0) &= P(R_n = \frac{1}{2}(n + x)) \\ &= \binom{n}{\frac{1}{2}(n + x)} p^{\frac{1}{2}(n+x)} q^{\frac{1}{2}(n-x)}. \end{aligned}$$

□

En particular, cuando la caminata es simétrica, es decir, cuando $p = 1/2$, y con las mismas restricciones para n y x ($-n \leq x \leq n$, ambos pares o ambos impares) se tiene la expresión

$$P(X_n = x \mid X_0 = 0) = \binom{n}{\frac{1}{2}(n + x)} \frac{1}{2^n}. \quad (2.2)$$

Esta fórmula puede también justificarse mediante argumentos de análisis combinatorio de la siguiente forma: En total hay 2^n posibles trayectorias que la caminata puede seguir al efectuar n pasos, todas ellas con la misma probabilidad de ocurrir debido a la simetría. Ahora, ¿Cuántas de estas trayectorias terminan en $x \geq 0$, por ejemplo? Como se ha argumentado antes, el número de pasos a la derecha debe ser $\frac{1}{2}(n + x)$, y el número de trayectorias que cumplen la condición es el número de formas en que los $\frac{1}{2}(n + x)$ pasos a la derecha pueden escogerse de los n pasos totales. La respuesta es entonces el cociente que aparece en (2.2).

■ Ejercicio. Demuestre nuevamente la identidad (2.1) analizando ahora la variable L_n , es decir, demuestre primero las igualdades $L_n = (n - X_n)/2 = \sum_{i=1}^n (1 - \xi_i)/2$. Concluya que L_n tiene distribución binomial(n, q). A partir de aquí obtenga (2.1). ■

■ Ejercicio. Demuestre que la probabilidad (2.1) es una función simétrica de x si, y sólo si, la caminata es simétrica. ■

La fórmula (2.1) puede extenderse fácilmente al caso general de pasar de un estado cualquiera y a otro estado x en n pasos.

Proposición. Si los números n y $x - y$ son ambos pares o ambos impares, entonces para $-n \leq x - y \leq n$,

$$P(X_n = x | X_0 = y) = \binom{n}{\frac{1}{2}(n+x-y)} p^{\frac{1}{2}(n+x-y)} q^{\frac{1}{2}(n-x+y)}. \quad (2.3)$$

Para valores de $x - y$ y n que no cumplen las condiciones indicadas la probabilidad en cuestión vale cero.

Demostración. Tenemos como hipótesis que $X_0 = y$. Consideremos el proceso $Z_n = X_n - y$. Entonces $\{Z_n\}$ es ahora una caminata aleatoria que inicia en cero como en el caso antes demostrado. El resultado enunciado se obtiene de la identidad $P(X_n = x | X_0 = y) = P(Z_n = x - y | Z_n = 0)$. □

Probabilidad de regreso a la posición de origen. Nos plantearemos ahora el problema de encontrar la probabilidad de que una caminata aleatoria que inicia en el origen, regrese eventualmente al punto de partida. Demostraremos que en el caso asimétrico, $p \neq 1/2$, la probabilidad de tal evento es menor que uno, es decir, no es seguro que ello ocurra, pero en el caso simétrico, $p = 1/2$, se cumple que con probabilidad uno la caminata regresa eventualmente al origen. Para el último caso demostraremos además que el número de pasos promedio para regresar al origen es, sin embargo, infinito. La demostración es un tanto técnica y hace uso de las funciones generadoras de probabilidad. Dado que este capítulo es introductorio, tal vez sea mejor recomendar al lector, cuando se trate de una primera lectura, omitir los detalles de esta demostración.

Proposición. Para una caminata aleatoria sobre \mathbb{Z} , la probabilidad de un eventual regreso al punto de partida es

$$1 - |p - q| = \begin{cases} 1 & \text{si } p = q, \\ < 1 & \text{si } p \neq q. \end{cases}$$

Es decir, sólo en el caso simétrico, $p = q$, se tiene la certeza de un eventual retorno, sin embargo el tiempo promedio de regreso en tal caso es infinito.

Demostración. Para demostrar estas afirmaciones utilizaremos los siguientes elementos:

a) Para cada $n \geq 0$ denotaremos por p_n a la probabilidad de que la caminata se encuentre en el estado 0 al tiempo n , es decir, $p_n = P(X_n = 0 | X_0 = 0)$. Esta probabilidad es distinta de cero sólo cuando n es un número par. Naturalmente $p_0 = 1$. Denotaremos también por f_k a la probabilidad de que la caminata visite el estado 0 por primera vez en el paso $k \geq 0$. El uso de la letra f proviene el término en Inglés *first*. Por conveniencia se define $f_0 = 0$. Observe que en términos de las probabilidades f_k , la probabilidad de que la caminata regrese eventualmente al origen es $\sum_{k=0}^{\infty} f_k$. Esta serie es convergente pues se trata de la suma de probabilidades de eventos disjuntos, y por lo tanto a lo sumo vale uno. Demostraremos que en el caso simétrico la suma vale uno. Recordemos nuevamente que los valores de f_k son estrictamente positivos sólo para valores pares de k distintos de cero.

b) No es difícil comprobar que se cumple la siguiente igualdad

$$p_n = \sum_{k=0}^n f_k p_{n-k}. \quad (2.4)$$

En esta expresión simplemente se descompone la probabilidad de regreso al origen, p_n , en las distintas posibilidades en donde se efectúa el primer regreso al origen. Este primero regreso puede darse en el paso 1, o en el paso 2, ..., o en el último momento, el paso n . Después de efectuado el primer regreso se multiplica por la probabilidad de regresar al origen en el número de pasos restantes. Observe que el primer sumando es cero. Esta fórmula será demostrada más adelante en el contexto de las cadenas de Markov, véase la fórmula (3.2) en la página 44. Usaremos (2.4) para encontrar la función generadora de probabilidad de la colección de números

f_0, f_1, f_2, \dots , es decir, encontraremos

$$G(t) = \sum_{k=0}^{\infty} f_k t^k.$$

Multiplicando (2.4) por t^n , sumando y cambiando el orden de las sumas se obtiene

$$\begin{aligned} \sum_{n=1}^{\infty} p_n t^n &= \sum_{n=1}^{\infty} \left(\sum_{k=0}^n f_k p_{n-k} \right) t^n \\ &= \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} f_k p_{n-k} t^n \\ &= \sum_{k=0}^{\infty} f_k t^k \sum_{n=k}^{\infty} p_{n-k} t^{n-k} \\ &= G(t) \sum_{n=0}^{\infty} p_n t^n. \end{aligned}$$

Por lo tanto

$$\left(\sum_{n=0}^{\infty} p_n t^n \right) - 1 = G(t) \sum_{n=0}^{\infty} p_n t^n. \quad (2.5)$$

Para encontrar $G(t)$ se necesita encontrar una expresión para la suma que aparece en la última ecuación, y que no es otra cosa sino el función generadora de los números p_0, p_1, p_2, \dots . Haremos esto a continuación.

c) Para el análisis que sigue necesitamos recordar que para cualquier número real a y para cualquier entero n , se tiene el coeficiente binomial

$$\binom{a}{n} = \frac{a(a-1)\cdots(a-(n-1))}{n!}. \quad (2.6)$$

Observe que en el numerador hay n factores. Este número es una generalización del coeficiente binomial usual y aparece en la siguiente expansión binomial infinita válida para $|t| < 1$,

$$(1+t)^a = \sum_{n=0}^{\infty} \binom{a}{n} t^n. \quad (2.7)$$

En particular y como una aplicación de (2.6) se tiene que

$$\begin{aligned}
\binom{2n}{n} &= \frac{2n(2n-1)(2n-2)\cdots 3\cdot 2\cdot 1}{n!n!} \\
&= \frac{2^n n! (2n-1)(2n-3)\cdots 5\cdot 3\cdot 1}{n!n!} \\
&= \frac{2^n 2^n}{n!} \left(\frac{2n-1}{2}\right)\left(\frac{2n-3}{2}\right)\cdots\left(\frac{5}{2}\right)\left(\frac{3}{2}\right)\left(\frac{1}{2}\right) \\
&= \frac{2^n 2^n}{n!} (-1)^n \left(-n + \frac{1}{2}\right)\left(-n + \frac{3}{2}\right)\cdots\left(-\frac{5}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{1}{2}\right) \\
&= (-4)^n \binom{-1/2}{n}.
\end{aligned} \tag{2.8}$$

d) Usando (2.7) y (2.8) podemos ahora encontrar una expresión cerrada para la función generadora de la colección de números p_0, p_1, p_2, \dots

$$\begin{aligned}
\sum_{n=0}^{\infty} p_n t^n &= \sum_{n=0}^{\infty} \binom{2n}{n} p^n q^n t^{2n} \\
&= \sum_{n=0}^{\infty} (-4)^n \binom{-1/2}{n} p^n q^n t^{2n} \\
&= \sum_{n=0}^{\infty} \binom{-1/2}{n} (-4pqt^2)^n \\
&= (1 - 4pqt^2)^{-1/2}.
\end{aligned} \tag{2.9}$$

e) Substituyendo (2.9) en (2.5) se llega a la igualdad

$$(1 - 4pqt^2)^{-1/2} - 1 = G(t) (1 - 4pqt^2)^{-1/2}.$$

De donde se obtiene finalmente

$$G(t) = 1 - (1 - 4pqt^2)^{1/2}. \tag{2.10}$$

Usando esta expresión podemos ahora calcular la probabilidad de un eventual regreso al estado inicial. En efecto, por el lema de Abel,

$$\sum_{n=0}^{\infty} f_n = \lim_{t \nearrow 1} G(t) = 1 - (1 - 4pq)^{1/2} = 1 - |p - q|.$$

En el caso asimétrico, $p \neq 1/2$, esta cantidad es estrictamente menor a uno y por lo tanto no hay seguridad de que la caminata sea capaz de regresar al origen. En cambio, en el caso simétrico, $p = 1/2$, esta cantidad vale uno, es decir, con probabilidad uno la cadena aleatoria simétrica regresa eventualmente a su posición de origen. Además, el tiempo promedio de regreso se obtiene derivando la función generadora $G(t) = 1 - (1 - t^2)^{1/2}$, es decir,

$$\sum_{n=0}^{\infty} n f_n = \lim_{t \nearrow 1} G'(t) = \lim_{t \nearrow 1} \frac{t}{\sqrt{1 - t^2}} = \infty.$$

□

Puede considerarse también el caso de una caminata en donde sea posible permanecer en el mismo estado después de una unidad de tiempo. Esta situación se ilustra en la Figura 2.3(a), en donde p , q y r son probabilidades tales que $p + q + r = 1$. También pueden considerarse caminatas aleatorias en \mathbb{Z}^2 como la que se muestra en la Figura 2.3(b), en donde $p + q + r + s = 1$, o más generalmente en \mathbb{Z}^n o cualquier otro conjunto reticulado.

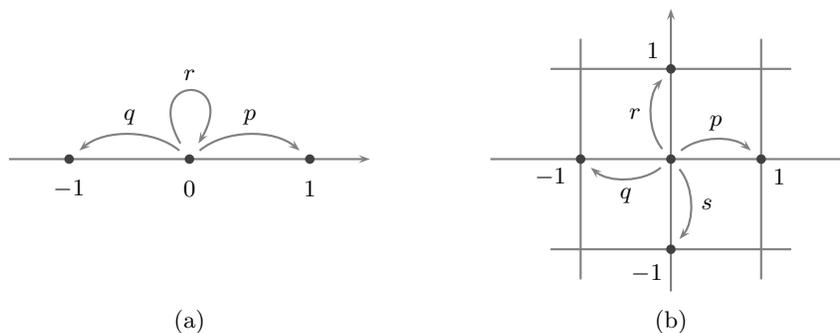


Figura 2.3:

En el siguiente capítulo se estudiarán algunas otras propiedades de las caminatas aleatorias en el contexto de las cadenas de Markov.

2.2. El problema del jugador

En esta sección consideraremos un ejemplo particular de una caminata aleatoria puesta en el contexto de un juego de apuestas.

Planteamiento del problema. Suponga que un jugador A apuesta sucesivamente una unidad monetaria a un jugador B . Inicialmente el jugador A tiene k unidades, y B tiene $N - k$ unidades, es decir, el capital conjunto entre los dos jugadores es de N unidades monetarias. En cada apuesta el jugador A tiene probabilidad de ganar p , y probabilidad de perder $q = 1 - p$, se asume además que no hay empates. Sea X_n la fortuna del jugador A al tiempo n . Entonces $\{X_n : n = 0, 1, \dots\}$ es una caminata aleatoria que inicia en el estado k y eventualmente puede terminar en el estado 0 cuando el jugador A ha perdido todo su capital, o bien puede terminar en el estado N que corresponde a la situación en donde el jugador A ha ganado todo el capital.

Este proceso es entonces una caminata aleatoria sobre el conjunto $\{0, 1, \dots, N\}$, en donde los estados 0 y N son absorbentes, pues una vez que la cadena llega a alguno de ellos, jamás lo abandona. Una posible trayectoria cuando la caminata se absorbe en el estado 0 se muestra en la Figura 2.4.

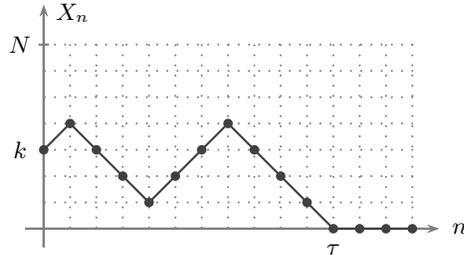


Figura 2.4:

Una de las preguntas que resolveremos para esta caminata es la siguiente: ¿Cuál es la probabilidad de que eventualmente el jugador A se arruine? Es

decir, ¿Cuál es la probabilidad de que la caminata se absorba en el estado 0 y no en el estado N , u oscile entre estos dos estados? Este problema se conoce como el *problema de la ruina del jugador*, y encontraremos a continuación su solución. Usando probabilidad condicional es posible transformar este problema en resolver una ecuación en diferencias.

Solución al problema. Sea τ es el primer momento en el que la caminata visita alguno de los dos estados absorbentes, es decir, $\tau = \min\{n \geq 0 : X_n = 0 \text{ ó } X_n = N\}$. Puede demostrarse que esta variable aleatoria es finita casi seguramente. La

pregunta planteada se traduce en encontrar la probabilidad

$$u_k = P(X_\tau = 0 \mid X_0 = k).$$

Por el teorema de probabilidad total se obtiene la ecuación en diferencias

$$u_k = p u_{k+1} + q u_{k-1}, \quad (2.11)$$

válida para $k = 1, 2, \dots, N - 1$. La interpretación intuitiva de esta identidad es sencilla: A partir del estado k se busca la probabilidad de ruina analizando lo que sucede en la siguiente apuesta. Se tienen dos casos. El jugador gana con probabilidad p y ahora se busca la probabilidad de ruina a partir del estado $k + 1$, o bien el jugador pierde con probabilidad q y se busca la probabilidad de ruina ahora a partir del estado $k - 1$. Las condiciones de frontera son $u_0 = 1$ y $u_N = 0$. Esta ecuación es una ecuación en diferencias, lineal, de segundo orden y homogénea. Puede encontrarse su solución de la siguiente forma: Multiplicando el lado izquierdo de (2.11) por $(p + q)$ y agrupando términos se llega a la expresión equivalente

$$u_{k+1} - u_k = (q/p)(u_k - u_{k-1}). \quad (2.12)$$

Resolviendo iterativamente, las primeras $k - 1$ ecuaciones se escriben de la forma siguiente

$$\begin{aligned} u_2 - u_1 &= (q/p)(u_1 - 1) \\ u_3 - u_2 &= (q/p)^2(u_1 - 1) \\ &\vdots \\ u_k - u_{k-1} &= (q/p)^{k-1}(u_1 - 1). \end{aligned}$$

Hemos usado aquí la condición de frontera $u_0 = 1$. Conviene ahora definir $S_k = 1 + (q/p) + \dots + (q/p)^k$ pues al sumar las $k - 1$ ecuaciones anteriores se obtiene $u_k - u_1 = (S_{k-1} - 1)(u_1 - 1)$. O bien

$$u_k = 1 + S_{k-1}(u_1 - 1). \quad (2.13)$$

De manera análoga pero ahora sumando todas las ecuaciones de (2.12) se obtiene $u_N = 1 + S_{N-1}(u_1 - 1)$. Usando la segunda condición de frontera $u_N = 0$ se llega a $u_1 - 1 = -1/S_{N-1}$. Substituyendo en (2.13) y simplificando se llega a la solución

$$u_k = 1 - \frac{S_{k-1}}{S_{N-1}}.$$

Es necesario ahora distinguir los siguientes dos casos:

$$S_k = 1 + (q/p) + \cdots + (q/p)^k = \begin{cases} k + 1 & \text{si } p = q, \\ \frac{1 - (q/p)^{k+1}}{1 - (q/p)} & \text{si } p \neq q. \end{cases}$$

Por lo tanto

$$u_k = \begin{cases} (N - k)/N & \text{si } p = 1/2, \\ \frac{(q/p)^k - (q/p)^N}{1 - (q/p)^N} & \text{si } p \neq 1/2. \end{cases} \quad (2.14)$$

En la Figura 2.5 se muestra la gráfica de la probabilidad u_k como función del parámetro k para varios valores de p y con $N = 50$. En el caso simétrico la solución es la línea recta que une la probabilidad uno con la probabilidad cero. Naturalmente la probabilidad de ruina decrece cuando el capital inicial k aumenta. En la sección de ejercicios se sugiere otra forma de resolver la ecuación en diferencias (2.11).

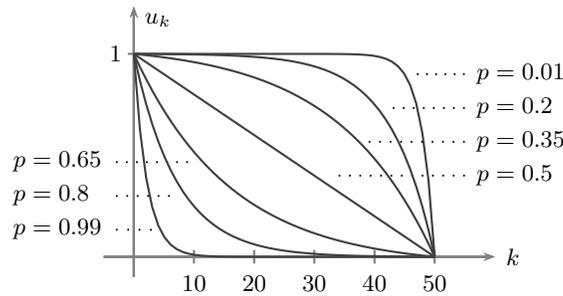


Figura 2.5:

Análisis de la solución. Es interesante analizar la fórmula (2.14) en sus varios aspectos. Por ejemplo, para el caso simétrico $p = 1/2$, es decir, cuando el juego es justo, la probabilidad de ruina es muy cercana a 1 cuando k es muy pequeño comparado con N . Esto sugiere que no conviene jugar esta serie de apuestas contra adversarios demasiado ricos, aún cuando el juego sea justo. Es también un tanto inesperado observar que la probabilidad u_k es muy sensible a los valores de p cercanos a $1/2$. Esto puede apreciarse en la Figura 2.6. Cuando p es distante de $1/2$ la probabilidad u_k es casi constante, pero para valores de p cercanos a $1/2$

la probabilidad cambia rápidamente de un extremo a otro. Estas gráficas fueron elaboradas tomando $N = 50$.

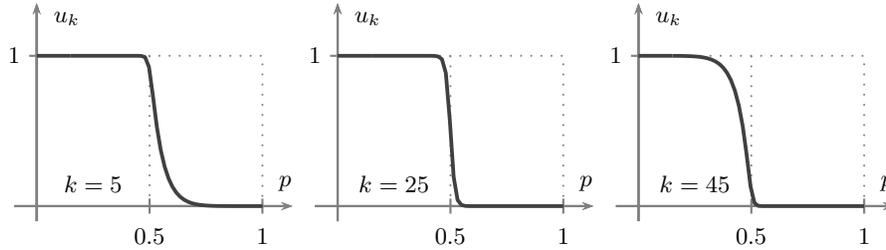


Figura 2.6:

■ Ejercicio. (Probabilidad de ruina del otro jugador). Si el juego se considera desde el punto de vista del jugador B , entonces se trata de la misma caminata aleatoria sólo que ahora el capital inicial es $N - k$ y la probabilidad de ganar en cada apuesta es q . Substituya estos parámetros en la solución (2.14) y compruebe que la probabilidad de ruina del jugador B , denotada por v_{N-k} , es la que aparece abajo. Verifique además que $u_k + v_{N-k} = 1$, es decir, la probabilidad de que eventualmente el juego termine con la ruina de alguno de los jugadores es uno.

$$v_{N-k} = \begin{cases} k/N & \text{si } q = 1/2, \\ \frac{(q/p)^k - 1}{(q/p)^N - 1} & \text{si } q \neq 1/2. \end{cases}$$

■

Número esperado de apuestas antes de la ruina. Hemos comprobado en el ejercicio anterior que con probabilidad uno ocurre que eventualmente alguno de los dos jugadores se arruina. Es natural entonces plantearse el problema de encontrar el tiempo promedio que transcurre antes de observar la eventual ruina de alguna de las partes. Este número de apuestas promedio antes de la ruina puede encontrarse explícitamente, y el método que usaremos para encontrarlo es nuevamente el planteamiento de una ecuación en diferencias que resolveremos del modo mostrado antes. Sea entonces m_k el número esperado de apuesta antes de que termine el juego, en donde el jugador A tiene un capital inicial de k unidades, y B tiene $N - k$.

Proposición. El número esperado de apuestas antes de la ruina es

$$m_k = \begin{cases} k(N-k) & \text{si } p = q, \\ \frac{1}{q-p} \left(k - N \frac{1 - (q/p)^k}{1 - (q/p)^N} \right) & \text{si } p \neq q. \end{cases}$$

Demostración. Condicionando sobre el resultado de la primera apuesta se obtiene que m_k satisface la ecuación

$$m_k = 1 + p m_{k+1} + q m_{k-1},$$

válida para $k = 1, 2, \dots, N-1$. Esta es una ecuación en diferencias, de segundo orden, lineal y no homogénea. Las condiciones de frontera son ahora $m_0 = 0$ y $m_N = 0$. Substituyendo $(p+q)m_k$ por m_k y agrupando términos convenientemente la ecuación anterior puede reescribirse del siguiente modo

$$m_{k+1} - m_k = \frac{q}{p}(m_k - m_{k-1}) - \frac{1}{p}. \quad (2.15)$$

Recordando la notación $S_k = 1 + (q/p) + \dots + (q/p)^k$, y substituyendo iterativamente, las primeras $k-1$ ecuaciones son

$$\begin{aligned} m_2 - m_1 &= (q/p)m_1 - \frac{1}{p}S_0, \\ m_3 - m_2 &= (q/p)^2 m_1 - \frac{1}{p}S_1, \\ &\vdots \\ m_k - m_{k-1} &= (q/p)^{k-1} m_1 - \frac{1}{p}S_{k-2}. \end{aligned}$$

Aquí se ha hecho uso de la condición de frontera $m_0 = 0$. Sumando todas estas ecuaciones se obtiene

$$m_k - m_1 = m_1(S_{k-1} - 1) - \frac{1}{p} \sum_{i=0}^{k-2} S_i.$$

Es decir,

$$m_k = m_1 S_{k-1} - \frac{1}{p} \sum_{i=0}^{k-2} S_i. \quad (2.16)$$

En particular, sumando todas las ecuaciones de (2.15) se obtiene

$$m_N = m_1 S_{N-1} - \frac{1}{p} \sum_{i=0}^{N-2} S_i.$$

Ahora se hace uso de la condición $m_N = 0$, y se obtiene

$$m_1 = \frac{1}{S_{N-1}} \frac{1}{p} \sum_{i=0}^{k-2} S_i.$$

Substituyendo en (2.16),

$$m_k = \frac{S_{k-1}}{S_{N-1}} \frac{1}{p} \sum_{i=0}^{N-2} S_i - \frac{1}{p} \sum_{i=0}^{k-2} S_i. \quad (2.17)$$

Nuevamente se deben distinguir los siguientes dos casos:

$$S_k = 1 + (q/p) + \dots + (q/p)^k = \begin{cases} k + 1 & \text{si } p = q, \\ \frac{1 - (q/p)^{k+1}}{1 - (q/p)} & \text{si } p \neq q. \end{cases}$$

Substituyendo en (2.17) y simplificando, después de algunos cálculos se llega a la respuesta enunciada. \square

La forma en la que m_k cambia al variar el parámetro k se muestra en la Figura 2.7 cuando $N = 50$. La duración máxima promedio del juego se obtiene cuando el juego es justo, $p = 1/2$, y ambos jugadores tienen el mismo capital inicial, es decir, $k = N/2$. Esto arroja un promedio máximo de $(N/2)(N - N/2) = (N/2)^2$ apuestas antes del fin del juego. En el caso ilustrado, la duración máxima promedio del juego es de $(50/2)^2 = 625$ apuestas.

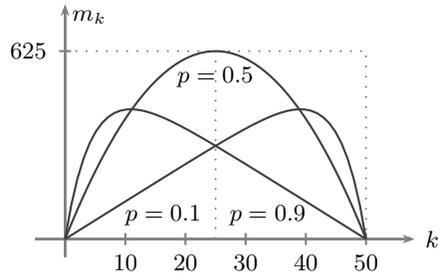


Figura 2.7:

■ Ejercicio. Demuestre que la duración promedio del juego es siempre menor o igual a $(N/2)^2$, es decir, para cada $k = 1, \dots, N$, se cumple que $m_k \leq (N/2)^2$,

tanto en el caso simétrico como no simétrico. ■

Notas y referencias. El tema de caminatas aleatorias puede ser encontrado en la mayoría de los textos sobre procesos estocásticos, ya sea de una manera explícita o como un ejemplo importante de cadena de Markov. En particular el libro de Jones y Smith [17], y el de Basu [1] contienen un capítulo completo sobre el tema. Como lectura más avanzada véase el texto de Resnick [28], y el de Spitzer [34].

Ejercicios

Caminatas aleatorias

1. Demuestre que una caminata aleatoria simple sobre \mathbb{Z} cumple la propiedad de Markov, es decir, demuestre que para cualesquiera enteros x_0, x_1, \dots, x_{n+1} ,

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n).$$

2. Para una caminata aleatoria simple $\{X_n\}$ sobre \mathbb{Z} demuestre que

$$P(X_{n+1} = x) = pP(X_n = x - 1) + qP(X_n = x + 1).$$

3. Una partícula realiza una caminata aleatoria simétrica sobre \mathbb{Z} empezando en cero. Encuentre la probabilidad de que la partícula se encuentre nuevamente en el origen en el sexto paso.
4. Una partícula realiza una caminata aleatoria simétrica sobre \mathbb{Z} empezando en cero. ¿Cuál es la probabilidad de que la partícula regrese al estado cero por primera vez en el sexto paso?
5. Considere una caminata aleatoria simple sobre \mathbb{Z} que empieza en cero y tal que la probabilidad de pasar al estado de la derecha es p y la probabilidad de pasar al estado de la izquierda es $q = 1 - p$. Sea τ_n el primer momento en el que la caminata visita el estado $n \geq 1$. Usando análisis del primer paso (condicionando sobre si primer paso se efectúa a la izquierda o a la derecha) demuestre que

$$P(\tau_n < \infty) = \begin{cases} (p/q)^n & \text{si } p < 1/2, \\ 1 & \text{si } p \geq 1/2. \end{cases}$$

En particular, la probabilidad de que la caminata se mantenga siempre en el conjunto $\{\dots, -1, 0, 1, \dots, n - 1\}$ es $1 - (p/q)^n$ en el caso $p < 1/2$, y con probabilidad uno llegará al estado n en el caso $p \geq 1/2$. Generalice la fórmula anterior en el caso cuando el estado inicial es m , menor o mayor a n .

6. Un algoritmo aleatorio de búsqueda del estado cero opera del siguiente modo: si se encuentra en el estado k en algún momento, entonces el estado del algoritmo al siguiente paso es aleatorio con distribución uniforme en el conjunto $\{0, 1, \dots, k - 1\}$. Encuentre el número esperado de pasos en los que el algoritmo alcanza el estado cero cuando inicia en el estado k .

El problema de la ruina del jugador

7. Siguiendo la notación usada en el problema de la ruina del jugador, demuestre la validez de la ecuación $u_k = p u_{k+1} + q u_{k-1}$, para $k = 1, 2, \dots, N - 1$.
8. Para resolver el problema del jugador se requiere resolver la ecuación en diferencias $u_k = p u_{k+1} + q u_{k-1}$. Resuelva nuevamente esta ecuación siguiendo los siguientes pasos [18]:
 - a) Proponga la solución $u_k = a m^k$, con a y m constantes distintas de cero, y que encontraremos a continuación.
 - b) Substituya la solución propuesta en la ecuación en diferencias y encuentre la ecuación cuadrática $pm^2 - m + q = 0$. Esta ecuación se conoce como la ecuación característica de la ecuación en diferencias.
 - c) Suponiendo el caso $p \neq q$, esta ecuación cuadrática tiene dos soluciones distintas: $m_1 = 1$ y $m_2 = q/p$. Dada la linealidad de la ecuación en diferencias, la solución general puede escribirse como $u_k = a_1 m_1^k + a_2 m_2^k = a_1 + a_2 (q/p)^k$.
 - d) Use las condiciones de frontera $u_0 = 1$ y $u_N = 0$ para encontrar los valores de las constantes a_1 y a_2 , y obtener nuevamente la solución (2.14).
 - e) Cuando $p = q$ la ecuación característica tiene una única raíz: $m_1 = 1$. Como segundo valor para m se propone k veces el valor de la primera solución, es decir, $m_2 = k$. Proceda como en el inciso anterior para encontrar (2.14) en este caso.
9. Siguiendo la notación usada para encontrar el número esperado de apuestas antes de la ruina en el problema del jugador, demuestre la igualdad $m_k = 1 + p m_{k+1} + q m_{k-1}$, para $k = 1, 2, \dots, N - 1$.
10. Considere el problema de la ruina del jugador permitiendo ahora que existan empates en cada una de las apuestas. Las probabilidades de ganar, perder o empatar para el jugador A son p , q y r , respectivamente, con $p + q + r = 1$. Demuestre que la probabilidad de ruina para el jugador A sigue siendo la misma expresión (2.14). Es decir, la posibilidad de empate en cada apuesta extiende posiblemente la duración del juego pero no modifica las probabilidades de ruina.

CAPÍTULO 3

Cadenas de Markov

Las cadenas de Markov fueron introducidas por el matemático ruso Andrey Markov alrededor de 1905. Su intención era crear un modelo probabilístico para analizar la frecuencia con la que aparecen las vocales en poemas y textos literarios. El éxito del modelo propuesto por Markov radica en que es lo suficientemente complejo como para describir ciertas características no triviales de algunos sistemas, pero al mismo tiempo es lo suficientemente sencillo para ser analizado matemáticamente. Las cadenas de Markov pueden aplicarse a una amplia gama de fenómenos científicos y sociales, y se cuenta con una teoría matemática extensa al respecto. En este capítulo presentaremos una introducción a algunos aspectos básicos de este modelo.



Andrey Markov
(Rusia 1856-1922)

3.1. Propiedad de Markov

Vamos a considerar procesos estocásticos a tiempo discreto $\{X_n\}$ que cumplen la propiedad de Markov. Para describir a esta propiedad y varias de sus condiciones equivalentes de una manera simple, a la probabilidad $P(X_n = x_n)$ se le escribirá como $p(x_n)$, es decir, el subíndice indica también la variable a la que se hace referencia. El significado de la probabilidad condicional $p(x_{n+1} | x_n)$ es análogo.

Definición. Una *cadena de Markov* es un proceso estocástico a tiempo discreto $\{X_n : n = 0, 1, \dots\}$, con espacio de estados discreto, y que satisface la propiedad de Markov, esto es, para cualquier entero $n \geq 0$, y para cualesquiera estados x_0, \dots, x_{n+1} , se cumple

$$p(x_{n+1} | x_0, \dots, x_n) = p(x_{n+1} | x_n). \quad (3.1)$$

Si el tiempo $n+1$ se considera como un tiempo futuro, el tiempo n como el presente y los tiempos $0, 1, \dots, n-1$ como el pasado, entonces la condición (3.1) establece que la distribución de probabilidad del estado del proceso al tiempo futuro $n+1$ depende únicamente del estado del proceso al tiempo n , y no depende de los estados en los tiempos pasados $0, 1, \dots, n-1$. Existen otras formas equivalentes de expresar esta propiedad, algunas de ellas se encuentran enunciadas en la sección de ejercicios. Por ejemplo, es posible demostrar que la condición (3.1) es equivalente a poder calcular la distribución conjunta de las variables X_0, X_1, \dots, X_n de la siguiente forma

$$p(x_0, x_1, \dots, x_n) = p(x_0)p(x_1 | x_0) \cdots p(x_n | x_{n-1}).$$

Sin pérdida de generalidad tomaremos como espacio de estados de una cadena de Markov al conjunto discreto $\{0, 1, 2, \dots\}$, o cualquier subconjunto finito que conste de los primeros elementos de este conjunto. Cuando el espacio de estados de una cadena de Markov es un conjunto finito se dice que la cadena es *finita*.

Probabilidades de transición. A la probabilidad $P(X_{n+1} = j | X_n = i)$ se le denota por $p_{ij}(n, n+1)$, y representa la probabilidad de transición del estado i en el tiempo n , al estado j en el tiempo $n+1$. Estas probabilidades se conocen como las *probabilidades de transición en un paso*. Cuando los números $p_{ij}(n, n+1)$ no dependen de n se dice que la cadena es *estacionaria* u *homogénea en el tiempo*. Por

simplicidad se asume tal situación de modo que las probabilidades de transición en un paso se escriben simplemente como p_{ij} . Variando los índices i y j , por ejemplo sobre el conjunto de estados $\{0, 1, 2, \dots\}$, se obtiene la matriz de probabilidades de transición en un paso que aparece en la Figura 3.1. La entrada (i, j) de esta matriz es la probabilidad de transición p_{ij} , es decir, la probabilidad de pasar del estado i al

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \cdots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} & \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \end{matrix}$$

Figura 3.1:

estado j en una unidad de tiempo. En general, al escribir estas matrices omitiremos escribir la identificación de los estados en los renglones y columnas como aparece en la Figura 3.1, tal identificación será evidente a partir de conocer el espacio de estados del proceso. El índice i se refiere al renglón de la matriz, y el índice j a la columna.

Proposición. La matriz de probabilidades de transición $P = (p_{ij})$ cumple las siguientes dos propiedades.

a) $p_{ij} \geq 0$.

b) $\sum_j p_{ij} = 1$.

Demostración. La primera condición es evidente a partir del hecho de que estos números son probabilidades. Para la segunda propiedad se tiene que para cualquier estado i y cualquier entero $n \geq 0$,

$$\begin{aligned} 1 &= P(X_{n+1} \in \{0, 1, \dots\}) \\ &= P(X_{n+1} \in \{0, 1, \dots\} | X_n = i) \\ &= P\left(\bigcup_j (X_{n+1} = j) | X_n = i\right) \\ &= \sum_j P(X_{n+1} = j | X_n = i) \\ &= \sum_j p_{ij}. \end{aligned}$$

□

Esto último significa que a partir de cualquier estado i con probabilidad uno la cadena pasa necesariamente a algún elemento del espacio de estados al siguiente momento. En general toda matriz cuadrada que cumpla estas dos propiedades se dice que es una *matriz estocástica*. Debido a la propiedad de Markov, esta matriz captura la esencia del proceso y determina el comportamiento de la cadena en cualquier tiempo futuro. Si además la matriz satisface la condición $\sum_i p_{ij} = 1$, es decir, cuando la suma por columnas también es uno, entonces se dice que es *doblemente estocástica*.

Distribución de probabilidad inicial. En general puede considerarse que una cadena de Markov inicia su evolución partiendo de un estado i cualquiera, o más generalmente considerando una distribución de probabilidad inicial sobre el espacio de estados. Una distribución inicial para una cadena de Markov con espacio de estados $\{0, 1, 2, \dots\}$ es simplemente una distribución de probabilidad sobre este conjunto, es decir, es una colección de números $p_0, p_1, p_2 \dots$ que son no negativos y que suman uno. El número p_i corresponde a la probabilidad de que la cadena inicie en el estado i . En general, la distribución inicial juega un papel secundario en el estudio de las cadenas de Markov.

Existencia. Hemos mencionado que la propiedad de Markov es equivalente a la igualdad $p(x_0, x_1, \dots, x_n) = p(x_0)p(x_1 | x_0) \cdots p(x_n | x_{n-1})$. Esta identidad establece que las distribuciones conjuntas $p(x_0, x_1, \dots, x_n)$ se encuentran completamente especificadas por la matriz de probabilidades de transición y una distribución inicial. En el texto de Chung [7] puede encontrarse una demostración del hecho de que dada una matriz estocástica y una distribución de probabilidad inicial, existe un espacio de probabilidad y una cadena de Markov con matriz de probabilidades de transición y distribución inicial las especificadas. Es por ello que a la matriz misma se le llama a veces cadena de Markov.

■ Ejercicio. Sea $\{X_n\}$ una cadena de Markov con tres estados: 0,1,2, y con matriz de probabilidades de transición como aparece abajo. Calcule $P(X_2 = 0, X_1 = 0 | X_0 = 1)$ y $P(X_3 = 1, X_2 = 1 | X_1 = 2)$.

$$P = \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.2 & 0.5 \\ 0.7 & 0 & 0.3 \end{pmatrix}.$$

■

■ Ejercicio. Sea $\{X_n\}$ una cadena de Markov con dos estados 0 y 1, con distribución inicial $P(X_0 = 0) = 0.5$, $P(X_0 = 1) = 0.5$, y con matriz de probabilidades de transición como aparece abajo. Calcule $P(X_0 = 0, X_1 = 1, X_2 = 0)$, $P(X_0 = 0 | X_1 = 1)$, $P(X_1 = 0)$, $P(X_2 = 1)$ y $P(X_0 = 0, X_1 = 0, X_2 = 1, X_3 = 1)$.

$$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.9 & 0.1 \end{pmatrix}.$$

■

■ Ejercicio. Sea $\{X_n\}$ una cadena de Markov con dos estados: 0,1, con distribución inicial $P(X_0 = 0) = 0.2$, $P(X_0 = 1) = 0.8$, y con matriz de probabilidades de transición como aparece abajo. Encuentre la distribución de las variables X_1 , X_2 , X_1 y X_2 conjuntamente.

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 2/3 & 1/3 \end{pmatrix}.$$

Probabilidades de transición en n pasos. La probabilidad $P(X_{n+m} = j | X_m = i)$ corresponde a la probabilidad de pasar del estado i al tiempo m , al estado j al tiempo $m + n$. Dado que hemos supuesto la condición de homogeneidad en el tiempo, esta probabilidad no depende realmente de m , por lo tanto coincide con $P(X_n = j | X_0 = i)$, y se le denota por $p_{ij}(n)$. A esta probabilidad también se le denota como $p_{ij}^{(n)}$, en donde el número de pasos n se escribe entre paréntesis para distinguirlo de algún posible exponente, y se le llama *probabilidad de transición en n pasos*. Cuando $n = 1$ simplemente se omite su escritura, a menos que se quiera hacer énfasis en ello. También cuando $n = 0$ es natural definir

$$p_{ij}(0) = \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases}$$

Es decir, después de realizar cero pasos la cadena no puede estar en otro estado mas que en su lugar de origen. Esta es la función *delta de Kronecker*.

3.2. Ejemplos

Revisaremos a continuación algunos ejemplos particulares clásicos de cadenas de Markov. Retomaremos más adelante varios de estos ejemplos para ilustrar otros conceptos y resultados generales de la teoría general.

Cadena de dos estados. Considere una cadena de Markov con espacio de estados $\{0, 1\}$, con matriz y diagrama de transición como aparece en la Figura 3.2, en donde $0 \leq a \leq 1$, y $0 \leq b \leq 1$. Suponga que la distribución inicial está dada por $p_0 = P(X_0 = 0)$ y $p_1 = P(X_0 = 1)$.

Aunque de aspecto sencillo, esta cadena es susceptible de muchas aplicaciones pues es común encontrar situaciones en donde se presenta siempre la dualidad de ser o

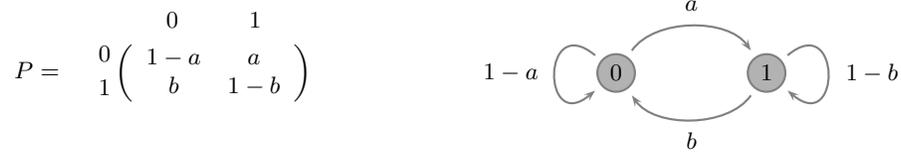


Figura 3.2:

no ser, estar o no estar, tener o no tener, siempre en una constante alternancia entre un estado y el otro. Cuando $a = 1 - b$, las variables X_1, X_2, \dots son independientes e idénticamente distribuidas con $P(X_n = 0) = 1 - a$ y $P(X_n = 1) = a$, para cada $n \geq 1$. Cuando $a \neq 1 - b$, X_n depende de X_{n-1} . Mas adelante demostraremos que las probabilidades de transición en n pasos son, para $a + b > 0$,

$$\begin{pmatrix} p_{00}(n) & p_{01}(n) \\ p_{10}(n) & p_{11}(n) \end{pmatrix} = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{(1-a-b)^n}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}.$$

■ Ejercicio. Para la cadena de Markov de dos estados, compruebe directamente que

- $P(X_0 = 0, X_1 = 1, X_2 = 0) = ab p_0$.
- $P(X_0 = 0, X_2 = 1) = (ab + (1-a)a) p_0$.
- $P(X_2 = 0) = ((1-a)^2 + ab) p_0 + ((1-b)b + b(1-a)) p_1$.

Cadena de variables aleatorias independientes. Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes con valores en el conjunto $\{0, 1, \dots\}$, y con idéntica distribución dada por las probabilidades a_0, a_1, \dots . Definiremos varias cadenas de Markov a partir de esta sucesión.

- La sucesión $X_n = \xi_n$ es una cadena de Markov con probabilidades de transición $p_{ij} = P(X_n = j | X_{n-1} = i) = P(X_n = j) = a_j$, es decir, la matriz de

probabilidades de transición es de la siguiente forma

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots \\ a_0 & a_1 & a_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Esta cadena tiene la cualidad de poder pasar a un estado cualquiera siempre con la misma probabilidad, sin importar el estado de partida. Puede modelar, por ejemplo, una sucesión de lanzamientos de una moneda.

b) El proceso $X_n = \max\{\xi_1, \dots, \xi_n\}$ es una cadena de Markov con matriz

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 + a_1 & a_2 & a_3 & \cdots \\ 0 & 0 & a_0 + a_1 + a_2 & a_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

c) El proceso $X_n = \xi_1 + \dots + \xi_n$ es una cadena de Markov con matriz

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots \\ 0 & a_0 & a_1 & \cdots \\ 0 & 0 & a_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

■ Ejercicio. ¿Es el proceso $X_n = \min\{\xi_1, \dots, \xi_n\}$ una cadena de Markov? En caso afirmativo encuentre la matriz de probabilidades de transición. ■

Cadena de rachas de éxitos. Sea ξ_1, ξ_2, \dots una sucesión de ensayos independientes Bernoulli con probabilidad de éxito p , y probabilidad de fracaso $q = 1 - p$. Sea X_n el número de éxitos consecutivos previos al tiempo n , incluyendo el tiempo n . Se dice que una *racha de éxitos* de longitud r ocurre al tiempo n si en el ensayo $n - r$ se obtiene un fracaso y los resultados de los ensayos $n - r + 1$ al n son todos éxitos. Gráficamente esta situación se ilustra en la Figura 3.3.

La colección de variables aleatorias $\{X_n : n = 1, 2, \dots\}$ es una cadena de Markov con espacio de estados $\{0, 1, \dots\}$. Las probabilidades de transición y la matriz correspondiente se muestran en la Figura 3.4.

Las posibles transiciones de un estado a otro para esta cadena de Markov se pueden observar en la Figura 3.5.

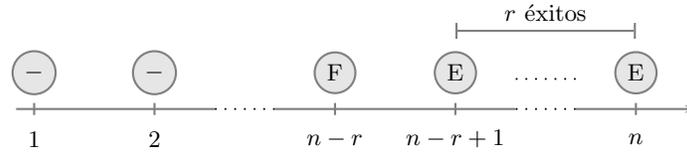


Figura 3.3:

$$p_{ij} = \begin{cases} p & \text{si } j = i + 1, \\ q & \text{si } j = 0, \\ 0 & \text{otro caso.} \end{cases} \quad P = \begin{matrix} & 0 & 1 & 2 & 3 & \dots \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} & \begin{pmatrix} q & p & 0 & 0 & \dots \\ q & 0 & p & 0 & \dots \\ q & 0 & 0 & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{matrix}.$$

Figura 3.4:

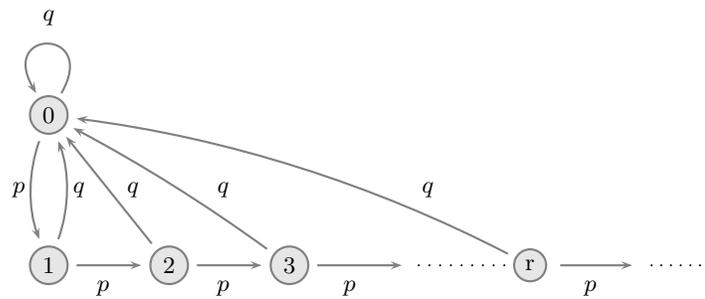


Figura 3.5:

Cadena de la caminata aleatoria. Una caminata aleatoria simple sobre el conjunto de números enteros constituye una cadena de Markov con espacio de estados el conjunto \mathbb{Z} , y con probabilidades de transición

$$p_{ij} = \begin{cases} p & \text{si } j = i + 1, \\ q & \text{si } j = i - 1, \\ 0 & \text{otro caso,} \end{cases}$$

en donde $p+q = 1$. Hemos demostrado en el capítulo anterior que las probabilidades de transición en n pasos son las siguientes: Si $n + j - i$ es par con $-n \leq j - i \leq n$, entonces

$$p_{ij}(n) = \binom{n}{\frac{1}{2}(n+j-i)} p^{(n+j-i)/2} q^{(n-j+i)/2}.$$

En otro caso $p_{ij}(n) = 0$, excepto cuando $n = 0$ e $i = j$. Más adelante usaremos este modelo para ilustrar algunos conceptos generales de cadenas de Markov.

Cadena del jugador. El modelo usado para el problema del jugador estudiado en el primer capítulo es una caminata aleatoria sobre el conjunto de números enteros $\{0, 1, \dots, N\}$, en donde los estados 0 y N son absorbentes. Las probabilidades de transición son como las de la caminata aleatoria simple sólo que ahora $p_{00} = 1$ y $p_{NN} = 1$. Este proceso es otro ejemplo de cadena de Markov con espacio de estados finito. Hemos demostrado antes que con probabilidad uno la cadena eventualmente se absorbe en alguno de los dos estados absorbentes, hemos calculado la probabilidad de ruina, es decir, la probabilidad de que la absorción se observe en el estado 0 , y hemos además encontrado el tiempo medio de absorción.

Cadena de Ehrenfest. Sean A y B dos urnas dentro de las cuales se encuentran distribuidas N bolas de acuerdo a cierta configuración inicial. En cada unidad de tiempo se escoge una bola al azar y se cambia de urna. Para tal efecto puede considerarse que las bolas se encuentran numeradas y que se escoge un número al azar, se busca la bola con ese número y se cambia de urna. Véase la Figura 3.6.

Sea X_n el número de bolas en la urna A después de n extracciones. Entonces la colección $\{X_n : n = 0, 1, \dots\}$ constituye una cadena de Markov con espacio de estados finito $\{0, 1, \dots, N\}$. Es claro que a partir del estado i la cadena sólo puede pasar al estado $i - 1$ o al estado $i + 1$ al siguiente momento, de modo que las probabilidades son $p_{i,i-1} = i/N$, y $p_{i,i+1} = (N - i)/N$, válidas para $i = 1, 2, \dots, N - 1$. Naturalmente $p_{01} = 1$ y $p_{N,N-1} = 1$. En este caso se dice que los estados 0 y N son *reflejantes*. La matriz de probabilidades de transición aparece más abajo. Este modelo fue propuesto por Ehrenfest para describir el intercambio aleatorio de moléculas (bolas) en dos regiones separadas por una membrana porosa. La región con mayor número de moléculas tenderá a

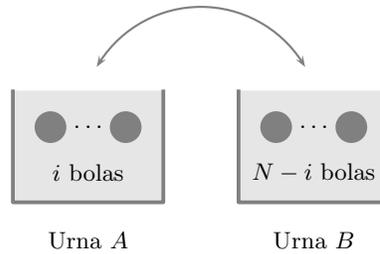


Figura 3.6:

liberar mas moléculas.

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1/N & 0 & (N-1)/N & 0 & \cdots & 0 & 0 \\ 0 & 2/N & 0 & (N-2)/N & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1/N \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Cadena de ramificación. Considere una partícula u objeto que es capaz de generar otras partículas del mismo tipo al final de un periodo establecido de tiempo. El conjunto de partículas iniciales constituye la generación 0. Cada una de estas partículas simultáneamente y de manera independiente genera un número de descendientes dentro del conjunto $\{0, 1, \dots\}$, y el total de estos descendientes pertenece a la generación 1, éstos a su vez son los progenitores de la generación 2, y así sucesivamente. Una posible sucesión de generaciones se muestra en la Figura 3.7. El posible evento cuando una partícula no genera ningún descendiente se interpreta en el sentido de que la partícula se ha muerto o extinguido.

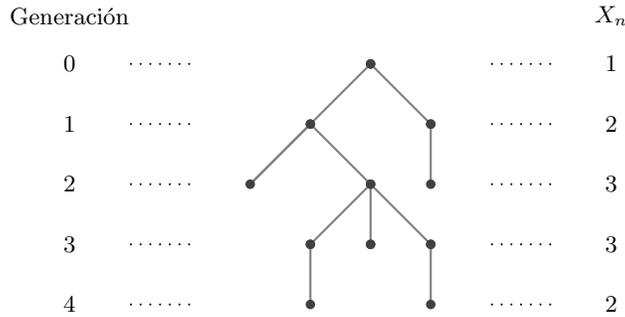


Figura 3.7:

Sea ξ_k la variable aleatoria que modela el número de descendientes de la k -ésima partícula. Para cada $n \geq 0$ defina X_n como el número de partículas en la generación n . Entonces $\{X_n : n = 0, 1, \dots\}$ es una cadena de Markov con espacio de estados $\{0, 1, \dots\}$, y probabilidades de transición $p_{ij} = P(\xi_1 + \dots + \xi_i = j)$, para $i \geq 1$. Si en algún momento $X_n = 0$, entonces se dice que la población de partículas se ha extinguido. Naturalmente el estado 0 es un estado absorbente. Este modelo ha

sido usado para determinar la probabilidad de extinción de la descendencia de una persona.

Cadena de la fila de espera. Considere una cola o línea de espera de “clientes” que solicitan algún tipo de servicio de un “servidor”. Suponga que el sistema es observado en los tiempos discretos $n = 0, 1, \dots$, y que la fila funciona del siguiente modo: cuando hay algún cliente esperando servicio al inicio de un periodo, el cliente al frente de la fila es atendido terminando el servicio al final del periodo. Naturalmente si no existiera ningún cliente en la fila al inicio de algún periodo, entonces ningún cliente es atendido. Para cada entero $n \geq 1$ defina ξ_n como el número de nuevos clientes que se incorporan a la fila durante el periodo n . Bajo ciertas condiciones es natural suponer que estas variables aleatorias son independientes, idénticamente distribuidas y con valores enteros no negativos. Suponga que $P(\xi_n = k) = a_k$, con $a_k \geq 0$ y $a_0 + a_1 + \dots = 1$. Sea X_0 el número de clientes iniciales, y para cada $n \geq 1$ defina a X_n como el número de clientes en la fila al final del periodo n . Las dos reglas de operación mencionadas se pueden escribir de la siguiente forma:

$$X_{n+1} = \begin{cases} \xi_{n+1} & \text{si } X_n = 0, \\ X_n + \xi_{n+1} - 1 & \text{si } X_n \geq 1. \end{cases}$$

Esto puede escribirse como $X_{n+1} = (X_n - 1)^+ + \xi_{n+1}$, en donde $x^+ = \max\{x, 0\}$. Por lo tanto, el proceso $\{X_n : n = 0, 1, \dots\}$ es una cadena de Markov con espacio de estados $\{0, 1, \dots\}$, y probabilidades de transición

$$p_{ij} = \begin{cases} P(\xi = j) & \text{si } i = 0, \\ P(\xi = j - i + 1) & \text{si } i \geq 1. \end{cases}$$

Es decir,

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots \\ a_0 & a_1 & a_2 & \cdots \\ 0 & a_0 & a_1 & \cdots \\ 0 & 0 & a_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Cadena de inventarios. Suponga que se almacena un cierto número de bienes en una bodega, y que se requiere una demanda aleatoria ξ_n del bien en el periodo n . Suponga que $P(\xi_n = k) = a_k$, para $k = 0, 1, \dots$ con $a_k \geq 0$ y $\sum_k a_k = 1$, es decir, la distribución de ξ_n es la misma para cualquier n . El número de bienes en

el almacén es revisado al final de cada periodo y se aplica la siguiente política de reabastecimiento: Si al final del periodo la cantidad del bien es menor o igual a un cierto nivel s , entonces se reabastece la bodega inmediatamente hasta un nivel máximo S . Si al final del periodo el número de bienes es mayor a s , entonces no hay ningún reabastecimiento. Naturalmente $s < S$. Sea X_n el número de bienes al final del periodo n y justo antes de aplicar la política de reabastecimiento. Entonces X_n es una cadena de Markov con espacio de estados $\{\dots, -2, -1, 0, 1, 2, \dots, S\}$. Se permiten valores negativos para X_n los cuales corresponden a demandas del bien no satisfechas en su momento pero que serán cubiertas en el siguiente reabastecimiento. El valor de X_{n+1} en términos de X_n se escribe de la forma siguiente:

$$X_{n+1} = \begin{cases} X_n - \xi_{n+1} & \text{si } s \leq X_n < S, \\ S - \xi_{n+1} & \text{si } X_n \leq s. \end{cases}$$

Las probabilidades de transición para esta cadena son

$$p_{ij} = P(X_{n+1} = j | X_n = i) = \begin{cases} P(\xi_{n+1} = i - j) = a_{i-j} & \text{si } s < i \leq S, \\ P(\xi_{n+1} = S - j) = a_{S-j} & \text{si } i \leq s. \end{cases}$$

La primera expresión corresponde al caso cuando no hay reabastecimiento, la probabilidad de pasar de i a j es la probabilidad de que la demanda al final de ese periodo sea de $j - i$ pues el nuevo nivel del almacén será de $i - (i - j) = j$. La segunda expresión corresponde al caso cuando hay reabastecimiento y el nuevo nivel será de $S - (S - j) = j$, cuando la demanda sea $S - j$.

3.3. Ecuación de Chapman-Kolmogorov

Esta ecuación es una fórmula sencilla y muy útil que permite descomponer la probabilidad de pasar del estado i al estado j en n pasos, en la suma de probabilidades de las trayectorias que van de i a j , y que atraviesan por un estado k cualquiera en un tiempo intermedio r .

Ecuación de Chapman-Kolmogorov. Para cualesquiera números enteros r y n tales que $0 \leq r \leq n$, y para cualesquiera estados i y j se cumple

$$p_{ij}(n) = \sum_k p_{ik}(r) p_{kj}(n-r).$$

Demostración. Por el teorema de probabilidad total y la propiedad de Markov,

$$\begin{aligned}
 p_{ij}(n) &= P(X_n = j | X_0 = i) \\
 &= \sum_k P(X_n = j, X_r = k, X_0 = i) / P(X_0 = i) \\
 &= \sum_k P(X_n = j | X_r = k) P(X_r = k | X_0 = i) \\
 &= \sum_k p_{kj}(n-r) p_{ik}(r).
 \end{aligned}$$

□

Gráficamente, las trayectorias que van del estado i al estado j en n pasos se descomponen como se muestra en la Figura 3.8. Para fines ilustrativos se dibujan las trayectorias de manera continua pero en realidad no lo son.

Esta ecuación es importante para hacer ciertos cálculos, y se usa con regularidad en el estudio de las cadenas de Markov. En particular, la siguiente desigualdad será utilizada más adelante: Para cualquier estado k y para $0 \leq r \leq n$, se tiene que $p_{ij}(n) \geq p_{ik}(r) p_{kj}(n-r)$. Como una consecuencia importante de la ecuación de Chapman-Kolmogorov se tiene el siguiente resultado.

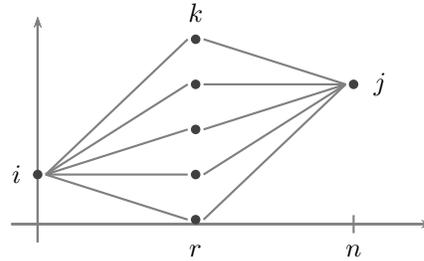


Figura 3.8:

Proposición. La probabilidad de transición en n pasos, $p_{ij}(n)$, está dada por la entrada (i, j) de la n -ésima potencia de la matriz P , es decir,

$$p_{ij}(n) = (P^n)_{ij}.$$

Demostración. Esta identidad es consecuencia de la ecuación de Chapman-Kolmogorov aplicada $n-1$ veces. La suma que aparece abajo corresponde a la entrada (i, j) de

la matriz resultante de multiplicar P consigo misma n veces.

$$\begin{aligned}
 p_{ij}(n) &= \sum_{i_1} p_{i,i_1}(1) p_{i_1,j}(n-1) \\
 &= \sum_{i_1, i_2} p_{i,i_1}(1) p_{i_1,i_2}(1) p_{i_2,j}(n-2) \\
 &\quad \vdots \\
 &= \sum_{i_1, \dots, i_{n-1}} p_{i,i_1}(1) p_{i_1,i_2}(1) \cdots p_{i_{n-1},j}(1) \\
 &= (P^n)_{ij}.
 \end{aligned}$$

□

En palabras este resultado establece que el problema de calcular las probabilidades de transición en n pasos se transforma en obtener la n -ésima potencia de la matriz de probabilidades de transición en un paso, es decir,

$$\begin{pmatrix} p_{00}(n) & p_{01}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}^n.$$

Si se conoce esta matriz y si $p_i = P(X_0 = i)$ es una distribución inicial, entonces la distribución de X_n es $P(X_n = j) = \sum_i p_i p_{ij}(n)$.

Cuando una matriz estocástica P es diagonalizable, es decir, cuando puede ser escrita en la forma QDQ^{-1} en donde D es una matriz diagonal, las potencias de P se calculan fácilmente pues $P^n = QD^nQ^{-1}$. Como D es diagonal, D^n es la matriz con cada elemento de la diagonal elevado a la n -ésima potencia.

Por ejemplo, vamos a ilustrar el proceso de diagonalización de una matriz estocástica de dimensión dos. Consideremos nuevamente la cadena general de dos estados

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}.$$

Los eigenvalores de esta matriz están dados por la ecuación $|P - \lambda I| = 0$, y resultan ser $\lambda_1 = 1$ y $\lambda_2 = 1 - a - b$. Los correspondientes eigenvectores escritos como vectores

renglón son $(1, 1)$ y $(a, -b)$, respectivamente. La matriz Q esta compuesta por los eigenvectores como columnas, y si se cumple que $a + b > 0$, entonces es invertible, es decir,

$$Q = \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix}, \quad Q^{-1} = \frac{1}{a+b} \begin{pmatrix} b & a \\ 1 & -1 \end{pmatrix}.$$

La matriz D es la matriz diagonal que contiene a los dos eigenvalores. Puede entonces comprobarse la identidad $P = QDQ^{-1}$, es decir,

$$\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1-a-b \end{pmatrix} \frac{1}{a+b} \begin{pmatrix} b & a \\ 1 & -1 \end{pmatrix}.$$

Por lo tanto,

$$\begin{aligned} P^n &= QD^nQ^{-1} \\ &= \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-a-b)^n \end{pmatrix} \begin{pmatrix} b & a \\ 1 & -1 \end{pmatrix} \frac{1}{a+b} \\ &= \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{(1-a-b)^n}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}. \end{aligned}$$

■ Ejemplo. Toda matriz estocástica P con renglones idénticos es idempotente, es decir, para cualquier entero $n \geq 1$, se cumple que $P^n = P$. Este es el caso de la cadena de variables aleatorias independientes. ◦

3.4. Comunicación

Definiremos a continuación a la comunicación entre dos estados de una cadena de Markov como la posibilidad de pasar de un estado a otro en algún número finito de transiciones.

Definición. Se dice que el estado j es *accesible* desde el estado i si existe un entero $n \geq 0$ tal que $p_{ij}(n) > 0$, esto se escribe simplemente como $i \rightarrow j$. Se dice además que los estados i y j son *comunicantes*, y se escribe $i \leftrightarrow j$, si se cumple que $i \rightarrow j$ y $j \rightarrow i$.

Observe que siempre se cumple que $i \rightarrow i$, pues por definición $p_{ii}(0) = 1$. Además observe que si ocurre que $i \leftrightarrow j$, la accesibilidad de i a j puede darse en un número de pasos distinto que la accesibilidad de j a i . Gráficamente la accesibilidad y la comunicación se representan, como lo hemos hecho antes en los ejemplos, mediante flechas entre nodos como se muestra en la Figura 3.9.

Es sencillo verificar que la comunicación es una relación de equivalencia, es decir, cumple las siguientes propiedades.

- Es reflexiva: $i \leftrightarrow i$.
- Es simétrica: si $i \leftrightarrow j$, entonces $j \leftrightarrow i$.
- Es transitiva: si $i \leftrightarrow j$ y $j \leftrightarrow k$, entonces $i \leftrightarrow k$.

En consecuencia, la comunicación induce una partición del espacio de estados de una cadena de Markov dada por los subconjuntos de estados comunicantes, es decir, dos estados pertenecen al mismo elemento de la partición si, y sólo si, tales estados se comunican. De este modo el espacio de estados de una cadena de Markov se subdivide en *clases de comunicación*. A la clase de comunicación de un estado i se le denotará por $C(i)$. Por lo tanto, $i \leftrightarrow j$ si, y sólo si, $C(i) = C(j)$.

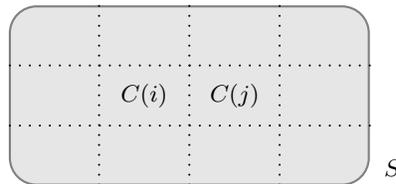
Accesibilidad



Comunicación



Figura 3.9:



Partición de un espacio de estados en clases de comunicación

Figura 3.10:

■ Ejemplo. La cadena de Markov con espacio de estados $\{0, 1, 2, 3\}$ y matriz de probabilidades de transición que se muestra en la Figura 3.11 tiene tres clases de

comunicación que son $C(0) = \{0\}$, $C(1) = \{1, 2\}$ y $C(3) = \{3\}$. Es evidente que el estado 0 se comunica consigo mismo pues existe una flecha que parte de tal estado y llega a él mismo. Visualmente tampoco hay duda de que la segunda colección de estados es una clase de comunicación. Para la clase $C(3)$ no existe una conexión física del estado 3 consigo mismo, sin embargo, por definición $p_{33}(0) = 1$, y ello hace que esta clase de únicamente un estado sea de comunicación. Observe que estas clases de comunicación conforman una partición del espacio de estados.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$$

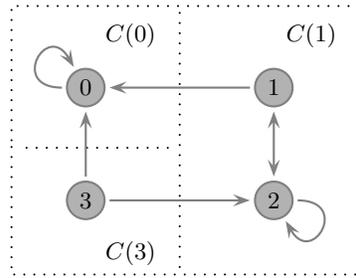


Figura 3.11:

◦

■ Ejercicio. Dibuje un diagrama de transición e identifique las clases de comunicación para las siguientes cadenas de Markov.

a) $P = \begin{pmatrix} 0.3 & 0 & 0.7 \\ 0.4 & 0.6 & 0 \\ 0.5 & 0 & 0.5 \end{pmatrix}$

b) $P = \begin{pmatrix} 0 & 0.2 & 0 & 0.8 \\ 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0.3 & 0.7 \end{pmatrix}$

▪

■ Ejercicio. ¿Cuál es el número máximo y mínimo de clases de comunicación que puede existir para una cadena de Markov de n estados? Dibuje un diagrama de transición ilustrando cada una de estas dos situaciones.

▪

El estado i de una cadena de Markov se llama *absorbente* si $p_{ii}(1) = 1$. Por ejemplo, el estado 0 de la cadena de la Figura 3.11 es absorbente.

Definición. Se dice que una cadena de Markov es *irreducible* si todos los estados se comunican entre sí.

En otras palabras, una cadena de Markov es irreducible si existe sólo una clase de comunicación, es decir, si la partición generada por la relación de comunicación es trivial. Por ejemplo, la cadena de racha de éxitos o la cadena de la caminata aleatoria son cadenas irreducibles pues todos los estados se comunican entre sí. La cadena de la Figura 3.11 no es irreducible pues no se cumple que todos los estados se comuniquen.

- Ejercicio. Dibuje el diagrama de transición de una cadena de Markov que sea
- finita e irreducible.
 - finita y reducible.
 - infinita e irreducible.
 - infinita y reducible. ■

3.5. Periodo

El periodo es un número entero no negativo que se calcula para cada estado de una cadena. Una interpretación de este número será mencionada más adelante y aparecerá también dentro de los enunciados generales sobre el comportamiento límite de cadenas de Markov.

Definición. El *periodo* de un estado i es un número entero no negativo denotado por $d(i)$, y definido como sigue:

$$d(i) = \text{m.c.d. } \{n \geq 1 : p_{ii}(n) > 0\},$$

en donde m.c.d. significa *máximo común divisor*. Cuando $p_{ii}(n) = 0$ para toda $n \geq 1$, se define $d(i) = 0$. En particular, se dice que un estado i es *aperiódico* si $d(i) = 1$. Cuando $d(i) = k \geq 2$ se dice que i es *periódico* de periodo k .

- Ejemplo. Considere una cadena de Markov con diagrama de transición como en la Figura 3.12. Es fácil comprobar que $d(0) = 1$, $d(1) = 2$, $d(2) = 2$ y $d(3) = 0$.

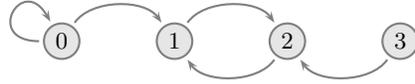


Figura 3.12:

o

■ Ejercicio. Dibuje un diagrama de transición, determine las clases de comunicación y calcule el periodo de cada uno de los estados de las siguientes cadenas de Markov.

$$a) \quad P = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad b) \quad P = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{pmatrix}$$

▪

Demostremos a continuación que el periodo es una propiedad de clase, es decir, todos los estados de una misma clase de comunicación tienen el mismo periodo. De este modo uno puede hablar de clases de comunicación *periódicas*, o *clases aperiódicas*.

Proposición. Si los estados i y j pertenecen a la misma clase de comunicación, entonces tienen el mismo periodo.

Demostración. Claramente el resultado es válido para $i = j$. Suponga entonces que i y j son distintos. Como los estados i y j están en la misma clase de comunicación, existen enteros $n \geq 1$ y $m \geq 1$ tales que $p_{ij}(n) > 0$ y $p_{ji}(m) > 0$. Sea $s \geq 1$ un entero cualquiera tal que $p_{ii}(s) > 0$. Tal entero existe pues $p_{ii}(n+m) \geq p_{ij}(n)p_{ji}(m) > 0$. Esto quiere decir que $d(i)|s$. Además $p_{jj}(n+m+s) \geq p_{ji}(m)p_{ii}(s)p_{ij}(n) > 0$. Análogamente, $p_{jj}(n+m+2s) \geq p_{ji}(m)p_{ii}(2s)p_{ij}(n) > 0$. Por lo tanto $d(j)|(n+m+s)$ y $d(j)|(n+m+2s)$. Entonces $d(j)$ divide a la diferencia $(n+m+2s) - (n+m+s)$.

$m + s) = s$. Por lo tanto, todo entero $s \geq 1$ tal que $p_{ii}(s) > 0$ cumple $d(j)|s$. Pero $d(i)$ divide a s de manera máxima, por lo tanto $d(i) \geq d(j)$. De manera análoga, escribiendo i por j , y j por i , se obtiene $d(j) \geq d(i)$. Se concluye entonces que $d(i) = d(j)$. \square

El recíproco del resultado anterior es en general falso, es decir, dos estados pueden tener el mismo periodo y sin embargo no ser comunicantes. ¿Puede usted dar un ejemplo de tal situación? El siguiente resultado establece que después de un número suficientemente grande de pasos, con probabilidad positiva toda cadena puede regresar a cada estado i cada $d(i)$ pasos. Esta es la razón por la que a tal número se le llama periodo.

Proposición. Para cada estado i , existe un entero N tal que para toda $n \geq N$, se cumple $p_{ii}(nd(i)) > 0$.

Demostración. Si $p_{ii}(n) = 0$ para cada $n \geq 1$, entonces $d(i) = 0$ y por lo tanto la afirmación es válida pues $p_{ii}(0) = 1$, sin embargo la interpretación de recurrencia periódica no se aplica en este caso. Suponga entonces que n_1, \dots, n_k son enteros tales que $p_{ii}(n_1) > 0, \dots, p_{ii}(n_k) > 0$. Sea $d = \text{m.c.d.}\{n_1, \dots, n_k\} \geq d(i)$. Como $d(i)$ es divisor de cada entero n_1, \dots, n_k , se tiene que $d(i)|d$, y por lo tanto existe un entero q tal que $qd(i) = d$. Ahora se hace uso del siguiente resultado cuya demostración puede encontrarse en [19].

Sean n_1, \dots, n_k enteros no negativos y sea $d = \text{m.c.d.}\{n_1, \dots, n_k\}$. Entonces existe un entero M tal que para cada $m \geq M$ existen enteros no negativos c_1, \dots, c_k tales que $md = c_1n_1 + \dots + c_kn_k$.

Entonces existe un entero no negativo M tal que para cada $m \geq M$, $md = c_1n_1 + \dots + c_kn_k$, para algunos enteros c_1, \dots, c_k , y por lo tanto

$$p_{ii}(md) = p_{ii}(c_1n_1 + \dots + c_kn_k) \geq p_{ii}(c_1n_1) \cdots p_{ii}(c_kn_k) > 0.$$

Por lo tanto, para cada $m \geq M$, $p_{ii}(md) = p_{ii}(mqd(i)) > 0$. Defina $N = Mq$. Se puede entonces concluir que para toda $n \geq N$, $p_{ii}(nd(i)) > 0$. \square

Como corolario de la proposición anterior se tiene el siguiente resultado: Si es posible pasar de i a j en m pasos, entonces también es posible tal transición en $m + nd(j)$ pasos con n suficientemente grande, suponiendo $d(j) > 0$.

Proposición. Si $p_{ij}(m) > 0$ para algún entero m , entonces existe un entero N tal que para toda $n \geq N$ se cumple $p_{ij}(m + nd(j)) > 0$.

Demostración. Por el resultado anterior, para n suficientemente grande, se tiene que $p_{ij}(m + nd(j)) \geq p_{ij}(m)p_{jj}(nd(j)) > 0$. \square

3.6. Primeras visitas

En ocasiones interesa estudiar el primer momento en el que una cadena de Markov visita un estado particular o un conjunto de estados. Definiremos a continuación este tiempo aleatorio y después demostraremos una fórmula útil que lo relaciona con las probabilidades de transición.

Definición. Sea A un subconjunto del espacio de estados de una cadena de Markov X_n . El *tiempo de primera visita* al conjunto A es la variable aleatoria

$$\tau_A = \begin{cases} \min \{n \geq 1 : X_n \in A\} & \text{si } X_n \in A \text{ para algún } n \geq 1, \\ \infty & \text{otro caso.} \end{cases}$$

Es decir, τ_A es el primer momento positivo en el cual la cadena toma un valor dentro de la colección de estados A , si ello eventualmente sucede. Estaremos interesados principalmente en el caso cuando el conjunto A consta de un solo estado j , y si suponemos que la cadena inicia en i , entonces el tiempo de primera visita al estado j se escribe τ_{ij} . Cuando los estados i y j coinciden se escribe simplemente τ_i . En general no es fácil encontrar la distribución de probabilidad de esta variable aleatoria, los siguientes ejemplos, sin embargo, son casos particulares sencillos.

■ **Ejemplo.** Considere la cadena de Markov de dos estados. Es inmediato comprobar que

a) $P(\tau_{01} = n) = (1 - a)^{n-1}a$, para $n = 1, 2, \dots$

b) $P(\tau_{00} = n) = a(1 - b)^{n-2}b$, para $n = 2, 3, \dots$ \circ

■ Ejemplo. Para la cadena de racha de éxitos se tiene que

a) $P(\tau_{01} = n) = q^{n-1}p$ para $n = 1, 2, \dots$

b) $P(\tau_{00} = n) = p^{n-1}q$ para $n = 1, 2, \dots$ ○

■ Ejercicio. Demuestre las siguientes igualdades.

a) $P(\tau_{ij} = 1) = p_{ij}(1)$.

b) $P(\tau_{ij} = 2) = \sum_{k \neq j} p_{ik}(1) p_{kj}(1)$.

c) $P(\tau_{ij} = n + 1) = \sum_{k \neq j} p_{ik}(1) P(\tau_{kj} = n)$, para $n \geq 1$. ■

Definición. Para cada $n \geq 1$, el número $f_{ij}(n)$ denota la probabilidad de que una cadena que inicia en el estado i , llegue al estado j por *primera vez* en exactamente n pasos, es decir,

$$f_{ij}(n) = P(\tau_{ij} = n).$$

Adicionalmente se define $f_{ij}(0) = 0$, incluyendo el caso $i = j$.

Otra forma equivalente de escribir a la probabilidad de primera visita es a través de la expresión $f_{ij}(n) = P(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j \mid X_0 = i)$. En particular, observe que $f_{ii}(n)$ es la probabilidad de regresar por primera vez al mismo estado i en el n -ésimo paso, y que $f_{ii}(1)$ es simplemente p_{ii} . El uso de la letra f para esta probabilidad proviene el término en inglés *first* para indicar que es la probabilidad de primera visita, saber esto ayuda a recordar su significado.

El siguiente resultado establece que la probabilidad de visitar el estado j , a partir de i , en n pasos, puede descomponerse en las probabilidades de los eventos disjuntos en los que se presenta la primera visita, la cual puede efectuarse en el primer paso, o en el segundo paso, y así sucesivamente, hasta el último momento posible n .

Proposición. Para cada $n \geq 1$,

$$p_{ij}(n) = \sum_{k=1}^n f_{ij}(k) p_{jj}(n-k). \quad (3.2)$$

Demostración. La prueba se basa en la siguiente descomposición que involucra eventos disjuntos. Se usa además la propiedad de Markov.

$$\begin{aligned}
 p_{ij}(n) &= P(X_n = j | X_0 = i) \\
 &= \sum_{k=1}^n P(X_n = j, X_k = j, X_{k-1} \neq j, \dots, X_1 \neq j | X_0 = i) \\
 &= \sum_{k=1}^n P(X_n = j | X_k = j) P(X_k = j, X_{k-1} \neq j, \dots, X_1 \neq j | X_0 = i) \\
 &= \sum_{k=1}^n p_{jj}(n-k) f_{ij}(k).
 \end{aligned}$$

□

■ Ejercicio. Use (3.2) para demostrar que para cada $n \geq 1$, se cumple la desigualdad $p_{ij}(n) \leq P(\tau_{ij} \leq n)$. .

3.7. Recurrencia y transitoriedad

Veremos a continuación que los estados de una cadena de Markov pueden ser clasificados, en una primera instancia, en dos tipos, dependiendo si la cadena es capaz de regresar con certeza al estado de partida.

Definición. (I) Se dice que un estado i es *recurrente* si la probabilidad de eventualmente regresar a i , partiendo de i , es uno, es decir, si

$$P(X_n = i \text{ para alguna } n \geq 1 | X_0 = i) = 1.$$

Un estado que no es recurrente se llama *transitorio*, y en tal caso la probabilidad anterior es estrictamente menor a uno.

De manera intuitiva, un estado es recurrente si con probabilidad uno la cadena es capaz de regresar eventualmente a ese estado, y cuando ello ocurre en algún momento finito, por la propiedad de Markov, se puede regresar a él una y otra vez con probabilidad uno. Debido a este comportamiento es que al estado en cuestión se le

llama recurrente. En cambio, el estado se llama transitorio si existe una probabilidad positiva de que la cadena, iniciando en él, ya no regrese nunca a ese estado. En términos de las probabilidades de primera visita, la probabilidad de una eventual visita al estado j , a partir del estado i , es el número

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n).$$

Recordemos que $f_{ij}(n)$ es la probabilidad de que la primera visita al estado j , a partir de i , se efectúe exactamente en el paso n . Siendo estos eventos disjuntos para valores distintos de n , esta suma representa la probabilidad de una posible visita al estado j . De este modo la definición de recurrencia y transitoriedad puede enunciarse de manera equivalente de la siguiente forma.

Definición. (II) Un estado i es *recurrente* si $f_{ii} = 1$, es decir, si la probabilidad de regresar a él en un tiempo finito es uno. Análogamente, un estado i es *transitorio* si $f_{ii} < 1$.

Además de la definición, tenemos el siguiente criterio útil para determinar si un estado es recurrente o transitorio.

Proposición. (Criterio para la recurrencia) El estado i es

1. recurrente si, y sólo si $\sum_{n=1}^{\infty} p_{ii}(n) = \infty$.
2. transitorio si, y sólo si $\sum_{n=1}^{\infty} p_{ii}(n) < \infty$.

Demostración. Sea N_i la variable aleatoria que cuenta el número de veces que el proceso regresa al estado i a partir del primer paso, es decir, $N_i = \sum_{n=1}^{\infty} 1_{(X_n=i)}$,

cuando $X_0 = i$. Entonces N_i tiene una distribución geométrica pues para $k \geq 1$,

$$\begin{aligned}
 P(N_i \geq k | X_0 = i) &= \sum_{m=1}^{\infty} P(N_i \geq k, X_m = i, X_{m-1} \neq i, \dots, X_1 \neq i | X_0 = i) \\
 &= \sum_{m=1}^{\infty} P(N_i \geq k | X_m = i, X_{m-1} \neq i, \dots, X_1 \neq i, X_0 = i) \\
 &\quad P(X_m = i, X_{m-1} \neq i, \dots, X_1 \neq i | X_0 = i) \\
 &= \sum_{m=1}^{\infty} P(N_i \geq k-1 | X_0 = i) f_{ii}(m) \\
 &= P(N_i \geq k-1 | X_0 = i) f_{ii} \\
 &\quad \vdots \\
 &= P(N_i \geq 1 | X_0 = i) (f_{ii})^{k-1} \\
 &= (f_{ii})^k.
 \end{aligned}$$

La esperanza de N_i , posiblemente infinita, puede calcularse de las siguientes dos formas. Primero,

$$\begin{aligned}
 E(N_i | X_0 = i) &= \sum_{k=1}^{\infty} P(N_i \geq k | X_0 = i) \\
 &= \sum_{k=1}^{\infty} (f_{ii})^k \\
 &= \begin{cases} \frac{f_{ii}}{1-f_{ii}} & \text{si } 0 \leq f_{ii} < 1, \\ \infty & \text{si } f_{ii} = 1. \end{cases}
 \end{aligned}$$

Por otro lado, por el teorema de convergencia monótona,

$$\begin{aligned}
 E(N_i | X_0 = i) &= \sum_{n=1}^{\infty} E(1_{(X_n=i)} | X_0 = i) \\
 &= \sum_{n=1}^{\infty} P(X_n = i | X_0 = i) \\
 &= \sum_{n=1}^{\infty} p_{ii}(n).
 \end{aligned}$$

El resultado se sigue de igualar estas dos expresiones. \square

Siguiendo con la notación de la demostración anterior, observe que la esperanza $E(N_i | X_0 = i)$ es el número promedio de retornos al estado i , de modo que un estado i es recurrente si, y sólo si, el número promedio de retornos a él es infinito. En contraparte, un estado i es transitorio si, y sólo si, el número promedio de retornos a él es finito.

■ **Ejemplo.** Considere una caminata aleatoria sobre \mathbb{Z} . En el primer capítulo demostramos que $\sum_{n=0}^{\infty} f_{00}(n) = 1 - |p - q|$. De aquí hemos concluido que el estado 0 es recurrente en el caso simétrico. Siendo la cadena irreducible, la cadena toda es recurrente. Alternativamente, hemos encontrado también la fórmula $p_{ii}(n) = \binom{2}{\frac{1}{2}n} \frac{1}{2^n}$, para n par. Estimando los factoriales mediante la fórmula de Stirling: $n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}$, puede comprobarse que $\sum_{n=0}^{\infty} p_{00}(n) = \infty$, confirmando nuevamente la recurrencia del estado 0 y de la cadena, en el caso simétrico. La cadena es transitoria cuando es asimétrica. ◻

Mostraremos a continuación que la recurrencia y la transitoriedad son propiedades de clase, es decir, si dos estados están en una misma clase de comunicación, entonces ambos estados son recurrentes o ambos son transitorios.

Proposición. La recurrencia es una propiedad de clase, es decir,

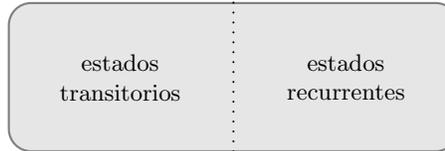
1. Si i es recurrente e $i \leftrightarrow j$, entonces j es recurrente.
2. Si i es transitorio e $i \leftrightarrow j$, entonces j es transitorio.

Demostración. Como $i \leftrightarrow j$, existen enteros $n \geq 1$ y $m \geq 1$ tales que $p_{ij}(n) > 0$ y $p_{ji}(m) > 0$. Entonces $p_{jj}(m+n+r) \geq p_{ji}(m) p_{ii}(r) p_{ij}(n)$. De modo que, por la ecuación de Chapman-Kolmogorov,

$$\sum_{r=1}^{\infty} p_{jj}(m+n+r) \geq p_{ji}(m) \sum_{r=1}^{\infty} p_{ii}(r) p_{ij}(n).$$

Si i es recurrente, la suma del lado derecho es infinita. Se sigue entonces que la suma del lado izquierdo también lo es, es decir, j es recurrente. La segunda afirmación se demuestra fácilmente por contradicción usando el primer resultado. ◻

En consecuencia, cuando una cadena es irreducible y algún estado es recurrente, todos los estados lo son, y se dice que la cadena es *recurrente*. También puede presentarse la situación en donde el espacio de estados conste de varias clases de comunicación recurrentes, en tal



Descomposición del espacio de estados

Figura 3.13:

caso la cadena también se llama *recurrente*. En contraparte, una cadena es transitoria si todos los estados lo son, ya sea conformando una sola clase de comunicación de estados transitorios o varias de ellas. De este modo el espacio de estados de una cadena de Markov puede descomponerse en dos subconjuntos ajenos de estados, aquellos que son transitorios y aquellos que son recurrentes. Tal partición se muestra en la Figura 3.13. Cada uno de estos subconjuntos puede constar de ninguna, una o varias clases de comunicación.

■ Ejemplo. La cadena de dos estados es irreducible y recurrente cuando $a, b \in (0, 1)$. En efecto, tenemos que $f_{00}(1) = 1 - a$, y $f_{00}(n) = a(1 - b)^{n-2}b$ para $n \geq 2$. Por lo tanto,

$$f_{00} = \sum_{n=1}^{\infty} f_{00}(n) = (1 - a) + ab \sum_{n=2}^{\infty} (1 - b)^{n-2} = (1 - a) + \frac{ab}{b} = 1.$$

◦

■ Ejemplo. Considere la cadena de rachas de éxitos. En este caso es sencillo demostrar que el estado 0 es recurrente pues

$$f_{00} = \sum_{n=1}^{\infty} f_{00}(n) = q(1 + p + p^2 + \dots) = \frac{q}{1 - p} = 1.$$

Dado que la cadena es irreducible y la recurrencia es una propiedad de clase, cualquier otro estado es recurrente. Por lo tanto, la cadena es recurrente. ◦

■ Ejercicio. Determine las clases de comunicación de las siguientes cadenas de Markov y clasifique éstas como recurrentes o transitorias.

$$a) \quad P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad b) \quad P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \quad .$$

■ Ejercicio. Dibuje un diagrama de transición de una cadena de Markov tal que:

- a) todos sus estados sean recurrentes.
- b) todos sus estados sean transitorios.
- c) tenga igual número de estados transitorios y recurrentes. .

Veremos a continuación algunos ejemplos de aplicación del criterio anterior.

Proposición. Sea j un estado transitorio. Para cualquier estado inicial i , se cumple que $\sum_{n=1}^{\infty} p_{ij}(n) < \infty$. En consecuencia, $\lim_{n \rightarrow \infty} p_{ij}(n) = 0$.

Demostración. Usando (3.2), por el teorema de Fubini,

$$\begin{aligned} \sum_{n=1}^{\infty} p_{ij}(n) &= \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} f_{ij}(n-k) p_{jj}(k) = \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} f_{ij}(n-k) p_{jj}(k) \\ &= \sum_{k=0}^{\infty} \sum_{m=1}^{\infty} f_{ij}(m) p_{jj}(k) = f_{ij} \sum_{k=0}^{\infty} p_{jj}(k) \leq \sum_{k=0}^{\infty} p_{jj}(k) < \infty. \end{aligned}$$

□

Proposición. Toda cadena de Markov finita tiene por lo menos un estado recurrente.

Demostración. Por contradicción, suponga que todos los estados son transitorios. Entonces para cualesquiera estados i y j , se cumple $\sum_{n=1}^{\infty} p_{ij}(n) < \infty$. Sumando sobre el conjunto finito de todos los posibles estados j se obtiene $\sum_j \sum_{n=1}^{\infty} p_{ij}(n) < \infty$. Por otro lado, intercambiando el orden de las sumas se llega a la afirmación contraria, $\sum_{n=1}^{\infty} \sum_j p_{ij}(n) = \sum_{n=1}^{\infty} 1 = \infty$. Por lo tanto es erróneo suponer que todos los estados son transitorios, debe existir por lo menos uno que es recurrente. □

En consecuencia, toda cadena finita e irreducible es recurrente. Más adelante demostraremos que en tal caso, con probabilidad uno la cadena visita cada uno de sus estados una infinidad de veces.

3.8. Tiempo medio de recurrencia

Hemos visto que si una cadena de Markov inicia en un estado recurrente, entonces regresa a él una infinidad de veces con probabilidad uno. Y hemos definido el tiempo de primera visita a un estado j , a partir de cualquier estado i , como la variable aleatoria discreta $\tau_{ij} = \min\{n \geq 1 : X_n = j \mid X_0 = i\}$, con la posibilidad de que tome el valor infinito. Vamos a definir el tiempo medio de recurrencia como la esperanza de esta variable aleatoria en el caso cuando el estado a visitar es recurrente.

Definición. El *tiempo medio de recurrencia* de un estado recurrente j , a partir del estado i , se define como la esperanza de τ_{ij} , y se denota por μ_{ij} , es decir,

$$\mu_{ij} = E(\tau_{ij}) = \sum_{n=1}^{\infty} n f_{ij}(n).$$

Recordemos que cuando el tiempo de primera visita se refiere al mismo estado de inicio y de llegada j , se escribe τ_j en lugar de τ_{jj} . En este caso el tiempo medio de recurrencia se escribe simplemente como μ_j . Esta esperanza puede ser finita o infinita, y representa el número de pasos promedio que a la cadena le toma regresar al estado recurrente j .

■ **Ejemplo.** La cadena de Markov de dos estados es irreducible y recurrente cuando $a, b \in (0, 1)$. Vamos a calcular los tiempos medios de recurrencia de estos dos estados. Tenemos que $f_{00}(1) = 1 - a$, y $f_{00}(n) = a(1 - b)^{n-2}b$ para $n \geq 2$. Por

lo tanto,

$$\begin{aligned}\mu_0 &= \sum_{n=1}^{\infty} n f_{00}(n) = (1-a) + ab \sum_{n=2}^{\infty} n(1-b)^{n-2} \\ &= (1-a) + ab \left(\frac{b+1}{b^2} \right) \\ &= \frac{a+b}{b}.\end{aligned}$$

De manera análoga, o bien intercambiando los valores de a y b , se encuentra que $\mu_1 = (a+b)/a$. Observe que $1/\mu_0 + 1/\mu_1 = 1$, más adelante explicaremos la razón de ello. Observe también que estos dos tiempos medios de recurrencia son, en general, distintos. Esto ejemplifica el hecho de que los tiempos medios de recurrencia no son necesariamente idénticos para cada elemento de una misma clase de comunicación recurrente. \circ

3.9. Clases cerradas

Las clases cerradas son subconjuntos de estados de una cadena de Markov que cumplen la propiedad de que partiendo de cualquiera estado de este subconjunto, no se puede pasar a cualquier otro estado fuera del subconjunto. Esa propiedad hace a este subconjunto de estados un subsistema propio de la cadena de Markov.

Definición. Una colección de estados no vacía \mathcal{C} es *cerrada* si ningún estado fuera de \mathcal{C} es accesible desde algún estado dentro de \mathcal{C} , es decir, si para cualquier $i \in \mathcal{C}$ y $j \notin \mathcal{C}$, $i \not\rightarrow j$.

Por ejemplo, si i es un estado absorbente, entonces la colección $\{i\}$ es claramente una clase cerrada. Como un ejemplo adicional considere cualquier clase de comunicación recurrente. Es claro que esta clase es cerrada pues de lo contrario, si $i \in \mathcal{C}$ y $j \notin \mathcal{C}$ con \mathcal{C} recurrente, e $i \rightarrow j$, entonces necesariamente $j \rightarrow i$ pues hemos supuesto que i es recurrente. Por lo tanto $i \leftrightarrow j$, y entonces $j \in \mathcal{C}$, contrario a la hipótesis $j \notin \mathcal{C}$. Por lo tanto no es posible salir de una clase de comunicación recurrente. El siguiente resultado es una especie de recíproco de lo que acabamos de mencionar.

Proposición. Toda colección de estados que es cerrada e irreducible es una clase de comunicación.

Demostración. Sea \mathcal{C} una colección no vacía de estados que es irreducible y cerrada, y sea $i \in \mathcal{C}$. Entonces $\mathcal{C} \subseteq C(i)$ pues como \mathcal{C} es irreducible, todos sus estados se comunican y por lo tanto deben pertenecer a la misma clase de comunicación. Como \mathcal{C} es cerrada, no es posible salir de tal colección, de modo que la diferencia $C(i) - \mathcal{C}$ es vacía, si existiera $j \in C(i) - \mathcal{C}$, entonces $i \rightarrow j$, lo cual contradice el supuesto de que \mathcal{C} es cerrada. Por lo tanto $\mathcal{C} = C(i)$. \square

3.10. Número de visitas

En esta sección vamos a estudiar la variable aleatoria que registra el número de visitas que una cadena realiza sobre un estado j a partir del estado i , es decir, para cualquier tiempo finito n se define la variable aleatoria

$$N_{ij}(n) = \sum_{k=1}^n 1_{(X_k=j)}, \quad \text{cuando } X_0 = i.$$

Cuando los estados i y j coinciden, se escribe $N_i(n)$ en lugar de $N_{ii}(n)$. Observe que $0 \leq N_{ij}(1) \leq N_{ij}(2) \leq \dots$, es decir, se trata de una sucesión monótona creciente de variables aleatorias no negativas que converge casi seguramente a la variable

$$N_{ij} = \sum_{k=1}^{\infty} 1_{(X_k=j)}, \quad \text{cuando } X_0 = i.$$

Los siguientes resultados acerca de estas variables aleatorias permiten distinguir la diferencia cualitativa en el comportamiento de los estados transitorios respecto de los recurrentes.

Proposición. Para cualesquiera estados i y j ,

1. $P(N_{ij} \geq k) = \begin{cases} 1 & \text{si } k = 0, \\ f_{ij}(f_{jj})^{k-1} & \text{si } k \geq 1. \end{cases}$
2. $P(N_{ij} = k) = \begin{cases} 1 - f_{ij} & \text{si } k = 0, \\ f_{ij}(f_{jj})^{k-1}(1 - f_{jj}) & \text{si } k \geq 1. \end{cases}$
3. $E(N_{ij}) = \sum_{n=1}^{\infty} p_{ij}(n) = \begin{cases} 0 & \text{si } f_{ij} = 0, \\ \frac{f_{ij}}{1 - f_{jj}} & \text{si } 0 \leq f_{jj} < 1, \\ \infty & \text{si } f_{ij} \neq 0 \text{ y } f_{jj} = 1. \end{cases}$
4. $P(N_{ij} = \infty) = \begin{cases} 0 & \text{si } j \text{ es transitorio,} \\ f_{ij} & \text{si } j \text{ es recurrente.} \end{cases}$
5. $P(N_{ij} < \infty) = \begin{cases} 1 & \text{si } j \text{ es transitorio,} \\ 1 - f_{ij} & \text{si } j \text{ es recurrente.} \end{cases}$

Demostración.

1. La primera parte de esta igualdad es evidente. Para demostrar el caso $k \geq 1$ se usa análisis del primer paso,

$$\begin{aligned} P(N_{ij} \geq k) &= \sum_{n=1}^{\infty} f_{ij}(n) P(N_{jj} \geq k - 1) \\ &= f_{ij} P(N_{jj} \geq k - 1) \\ &= f_{ij}(f_{jj})^{k-1}. \end{aligned}$$

2. Este resultado se sigue de la primera fórmula y de la igualdad $P(N_{ij} = k) = P(N_{ij} \geq k) - P(N_{ij} \geq k + 1)$.

3. Por el teorema de convergencia monótona,

$$\begin{aligned} E(N_{ij}) &= E\left(\sum_{n=1}^{\infty} 1_{(X_n=j)} \mid X_0 = i\right) \\ &= \sum_{n=1}^{\infty} E(1_{(X_n=j)} \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} p_{ij}(n). \end{aligned}$$

Por otro lado, usando la primera fórmula

$$\begin{aligned} E(N_{ij}) &= \sum_{k=1}^{\infty} P(N_{ij} \geq k) \\ &= \sum_{k=1}^{\infty} f_{ij}(f_{jj})^{k-1} \\ &= \begin{cases} 0 & \text{si } f_{ij} = 0, \\ \frac{f_{ij}}{1 - f_{jj}} & \text{si } 0 \leq f_{jj} < 1, \\ \infty & \text{si } f_{ij} \neq 0 \text{ y } f_{jj} = 1. \end{cases} \end{aligned}$$

4. Por la primera fórmula,

$$\begin{aligned} P(N_{ij} = \infty) &= \lim_{k \rightarrow \infty} P(N_{ij} \geq k) \\ &= \lim_{k \rightarrow \infty} f_{ij}(f_{jj})^{k-1}. \\ &= \begin{cases} 0 & \text{si } j \text{ es transitorio,} \\ f_{ij} & \text{si } j \text{ es recurrente.} \end{cases} \end{aligned}$$

5. Esta probabilidad es el complemento de la anterior.

□

Distribución de probabilidad de la variable N_{ij} . La variable aleatoria discreta N_{ij} toma valores en el conjunto $\{0, 1, \dots, \infty\}$. La función de probabilidad que hemos encontrado para esta variable incluye los siguientes casos:

- a) Si $f_{ij} = 0$, entonces no es posible visitar j a partir de i , y por lo tanto $P(N_{ij} = 0) = 1$, es decir la probabilidad se concentra completamente en el valor 0.
- b) Si $f_{ij} > 0$ y $f_{jj} = 1$, es decir, si se puede pasar de i a j y j es recurrente, entonces para cualquier valor de $k \geq 1$, $P(N_{ij} \geq k) = f_{ij}$, y por lo tanto $P(N_{ij} = \infty) = f_{ij}$. Mientras que $P(N_{ij} = 0) = 1 - f_{ij}$. Se trata entonces de una medida de probabilidad concentrada en los valores 0 e ∞ .
- c) Si $f_{ij} > 0$ y $f_{jj} < 1$, es decir, si se puede pasar de i a j y j es transitorio, entonces la probabilidad se distribuye sobre los valores finitos $0, 1, \dots$ como indica la fórmula 2.

Recurrencia, transitoriedad y número esperado de visitas. A partir de estas fórmulas podemos ahora distinguir el comportamiento del número de visitas en los casos cuando el estado j es transitorio o recurrente.

- a) Si j es transitorio, entonces sin importar el estado inicial i , con probabilidad uno la cadena realiza sólo un número finito de visitas al estado j , esto es lo que dice la fórmula 5, y el número esperado de visitas a tal estado es siempre finito por la fórmula 3 con $f_{jj} < 1$, es decir, $E(N_{ij}) = \sum_{n=1}^{\infty} p_{ij}(n) < \infty$. Por lo tanto, encontramos nuevamente que $\lim_{n \rightarrow \infty} p_{ij}(n) = 0$.
- b) Si j es recurrente, y si se inicia en j , entonces con probabilidad uno la cadena regresa a j una infinidad de veces, esto es lo que dice la fórmula 4 con $f_{ij} = f_{jj} = 1$, y el número esperado de visitas al estado j es infinito. Si la cadena inicia en cualquier otro estado i , entonces existe la posibilidad de que la cadena nunca visite j ($f_{ij} = 0$), y el número esperado de visitas es naturalmente cero (fórmula 3 con $f_{ij} = 0$). Pero si la cadena visita j alguna vez ($f_{ij} > 0$), entonces regresará a j una infinidad de veces, y el número esperado de visitas al estado j es infinito por la fórmula 3 con $f_{ij} > 0$ y $f_{jj} = 1$.

Anteriormente habíamos demostrado un criterio para la transitoriedad y la recurrencia del estado i en términos de la convergencia o divergencia de la serie $\sum_{n=1}^{\infty} p_{ii}(n)$. En vista de la fórmula 3 ahora podemos corroborar que un estado i es recurrente si, y sólo si, el número promedio de regresos a él es infinito, y es transitorio si, y sólo si, el número promedio de regresos es finito. También en particular, la fórmula 4 demuestra que toda cadena de Markov irreducible y recurrente, visita cada uno de sus estados una infinidad de veces con probabilidad uno.

■ Ejemplo. La cadena de racha de éxitos es irreducible y recurrente. Por lo tanto con probabilidad uno visita cada uno de sus estados una infinidad de veces. ◦

■ Ejemplo: El problema del mono. Suponga que un mono escribe caracteres al azar en una máquina de escribir. ¿Cuál es la probabilidad de que eventualmente el mono escriba exactamente, y sin ningún error, las obras completas de Shakespeare? Puede encontrarse la respuesta a esta pregunta de varias formas [30]. Usaremos la teoría de cadenas de Markov para demostrar que esta probabilidad es uno. Imaginemos entonces que un mono escribe caracteres al azar en una máquina de escribir, y que lo hace de manera continua generando una sucesión lineal de caracteres. Cada uno de los caracteres tiene la misma probabilidad de aparecer y se genera un caracter independientemente de otro. Sea m el total de caracteres disponibles que se pueden imprimir, y sea N la longitud de caracteres de los cuales consta las obras completas de Shakespeare. Sea X_n el número de caracteres correctos obtenidos inmediatamente antes e incluyendo el último momento observado n , es decir, se trata de un modelo de rachas de éxitos. Es claro que las variables X_n toman valores en el conjunto $\{0, 1, 2, \dots, N\}$, y dado que los caracteres se generan de manera independiente, el valor de X_{n+1} depende únicamente del valor de X_n y no de los anteriores, es decir, se trata efectivamente de una cadena de Markov. Considerando entonces un conjunto de símbolos de m caracteres se tiene que $P(X_{n+1} = x + 1 \mid X_n = x) = 1/m$, y $P(X_{n+1} = 0 \mid X_n = x) = (m - 1)/m$. El primer caso corresponde a obtener el caracter correcto al siguiente tiempo $n + 1$, y denotaremos a la probabilidad de tal evento por p . La segunda igualdad refleja la situación de cometer un error en el siguiente caracter generado cuando ya se habían obtenido x caracteres correctos, tal probabilidad será denotada por q . Las posibles transiciones de un estado a otro y las correspondientes probabilidades se muestran en la Figura 3.14. Técnicamente existen algunas otras posibles transiciones de algunos estados en otros, pero ello no modifica substancialmente el comportamiento cualitativo del modelo.

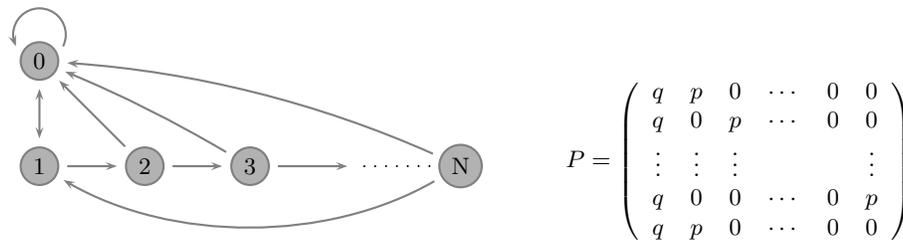


Figura 3.14:

Como puede observarse, se trata de una matriz finita e irreducible pues todos los

estados se comunican. Por lo tanto es recurrente. Entonces con probabilidad uno la cadena visita cada uno de sus estados una infinidad de veces. En particular, cada vez que la cadena visita el estado N el mono concluye una sucesión exitosa de caracteres, y ello sucederá una infinidad de veces con probabilidad uno. \circ

Teorema ergódico para cadenas de Markov. Para cualesquiera estados i y j de una cadena de Markov irreducible se cumple que

$$\lim_{n \rightarrow \infty} \frac{N_{ij}(n)}{n} = \frac{1}{\mu_j} \quad c.s. \quad (3.3)$$

siendo este límite cero cuando $\mu_j = \infty$.

Demostración. Si la cadena es transitoria, entonces ambos lados de la igualdad se anulan. Suponga que la cadena es recurrente. El tiempo de primera visita al estado j a partir de i es $\tau_{ij} = \min\{n \geq 1 : X_n = j \mid X_0 = i\}$. Dada la recurrencia e irreducibilidad, $P(\tau_{ij} < \infty) = 1$, y entonces para cualquier $n \geq 1$ se cumple la identidad

$$N_{ij}(\tau_{ij} + n) = 1 + N_{jj}(n).$$

Por lo tanto es suficiente demostrar la convergencia para $N_{jj}(n)/n$ pues

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{N_{ij}(n)}{n} &= \lim_{n \rightarrow \infty} \frac{N_{ij}(\tau_{ij} + n)}{\tau_{ij} + n} \\ &= \lim_{n \rightarrow \infty} \frac{1 + N_{jj}(n)}{\tau_{ij} + n} \\ &= \lim_{n \rightarrow \infty} \frac{N_{jj}(n)}{n} \frac{n}{\tau_{ij} + n} \\ &= \lim_{n \rightarrow \infty} \frac{N_{jj}(n)}{n}. \end{aligned}$$

Sea $Y(k)$ la variable que registra el número de pasos que transcurren entre la visita $k - 1$ y la visita k que la cadena realiza al estado j . Sabemos que el tiempo medio de recurrencia es $E(Y(k)) = \mu_j$, para $j = 1, 2, \dots$, y usando la propiedad fuerte de Markov puede demostrarse que las variables $Y(1), Y(2), \dots$ son independientes. Se tienen entonces las siguientes estimaciones

$$\frac{Y(1) + \dots + Y(N_{jj}(n))}{N_{jj}(n)} \leq \frac{n}{N_{jj}(n)} \leq \frac{Y(1) + \dots + Y(N_{jj}(n) + 1)}{N_{jj}(n)}.$$

Por la recurrencia, $N_{jj}(n) \rightarrow \infty$, cuando $n \rightarrow \infty$, de modo que por la ley de los grandes números, los dos extremos de esta desigualdad convergen a μ_j casi seguramente. Por lo tanto,

$$\lim_{n \rightarrow \infty} \frac{N_{jj}(n)}{n} = \frac{1}{\mu_j} \quad \text{c.s.}$$

□

Interpretación: Para una cadena de Markov irreducible, el número $\pi_j = 1/\mu_j$ es el tiempo promedio que la cadena permanece en el estado j a largo plazo.

Tomando esperanza en (3.3), por el teorema de convergencia dominada, y para una cadena irreducible, se cumple que

$$\begin{aligned} E\left(\lim_{n \rightarrow \infty} \frac{N_{ij}(n)}{n}\right) &= \lim_{n \rightarrow \infty} \frac{1}{n} E(N_{ij}(n)) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) \\ &= \frac{1}{\mu_j}. \end{aligned}$$

En particular, cuando el tiempo medio de recurrencia μ_j es infinito, y aún más particularmente cuando el estado j es transitorio, se tiene que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) = 0.$$

3.11. Recurrencia positiva y nula

Hemos visto que si una cadena de Markov inicia en un estado recurrente, entonces regresa a él una infinidad de veces con probabilidad uno. Sin embargo, esta recurrencia puede presentarse de dos formas: cuando el tiempo promedio de retorno es finito o cuando es infinito. Esto lleva a la definición de recurrencia positiva y recurrencia nula respectivamente. Consideremos entonces que j es un estado recurrente. El tiempo de primera visita a este estado, a partir de cualquier otro estado

i , es la variable aleatoria discreta $\tau_{ij} = \text{mín} \{n \geq 1 : X_n = j \mid X_0 = i\}$. Recordemos que cuando el tiempo de primera visita se refiere al mismo estado recurrente de inicio y de llegada i , se escribe simplemente como τ_i en lugar de τ_{ii} . La esperanza de esta variable aleatoria es naturalmente el tiempo medio de recurrencia.

Definición. El *tiempo medio de recurrencia* de un estado recurrente j , a partir del estado i , se define como la esperanza de τ_{ij} , y se denota por μ_{ij} , es decir,

$$\mu_{ij} = E(\tau_{ij}) = \sum_{n=1}^{\infty} n f_{ij}(n).$$

Nuevamente cuando el tiempo medio de recurrencia se refiere al mismo estado recurrente de inicio y de llegada i , se escribe simplemente como μ_i . Como hemos mencionado, esta esperanza puede ser finita o infinita, y ello lleva a la siguiente clasificación de estados recurrentes.

Definición. Se dice que un estado recurrente i es

- a) *recurrente positivo* si $\mu_i < \infty$.
- b) *recurrente nulo* si $\mu_i = \infty$.

Demostraremos a continuación que la recurrencia positiva y la recurrencia nula son propiedades de las clases de comunicación. Es decir, dos estados en una misma clase de comunicación recurrente, son ambos recurrentes positivos o recurrente nulos.

Proposición. (La recurrencia positiva o nula es una propiedad de clase).
Sea i un estado recurrente. Entonces

- a) si i es recurrente positivo e $i \leftrightarrow j$, entonces j es recurrente positivo.
- b) si i es recurrente nulo e $i \leftrightarrow j$, entonces j es recurrente nulo.

Demostración. Observe que es suficiente demostrar cualquiera de estas afirmaciones. Demostraremos la primera. Suponga que i es un estado recurrente positivo, es decir, i es recurrente y es tal que $\mu_i < \infty$. Como $i \leftrightarrow j$, se tiene que j es también un estado recurrente. Además existen enteros no negativos n y m tales que $p_{ij}(n) > 0$

y $p_{ji}(m) > 0$. Entonces para cualquier entero natural k ,

$$p_{jj}(n + m + k) \geq p_{ji}(m) p_{ii}(k) p_{ij}(n).$$

Sumando para $k = 1, \dots, N$, y dividiendo entre N ,

$$\frac{1}{N} \sum_{k=1}^N p_{jj}(n + m + k) \geq p_{ji}(m) \frac{1}{N} \sum_{k=1}^N p_{ii}(k) p_{ij}(n).$$

Haciendo $N \rightarrow \infty$ se obtiene

$$\frac{1}{\mu_j} \geq p_{ji}(m) \frac{1}{\mu_i} p_{ij}(n) > 0.$$

Por lo tanto también μ_j es finito, es decir, j es recurrente positivo. \square

De esta forma, el espacio de estados de toda cadena de Markov puede descomponerse en tres grandes subconjuntos de estados: transitorios, recurrentes positivos y recurrentes nulos. Esto se muestra en la Figura 3.15. Cada una de estas colecciones de estados puede estar constituida por ninguna, una, o varias clase de comunicación.



Descomposición del espacio de estados

Figura 3.15:

■ Ejemplo. Anteriormente demostramos que para la caminata aleatoria sobre \mathbb{Z} , el tiempo promedio de regreso al estado 0 es

$$\mu_0 = \sum_{n=0}^{\infty} n f_{00}(n) = \frac{4pq}{\sqrt{1-4pq}}.$$

En el caso simétrico, es decir, en el caso en el que la cadena es recurrente, este cociente se hace infinito. Esto demuestra que el estado 0 es recurrente nulo, y por lo tanto la cadena entera es recurrente nula. \circ

■ **Ejemplo.** Demostraremos ahora que la cadena de Markov de rachas de éxitos es recurrente positiva. Recordemos que dicha cadena es irreducible y recurrente. Comprobaremos que el tiempo medio de recurrencia del estado 0 es finito. En efecto,

$$\mu_0 = \sum_{n=1}^{\infty} n f_{00}(n) = \sum_{n=1}^{\infty} n (1-p)p^{n-1} = (1-p) \sum_{n=1}^{\infty} n p^{n-1} = \frac{1}{1-p} < \infty.$$

Esto demuestra que el estado 0 es recurrente positivo y siendo la cadena irreducible, es recurrente positiva. Por lo tanto el tiempo medio de recurrencia de cada estado es finito. Hemos aprovechado la facilidad del cálculo de las probabilidades de primer regreso al estado 0. ◻

Se ha demostrado antes que toda cadena finita tiene por lo menos un estado recurrente. Demostraremos ahora que para cadenas finitas sólo puede haber dos tipos de estados: transitorios o recurrentes positivos.

Proposición. No existen estados recurrentes nulos en cadenas de Markov finitas.

Demostración. Sea j un estado recurrente y sea C su clase de comunicación. La clase C es cerrada y además es finita pues la cadena completa lo es. Demostraremos que $\mu_j < \infty$. Para cualquier $i \in C$, y k natural,

$$\sum_{j \in C} p_{ij}(k) = 1.$$

Entonces

$$\frac{1}{n} \sum_{k=1}^n \sum_{j \in C} p_{ij}(k) = \sum_{j \in C} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) = 1.$$

Haciendo $n \rightarrow \infty$, por el teorema ergódico aplicado a la clase cerrada C se obtiene

$$\sum_{j \in C} \frac{1}{\mu_j} = 1.$$

Para que esta suma sea uno debe existir por lo menos un valor de j en C tal que $\mu_j < \infty$, pues de lo contrario cada sumando sería cero y la suma total no

podría ser uno. Por lo tanto existe un estado j que es recurrente positivo. Dado que la recurrencia positiva es una propiedad de clase, todos los elementos de C son recurrentes positivos. \square

Observe que en particular, todos los estados de una cadena finita e irreducible son recurrentes positivos.

3.12. Evolución de distribuciones

Una matriz estocástica establece una dinámica en el conjunto de las distribuciones de probabilidad definidas sobre el espacio de estados de la correspondiente cadena de Markov. Para explicar la situación de manera simple consideraremos un espacio de estados finito $\{0, 1, \dots, N\}$, y una distribución de probabilidad inicial $\pi(0) = (\pi_0(0), \pi_1(0), \dots, \pi_N(0))$. Después de transcurrida la primera unidad de tiempo, la cadena se encuentre en cualquiera de sus posibles estados de acuerdo a la distribución $\pi(1) = (\pi_0(1), \pi_1(1), \dots, \pi_N(1))$, en donde la j -ésima entrada de este vector es

$$\begin{aligned} \pi_j(1) &= P(X_1 = j) \\ &= \sum_{i=0}^N P(X_1 = j | X_0 = i) P(X_0 = i) \\ &= \pi_0(0) p_{0j} + \pi_1(0) p_{1j} + \dots + \pi_N(0) p_{Nj} \\ &= (\pi(0) P)_j. \end{aligned}$$

Es decir, $\pi(1)$ se obtiene a partir de $\pi(0)$ y de la matriz de probabilidades de transición P a través de la fórmula $\pi(1) = \pi(0) P$,

$$(\pi_0(1), \dots, \pi_N(1)) = (\pi_0(0), \dots, \pi_N(0)) \begin{pmatrix} p_{00} & \cdots & p_{0N} \\ \vdots & & \vdots \\ p_{N0} & \cdots & p_{NN} \end{pmatrix}.$$

A su vez la distribución $\pi(1)$ se transforma en $\pi(2)$ a través de la ecuación $\pi(2) = \pi(1) P = \pi(0) P^2$, y así sucesivamente. En general, $\pi(n+1) = \pi(n) P = \pi(0) P^{n+1}$. De esta forma se obtiene una sucesión infinita de distribuciones de probabilidad $\pi(0), \pi(1), \pi(2), \dots$, en donde cada una de ellas, excepto la primera, es obtenida de

la anterior multiplicada por la derecha por la matriz de probabilidades de transición en un paso. Por ejemplo, considere la matriz estocástica

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

con distribución inicial $\pi(0) = (0.1, 0, 0.9)$. Las subsecuentes distribuciones se calculan a continuación y las gráficas aparecen en la Figura 3.16.

$$\begin{aligned} \pi(1) &= \pi(0) P^1 = (0.45, 0.55, 0) \\ \pi(2) &= \pi(0) P^2 = (0, 0.45, 0.55) \\ \pi(3) &= \pi(0) P^3 = (0.275, 0.275, 0.45) \\ \pi(4) &= \pi(0) P^4 = (0.225, 0.5, 0.275) \\ &\vdots \end{aligned}$$

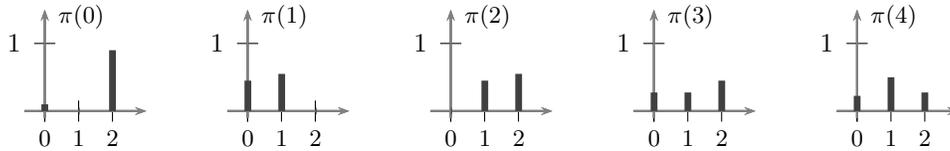


Figura 3.16:

Es natural preguntarse si existe algún límite para esta sucesión de distribuciones. En las siguientes secciones estudiaremos tal problema y encontraremos condiciones bajo las cuales existe un único límite para esta sucesión. Estudiaremos a continuación el caso particular cuando la distribución inicial no cambia al ser multiplicada por la derecha por P . A tales distribuciones se les llama estacionarias o invariantes en el tiempo.

3.13. Distribuciones estacionarias

Definición. Una distribución de probabilidad $\pi = (\pi_0, \pi_1, \dots)$ es *estacionaria* o *invariante* para una cadena de Markov con probabilidades de transición p_{ij} si

$$\pi_j = \sum_i \pi_i p_{ij}.$$

En términos matriciales, π es *estacionaria* si $\pi = \pi P$. Esta identidad tiene como consecuencia el hecho de que para cualquier número natural n , se cumpla que $\pi = \pi P^n$, es decir, π es también una distribución estacionaria para la matriz P^n . Esto significa que si la variable aleatoria inicial X_0 tiene esa distribución π , entonces la distribución de X_n también es π pues $P(X_n = j) = \sum_i \pi_i p_{ij}(n) = \pi_j$, es decir, esta distribución no cambia con el paso del tiempo y por ello es que se le llama estacionaria o invariante. Observe que el vector de ceros cumple la condición $\pi = \pi P$, sin embargo no corresponde a una distribución de probabilidad. Los siguientes ejemplos muestran que las distribuciones estacionarias pueden no ser únicas y pueden incluso no existir.

■ **Ejemplo: Existencia múltiple.** Considere una cadena de Markov sobre el conjunto de estados $\{0, 1, 2\}$ y con probabilidades de transición dada por la siguiente matriz

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Es inmediato comprobar que el vector $\pi = (1 - \alpha, 0, \alpha)$ satisface el sistema de ecuaciones $\pi = \pi P$ para cada $\alpha \in [0, 1]$. Existen entonces una infinidad de distribuciones estacionarias para esta cadena. Observe que el vector $(1 - \alpha, 0, \alpha)$ se puede escribir como la combinación lineal $(1 - \alpha)(1, 0, 0) + \alpha(0, 0, 1)$. ◻

■ **Ejemplo: No existencia.** No existe ninguna distribución estacionaria para la caminata aleatoria simétrica simple sobre \mathbb{Z} pues la condición $\pi = \pi P$ se traduce en el sistema de ecuaciones en diferencias $\pi_j = \pi_{j-1}/2 + \pi_{j+1}/2$. O bien $\pi_{j+1} - \pi_j = \pi_j - \pi_{j-1}$. Sumando estas ecuaciones puede encontrarse que para todo entero $n \geq 1$, $\pi_{n+1} - \pi_0 = n(\pi_1 - \pi_0)$. El lado izquierdo es acotado mientras el lado derecho crece sin límite cuando n es grande, a menos que $\pi_1 - \pi_0 = 0$. Es-

to demuestra que todas estas diferencias son cero, y por lo tanto π_j es constante para cualquier valor de j . Es decir, el vector constante es la solución al sistema de ecuaciones en diferencias planteado, pero ello es incompatible con la restricción $\sum_j \pi_j = 1$. Por lo tanto no existe ninguna distribución de probabilidad π que cumpla la igualdad $\pi = \pi P$ para esta cadena. \circ

■ Ejemplo: Existencia única. La cadena de Markov de dos estados dada por la matriz

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

tiene una única distribución estacionaria dada por $\pi = (\pi_0, \pi_1) = (\frac{b}{a+b}, \frac{a}{a+b})$, cuando $a + b > 0$. En cambio, cuando $a = b = 0$, la matriz resultante es la matriz identidad que acepta como distribución estacionaria a cualquier distribución de probabilidad sobre $\{0, 1\}$, es decir, en este caso existen una infinidad de distribuciones estacionarias para esta cadena. \circ

Con base en los ejemplos anteriores haremos algunas observaciones sobre las distribuciones estacionarias. Primeramente observe que para encontrar una posible distribución estacionaria de una cadena con matriz P , un primer método consiste en resolver el sistema de ecuaciones $\pi = \pi P$, sujeto a la condición $\sum_j \pi_j = 1$. Más adelante expondremos una forma alternativa de buscar distribuciones estacionarias para cadenas reversibles. Por otro lado, suponga que π y π' son dos distribuciones estacionarias distintas para una matriz P . Entonces la combinación lineal convexa $\alpha\pi + (1 - \alpha)\pi'$, para $\alpha \in [0, 1]$, también es una distribución estacionaria pues

$$(\alpha\pi + (1 - \alpha)\pi')P = \alpha\pi P + (1 - \alpha)\pi'P = \alpha\pi + (1 - \alpha)\pi'.$$

Por lo tanto, si existen dos distribuciones estacionarias distintas para una cadena, entonces existe una infinidad de ellas. Esto demuestra que el conjunto de distribuciones estacionarias es un conjunto convexo. Tomando en cuenta las observaciones anteriores y de acuerdo a los ejemplos mostrados, sólo hay tres situaciones sobre la existencia de distribuciones estacionarias para una cadena de Markov cualquiera: no existe ninguna distribución estacionaria, existe una distribución estacionaria y es única, o existe una infinidad de distribuciones estacionarias. Dadas estas consideraciones, es natural plantearse el problema de encontrar condiciones necesarias y suficientes para que una cadena tenga alguna distribución estacionaria. Primeramente demostraremos que cuando existe una distribución estacionaria, ésta tiene como soporte el conjunto de estados recurrentes positivos.

Proposición. (Soporte de una distribución estacionaria). Sea π una distribución estacionaria. Si j es un estado transitorio o recurrente nulo, entonces $\pi_j = 0$.

Demostración. Usaremos el hecho de que si j es un estado transitorio o recurrente nulo, entonces para cualquier estado i ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) = 0.$$

Como π es una distribución estacionaria,

$$\begin{aligned} \pi_j &= \sum_i \pi_i p_{ij} \\ &= \sum_i \pi_i p_{ij}(k) \\ &= \frac{1}{n} \sum_{k=1}^n \sum_i \pi_i p_{ij}(k) \\ &= \sum_i \pi_i \left(\frac{1}{n} \sum_{k=1}^n p_{ij}(k) \right) \end{aligned}$$

Tomando el límite cuando $n \rightarrow \infty$, por el teorema de convergencia dominada, se obtiene

$$\pi_j = \sum_i \pi_i \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) \right) = 0.$$

□

En particular, si $\pi_j > 0$, entonces j es un estado recurrente positivo. Esto es una consecuencia inmediata del resultado anterior y para ello puede usarse un argumento por contradicción. Por otro lado, la proposición recién demostrada también nos ayuda a corroborar nuevamente, por ejemplo, que una caminata aleatoria simétrica simple no tiene distribución estacionaria pues se trata de una cadena cuyos estados son todos recurrentes nulos. Se presenta a continuación una solución al problema de encontrar condiciones suficientes que garanticen la existencia y unicidad de la distribución estacionaria.

Proposición. (Existencia y unicidad de la distribución estacionaria). Toda cadena de Markov que es irreducible y recurrente positiva tiene una única distribución estacionaria dada por

$$\pi_j = \frac{1}{\mu_j} > 0.$$

En particular, toda cadena finita e irreducible tiene una única distribución estacionaria.

Demostración. Sea $\pi_j = 1/\mu_j$. Demostraremos que

$$(1) \quad \pi_j = \sum_i \pi_i p_{ij}.$$

$$(2) \quad \sum_j \pi_j = 1.$$

(3) π_j es única.

Como la cadena es irreducible y recurrente positiva, se tiene que $\mu_j < \infty$, para cualquier estado j . Por lo tanto el cociente $1/\mu_j$ es estrictamente positivo. Por el teorema ergódico para cadenas de Markov se tiene que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{ij}(m) = \frac{1}{\mu_j}.$$

(1) Estacionariedad. Para cualquier natural N ,

$$\begin{aligned} \sum_{i=0}^N \pi_i p_{ij} &= \sum_{i=0}^N \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{ki}(m) \right) p_{ij} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n \sum_{i=0}^N p_{ki}(m) p_{ij} \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{kj}(m+1) \\ &= \pi_j. \end{aligned}$$

Haciendo $N \rightarrow \infty$,

$$\sum_i \pi_i p_{ij} \leq \pi_j. \quad (3.4)$$

Suponga que para algún valor de j la desigualdad anterior es estricta. Sumando sobre todos los valores j , por el teorema de Fubini,

$$\sum_j \pi_j > \sum_j \sum_i \pi_i p_{ij} = \sum_i \pi_i \sum_j p_{ij} = \sum_i \pi_i.$$

Lo cual es una contradicción. Por lo tanto (3.4) es una igualdad. Esto demuestra que π es estacionaria, y para cualquier m natural,

$$\pi_j = \sum_i \pi_i p_{ij}(m). \quad (3.5)$$

(2) Distribución de probabilidad. Para cualquier natural N ,

$$\begin{aligned} \sum_{j=0}^N \pi_j &= \sum_{j=0}^N \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \sum_{j=0}^N p_{ij}(k) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n 1 \\ &= 1. \end{aligned}$$

Por otra parte, por (3.5), para cualquier n natural,

$$\pi_j = \sum_i \pi_i \left(\frac{1}{n} \sum_{m=1}^n p_{ij}(m) \right).$$

Haciendo $n \rightarrow \infty$, por el teorema de convergencia dominada de Lebesgue, y usando el hecho de que $\sum_j \pi_j \leq 1$,

$$\pi_j = \sum_{i=0}^{\infty} \pi_i \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{ij}(m) \right) = \sum_{i=0}^{\infty} \pi_i \pi_j.$$

Dado que $\pi_j > 0$, se obtiene $\sum_i \pi_i = 1$.

(3) Unicidad. Sean π y π' dos distribuciones estacionarias para la matriz P . Entonces

$$\pi'_j = \sum_i \pi'_i p_{ij}(m) = \sum_i \pi'_i \left(\frac{1}{n} \sum_{m=1}^n p_{ij}(m) \right) \geq \sum_{i=0}^N \pi'_i \left(\frac{1}{n} \sum_{m=1}^n p_{ij}(m) \right).$$

Haciendo $n \rightarrow \infty$,

$$\pi'_j \geq \sum_{i=0}^N \pi'_i \pi_j.$$

Ahora hacemos $N \rightarrow \infty$ para obtener

$$\pi'_j \geq \pi_j. \quad (3.6)$$

Si esta desigualdad fuera estricta para algún valor de j , entonces sumando sobre todos los valores de j se obtiene $1 = \sum_j \pi'_j > \sum_j \pi_j = 1$, lo cual es una contradicción. Por lo tanto (3.6) es una igualdad para cada valor de j , y ello demuestra la unicidad. \square

■ Ejemplo. La cadena de dos estados es finita e irreducible cuando $a + b > 0$, y por lo tanto es recurrente positiva. Por el resultado recién demostrado esta cadena tiene una única distribución estacionaria. Resolviendo el sistema de ecuaciones $\pi = \pi P$ se encuentra que $\pi = (\pi_0, \pi_1) = (b/a + b, a/a + b)$. Como una aplicación de la última proposición demostrada encontramos nuevamente que los tiempos medios de recurrencia son $(\mu_0, \mu_1) = (1/\pi_0, 1/\pi_1) = ((a + b)/b, (a + b)/a)$. \circ

■ Ejemplo. La cadena de Ehrenfest es finita e irreducible, en consecuencia es recurrente positiva. Por lo tanto tiene una única distribución estacionaria. Resolviendo el sistema de ecuaciones $\pi = \pi P$ se encuentra que el vector π tiene distribución $\text{bin}(N, 1/2)$, es decir,

$$\pi_j = \binom{N}{j} \frac{1}{2^N}, \quad \text{para } j = 0, 1, \dots, N.$$

En vista de la proposición demostrada, los tiempos medios de recurrencia para esta cadena son

$$\mu_j = \frac{1}{\pi_j} = \frac{2^N}{\binom{N}{j}}, \quad \text{para } j = 0, 1, \dots, N.$$

\circ

■ Ejemplo. La cadena de racha de éxitos, aunque no es finita, es irreducible y recurrente positiva. Por lo tanto tiene una única distribución estacionaria dada por la distribución $\text{geo}(1 - p)$, es decir, el sistema de ecuaciones $\pi = \pi P$ tiene como única solución la distribución

$$\pi_j = (1 - p) p^j, \quad \text{para } j = 0, 1, 2, \dots$$

Como consecuencia de la proposición anterior, se obtiene que los tiempos medios de recurrencia para cada uno de los estados de esta cadena son

$$\mu_j = \frac{1}{\pi_j} = \frac{1}{(1-p)p^j}, \quad \text{para } j = 0, 1, 2 \dots$$

En particular, $\mu_0 = 1/(1-p)$. Esto confirma los cálculos realizados antes para μ_0 a partir de la fórmula $\mu_0 = \sum_{n=1}^{\infty} n f_{00}(n)$. \circ

3.14. Distribuciones límite

Como hemos mencionado antes, toda matriz de probabilidades de transición P determina una sucesión de distribuciones de probabilidad π^0, π^1, \dots sobre el espacio de estados en donde

$$\pi^{n+1} = \pi^n P = \pi^0 P^{n+1}. \quad (3.7)$$

Bajo ciertas condiciones tal sucesión es convergente a una distribución límite π . Imaginemos por ahora que tal es el caso y examinemos algunas propiedades de esta distribución límite. Por (3.7) se tiene que

$$\pi = \lim_{n \rightarrow \infty} \pi^n = \pi^0 \lim_{n \rightarrow \infty} P^{n+1} = \pi^1 \lim_{n \rightarrow \infty} P^n.$$

Estas igualdades revelan varias cosas. Primero, la distribución límite no depende de la distribución inicial pues el lado derecho de estas ecuaciones debe ser el mismo. Segundo, la distribución límite está dada por el límite de las potencias de P pues si se toma como distribución inicial aquella concentrada en el i -ésimo estado, entonces el j -ésimo elemento de la distribución límite es $\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n)$. Tercero, el límite de las potencias de P es una matriz con todos sus renglones idénticos, siendo este renglón la distribución límite. Por último, a partir de (3.7) se obtiene que $\lim_{n \rightarrow \infty} \pi^{n+1} = \lim_{n \rightarrow \infty} \pi^n P$, esto es, $\pi = \pi P$, es decir, la distribución límite es una distribución estacionaria, suponiendo un espacio de estados finito. En esta sección se establecen condiciones para obtener rigurosamente estos resultados.

Como se ha visto, toda distribución límite es estacionaria pero el recíproco es en general falso como se ilustra en el siguiente ejemplo.

■ **Ejemplo.** Es inmediato comprobar que la distribución uniforme $(1/2, 1/2)$ es estacionaria para la cadena con matriz $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, y sin embargo las potencias

P^n no convergen pues para cualquier entero natural n , $P^{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, mientras que $P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. \circ

El siguiente resultado es válido para espacios de estados finitos o infinitos, y establece que si el límite de las probabilidades $p_{ij}(n)$, cuando $n \rightarrow \infty$, existen y no dependen de i , entonces la distribución límite podría ser una distribución estacionaria. Esto es solamente una posibilidad pues los límites podrían ser todos cero. En el caso finito, sin embargo, demostraremos que tales límites conforman una distribución de probabilidad verdadera.

Proposición. Considere una cadena de Markov con probabilidades de transición p_{ij} tales que los límites $\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n)$ existen para cada j , y no dependen del estado i . Entonces

1. $\sum_j \pi_j \leq 1$.
2. $\pi_j = \sum_i \pi_i p_{ij}(n)$.

Cuando el espacio de estados es finito se cumple la igualdad en el primer resultado obteniéndose una distribución de probabilidad verdadera.

Demostración. Suponga primero el caso cuando el espacio de estados es el conjunto finito $\{0, 1, \dots, N\}$. Entonces la primera afirmación se cumple con igualdad pues

$$\sum_{j=0}^N \pi_j = \sum_{j=0}^N \lim_{n \rightarrow \infty} p_{ij}(n) = \lim_{n \rightarrow \infty} \sum_{j=0}^N p_{ij}(n) = 1.$$

Para la segunda afirmación se tiene que para cualquier $n \geq 1$,

$$\begin{aligned} \sum_{i=0}^N \pi_i p_{ij}(n) &= \sum_{i=0}^N \lim_{m \rightarrow \infty} p_{ki}(m) p_{ij}(n) \\ &= \lim_{m \rightarrow \infty} \sum_{i=0}^N p_{ki}(m) p_{ij}(n) \\ &= \lim_{m \rightarrow \infty} p_{kj}(m+n) \\ &= \pi_j. \end{aligned}$$

Suponga ahora que el espacio de estados es infinito. En este caso no es posible garantizar la validez del intercambio de límite y suma efectuado antes. Para la primera afirmación se tiene que para cualquier número natural $N \geq 1$,

$$\sum_{j=0}^N \pi_j = \sum_{j=0}^N \lim_{n \rightarrow \infty} p_{ij}(n) = \lim_{n \rightarrow \infty} \sum_{j=0}^N p_{ij}(n) \leq \lim_{n \rightarrow \infty} 1 = 1.$$

Haciendo $N \rightarrow \infty$ se obtiene el resultado buscado. Para la segunda afirmación, nuevamente para cualquier número natural $N \geq 1$,

$$\begin{aligned} \sum_{i=0}^N \pi_i p_{ij} &= \sum_{i=0}^N \lim_{n \rightarrow \infty} p_{ki}(n) p_{ij} \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^N p_{ki}(n) p_{ij} \\ &\leq \lim_{n \rightarrow \infty} p_{kj}(n+1) \\ &= \pi_j. \end{aligned}$$

Si $\pi_j = 0$ para cualquier j , entonces estas desigualdades se convierten en identidades como dice el enunciado. Suponga que $\sum_j \pi_j > 0$. Demostraremos que las desigualdades estrictas no son posibles. Suponga que para algún estado j , $\sum_i \pi_i p_{ij} < \pi_j$. Entonces

$$\sum_j \pi_j > \sum_{i,j} \pi_i p_{ij} = \sum_i \pi_i \sum_j p_{ij} = \sum_i \pi_i.$$

Esto es una contradicción, por lo tanto la igualdad en el párrafo anterior se cumple. \square

Ahora se establecen condiciones suficientes para que exista el límite de las probabilidades de transición en n pasos. Este resultado es una especie de recíproco del resultado anterior pues supone la existencia de una distribución estacionaria para concluir que los límites de las probabilidades existen.

Teorema de convergencia. Considere una cadena de Markov que es

- a) irreducible,
- b) aperiódica, y
- c) con distribución estacionaria π .

Entonces para cualesquiera estados i y j , $\lim_{n \rightarrow \infty} p_{ij}(n) = \pi_j$.

Demostración. El método de esta demostración se conoce como *técnica de acople*. Sea $\{Y_n\}$ una cadena de Markov independiente de la original $\{X_n\}$, pero con la misma matriz de probabilidades de transición. Entonces el proceso $\{Z_n\}$ definido por $Z_n = (X_n, Y_n)$ es una cadena de Markov con probabilidades de transición

$$P(Z_{n+1} = (x_{n+1}, y_{n+1}) | Z_n = (x_n, y_n)) = p_{x_n, x_{n+1}} p_{y_n, y_{n+1}},$$

y puede fácilmente comprobarse que tiene distribución estacionaria $\pi_{x_n, y_n} = \pi_{x_n} \pi_{y_n}$. Puede además verificarse que la cadena $\{Z_n\}$ es recurrente positiva. Además es irreducible pues como $\{X_n\}$ y $\{Y_n\}$ son aperiódicas, existe un número natural N tal que $p_{ij}(n) p_{kl}(n) > 0$, para toda $n > N$. Por lo tanto $p_{(i,k)(j,l)}(n) > 0$.

Sea j un estado cualquiera de la cadena original. Defina el primer momento en el que la cadena $\{Z_n\}$ visita el estado (j, j) como $\tau_j = \min\{n \geq 1 : Z_n = (j, j)\}$. Sea además $\tau = \min\{n \geq 1 : X_n = Y_n\}$. Este es el primer momento de acople de las dos cadenas. Como $\{Z_n\}$ es recurrente, $P(\tau < \infty) = 1$. Además $\tau \leq \tau_j$. Por la propiedad de Markov

$$\begin{aligned} P(X_n = x, \tau \leq n) &= \sum_j \sum_{r=1}^n P(X_n = x, X_r = j, \tau = r) \\ &= \sum_j \sum_{r=1}^n P(X_n = x | X_r = j, \tau = r) P(X_r = j, \tau = r) \\ &= \sum_j \sum_{r=1}^n P(Y_n = x | Y_r = j, \tau = r) P(Y_r = j, \tau = r) \\ &= \sum_j \sum_{r=1}^n P(Y_n = x | Y_r = j) P(Y_r = j, \tau = r) \\ &= P(Y_n = x, \tau \leq n), \end{aligned}$$

es decir, sobre el evento $(\tau \leq n)$, las variables X_n y Y_n tienen la misma distribución de probabilidad. Por otro lado

$$\begin{aligned} P(X_n = j) &= P(X_n = j, \tau \leq n) + P(X_n = j, \tau > n) \\ &= P(Y_n = j, \tau \leq n) + P(X_n = j, \tau > n) \\ &\leq P(Y_n = j) + P(\tau > n). \end{aligned}$$

De manera análoga, $P(Y_n = j) \leq P(X_n = j) + P(\tau > n)$. Por lo tanto,

$$|P(X_n = j) - P(Y_n = j)| \leq P(\tau > n) \rightarrow 0, \quad (3.8)$$

cuando $n \rightarrow \infty$. Si ahora se toma $X_0 = i$ con probabilidad uno, y Y_0 con la distribución estacionaria π , entonces $P(X_n = j) = P(X_n = j | X_0 = i)P(X_0 = i) = p_{ij}(n)P(X_0 = i) = p_{ij}(n)$, y $P(Y_n = j) = \sum_i P(Y_n = j | Y_0 = i)\pi_i = \sum_i \pi_i p_{ij}(n) = \pi_j$. Por lo tanto (3.8) establece que $|p_{ij}(n) - \pi_j| \rightarrow 0$. \square

El siguiente resultado establece condiciones suficientes para la existencia del límite de las probabilidades de transición, asegurando además que se trata de una distribución de probabilidad.

Teorema de convergencia. Considere una cadena de Markov que es

- a) irreducible,
- b) recurrente positiva, y
- c) aperiódica.

Entonces las probabilidades límite $\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n) = \frac{1}{\mu_j}$ existen, y constituyen la única solución al sistema de ecuaciones

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}, \quad (3.9)$$

sujeto a las condiciones $\pi_j \geq 0$, y $\sum_{j=0}^{\infty} \pi_j = 1$.

Demostración. Como la cadena es irreducible y recurrente positiva, tiene una única distribución estacionaria dada por $\pi_j = 1/\mu_j$. Es decir, es la única solución al sistema de ecuaciones $\pi = \pi P$, con $\pi_j \geq 0$ y $\sum_j \pi_j = 1$. Además, por la aperiódicidad, $p_{ij}(n) \rightarrow \pi_j$. \square

3.15. Cadenas regulares

Las cadenas de Markov regulares son cadenas finitas que cumplen la propiedad de que a partir de un cierto momento, con probabilidad positiva se puede pasar de un estado a otro cualquiera en un paso. Demostraremos que para este tipo de cadenas existe siempre la distribución límite.

Definición. Se dice que una cadena finita o su matriz de probabilidades de transición es *regular* si existe un entero natural n tal que $p_{ij}(n) > 0$, para cualesquiera estados i y j .

En palabras, una cadena de Markov finita es regular si alguna potencia de su matriz de probabilidades de transición tiene todas sus entradas estrictamente positivas. Otra forma de definir a una cadena regular es a través del siguiente resultado.

Proposición. Una matriz estocástica es regular si, y sólo si, es finita, irreducible y aperiódica.

Demostración. Si la matriz es regular, entonces claramente es finita, irreducible y aperiódica. Recíprocamente, por la irreducibilidad, para cualesquiera dos estados i y j , existe un entero m tal que $p_{ij}(m) > 0$. Entonces existe un entero N tal que $p_{ij}(m + n d(j)) > 0$, para cada $n \geq N$. Como $d(j) = 1$ y siendo la matriz finita, esto implica la regularidad. \square

Para este tipo particular de cadenas finitas se conocen los siguientes resultados acerca de su comportamiento límite.

Proposición. Toda cadena finita que además es

1. regular, tiene como distribución límite la única solución no negativa del sistema (3.9).
2. regular y doblemente estocástica, tiene como distribución límite la distribución uniforme.
3. irreducible, aperiódica y doblemente estocástica, tiene como distribución límite la distribución uniforme.

Demostración.

1. Como la cadena es regular, es irreducible, aperiódica, y es recurrente positiva por ser finita. Por el teorema anterior la distribución límite existe y está dada por el sistema de ecuaciones (3.9).
2. Como la cadena es regular, es aperiódica, irreducible y recurrente positiva por ser finita. Entonces la distribución límite existe. Por la hipótesis de doble estocasticidad y suponiendo que el espacio de estados es $\{0, 1, \dots, N\}$, se tiene que $\sum_{i=0}^N p_{ij}(n) = 1$. Tomando el límite cuando n tiende a infinito se obtiene $\sum_{i=0}^N \pi_j = 1$. Por lo tanto $\pi_j = 1/(N + 1)$.
3. Este resultado es idéntico al anterior pues hemos demostrado que la regularidad es equivalente a la finitud, irreducibilidad y aperiodicidad conjuntamente.

□

■ Ejemplo. Sea S_n la suma de los resultados que se obtienen al lanzar un dado equilibrado n veces. Encontraremos

$$\lim_{n \rightarrow \infty} P(S_n \text{ es múltiplo de } 7).$$

Defina el proceso $X_n = S_n \bmod 7$ cuyo espacio de estados es $\{0, 1, \dots, 6\}$. No es difícil convencerse que X_n es una cadena de Markov con matriz de probabilidades

de transición

$$P = \begin{pmatrix} 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 \end{pmatrix}$$

El evento (S_n es múltiplo de 7) es idéntico al evento ($X_n = 0$), de modo que la probabilidad de que S_n sea múltiplo de 7 a largo plazo es el tiempo de estancia a largo plazo que la cadena X_n pasa en el estado 0. El problema se reduce a encontrar la distribución límite de P . Pero esta matriz es regular, y entonces su distribución límite es la uniforme. Por lo tanto $\lim_{n \rightarrow \infty} P(X_n = 0) = 1/7$. \circ

3.16. Cadenas reversibles

Sea $\{X_n\}$ una cadena de Markov con probabilidades de transición p_{ij} . Sea $m \geq 1$ un entero fijo y defina un nuevo proceso $Y_n = X_{m-n}$ para $n = 0, \dots, m$, es decir, Y_n es la cadena original pero vista en sentido inverso en el tiempo, ahora del tiempo m al tiempo 0. Este nuevo proceso resulta también ser una cadena de Markov pues cumple el criterio de independencia entre pasado y futuro cuando se conoce el presente, las nociones de pasado y futuro se intercambian debido al cambio en el sentido del tiempo. En efecto, para $1 \leq r < n \leq m$, considere la probabilidad condicional

$$P(y_1, \dots, y_{r-1}, y_{r+1}, \dots, y_n | y_r).$$

En términos del proceso $\{X_n\}$, esta probabilidad es

$$P(X_{m-1} = y_1, \dots, X_{m-r+1} = y_{r-1}, \\ X_{m-r-1} = y_{r+1}, \dots, X_{m-n} = y_n | X_{m-r} = y_r).$$

Por la propiedad de Markov del proceso $\{X_n\}$, esta probabilidad es el producto

$$P(X_{m-1} = y_1, \dots, X_{m-r+1} = y_{r-1} | X_{m-r} = y_r) \\ P(X_{m-r-1} = y_{r+1}, \dots, X_{m-n} = y_n | X_{m-r} = y_r),$$

que en términos del proceso $\{Y_n\}$ es la propiedad de Markov para este proceso

$$P(y_1, \dots, y_{r-1} | y_r) P(y_{r+1}, \dots, y_n | y_r).$$

Sin embargo las probabilidades de transición del nuevo proceso no son homogéneas pues para $0 \leq n < m$,

$$\begin{aligned} P(Y_{n+1} = j | Y_n = i) &= \frac{P(X_{m-n} = i, X_{m-n-1} = j)}{P(X_{m-n} = i)} \\ &= P(X_{m-n} = i | X_{m-n-1} = j) \frac{P(X_{m-n-1} = j)}{P(X_{m-n} = i)} \\ &= p_{ji} \frac{P(Y_{n+1} = j)}{P(Y_n = i)}, \end{aligned}$$

es decir, estas probabilidades dependen de n a través del cociente $P(Y_{n+1} = j)/P(Y_n = i)$. Tal dependencia desaparece cuando se toma como hipótesis la existencia de una distribución estacionaria π para $\{X_n\}$, pues en tal caso la igualdad anterior se reduce a

$$P(Y_{n+1} = j | Y_n = i) = p_{ji} \frac{\pi_j}{\pi_i}.$$

Bajo tal hipótesis las probabilidades de transición de la nueva cadena son ahora estacionarias. Si adicionalmente se pide que las probabilidades de transición son las mismas para ambas cadenas, entonces de la ecuación anterior se obtiene que debe satisfacerse la ecuación $p_{ij} = p_{ji} \frac{\pi_j}{\pi_i}$, es decir, $\pi_i p_{ij} = \pi_j p_{ji}$. Esto lleva a la siguiente definición de reversibilidad la cual añade la hipótesis de irreducibilidad.

Definición. Se dice que una cadena de Markov irreducible con probabilidades de transición p_{ij} y con distribución estacionaria π es *reversible* en el tiempo si para cualesquiera estados i y j ,

$$\pi_i p_{ij} = \pi_j p_{ji}. \quad (3.10)$$

A la ecuación (3.10) se llama *ecuación de balance detallado*. La utilidad de las cadenas reversibles está dada por el siguiente resultado, el cual ejemplificaremos más adelante con un par de aplicaciones.

Proposición. Considere una cadena irreducible para la cual existe una distribución π que cumple (3.10). Entonces π es una distribución estacionaria y la cadena es reversible.

Demostración. Si π cumple (3.10), entonces es una distribución estacionaria pues $\sum_i \pi_i p_{ij} = \sum_i \pi_j p_{ji} = \pi_j$. Además la cadena es reversible por definición. \square

■ **Ejemplo.** Considere nuevamente la cadena de Ehrenfest. El espacio de estados es $\{0, \dots, N\}$, y las probabilidades de transición son, para $i = 1, \dots, N - 1$,

$$p_{ij} = \begin{cases} (N - i)/N & \text{si } j = i + 1, \\ i/N & \text{si } j = i - 1, \\ 0 & \text{otro caso,} \end{cases}$$

con $p_{01} = 1$ y $p_{N,N-1} = 1$. Esta cadena es finita, irreducible y recurrente positiva. Estas propiedades garantizan que la cadena tiene una única distribución estacionaria. Si se desea encontrar esta distribución estacionaria a través de la ecuación $\pi = \pi P$, uno tendría que resolver el sistema

$$\begin{aligned} \pi_0 &= (1/N) \pi_1, \\ \pi_i &= ((N - i + 1)/N) \pi_{i-1} + ((i + 1)/N) \pi_{i+1}, \quad \text{para } i = 1, \dots, N - 1, \\ \pi_N &= (1/N) \pi_{N-1}. \end{aligned}$$

A partir de estas ecuaciones puede demostrarse que π es la distribución $\text{bin}(N, 1/2)$. Alternativamente, se puede buscar una posible solución al sistema (3.10), el cual se escribe como sigue

$$\pi_{i+1} = \frac{N - i}{i + 1} \pi_i, \quad \text{para } i = 0, 1, \dots, N - 1.$$

Resolviendo iterativamente hacia abajo se encuentra que

$$\pi_i = \binom{N}{i} \pi_0, \quad \text{para } i = 1, \dots, N.$$

Sumando todas estas cantidades,

$$1 = \pi_0 \sum_{i=0}^N \binom{N}{i} = \pi_0 2^N.$$

Por lo tanto $\pi_0 = 1/2^N$, y $\pi_i = \binom{N}{i}/2^N$ es la solución al sistema. Esta es la distribución $\text{bin}(N, 1/2)$. Por lo demostrado antes esta distribución es estacionaria y la cadena es reversible. \circ

■ Ejemplo. Considere una caminata aleatoria no homogénea sobre $\{0, 1, \dots\}$, con probabilidades de transición

$$p_{ij} = \begin{cases} p_i & \text{si } j = i + 1, \\ q_i & \text{si } j = i - 1, \\ 1 - p_i - q_i & \text{si } j = i, \\ 0 & \text{otro caso,} \end{cases}$$

en donde $q_0 = 0$. Esta es una cadena irreducible. Si uno busca una distribución estacionaria a través de la ecuación $\pi = \pi P$, uno tendría que resolver el sistema de ecuaciones

$$\begin{aligned} \pi_0 &= \pi_0(1 - p_0) + q_1 \pi_1, \\ \pi_i &= \pi_{i+1} q_{i+1} + \pi_i(1 - p_i - q_i) + \pi_{i-1} p_{i-1}, \quad \text{para } i \geq 1. \end{aligned}$$

Sin embargo la condición (3.10) se traduce en el sistema más simple $\pi_i p_i = \pi_{i+1} q_{i+1}$. Una posible solución de este sistema (su existencia dependerá de los parámetros p_i y q_i), será la distribución estacionaria para esta caminata y la cadena será entonces reversible en el tiempo. \circ

Resumen de la notación utilizada. Como referencia se presenta a continuación una lista de la notación utilizada en este capítulo.

- p_{ij} Probabilidad de pasar del estado i al estado j en un paso, es decir, $p_{ij} = P(X_1 = j | X_0 = i)$.
- $p_{ij}(n)$ Probabilidad de pasar del estado i al estado j en exactamente n pasos, es decir, $p_{ij}(n) = P(X_n = j | X_0 = i)$. En algunos textos aparece también como $p_{ij}^{(n)}$. En particular, se define $p_{ij}(0) = \delta_{ij}$, que vale uno cuando $i = j$, y cero cuando $i \neq j$.
- $f_{ij}(n)$ Probabilidad de llegar por primera vez al estado j a partir del estado i exactamente en el paso n , es decir, $f_{ij}(n) = P(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j | X_0 = i)$. A veces se escribe también como $f_{ij}^{(n)}$. En particular se define $f_{ij}(0) = 0$ para cualesquiera estados i y j , incluyendo el caso $i = j$.

- f_{ij} Probabilidad de una eventual visita al estado j a partir del estado i . En términos de probabilidades de primeras visitas, esta probabilidad es $f_{ij} = \sum_{k=0}^{\infty} f_{ij}(n)$.
- $N_{ij}(n)$ Variable aleatoria que registra el número de visitas realizadas al estado j durante las primeras n transiciones, cuando la cadena inicia en el estado i , es decir, $N_{ij}(n) = \sum_{k=1}^n 1_{(X_k=j)}$, cuando $X_0 = i$.
- N_{ij} Variable aleatoria que cuenta el número total de visitas realizadas al estado j cuando la cadena inicia en el estado i , es decir, $N_{ij} = \sum_{k=1}^{\infty} 1_{(X_k=j)}$, cuando $X_0 = i$. Puede tomar el valor infinito.
- τ_{ij} Tiempo en el que se logra la primera visita al estado j a partir del estado i , es decir, $\tau_{ij} = \min\{n \geq 1 : X_n = j\}$, cuando $X_0 = i$. Toma el valor infinito cuando nunca ocurre tal evento. Cuando $i = j$ se escribe τ_i , y corresponde al tiempo del primer regreso al estado i .
- μ_{ij} Tiempo medio de primera visita al estado j a partir del estado i , es decir, $\mu_{ij} = E(\tau_{ij})$. Puede ser infinito. Cuando $i = j$ se escribe μ_i y se le llama tiempo medio de recurrencia del estado i .

Notas y referencias. El tema de cadenas de Markov a tiempo discreto aparece en casi cualquier texto de procesos estocásticos en mayor o menor profundidad, e incluso puede encontrarse también en los últimos capítulos de algunos textos de probabilidad. El tema es regularmente la parte inicial y obligada de un curso elemental de procesos estocásticos. Las siguientes referencias son una muestra de algunos textos que contienen capítulos sobre el tema de cadenas de Markov a un nivel similar al presentado: Karlin y Taylor [19], Brzeźniak y Zastawniak [4], Jones y Smith [17], Hoel, Port y Stone [16], y Stirzaker [36]. Los textos de Caballero et al [6] y Norris [25], están dedicados enteramente al tema de cadenas de Markov.



A. A. Markov

Andrei Andreyevich Markov (Rusia 1856–1922).

Markov tuvo una salud muy precaria durante sus primeros años de vida, teniendo que caminar con muletas hasta la edad de 10 años. En 1874 ingresó a la facultad de física y matemáticas de la universidad de San Petersburgo, y asistió a las clases de reconocidos matemáticos de la época, entre ellos P. L. Chebyshev, quien tuvo una influencia decisiva en el quehacer futuro de Markov. Se graduó brillantemente en 1878, y continuó con sus estudios de maestría, los cuales concluyó en 1880. Trabajó como profesor en la universidad de San Petersburgo al mismo tiempo que estudiaba para su doctorado, el cual concluyó en 1884. Continuó trabajando en la misma universidad por prácticamente el resto de su vida. Después de 1900, y siguiendo los trabajos de P. L. Chebyshev, aplicó el método de fracciones continuas en la teoría de la probabilidad. Markov fue el mejor exponente y continuador de las ideas de Chebyshev y de sus temas de investigación en probabilidad. Especialmente sobresalientes son sus trabajos sobre la ley de los grandes números y el teorema central del límite, así como otros resultados fundamentales de la teoría de la probabilidad, y sobre el método de mínimos cuadrados. Markov fue un profesor muy estricto pero también muy claro en sus exposiciones, y demandaba mucho rigor matemático en los argumentos de sus estudiantes. Markov desarrolló su teoría de cadenas de Markov desde un punto de vista matemático, aunque también aplicó su modelo en el análisis de estilos de escritura en poemas. Con sus trabajos sobre cadenas de Markov, fundó una nueva rama de la probabilidad e inició la teoría de los procesos estocásticos. Fuente: Archivo MacTutor, Universidad de St. Andrews.

Ejercicios

Propiedad de Markov

Recordemos que para hacer la notación más corta, a la probabilidad $P(X_n = x_n)$ se le ha escrito como $p(x_n)$, es decir, el subíndice indica también la variable a la que se hace referencia. El significado de la probabilidad condicional $p(x_{n+1} | x_n)$ es análogo.

11. Demuestre que la propiedad de Markov (3.1) es equivalente a cada una de las siguientes condiciones.

- a) Esta condición establece que la distribución conjunta queda especificada a través de la distribución inicial y las probabilidades de transición en un paso: Para cualesquiera estados x_0, \dots, x_n ,

$$p(x_0, x_1, \dots, x_n) = p(x_0) p(x_1 | x_0) \cdots p(x_n | x_{n-1}).$$

- b) Esta condición establece que el futuro, sin importar lo distante que se encuentre, depende sólomente del último momento observado: Para cualesquiera enteros $n, m \geq 1$,

$$p(x_{n+m} | x_0, \dots, x_n) = p(x_{n+m} | x_n).$$

- c) Esta es la condición de Markov para tiempos no necesariamente consecutivos: Para cualesquiera enteros $0 \leq n_1 < \cdots < n_{m+1}$,

$$p(x_{n_{m+1}} | x_{n_1}, \dots, x_{n_m}) = p(x_{n_{m+1}} | x_{n_m}).$$

- d) Esta condición expresa la independencia entre el pasado y el futuro cuando se conoce el presente: Para cualesquiera enteros $0 < k < n$,

$$p(x_0, \dots, x_{k-1}, x_{k+1}, \dots, x_n | x_k) = \frac{p(x_0, \dots, x_{k-1} | x_k) p(x_{k+1}, \dots, x_n | x_k)}{p(x_k)}.$$

- e) En esta condición se consideran varios eventos en el futuro: Para cualesquiera enteros $n, m \geq 1$,

$$p(x_{n+m}, \dots, x_{n+1} | x_0, \dots, x_n) = p(x_{n+1} | x_n) \cdots p(x_{n+m} | x_{n+m-1}).$$

Matrices estocásticas

12. Demuestre que
- a) si P y Q dos matrices estocásticas (o doblemente estocásticas) de la misma dimensión, entonces PQ también es estocástica (o doblemente estocástica).
 - b) si P es estocástica (o doblemente estocástica), entonces para cualquier entero positivo n , la potencia P^n también es estocástica (o doblemente estocástica).
13. Demuestre que una matriz cuadrada P puede ser tal que la potencia P^n sea estocástica para algún entero $n \geq 2$, sin ser P misma estocástica.
14. Demuestre que si una matriz estocástica P es simétrica, entonces para cualquier $n \geq 1$, la potencia P^n también es simétrica.
15. Demuestre que toda matriz estocástica simétrica es doblemente estocástica.
16. Demuestre que toda matriz estocástica finita tiene siempre al número uno como valor propio.

Probabilidades de transición

17. Considere una cadena de Markov de dos estados: 0 y 1, con matriz de probabilidades de transición como aparece abajo. Suponga que la cadena tiene una distribución inicial dada por el vector $(0.6, 0.4)$. Encuentre $P(X_1 = 0)$, $P(X_1 = 1)$, $P(X_2 = 0)$ y $P(X_2 = 1)$.

$$P = \begin{pmatrix} 0.2 & 0.8 \\ 0.5 & 0.5 \end{pmatrix}.$$

18. Considere una cadena de Markov de dos estados: 0 y 1, con distribución inicial $(1/2, 1/2)$, y con matriz de probabilidades de transición dada como aparece abajo. Encuentre $P(X_1 = 0)$, $P(X_1 = 1)$, $P(X_5 = 1 | X_4 = 0)$, $P(X_3 = 0 | X_1 = 0)$, $P(X_{100} = 0 | X_{98} = 0)$, $P(X_1 = 0 | X_2 = 0)$, $P(X_1 = 1 | X_2 = 1, X_3 = 1)$, $P(X_3 = 0, X_2 = 0 | X_1 = 0)$.

$$P = \begin{pmatrix} 1/3 & 2/3 \\ 3/4 & 1/4 \end{pmatrix}.$$

19. *Ecuación de Chapman-Kolmogorov para cadenas no homogéneas.* Considere una cadena de Markov con probabilidades de transición no necesariamente homogéneas $p_{ij}(n, m) = P(X_m = j | X_n = i)$, para $n \leq m$. Demuestre que para cualquier tiempo u tal que $n < u < m$,

$$p_{ij}(n, m) = \sum_k p_{ik}(n, u) p_{kj}(u, m).$$

20. Demuestre o proporcione un contraejemplo.

a) $p_{ij}(n) \geq p_{ij}(1) p_{jj}(n-1)$.
 b) $p_{ij}(n) = 1 - p_{ji}(n)$.

21. Demuestre que

a) $p_{ii}(n) p_{ii}(m) \leq p_{ii}(n+m) \leq p_{ii}(n) p_{ii}(m) + (1 - p_{ii}(n))$.

b) $\sup_n \{p_{ij}(n)\} \leq f_{ij} \leq \sum_{n=1}^{\infty} p_{ij}(n)$.

c) $i \rightarrow j$ si, y sólo si, $f_{ij} > 0$.

d) $i \leftrightarrow j$ si, y sólo si, $f_{ij} f_{ji} > 0$.

e) $\sum_{n=1}^{\infty} p_{ij}(n) = f_{ij} \left[\sum_{n=1}^{\infty} p_{jj}(n) + 1 \right]$.

Cadenas de Markov

22. Para la cadena de Markov de dos estados con distribución inicial (p_0, p_1) , demuestre que para $a + b > 0$,

$$\begin{aligned} P(X_n = 0) &= \frac{b}{a+b} + \left(p_0 - \frac{b}{a+b}\right) (1-a-b)^n, \\ P(X_n = 1) &= \frac{a}{a+b} + \left(p_1 - \frac{a}{a+b}\right) (1-a-b)^n. \end{aligned}$$

23. Encuentre la matriz de probabilidades de transición de la cadena de Markov de cuatro estados dada por el diagrama de la Figura 3.17. A partir de cada posición sólo se permite pasar a los estados como indica el laberinto, con idéntica probabilidad para todas las posibilidades. Suponga que no se permite permanecer en la misma posición al efectuarse un movimiento.



Figura 3.17:

24. Se lanza un dado equilibrado repetidas veces. Sea X_n el número de lanzamientos realizados al tiempo n desde la última vez que apareció el número “6”. Demuestre que $\{X_n\}$ es una cadena de Markov y encuentre las probabilidades de transición.
25. Sea $\{X_n\}$ una cadena de Markov con probabilidades de transición p_{ij} , y sea $m \geq 1$ un entero fijo. Demuestre que los siguientes procesos son también cadenas de Markov y encuentre las correspondientes probabilidades de transición.
- a) X_{n+m} .
 - b) X_{nm} .
26. Demuestre que para cualquier entero $n \geq 1$,
- a) $P(X_n = j) = \sum_i P(X_{n-1} = i) p_{ij}(1)$.
 - b) $P(X_n = j) = \sum_i P(X_0 = i) p_{ij}(n)$.
27. Considere una cadena de Markov con espacio de estados $\{0, 1, \dots, N\}$, y probabilidades de transición tales que para cualquier estado i ,

$$E(X_{n+1} | X_n = i) = \sum_{j=0}^N j p_{ij} = i,$$

es decir, el estado promedio después de una transición es el estado inicial. Demuestre que esta cadena es una martingala y que los estados 0 y N son absorbentes.

28. *Renovaciones discretas.* Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes idénticamente distribuidas, y con valores en el conjunto $\{1, 2, \dots\}$.

Interpretaremos estas variables discretas como los tiempos de vida de artículos puestos en operación uno después de otro en un proceso de renovación a tiempo discreto. Sea $X_0 = 0$ y para cada $n \geq 1$ sea $X_n = \xi_1 + \dots + \xi_n$. Para cada $k = 1, 2, \dots$ defina

$$N_k = \text{máx} \{n \geq 1 : X_n \leq k\}.$$

Si el conjunto indicado es vacío, entonces el máximo se define como cero. La variable N_k cuenta el número de renovaciones efectuadas hasta el tiempo k . Sea $A_k = k - X_{N_k}$, es decir, A_k es la edad del artículo que se encuentra en operación al tiempo k , o bien el tiempo transcurrido desde la última renovación. La notación A proviene del término en inglés *age*. Demuestre que el proceso A_k es una cadena de Markov con espacio de estados $\{0, 1, \dots\}$, y con probabilidades de transición

$$p_{i,0} = \frac{P(X = i + 1)}{P(X \geq i + 1)}, \quad p_{i,i+1} = \frac{P(X \geq i + 2)}{P(X \geq i + 1)}.$$

29. Considere la cadena de inventarios en donde ξ_n tiene distribución uniforme en el conjunto $\{0, 1, 2, 3\}$, con $s = 1$ y $S = 4$. Encuentre la matriz de probabilidades de transición de esta cadena.
30. Sea X_n la cadena de Markov de dos estados. Demuestre que el proceso $Y_n = (X_{n-1}, X_n)$ es una cadena de Markov. Determine el espacio de estados de este nuevo proceso y encuentre la matriz de probabilidades de transición. Generalize este resultado para cualquier cadena.
31. Se colocan N bolas negras en una primera urna, y N bolas blancas en una segunda urna. Se selecciona al azar una bola de cada urna y se intercambian. Este ensayo se repite varias veces. Sea X_n el número de bolas blancas en la primera urna después del n -ésimo ensayo. Justifique que X_n es una cadena de Markov y encuentre la matriz de probabilidades de transición.
32. Sea ξ_0, ξ_1, \dots una sucesión de variables independientes con idéntica distribución $\text{Ber}(p)$. Determine si el proceso $\{X_n\}$ definido a continuación es una cadena de Markov. En caso afirmativo determine el espacio de estados y encuentre la matriz de probabilidades de transición.
 - a) $X_n = \xi_n + \xi_{n-1}$.
 - b) $X_n = X_{n-1} + \xi_n$.
 - c) $X_n = X_{n-1} + \xi_n \pmod{2}$.

33. Sea ξ_1, ξ_2, \dots una sucesión de variables independientes con valores en el conjunto $\{0, 1, 2, 3, 4\}$ y con idéntica distribución dada por $\{p_0, p_1, p_2, p_3, p_4\}$. Determine si el proceso $\{X_n\}$ definido a continuación es una cadena de Markov. En caso afirmativo determine el espacio de estados y encuentre la matriz de probabilidades de transición.

$$\begin{aligned} X_0 &= 0, \\ X_{n+1} &= X_n + \xi_{n+1} \pmod{5}, \text{ para } n \geq 0. \end{aligned}$$

34. *Modificación de la cadena de Ehrenfest.* Considere el esquema de dos urnas como en la cadena de Ehrenfest con un total de N bolas. Suponga ahora que un ensayo consiste en que cada una de las bolas se cambia de urna con una distribución de probabilidad especificada p_1, \dots, p_N , sin importar la posición de las bolas. Sea X_n el número de bolas en una de las urnas después del n -ésimo ensayo. ¿Es esta una cadena de Markov? En caso afirmativo encuentre la matriz de probabilidades de transición.
35. Sea $\{X_n\}$ una cadena de Markov con espacio de estados E . Defina el proceso $\{Y_n\}$ de la siguiente forma: $Y_n = (X_n, X_{n+1})$. Determine el espacio de estados del proceso $\{Y_n\}$, demuestre que es una cadena de Markov y encuentre sus probabilidades de transición en términos de las probabilidades de transición de $\{X_n\}$.

Comunicación

36. Dibuje un diagrama de transición y encuentre las clases de comunicación para cada una de las siguientes cadenas de Markov.

$$\begin{aligned} a) \quad P &= \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix} & b) \quad P &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} \\ c) \quad P &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \end{aligned}$$

37. Considere una cadena de Markov con n estados. Demuestre que si $i \rightarrow j$, entonces es posible pasar de i a j en a lo sumo $n - 1$ pasos. Suponga $i \neq j$.

Clases cerradas

38. Demuestre que toda cadena de Markov tiene por lo menos una clase de comunicación cerrada.
39. Demuestre que la colección de estados \mathcal{C} es cerrada si, y sólo si, alguna de las siguientes condiciones se cumple. Para cualesquiera $i \in \mathcal{C}$ y $j \notin \mathcal{C}$,
- $p_{ij}(n) = 0$, para cada $n \geq 1$.
 - $p_{ij}(1) = 0$.
40. Demuestre que
- toda clase de comunicación recurrente es cerrada.
 - la unión de dos clases de comunicación recurrentes es una clase cerrada.
 - la unión de todas las clases de comunicación recurrentes de una cadena de Markov es una clase cerrada.

Periodo

41. Dibuje un diagrama de transición, determine las clases de comunicación y calcule el periodo de cada uno de los estados de las siguientes cadenas de Markov.

$$a) \quad P = \begin{pmatrix} 1/2 & 1/2 & 0 & \\ 1/2 & 1/2 & 0 & \\ 0 & 1/2 & 1/2 & \end{pmatrix} \quad b) \quad P = \begin{pmatrix} 1/3 & 1/3 & 1/3 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix}$$

$$c) \quad P = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 1/2 & 0 & 1/2 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

42. Hemos demostrado que el periodo es una propiedad de clase, es decir, dos estados que pertenecen a la misma clase de comunicación tienen el mismo periodo. El recíproco de tal afirmación es en general falso, es decir, dos estados pueden tener el mismo periodo y sin embargo no ser comunicantes. Proporcione un ejemplo de tal situación.

Tiempo y probabilidad de primera visita

43. Demuestre que $i \rightarrow j$, si y sólo si, $f_{ij} > 0$.
44. Sea j un estado absorbente. Demuestre que para cada $n \geq 1$, $p_{ij}(n) = P(\tau_{ij} \leq n)$.

Tiempo medio de absorción

45. Se efectúa una sucesión de lanzamientos independientes de una moneda equilibrada con resultados A y S . Encuentre el número esperado de lanzamientos para que aparezca la secuencia particular
- “ S ”
 - “ SA ”
 - “ SAS ”

Recurrencia y transitoriedad

46. Encuentre las clases de comunicación de la siguiente cadena de Markov. Encuentre además los periodos, y clasifique cada clase de comunicación como transitoria o recurrente.

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

47. Demuestre que todo estado absorbente es recurrente.
48. Demuestre que el estado 0 de la cadena de Markov dada por el diagrama de transición de la Figura 3.18(a) es transitorio.
49. Demuestre que el estado 0 de la cadena de Markov dada por el diagrama de transición de la Figura 3.18(b) es recurrente. Suponga que $0 < \alpha < 1$ y $0 < \beta < 1$. Concluya que todos los estados de esta cadena son recurrentes.
50. Demuestre que si i es un estado recurrente e $i \rightarrow j$, entonces $j \rightarrow i$.

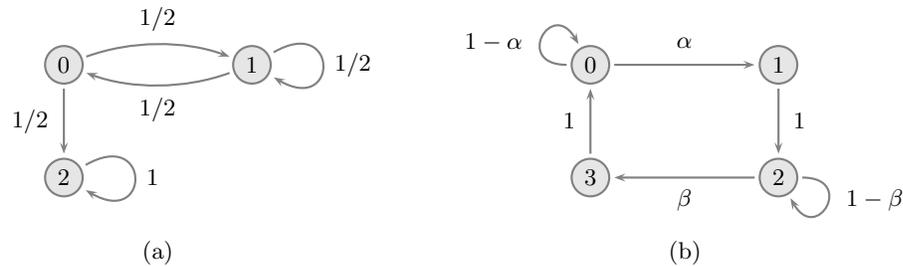


Figura 3.18:

Tiempo medio de recurrencia

51. Considere una sucesión de ensayos independientes Bernoulli con resultados A y B , y con probabilidades $P(A) = p$ y $P(B) = q = 1 - p$. Calcule el número promedio de ensayos para la primera ocurrencia de la secuencia “ AB ”.

Recurrencia positiva y nula

52. Determine si la cadena de racha de éxitos es recurrente positiva o nula.
53. Sea $\{X_n\}$ una cadena de Markov con espacio de estados $\{0, 1, 2, \dots\}$ y probabilidades de transición

$$\begin{aligned} p_{0,i} &= a_i \quad \text{para } i = 0, 1, 2, \dots \\ p_{i,i-1} &= 1 \quad \text{para } i = 1, 2, 3, \dots \end{aligned}$$

en donde $a_0 + a_1 + \dots = 1$. Encuentre condiciones suficientes sobre estas probabilidades para que la cadena sea

- irreducible.
 - recurrente.
 - recurrente positiva.
54. Sea P una matriz doblemente estocástica. Demuestre directamente que
- si P es finita, entonces todos los estados son recurrentes positivos.
 - si P es infinita e irreducible, entonces todos los estados son transitorios o recurrentes nulos.

Distribuciones estacionarias

55. Encuentre todas las distribuciones estacionarias, si existen, para cada una de las siguientes cadenas de Markov.

$$a) \quad P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad b) \quad P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

56. Determine si existe alguna distribución estacionaria para la siguiente matriz estocástica. En caso afirmativo encuentre todas las distribuciones estacionarias.

$$P = \begin{pmatrix} 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \\ 1-p & p & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

57. Demuestre que la siguiente cadena de Markov tiene un número infinito de distribuciones estacionarias.

$$P = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{pmatrix}$$

58. Demuestre que la cadena del jugador tiene una infinidad de distribuciones estacionarias.
59. Demuestre que la cadena de variables aleatorias independientes tiene una única distribución estacionaria dada por $(\pi_0, \pi_1, \dots) = (a_0, a_1, \dots)$.
60. Demuestre que toda matriz doblemente estocástica, aperiódica, finita e irreducible tiene una única distribución estacionaria dada por la distribución uniforme.
61. Demuestre que si la distribución uniforme es una distribución estacionaria para una cadena de Markov finita, entonces la correspondiente matriz de probabilidades de transición es doblemente estocástica.
62. Considere una cadena de Markov a tiempo discreto con espacio de estados finito y tal que para cada estado j los siguientes límites existen y no dependen del estado i ,

$$\lim_{n \rightarrow \infty} p_{ij}(n) = \pi_j.$$

Demuestre que $\pi = \{\pi_j\}$ es una distribución estacionaria.

63. Justifique la existencia y unicidad, y encuentre la distribución estacionaria para una caminata aleatoria simple no simétrica sobre el conjunto $\{0, 1, \dots, n\}$, en donde los estados 0 y n son reflejantes, es decir: $p_{00} = q$, $p_{01} = p$, $p_{nn} = p$, $p_{n,n-1} = q$. Las otras probabilidades de transición son: $p_{i,i+1} = p$ y $p_{i,i-1} = q$ para $i = 1, 2, \dots, n-1$.
64. Considere una caminata aleatoria $\{X_n\}$ sobre el conjunto $\{0, 1, \dots, N-1, N\}$ en donde los estados 0 y N son reflejantes, es decir, $P(X_{n+1} = 1 | X_n = 0) = 1$ y $P(X_{n+1} = N-1 | X_n = N) = 1$. El resto de las probabilidades de transición son $P(X_{n+1} = i+1 | X_n = i) = p$ y $P(X_{n+1} = i-1 | X_n = i) = q$, para $i = 1, 2, \dots, N-1$. Véase la Figura 3.19. Calcule el número promedio de visitas que la caminata realiza a cada uno de sus estados.

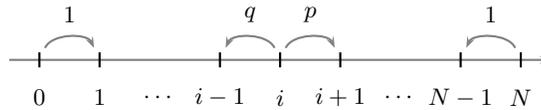


Figura 3.19:

Distribuciones límite

65. Calcule la distribución límite, cuando existe, de las siguientes cadenas de Markov.

$$a) \quad P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} \quad b) \quad P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \end{pmatrix}$$

$$c) \quad P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$$

66. Una persona se traslada todos los días de su casa a la oficina en la mañana, y de la oficina a su casa por la tarde. Esta persona tiene un coche que usa en cualquiera de estos dos viajes en caso de lluvia, siempre y cuando tenga

el coche disponible. No siempre tiene el coche disponible pues ha decidido dejarlo en la casa o en la oficina e irse caminando cuando al salir de alguno de estos lugares no está lloviendo. La probabilidad de lluvia por la mañana o por la tarde es $p > 0$, independiente un evento del otro.

- a) Calcule la proporción de viajes a largo plazo en los cuales la persona se moja por la lluvia.
- b) Conteste la misma pregunta cuando la persona posee r coches.

Cadenas regulares

67. Determine si las siguientes matrices estocásticas son regulares.

$$a) P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad b) P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

68. ¿Cuántas potencias de una matriz se necesitan calcular como máximo para verificar que es regular?
69. Sean $\{X_n\}$ y $\{Y_n\}$ dos cadenas de Markov independientes y regulares. Demuestre que $Z_n = (X_n, Y_n)$ es una cadena de Markov regular y encuentre sus probabilidades de transición.

Cadenas reversibles

70. Resolviendo la ecuación de balance detallado (3.10), demuestre que la cadena de dos estados es reversible y encuentre nuevamente la distribución estacionaria $(\pi_0, \pi_1) = (b/(a+b), a/(a+b))$, cuando $a+b > 0$.

CAPÍTULO 4

El proceso de Poisson

En el presente y en el siguiente capítulo estudiaremos el modelo de cadena de Markov a tiempo continuo. Como una introducción a la teoría general que se expondrá más adelante, en este capítulo estudiaremos uno de los ejemplos más importantes de este tipo de modelos: el proceso de Poisson. Definiremos este proceso de varias formas equivalentes, y estudiaremos algunas de sus propiedades, sus generalizaciones y algunas de sus aplicaciones. El proceso de Poisson es un modelo relevante tanto en las aplicaciones como en la teoría general de los procesos estocásticos.

4.1. Proceso de Poisson

Suponga que un mismo evento ocurre repetidas veces de manera aleatoria a lo largo del tiempo como se muestra en la Figura 4.1. Tal evento puede ser por ejemplo la llegada de una reclamación a una compañía aseguradora, o la recepción de una llamada a un conmutador, o la llegada de un cliente a una ventanilla para solicitar algún servicio, o los momentos en que una cierta maquinaria requiere reparación, etcétera. Suponga que las variables aleatorias T_1, T_2, \dots representan los tiempos que transcurren entre una ocurrencia del evento y la siguiente ocurrencia. Suponga que estos tiempos son independientes uno del otro, y que cada uno de ellos tiene distribución $\exp(\lambda)$. Se define el proceso de Poisson al tiempo t como el número



Figura 4.1:

de ocurrencias del evento que se han observado hasta ese instante t . Esta es una definición constructiva de este proceso y la formalizaremos a continuación. Más adelante enunciaremos otras definiciones axiomáticas equivalentes de este mismo proceso.

Definición. (I) Sea T_1, T_2, \dots una sucesión de variables aleatorias independientes cada una con distribución $\exp(\lambda)$. El *proceso de Poisson* de parámetro λ es el proceso a tiempo continuo $\{X_t : t \geq 0\}$ definido de la siguiente manera:

$$X_t = \max \{n \geq 1 : T_1 + \dots + T_n \leq t\}.$$

Se postula además que el proceso inicia en cero y para ello se define $\max \emptyset = 0$. En palabras, la variable X_t es el entero n máximo tal que $T_1 + \dots + T_n$ es menor o igual a t , y ello equivale a contar el número de eventos ocurridos hasta el tiempo t . A este proceso se le llama proceso de Poisson *homogéneo*, tal adjetivo se refiere a que el parámetro λ no cambia con el tiempo, es decir, es homogéneo en el tiempo.

Una trayectoria típica de este proceso puede observarse en la Figura 4.2, la cual es no decreciente, constante por partes, continua por la derecha y con límite por la izquierda. A los tiempos T_1, T_2, \dots se les llama *tiempos de estancia*, o *tiempos de interarribo*, y corresponden a los tiempos que transcurren entre un salto del proceso y el siguiente salto. Hemos supuesto que estos tiempos son independientes y que todos tienen distribución $\exp(\lambda)$. En consecuencia, la variable $W_n = T_1 + \dots + T_n$ tiene distribución $\text{gamma}(n, \lambda)$. Esta variable representa el tiempo real en el que se observa la ocurrencia del n -ésimo evento. Observe la igualdad de eventos $(X_t \geq n) = (W_n \leq t)$, esto equivale a decir que al tiempo t han ocurrido por lo menos n eventos si, y sólo si, el n -ésimo evento ocurrió antes de t .

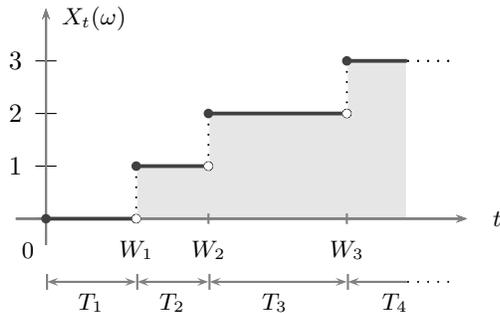


Figura 4.2: El proceso de Poisson y los tiempos de ocurrencia de eventos.

Una de las características sobresalientes de este proceso es que puede encontrarse explícitamente la distribución de probabilidad de la variable X_t para cualquier valor de t . La respuesta es la distribución Poisson, y es de allí de donde el proceso adquiere su nombre.

Proposición. La variable X_t tiene distribución Poisson(λt), es decir, para cualquier $t > 0$, y para $n = 0, 1, \dots$

$$P(X_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Demostración. Como W_n tiene distribución gamma(n, λ), su función de distribución es, para $t > 0$,

$$P(W_n \leq t) = 1 - e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}.$$

Entonces para cualquier $t > 0$ y para cada $n = 0, 1, \dots$

$$\begin{aligned} P(X_t = n) &= P(X_t \geq n) - P(X_t \geq n + 1) \\ &= P(W_n \leq t) - P(W_{n+1} \leq t) \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \end{aligned}$$

□

Entonces, dado que X_t tiene una distribución Poisson(λt), se tiene que $E(X_t) = \lambda t$, y $\text{Var}(X_t) = \lambda t$. Por lo tanto λt es el promedio de observaciones o registros del evento de interés en el intervalo $[0, t]$. Así, mientras mayor es la longitud del intervalo de observación, mayor es el promedio de observaciones realizadas, y mayor también la incertidumbre del número de observaciones.

Pérdida de memoria y sus consecuencias. Una de las propiedades que caracterizan de manera única a la distribución exponencial dentro del conjunto de distribuciones continuas es que satisface la propiedad de *pérdida de memoria*, esto es, si T tiene distribución $\exp(\lambda)$, entonces para cualesquiera tiempos $s, t > 0$ se cumple la igualdad

$$P(T > t + s | T > s) = P(T > t).$$

En otras palabras, condicionada al evento $(T > s)$, la variable $T - s$ sigue teniendo distribución $\exp(\lambda)$. Esto significa que el proceso de Poisson puede estudiarse a partir de cualquier tiempo $s \geq 0$, teniendo las mismas propiedades a partir de ese momento hacia adelante. Esto es, si uno toma cualquier tiempo $s \geq 0$ y empieza a contar desde cero las ocurrencias de eventos a partir de ese momento, lo que uno obtiene es nuevamente un proceso de Poisson. Entonces, para cualquier $t > 0$ y para $n = 0, 1, \dots$

$$P(X_{t+s} - X_s = n) = P(X_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \quad (4.1)$$

Esta propiedad caracteriza de manera única a este proceso y de ella pueden derivarse todas las propiedades del Proceso de Poisson, incluida la propiedad de Markov, la cual demostraremos a continuación.

Proposición. El proceso de Poisson satisface las siguientes propiedades.

- a) Es un proceso de Markov.
- b) Tiene incrementos independientes.
- c) Tiene incrementos estacionarios.
- d) Para cualesquiera $s, t > 0$, $X_{t+s} - X_s \sim \text{Poisson}(\lambda t)$.
- e) Para cualesquiera $s, t \geq 0$, y enteros $0 \leq i \leq j$, las probabilidades de transición son

$$P(X_{t+s} = j \mid X_s = i) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}. \quad (4.2)$$

Demostración.

a) Considere la probabilidad condicional $P(X_{t_n} = x_n \mid X_{t_1} = x_1, \dots, X_{t_{n-1}} = x_{n-1})$, para cualesquiera n tiempos $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, y estados $0 \leq x_1 \leq \dots \leq x_n$. Entonces al tiempo t_{n-1} ha habido x_{n-1} ocurrencias del evento de interés. A partir de ese momento inicia un nuevo proceso de Poisson y para que al tiempo t_n hayan x_n ocurrencias es necesario que en el intervalo de tiempo $(t_{n-1}, t_n]$ hayan ocurrido $x_n - x_{n-1}$ eventos. Esto es exactamente $P(X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1})$.

b) Considere cualesquiera n tiempos $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, y cualesquiera estados x_1, \dots, x_n . Por comodidad llamaremos s_n a la suma $x_1 + \dots + x_n$, para cada $n \geq 1$.

Por la propiedad de Markov y por (4.1),

$$\begin{aligned}
 &P(X_{t_1} = x_1, X_{t_2} - X_{t_1} = x_2, \dots, X_{t_n} - X_{t_{n-1}} = x_n) \\
 &= P(X_{t_1} = s_1, X_{t_2} = s_2, \dots, X_{t_n} = s_n) \\
 &= P(X_{t_1} = s_1) P(X_{t_2} = s_2 | X_{t_1} = s_1) \cdots \\
 &\quad \cdots P(X_{t_n} = s_n | X_{t_{n-1}} = s_{n-1}) \\
 &= P(X_{t_1} = x_1) P(X_{t_2} - X_{t_1} = x_2) \cdots P(X_{t_n} - X_{t_{n-1}} = x_n).
 \end{aligned}$$

Las propiedades c), d) y e) son consecuencia inmediata de (4.1). □

La expresión (4.2) establece de manera evidente que las probabilidades de transición son estacionarias en el tiempo, es decir, no dependen del parámetro s , y se escriben simplemente como $p_{ij}(t)$. Pero también es interesante observar que (4.2) dice que estas probabilidades son estacionarias en el espacio, es decir, dependen de los estados i y j únicamente a través de la diferencia $j - i$. En símbolos, $p_{ij}(t) = p_{0,j-i}(t)$, para $j \geq i$.

■ Ejemplo: Paradoja del autobús. Suponga que la llegada de autobuses a una estación se modela mediante un proceso de Poisson de parámetro λ , es decir, el tiempo que transcurre entre la llegada de un autobús y el siguiente es una variable aleatoria con distribución $\exp(\lambda)$. Suponga que el tiempo es medido en minutos. La propiedad de pérdida de memoria en este contexto puede interpretarse de la siguiente forma: Un persona ha llegado a la estación y ha esperado s minutos sin que un autobús aparezca. La probabilidad de que tenga que esperar más de t minutos adicionales es la misma que la probabilidad de espera de más de t minutos para una persona que acaba de llegar a la estación! ○

■ Ejemplo: La suma de dos procesos de Poisson independientes es un proceso de Poisson. Sean $\{X_t\}$ y $\{Y_t\}$ dos procesos de Poisson independientes de parámetros λ_1 y λ_2 respectivamente. Demostraremos que el proceso suma $\{X_t + Y_t\}$ es un proceso de Poisson de parámetro $\lambda_1 + \lambda_2$.

Denotaremos por T_1, T_2, \dots a los tiempos de interarribo del proceso suma. La forma en la que se obtienen estos tiempos se muestra en la Figura 4.3. Demostraremos que estas variables aleatorias son independientes con idéntica distribu-

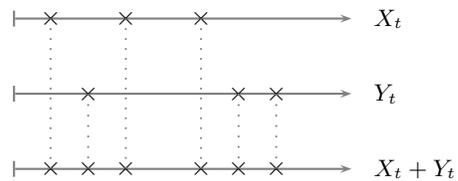


Figura 4.3: Suma de dos procesos de Poisson.

ción $\exp(\lambda_1 + \lambda_2)$. Para cualesquiera tiempos t_1, \dots, t_n , el evento $(T_1 > t_1, T_2 > t_2, \dots, T_n > t_n)$ puede expresarse como

$$((X + Y)_{[0, t_1]} = 0, (X + Y)_{[T_1, T_1 + t_2]} = 0, \dots, (X + Y)_{[T_{n-1}, T_{n-1} + t_n]} = 0),$$

esto es,

$$\begin{aligned} & (X_{[0, t_1]} = 0) \cap (Y_{[0, t_1]} = 0) \\ & \cap (X_{[T_1, T_1 + t_2]} = 0) \cap (Y_{[T_1, T_1 + t_2]} = 0) \\ & \dots \\ & \cap (X_{[T_{n-1}, T_{n-1} + t_n]} = 0) \cap (Y_{[T_{n-1}, T_{n-1} + t_n]} = 0). \end{aligned}$$

Por la independencia de los procesos y la propiedad de incrementos independientes de cada uno de ellos, la probabilidad de este evento es

$$\begin{aligned} & P(X_{[0, t_1]} = 0) P(Y_{[0, t_1]} = 0) \\ & P(X_{[T_1, T_1 + t_2]} = 0) P(Y_{[T_1, T_1 + t_2]} = 0) \\ & \dots \\ & P(X_{[T_{n-1}, T_{n-1} + t_n]} = 0) P(Y_{[T_{n-1}, T_{n-1} + t_n]} = 0). \end{aligned}$$

Por lo tanto,

$$P(T_1 > t_1, T_2 > t_2, \dots, T_n > t_n) = e^{-(\lambda_1 + \lambda_2)t_1} e^{-(\lambda_1 + \lambda_2)t_2} \dots e^{-(\lambda_1 + \lambda_2)t_n}.$$

Esta identidad demuestra que las variables T_1, T_2, \dots, T_n son independientes con idéntica distribución $\exp(\lambda_1 + \lambda_2)$. \circ

Distribuciones asociadas al proceso de Poisson. Además de las distribuciones exponencial y gamma ya mencionadas, existen otras distribuciones de probabilidad que surgen al estudiar ciertas características del proceso de Poisson. Mencionaremos a continuación algunas de ellas.

Proposición. Dado el evento $(X_t = n)$, el vector de tiempos reales (W_1, \dots, W_n) tiene la misma distribución que el vector de las estadísticas de orden $(Y_{(1)}, \dots, Y_{(n)})$ de una muestra aleatoria Y_1, \dots, Y_n de la distribución uniforme en el intervalo $[0, t]$, es decir,

$$f_{W_1, \dots, W_n | X_t}(w_1, \dots, w_n | n) = \begin{cases} \frac{n!}{t^n} & \text{si } 0 < w_1 < \dots < w_n < t, \\ 0 & \text{otro caso.} \end{cases}$$

Demostración. La fórmula general para la función de densidad conjunta de las estadísticas de orden $Y_{(1)}, \dots, Y_{(n)}$ de una muestra aleatoria Y_1, \dots, Y_n de una distribución con función de densidad $f(y)$ es, para $y_1 < \dots < y_n$,

$$f_{Y_{(1)}, \dots, Y_{(n)}}(y_1, \dots, y_n) = n! f(y_1) \cdots f(y_n).$$

Cuando la función de densidad $f(y)$ es la uniforme en el intervalo $[0, t]$, esta función de densidad conjunta es la que aparece en el enunciado. Demostraremos que la distribución conjunta de las variables W_1, \dots, W_n , condicionada al evento $(X_t = n)$ también tiene esta misma función de densidad. Usaremos nuevamente la identidad de eventos $(X_t \geq n) = (W_n \leq t)$. Para tiempos $0 < w_1 < \dots < w_n < t$, se tiene que

$$\begin{aligned} & f_{W_1, \dots, W_n | X_t}(w_1, \dots, w_n | n) \\ &= \frac{\partial^n}{\partial w_1 \cdots \partial w_n} P(W_1 \leq w_1, W_2 \leq w_2, \dots, W_n \leq w_n | X_t = n) \\ &= \frac{\partial^n}{\partial w_1 \cdots \partial w_n} P(X_{w_1} \geq 1, X_{w_2} \geq 2, \dots, X_{w_n} \geq n | X_t = n) \\ &= \frac{\partial^n}{\partial w_1 \cdots \partial w_n} P(X_t - X_{w_n} = 0, X_{w_n} - X_{w_{n-1}} = 1, \dots \\ &\quad \dots, X_{w_2} - X_{w_1} = 1, X_{w_1} = 1) / P(X_t = n) \\ &= \frac{\partial^n}{\partial w_1 \cdots \partial w_n} e^{-\lambda(t-w_n)} e^{-\lambda(w_n-w_{n-1})} \lambda(w_n - w_{n-1}) \cdots \\ &\quad \cdots e^{-\lambda(w_2-w_1)} \lambda(w_2 - w_1) e^{-\lambda w_1} \lambda w_1 / [e^{-\lambda t} (\lambda t)^n / n!] \\ &= \frac{\partial^n}{\partial w_1 \cdots \partial w_n} n! (w_n - w_{n-1}) \cdots (w_2 - w_1) w_1 / t^n \\ &= n! / t^n. \end{aligned}$$

Observe que bajo el signo de derivada, la probabilidad del evento $(X_{w_1} \geq 1, X_{w_2} \geq 2, \dots, X_{w_n} \geq n, X_t = n)$, que aparece en la segunda igualdad, es idéntica a la probabilidad de $(X_t - X_{w_n} = 0, X_{w_n} - X_{w_{n-1}} = 1, \dots, X_{w_2} - X_{w_1} = 1, X_{w_1} = 1)$, pues si alguna de estas identidades (exceptuando la primera) fuera distinta de uno, la derivada se anula. \square

La fórmula anterior nos provee de un mecanismo para obtener simulaciones por computadora de las trayectorias del proceso Poisson. El procedimiento es el siguiente: Fije un valor t y asigne un valor para λ . Genere un valor al azar de la variable X_t con distribución $\text{Poisson}(\lambda t)$. Suponga $X_t = n$. A continuación genere n valores u_1, \dots, u_n de la distribución $\text{unif}(0, t)$, y ordene estos valores de menor a mayor: $u_{(1)} \leq \dots \leq u_{(n)}$. Estos son los tiempos en donde la trayectoria tiene saltos. De esta forma pueden obtenerse trayectorias como la que se muestra en la Figura 4.2.

El siguiente resultado establece una forma de obtener la distribución binomial a partir del proceso de Poisson. Suponga que durante el intervalo de tiempo $[0, t]$ se han observado n ocurrencias del evento de interés, es decir, el evento $(X_t = n)$ ha ocurrido. La pregunta es ¿cuántos de estos eventos ocurrieron en el subintervalo $[0, s]$? Demostraremos a continuación que esta variable aleatoria tiene distribución binomial $(n, s/t)$.

Proposición. Sean s y t tiempos tales que $0 < s < t$, y sean k y n enteros tales que $0 \leq k \leq n$. Entonces

$$P(X_s = k | X_t = n) = \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}.$$

Demostración. Por la definición de probabilidad condicional,

$$P(X_s = k | X_t = n) = \frac{P(X_t = n | X_s = k) P(X_s = k)}{P(X_t = n)}.$$

Substituyendo estas probabilidades y simplificando se obtiene el resultado. \square

Recordando que la suma de dos procesos de Poisson independientes es nuevamente un proceso de Poisson, se puede comprobar fácilmente el siguiente resultado.

Proposición. Sean $X_1(t)$ y $X_2(t)$ dos procesos de Poisson independientes con parámetros λ_1 y λ_2 respectivamente, y sean k y n enteros tales que $0 \leq k \leq n$. Entonces

$$P(X_1(t) = k \mid X_1(t) + X_2(t) = n) = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^{n-k}.$$

Demostración. Por la hipótesis de independencia entre los procesos,

$$\begin{aligned} P(X_1(t) = k \mid X_1(t) + X_2(t) = n) &= P(X_1(t) = k, X_1(t) + X_2(t) = n) / P(X_1(t) + X_2(t) = n) \\ &= P(X_1(t) = k, X_2(t) = n - k) / P(X_1(t) + X_2(t) = n) \\ &= P(X_1(t) = k) P(X_2(t) = n - k) / P(X_1(t) + X_2(t) = n). \end{aligned}$$

Substituyendo estas probabilidades se obtiene el resultado. \square

4.2. Definiciones alternativas

La definición (I) de proceso de Poisson es constructiva pues a partir de los tiempos de interarribo se construye el proceso de conteo correspondiente. Existen otras formas equivalentes y un tanto axiomáticas de definir a este proceso. Revisaremos y comentaremos a continuación dos de ellas. Una de las ventajas de contar con estas definiciones alternativas es que para demostrar que un cierto proceso es de Poisson se puede tomar cualquiera de las definiciones a conveniencia.

Definición. (II) Un *proceso de Poisson* de parámetro $\lambda > 0$ es un proceso a tiempo continuo $\{X_t : t \geq 0\}$, con espacio de estados $\{0, 1, \dots\}$, y que cumple las siguientes propiedades:

- a) $X_0 = 0$.
- b) Tiene incrementos independientes y estacionarios.
- c) Para cualquier $t \geq 0$, y cuando $h \searrow 0$,
 - i) $P(X_{t+h} - X_t \geq 1) = \lambda h + o(h)$.
 - ii) $P(X_{t+h} - X_t \geq 2) = o(h)$.

Esta definición hace uso de las *probabilidades infinitesimales* del proceso y ello tiene algunas ventajas desde el punto de vista de la interpretación de lo que sucede en un intervalo infinitesimal de tiempo $(t, t+h]$. El proceso empieza en cero y por el tercer postulado la probabilidad de que pase al estado uno al final de un intervalo de tiempo pequeño $[0, h]$ es $\lambda h + o(h)$, la probabilidad de que el proceso no sufra ningún cambio en dicho intervalo es $1 - \lambda h + o(h)$, y finalmente la probabilidad de que el proceso tenga dos o más incrementos en tal intervalo es $o(h)$. Es decir, en un intervalo cualquiera de longitud infinitesimal h sólo pueden ocurrir dos situaciones: que haya un incremento o que no lo haya.

■ Ejemplo. (Demostración de que la variable $X_{t+s} - X_s$ tiene distribución $\text{Poisson}(\lambda t)$ a partir de los postulados de la definición (II)). Se define $p_n(t) = P(X_t = n)$ y se considera cualquier $h > 0$. Denotaremos por $p_n(t, t+h]$ a la probabilidad $P(X_{t+h} - X_t = n)$. Por la hipótesis de independencia, para $t \geq 0$ y cuando $h \searrow 0$,

$$\begin{aligned} p_0(t+h) &= p_0(t) p_0(t, t+h] \\ &= p_0(t) (1 - \lambda h + o(h)). \end{aligned}$$

Haciendo $h \rightarrow 0$ se obtiene la ecuación diferencial $p_0'(t) = -\lambda p_0(t)$, cuya solución es $p_0(t) = c e^{-\lambda t}$, en donde la constante c es uno por la condición inicial $p_0(0) = 1$. Ahora encontraremos $p_n(t)$ para $n \geq 1$. Nuevamente por independencia,

$$\begin{aligned} p_n(t+h) &= p_n(t) p_0(t, t+h] + p_{n-1}(t) p_1(t, t+h] + o(h) \\ &= p_n(t) (1 - \lambda h + o(h)) + p_{n-1}(t) (\lambda h + o(h)) + o(h). \end{aligned}$$

Entonces $p_n'(t) = -\lambda p_n(t) + \lambda p_{n-1}(t)$, con condición inicial $p_n(0) = 0$ para

$n \geq 1$. Definiendo $q_n(t) = e^{\lambda t} p_n(t)$ la ecuación diferencial se transforma en $q'_n(t) = \lambda q_{n-1}(t)$, con condiciones $q_n(0) = 0$ y $q_0(t) = 1$. Esta ecuación se resuelve iterativamente primero para $q_1(t)$, después para $q_2(t)$, y así sucesivamente, en general $q_n(t) = (\lambda t)^n/n!$. Por lo tanto $p_n(t) = e^{-\lambda t}(\lambda t)^n/n!$. Esto significa que X_t tiene distribución Poisson(λt). Debido al postulado de incrementos estacionarios, la variable $X_{t+s} - X_s$ tiene la misma distribución que X_t . \circ

Definición. (III) Un *proceso de Poisson* de parámetro $\lambda > 0$ es un proceso de tiempo continuo $\{X_t : t \geq 0\}$ con espacio de estados $\{0, 1, \dots\}$, con trayectorias no decrecientes y que cumple las siguientes propiedades:

- a) $X_0 = 0$.
- b) Tiene incrementos independientes.
- c) $X_{t+s} - X_s \sim \text{Poisson}(\lambda t)$, para cualesquiera $s \geq 0, t > 0$.

Esta es posiblemente la definición más frecuente del proceso de Poisson. A partir de ella inmediatamente sabemos que X_t tiene distribución Poisson(λt). La independencia de los incrementos es explícita, y la estacionariedad de los mismos aparece de manera implícita en el tercer postulado. Para demostrar la equivalencia con la definición (I), el reto consiste en definir los tiempos de interarribo y demostrar que éstos son variables aleatorias independientes con distribución exponencial(λ).

■ Ejercicio. Usando la definición (III), demuestre que la suma de dos procesos de Poisson independientes es nuevamente un proceso Poisson. \cdot

Proposición. Las definiciones (I), (II) y (III) son equivalentes.

Demostración. Hemos demostrado que (II) \Rightarrow (III). El recíproco es inmediato pues la propiedad de incrementos independientes y estacionarios aparecen en ambas

definiciones, y para cualquier $t \geq 0$ y $h > 0$,

$$\begin{aligned} P(X_{t+h} - X_t \geq 1) &= 1 - P(X_{t+h} - X_t = 0) \\ &= 1 - e^{-\lambda h} \\ &= 1 - (1 - \lambda h + o(h)) \\ &= \lambda h + o(h). \end{aligned}$$

Análogamente

$$\begin{aligned} P(X_{t+h} - X_t \geq 2) &= 1 - P(X_{t+h} - X_t = 0) - P(X_{t+h} - X_t = 1) \\ &= 1 - e^{-\lambda h} - e^{-\lambda h} \lambda h \\ &= 1 - (1 - \lambda h + o(h)) - (\lambda h + o(h)) \\ &= o(h). \end{aligned}$$

Estos cálculos y lo desarrollado antes demuestran que $(II) \Leftrightarrow (III)$. También antes hemos demostrado que $(I) \Rightarrow (III)$. Para demostrar el recíproco es necesario comprobar que los tiempos de interarribo T_1, T_2, \dots son independientes con distribución $\exp(\lambda)$. Daremos una demostración no completamente rigurosa de este resultado. Sean $t_1, \dots, t_n > 0$ tiempos cualesquiera y sean $\Delta t_1, \dots, \Delta t_n$ las longitudes que se muestran en la Figura 4.4.



Figura 4.4:

La probabilidad de que T_1 tome un valor en el intervalo Δt_1 , T_2 tome un valor en el intervalo Δt_2 , etcétera, es

$$\begin{aligned} &f_{T_1, \dots, T_n}(t_1, \dots, t_n) \Delta t_1 \cdots \Delta t_n \\ &= (e^{-\lambda t_1} \cdot e^{-\lambda \Delta t_1} \lambda \Delta t_1) (e^{-\lambda t_2} \cdot e^{-\lambda \Delta t_2} \lambda \Delta t_2) \cdots (e^{-\lambda t_n} \cdot e^{-\lambda \Delta t_n} \lambda \Delta t_n) \\ &= \lambda e^{-\lambda t_1} \cdot \lambda e^{-\lambda t_2} \cdots \lambda e^{-\lambda t_n} \cdot \Delta t_1 \cdots \Delta t_n \cdot e^{-\lambda(\Delta t_1 + \cdots + \Delta t_n)} \end{aligned}$$

Al hacer $\Delta t_1, \dots, \Delta t_n \rightarrow 0$ se obtiene

$$f_{T_1, \dots, T_n}(t_1, \dots, t_n) = \lambda e^{-\lambda t_1} \cdot \lambda e^{-\lambda t_2} \cdots \lambda e^{-\lambda t_n}.$$

Esto demuestra que las variables T_1, \dots, T_n son independientes y cada una de ellas tiene distribución $\exp(\lambda)$. En conjunto esto demuestra que $(I) \Leftrightarrow (III) \Leftrightarrow (II)$. \square

Observe que la pregunta acerca de la existencia de un proceso estocástico que cumpla los postulados de la definición (II) o (III) queda resuelta al verificar la equivalencia de tales postulados con la definición constructiva (I).

Presentaremos a continuación algunas generalizaciones del proceso de Poisson. Una de tales generalizaciones que estudiaremos con mayor detalle en un capítulo más adelante es aquella en la que se considera que las variables T_1, T_2, \dots no son necesariamente exponenciales, en tal caso se pierde la propiedad de Markov del proceso. A este tipo de procesos se les llama *procesos de renovación*.

4.3. Proceso de Poisson no homogéneo

Se considera ahora que el parámetro λ del proceso de Poisson no es necesariamente una constante sino una función del tiempo. A veces a este proceso se le llama también proceso de Poisson con parámetro dependiente del tiempo. Este modelo puede ser naturalmente más adecuado para algunas situaciones reales aunque deja de cumplir la propiedad de Markov.

Definición. Un *proceso de Poisson no homogéneo* es un proceso a tiempo continuo $\{X_t : t \geq 0\}$, con espacio de estados $\{0, 1, \dots\}$, con parámetro la función positiva y localmente integrable $\lambda(t)$, y que cumple las siguientes propiedades:

- a) $X_0 = 0$.
- b) Los incrementos son independientes.
- c) Para cualquier $t \geq 0$, y cuando $h \searrow 0$,
 - i) $P(X_{t+h} - X_t \geq 1) = \lambda(t)h + o(h)$.
 - ii) $P(X_{t+h} - X_t \geq 2) = o(h)$.

Comparando esta definición con la definición (II) de proceso de Poisson se observa mucha similitud excepto por dos aspectos: En lugar de la constante λ se escribe ahora la función $\lambda(t)$, y la hipótesis de incrementos estacionarios ya no aparece. Ello es consecuencia de que el parámetro varía con el tiempo, generalmente de manera decreciente. Es decir, la distribución de probabilidad de la variable incremento $X_{t+s} - X_s$ depende de los valores de la función λ en el intervalo $(s, s + t]$. Sin

embargo, y en completa analogía con el caso homogéneo, la variable X_t continúa teniendo distribución Poisson como a continuación demostraremos.

Proposición. La variable X_t en un proceso de Poisson no homogéneo de parámetro $\lambda(t)$ tiene distribución Poisson($\Lambda(t)$), en donde $\Lambda(t) = \int_0^t \lambda(s) ds$, es decir, para $n = 0, 1, \dots$

$$P(X_t = n) = e^{-\Lambda(t)} \frac{[\Lambda(t)]^n}{n!}.$$

Demostración. La prueba es análoga a una de las realizadas en el caso homogéneo. Se define nuevamente $p_n(t) = P(X_t = n)$ y se considera cualquier $h > 0$. Denotaremos por $p_n(t, t+h]$ a la probabilidad $P(X_{t+h} - X_t = n)$. Por la hipótesis de independencia, para $t \geq 0$ y cuando $h \searrow 0$,

$$\begin{aligned} p_0(t+h) &= p_0(t) p_0(t, t+h] \\ &= p_0(t) (1 - \lambda(t)h + o(h)). \end{aligned}$$

Calculando la derivada se obtiene la ecuación diferencial $p_0'(t) = -\lambda(t)p_0(t)$, cuya solución es $p_0(t) = ce^{-\Lambda(t)}$, en donde la constante c es uno debido a la condición inicial $p_0(0) = 1$. Ahora encontraremos $p_n(t)$ para $n \geq 1$. Nuevamente por independencia,

$$\begin{aligned} p_n(t+h) &= p_n(t) p_0(t, t+h] + p_{n-1}(t) p_1(t, t+h] + o(h) \\ &= p_n(t) (1 - \lambda(t)h + o(h)) + p_{n-1}(t) (\lambda(t)h + o(h)) + o(h). \end{aligned}$$

Entonces $p_n'(t) = -\lambda(t)p_n(t) + \lambda(t)p_{n-1}(t)$, con condición inicial $p_n(0) = 0$ para $n \geq 1$. Definiendo $q_n(t) = e^{\Lambda(t)}p_n(t)$ la ecuación diferencial se transforma en $q_n'(t) = \lambda(t)q_{n-1}(t)$, con condiciones $q_n(0) = 0$ y $q_0(t) = 1$. Esta ecuación se resuelve iterativamente primero para $q_1(t)$, después para $q_2(t)$, y así sucesivamente, en general $q_n(t) = (\Lambda(t))^n/n!$ y de aquí se obtiene $p_n(t)$. \square

Las trayectorias de un proceso de Poisson no homogéneo son semejantes a las trayectorias de un proceso de Poisson, es decir, son trayectorias no decrecientes y con saltos unitarios hacia arriba, pero la frecuencia promedio con la que aparecen los saltos cambia a lo largo del tiempo. De manera análoga al caso homogéneo, los incrementos de este proceso también tienen distribución Poisson.

Proposición. Para el proceso de Poisson no homogéneo, la variable incremento $X_{t+s} - X_s$ tiene distribución Poisson($\Lambda(t+s) - \Lambda(s)$).

Demostración. Se escribe $X_{t+s} = X_s + (X_{t+s} - X_s)$, en donde, por el axioma de incrementos independientes, en el lado derecho aparece la suma de dos variables aleatorias independientes. Recordando que la función generadora de momentos de la distribución Poisson(λ) es $M(r) = \exp[\lambda(e^r - 1)]$, al aplicar este resultado a la ecuación anterior se obtiene

$$\exp[\Lambda(t+s)(e^r - 1)] = \exp[\Lambda(s)(e^r - 1)] M_{X_{t+s}-X_s}(r).$$

Por lo tanto $M_{X_{t+s}-X_s}(r) = \exp[(\Lambda(t+s) - \Lambda(s))(e^r - 1)]$. \square

Si la función $\lambda(t)$ es constante igual a λ , entonces $\Lambda(t) = \lambda t$, y se recupera el proceso de Poisson homogéneo. Cuando $\lambda(t)$ es continua, $\Lambda(t)$ es diferenciable y por lo tanto $\Lambda'(t) = \lambda(t)$. A la función $\Lambda(t)$ se le llama *función de intensidad* y a $\lambda(t)$ se le conoce como *función de valor medio*. Un proceso de Poisson no homogéneo en donde $\{\Lambda(t) : t \geq 0\}$ es un proceso estocástico se le llama *proceso de Cox*.

El siguiente resultado establece que bajo una transformación del parámetro tiempo, un proceso de Poisson no homogéneo puede ser llevado a un proceso de Poisson homogéneo de parámetro uno. Antes de demostrar este resultado observemos que como la función $t \mapsto \lambda(t)$ es positiva, la función de intensidad $\Lambda(t)$ es continua y no decreciente, y en general no es invertible. Puede definirse, sin embargo, la inversa por la derecha

$$\Lambda^{-1}(t) = \inf\{u \geq 0 : \Lambda(u) = t\},$$

que cumple la identidad $\Lambda(\Lambda^{-1}(t)) = t$, y que es una función continua y creciente.

Proposición. Sea $\{X_t\}$ un proceso de Poisson no homogéneo de parámetro $\lambda(t)$, y función de intensidad $\Lambda(t) = \int_0^t \lambda(s) ds$. Defina la función

$$\Lambda^{-1}(t) = \inf\{u \geq 0 : \Lambda(u) = t\}.$$

Entonces el proceso $\{X_{\Lambda^{-1}(t)}\}$ es un proceso de Poisson homogéneo de parámetro $\lambda = 1$.

Demostración. Usaremos la definición (III) de proceso de Poisson. El proceso $X_{\Lambda^{-1}(t)}$ empieza en cero y tiene incrementos independientes pues si se consideran cualesquiera tiempos $0 \leq t_1 < t_2 < \dots < t_n$, bajo la función creciente $\Lambda^{-1}(t)$ estos tiempos se transforman en una nueva colección monótona de tiempos

$$0 \leq \Lambda^{-1}(t_1) < \Lambda^{-1}(t_2) < \dots < \Lambda^{-1}(t_n).$$

Por lo tanto las variables $X_{\Lambda^{-1}(t_1)}, X_{\Lambda^{-1}(t_2)} - X_{\Lambda^{-1}(t_1)}, \dots, X_{\Lambda^{-1}(t_n)} - X_{\Lambda^{-1}(t_{n-1})}$ son independientes. Finalmente, para cualesquiera tiempos $s, t \geq 0$ el incremento $X_{\Lambda^{-1}(t+s)} - X_{\Lambda^{-1}(s)}$ tiene distribución Poisson de parámetro $\Lambda(\Lambda^{-1}(t+s)) - \Lambda(\Lambda^{-1}(s)) = (t+s) - s = t$. \square

Pueden definirse los tiempos de interarribo T_1, T_2, \dots , y los tiempos de saltos W_1, W_2, \dots para un proceso de Poisson no homogéneo de manera análoga al caso homogéneo. Por hipótesis los incrementos de este procesos son independientes, sin embargo, debido a que el parámetro $\lambda(t)$ es una función del tiempo, los tiempos de interarribo T_1, T_2, \dots no son independientes pues, por ejemplo, T_2 depende de $\lambda(t)$ para valores de t mayores a T_1 . Y por las mismas razones las distribuciones de estas variables no son idénticas.

4.4. Proceso de Poisson compuesto

Esta es una de las generalizaciones del proceso de Poisson más conocidas y de amplia aplicación. La generalización consiste en que ahora los saltos ya no son necesariamente unitarios.

Definición. Sea $\{N_t\}$ un proceso de Poisson y sea Y_1, Y_2, \dots una sucesión de variables aleatorias independientes, idénticamente distribuidas e independientes del proceso Poisson. Sea $Y_0 = 0$. El *proceso de Poisson compuesto* se define de la siguiente forma:

$$X_t = \sum_{n=0}^{N_t} Y_n. \quad (4.3)$$

La variable X_t es una suma de variables aleatorias en donde el número de sumandos es aleatorio. Tal tipo de modelos encuentra aplicación natural en distintos

contextos. Por ejemplo, la variable X_t puede interpretarse como el monto total de reclamaciones recibidas por una compañía aseguradora al tiempo t . El proceso de Poisson determina el número de siniestros o reclamaciones efectuadas hasta un momento cualquiera. La variable Y_n representa el monto de la n -ésima reclamación y es natural suponer $Y_n > 0$.

En la Figura 4.5 se muestra una trayectoria de este proceso y en los siguientes ejercicios se presentan algunas de sus propiedades básicas. Cuando las variables Y_1, Y_2, \dots toman valores en el conjunto $\{1, 2, \dots\}$ se dice que este proceso es un *proceso de Poisson generalizado* pues los saltos ya no son necesariamente unitarios. Observe que si las variables Y_1, Y_2, \dots son todas idénticamente uno, este proceso se reduce al proceso de Poisson.

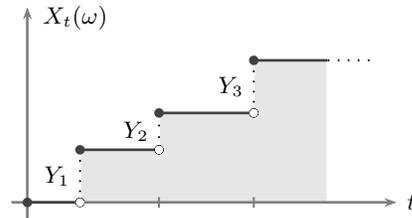


Figura 4.5: Trayectoria de un proceso de Poisson compuesto.

Proposición. El proceso de Poisson compuesto (4.3) cumple las siguientes propiedades.

- a) $E(X_t) = \lambda t E(Y)$.
- b) $\text{Var}(X_t) = \lambda t E(Y^2)$.
- c) $\text{Cov}(X_t, X_s) = \lambda E(Y^2) \min\{s, t\}$.
- d) Tiene incrementos independientes y estacionarios.
- e) Es un proceso de Markov a tiempo continuo.
- e) La función generadora de momentos de la variable X_t es

$$M_{X_t}(u) = E(e^{uX_t}) = \exp[\lambda t (M_Y(u) - 1)].$$

Notas y referencias. El estudio del proceso de Poisson generalmente se incluye en los cursos elementales de procesos estocásticos. La mayoría de los textos sobre procesos estocásticos que aparecen en la bibliografía cuentan por lo menos con una sección sobre este tema. Algunos textos específicos que dedican un capítulo entero

para el proceso de Poisson y que el lector puede consultar para profundizar en el tema son: Basu [1], Jones y Smith [17], y Taylor y Karlin [37].

Ejercicios

Proceso de Poisson

71. Los clientes de una tienda entran al establecimiento de acuerdo a un proceso de Poisson de parámetro $\lambda = 4$. Sea X_t el número de clientes que han ingresado hasta el instante t . Calcule
- a) $P(X_2 = 1)$.
 - b) $P(X_1 = 3, X_2 = 6)$.
 - c) $P(X_1 = 0 | X_3 = 4)$.
 - d) $P(X_2 = 4 | X_1 = 2)$.
 - e) $P(X_2 \leq 3)$.
 - f) $P(X_1 \geq 4 | X_1 \geq 2)$.
72. Los pasajeros llegan a una parada de autobús de acuerdo a un proceso de Poisson $\{X_t\}$ de parámetro $\lambda = 2$. Sea τ el momento en el que llega el primer autobús. Suponga que τ es una variable aleatoria con distribución $\text{unif}(0, 1)$ e independiente del proceso de Poisson. Calcule
- a) $E(X_\tau | \tau = t)$.
 - b) $E(X_\tau)$.
 - c) $E(X_\tau^2 | \tau = t)$.
 - d) $\text{Var}(X_\tau)$.
73. Sea $\{X_t\}$ un proceso de Poisson de parámetro $\lambda = 2$. Sea W_n el instante en el que ocurre el n -ésimo evento. Calcule
- a) $E(X_5)$.
 - b) $E(X_5 | X_2 = 1)$.
 - c) $E(W_7)$.
 - d) $E(W_7 | X_2 = 3)$.
 - e) $E(W_7 | W_5 = 4)$.

74. Sean T_1, T_2, \dots una sucesión de variables aleatorias independientes idénticamente distribuidas con distribución $\exp(\lambda)$. Defina la variable aleatoria N de la siguiente forma

$$N = \begin{cases} 0 & \text{si } T_1 > 1, \\ k & \text{si } T_1 + \dots + T_k \leq 1 < T_1 + \dots + T_{k+1}. \end{cases}$$

Demuestre que N tiene distribución $\text{Poisson}(\lambda)$. Este resultado puede ser usado para obtener valores al azar de la distribución Poisson .

75. Demuestre que si X tiene distribución $\text{unif}(0, 1)$, entonces $Y := -\frac{1}{\lambda} \ln(1 - X)$ tiene distribución $\exp(\lambda)$. Este resultado puede ser usado para generar valores al azar de la distribución $\exp(\lambda)$ a partir de valores de la distribución $\text{unif}(0, 1)$.
76. Sean T_1, \dots, T_n variables aleatorias independientes cada una con distribución $\exp(\lambda)$. Demuestre que la suma $W_n = T_1 + \dots + T_n$ tiene distribución $\text{gamma}(n, \lambda)$ y que la correspondiente función de distribución puede escribirse de la siguiente forma: para cada $t > 0$,

$$P(W_n \leq t) = \sum_{k=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

77. Sea $\{X_t\}$ un proceso de Poisson de parámetro $\lambda = 1$, e independiente de una variable aleatoria θ con distribución $\exp(\lambda)$ con $\lambda = 1$. Defina el proceso $Y_t = X_{\theta t}$. Demuestre que

$$a) P(Y_t = n) = \frac{1}{1+t} \left(\frac{t}{1+t}\right)^n, \text{ para } n = 0, 1, \dots$$

$$b) P(Y_t = n, Y_{t+s} = n+m) = \binom{n+m}{n} t^n s^m \left(\frac{1}{1+t+s}\right)^{n+m+1}.$$

Concluya que las variables Y_t y $Y_{t+s} - Y_t$ no son independientes.

78. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ . Demuestre los siguientes resultados de manera sucesiva.

$$a) (W_1 > t_1, W_2 > t_2) = (X_{t_1} = 0, X_{t_2} - X_{t_1} = 0 \text{ ó } 1).$$

$$b) P(W_1 > t_1, W_2 > t_2) = e^{-\lambda t_1} (1 + \lambda(t_2 - t_1)) e^{-\lambda(t_2 - t_1)}.$$

$$c) f_{W_1, W_2}(t_1, t_2) = \lambda^2 e^{-\lambda t_2}, \text{ para } 0 < t_1 < t_2.$$

- d) $f_{W_1}(t_1) = \lambda e^{-\lambda t_1}$.
- e) $f_{W_2}(t_2) = \lambda^2 t_1 e^{-\lambda t_1}$.
- f) Las variables $T_1 = W_1$ y $T_2 = W_2 - W_1$ son independientes.
79. Sea $\{X_t\}$ proceso de Poisson de parámetro λ , y sean s y t dos tiempos tales que $0 \leq s < t$. Demuestre que
- a) $P(X_s = 0, X_t = 1) = \lambda(t - s)e^{-\lambda t}$.
- b) $P(X_s = X_t) = e^{-\lambda(t-s)}$.
80. Para un proceso de Poisson de parámetro λ demuestre que
- $$\text{Cov}(X_t, X_s) = \lambda \min\{s, t\}.$$
81. Las funciones seno y coseno hiperbólico se definen de la siguiente forma: $\sinh(x) = (e^x - e^{-x})/2$, y $\cosh(x) = (e^x + e^{-x})/2$. Para un proceso de Poisson de parámetro λ demuestre que
- a) $P("X_t$ sea impar") = $e^{-\lambda t} \sinh(\lambda t)$.
- b) $P("X_t$ sea par") = $e^{-\lambda t} \cosh(\lambda t)$.
82. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ , y sea $s > 0$ un tiempo fijo. Demuestre que el proceso $Y_t = X_{t+s} - X_s$ también es un proceso de Poisson de parámetro λ . En la Figura 4.6 se ilustra gráficamente la definición de este nuevo proceso.

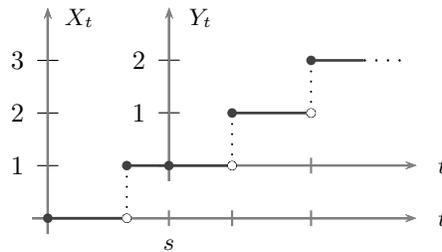


Figura 4.6:

83. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ . Demuestre que

$$\begin{aligned}
 a) \quad f_{W_1, W_2}(w_1, w_2) &= \begin{cases} \lambda^2 e^{-\lambda w_2} & \text{si } 0 < w_1 < w_2, \\ 0 & \text{otro caso.} \end{cases} \\
 b) \quad f_{W_1 | W_2}(w_1 | w_2) &= \begin{cases} 1/w_2 & \text{si } w_1 \in (0, w_2), \\ 0 & \text{otro caso.} \end{cases} \\
 c) \quad f_{W_2 | W_1}(w_2 | w_1) &= \begin{cases} \lambda e^{-\lambda(w_2 - w_1)} & \text{si } w_2 \in (w_1, \infty), \\ 0 & \text{otro caso.} \end{cases}
 \end{aligned}$$

84. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ . Sea T una variable aleatoria con distribución $\exp(\theta)$ e independiente del proceso de Poisson. Encuentre la función de densidad de la variable X_T .
85. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ . Sean r y n dos enteros tales que $1 \leq r \leq n$, y sea $t > 0$. Suponga que el evento $(X_t = n)$ ocurre. Encuentre la densidad condicional de W_r .
86. Sean $X_1(t)$ y $X_2(t)$ dos procesos de Poisson independientes con parámetros λ_1 y λ_2 respectivamente. Calcule la probabilidad de que
- $X_1(t) = 1$ antes que $X_2(t) = 1$.
 - $X_1(t) = 2$ antes que $X_2(t) = 2$.
87. Suponga que un cierto aparato eléctrico sufre desperfectos a lo largo del tiempo de acuerdo a un proceso de Poisson de parámetro λ . Suponga que cada vez que el aparato se descompone es enviado a reparación y después es puesto en funcionamiento nuevamente. Suponga además que el aparato se reemplaza completamente por uno nuevo cuando el tiempo que transcurre entre dos descomposturas sucesivas es menor o igual a una cierta constante $a > 0$, incluyendo el tiempo antes de la primera reparación. Encuentre la función de densidad del
- tiempo de vida útil del equipo antes de ser reemplazado.
 - número de reparaciones antes de realizar el reemplazo.
88. Suponga que un cierto circuito recibe impulsos eléctricos de acuerdo a un proceso de Poisson de parámetro λ . Suponga además que el circuito se descompone al recibir el k -ésimo impulso. Encuentre la función de densidad del tiempo de vida del circuito.

89. Sean $X_1(t), \dots, X_n(t)$ procesos de Poisson independientes, todos de parámetro λ . Encuentre la distribución de probabilidad del primer momento en el cual
- ha ocurrido en cada uno de los procesos al menos un evento.
 - ocurre el primer evento en cualquiera de los procesos.
90. Sean $\{X_1(t)\}$ y $\{X_2(t)\}$ dos procesos de Poisson independientes con parámetros λ_1 y λ_2 , respectivamente. Sea n cualquier número natural, y defina el tiempo aleatorio $\tau = \inf \{t \geq 0 : X_1(t) = n\}$. Demuestre que $X_2(\tau)$ tiene distribución binomial negativa, es decir, para $k = 0, 1, \dots$

$$P(X_2(\tau) = k) = \binom{n+k-1}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^n.$$

91. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ , y sea $a > 0$ una constante. Demuestre que $\{X_{at}\}$ es un proceso de Poisson de parámetro λa .
92. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ . Demuestre que, condicionado a que el proceso tiene exactamente un salto en el intervalo $[s, s+t]$, el momento en el que ese salto ocurre se distribuye de manera uniforme en dicho intervalo.
93. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ . Suponga que cada evento registrado es de tipo 1 con probabilidad α , o de tipo 2 con probabilidad $1 - \alpha$. Sea $X_1(t)$ el número de eventos del tipo 1 al tiempo t , y sea $X_2(t)$ lo correspondiente a eventos del tipo 2.
- Demuestre que $\{X_1(t)\}$ y $\{X_2(t)\}$ son procesos de Poisson de parámetro $\lambda\alpha$ y $\lambda(1 - \alpha)$ respectivamente.
 - Demuestre que $\{X_1(t)\}$ y $\{X_2(t)\}$ son independientes.
94. Suponga que los mensajes llegan a una cierta cuenta de correo electrónico de acuerdo a un proceso de Poisson de parámetro λ . Cada mensaje recibido es de tipo “basura” con probabilidad α y “no basura” con probabilidad $1 - \alpha$.
- Dado que hasta el tiempo t ha llegado n mensajes, encuentre la distribución del número de mensajes tipo “basura” que hasta ese momento han llegado.
 - Suponiendo que se recibe un mensaje tipo “basura” en algún momento, encuentre la distribución del número total de mensajes recibidos hasta la llegada del siguiente mensaje tipo “basura”.

- c) Encuentre la probabilidad de que al tiempo t hayan llegado mas mensajes tipo “basura” que “no basura”.
95. Sean T_1, T_2, \dots los tiempos de interarribo de un proceso de Poisson de parámetro λ , y sea c una constante positiva. Defina $N = \text{mín} \{n \geq 1 : T_n > c\}$. Calcule $E(W_{N-1} + c)$.

Distribuciones asociadas al proceso de Poisson

96. Sea t_0 un tiempo positivo fijo y suponga que hasta ese momento se ha observado un solo evento de un proceso de Poisson, es decir, $X_{t_0} = 1$. La pregunta es ¿cuándo ocurrió tal evento? Demuestre que la distribución condicional del momento en el que se ha registrado el evento, es decir T_1 , es uniforme en el intervalo $[0, t_0]$.
97. Sea $\{X_t\}$ un proceso Poisson de tasa λ y sean dos tiempos $0 < s < t$. Defina la variable $X_{(s,t]}$ como el número de eventos del proceso de Poisson que ocurren en el intervalo $(s, t]$. Demuestre que, condicionada a la ocurrencia del evento $(X_t = n)$, la variable $X_{(s,t]}$ tiene distribución binomial $(n, 1 - s/t)$.
98. Sea $\{X_t\}$ un proceso de Poisson de parámetro λ . Suponga que para un tiempo fijo $t > 0$ se observa el evento $(X_t = n)$, con $n \geq 1$.
- a) Demuestre que la función de densidad del momento en el que ocurre el k -ésimo evento $(1 \leq k \leq n)$, condicionada al evento $(X_t = n)$, es la siguiente: para $s \in (0, t)$,

$$f_{W_k | X_t}(s | n) = \binom{n}{k} \frac{k}{t} \left(\frac{s}{t}\right)^{k-1} \left(1 - \frac{s}{t}\right)^{n-k}.$$

- b) Demuestre que la función de densidad condicional del cociente W_k/t , dado el evento $(X_t = n)$, es la densidad beta $(k, n - k + 1)$.
- c) Encuentre nuevamente la función de densidad gamma (k, λ) de la variable W_k efectuando la siguiente suma

$$\sum_{n=k}^{\infty} f_{W_k | X_t}(s | n) P(X_t = n).$$

Proceso de Poisson no homogéneo

99. Sea $\{X_t\}$ un proceso Poisson no homogéneo de intensidad $\lambda(t)$, sean T_1, T_2, \dots los tiempos de interarribo, y sean W_1, W_2, \dots los tiempos reales de ocurrencia. Demuestre que para cualquier $t > 0$,

$$\text{a) } f_{T_1}(t) = e^{-\Lambda(t)} \lambda(t).$$

$$\text{b) } f_{T_2|T_1}(t|s) = e^{-\Lambda(t+s)+\Lambda(s)} \lambda(t+s).$$

$$\text{c) } f_{T_2}(t) = \int_0^\infty e^{-\Lambda(t+s)} \lambda(t+s) \lambda(s) ds.$$

$$\text{d) } f_{W_n}(t) = e^{-\Lambda(t)} \frac{[\Lambda(t)]^{n-1}}{(n-1)!} \lambda(t).$$

$$\text{e) } F_{T_k|W_{k-1}}(t|s) = 1 - e^{-\Lambda(t+s)+\Lambda(s)}.$$

$$\text{f) } F_{T_k}(t) = 1 - \int_0^\infty e^{-\Lambda(t+s)} \frac{[\Lambda(s)]^{k-2}}{(k-2)!} \lambda(s) ds, \quad k \geq 2.$$

Proceso de Poisson compuesto

100. Suponga que los sumandos Y de un proceso de Poisson compuesto $\{X_t\}$ de parámetro λ tienen distribución $\exp(\mu)$. Encuentre la distribución de la variable X_t .

101. Sea $\{X_t\}$ un proceso de Poisson compuesto de parámetro λ . Suponga que cada una de las variables Y de este proceso es constante igual a $k \in \mathbb{N}$. Encuentre la distribución de X_t .

102. Suponga que los usuarios de cuentas de correo electrónico solicitan acceso al servidor de acuerdo a un proceso de Poisson homogéneo de parámetro λ . Suponga además que cada usuario se mantiene conectado al servidor un tiempo aleatorio con función de distribución $F(x)$, e independiente uno del otro. Sea C_t el número de usuarios conectados al servidor tiempo al t , y defina la función $\alpha(t) = \int_0^t (1 - F(x)) dx$. Demuestre que para $k = 0, 1, \dots$

$$P(C_t = k) = \frac{[\lambda\alpha(t)]^k}{k!} e^{-\lambda\alpha(t)}.$$

CAPÍTULO 5

Cadenas de Markov a tiempo continuo

Vamos a considerar ahora cadenas de Markov $\{X_t : t \geq 0\}$ en donde el tiempo es continuo y las variables toman valores enteros. Varios de los conceptos que mencionaremos para este tipo de procesos son análogos al caso de tiempo discreto, pero existen diferencias en el comportamiento y desarrollo matemático entre los dos modelos, y no pretendemos presentar y justificar la teoría completa.

Usando la misma notación que en el caso de tiempos discretos, recordemos que el término $p(x_t)$ significa $P(X_t = x_t)$. La propiedad de Markov que consideraremos tiene la siguiente forma: Para cualesquiera tiempos $0 \leq t_1 < t_2 < \dots < t_n$,

$$p(x_{t_n} | x_{t_1}, \dots, x_{t_{n-1}}) = p(x_{t_n} | x_{t_{n-1}}).$$

Observe que no estamos suponiendo que se conoce la historia del proceso en todo el pasado a tiempo continuo, sino únicamente en una colección arbitraria pero finita de tiempos pasados t_1, \dots, t_{n-1} . Supondremos nuevamente que estas probabilidades de transición son *estacionarias* en el tiempo, esto significa que para cada $s \geq 0$ y $t \geq 0$, la probabilidad $P(X_{t+s} = j | X_s = i)$ es idéntica a $P(X_t = j | X_0 = i)$, es decir, no hay dependencia del valor de s . Esta probabilidad se escribe de manera breve mediante la expresión $p_{ij}(t)$, para i y j enteros no negativos. En particular para $t = 0$ se define nuevamente $p_{ij}(0)$ como la función delta de Kronecker, es decir,

$$p_{ij}(0) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Haciendo variar los índices i y j en el espacio de estados, se obtiene la matriz de probabilidades de transición al tiempo t :

$$P_t = \begin{pmatrix} p_{00}(t) & p_{01}(t) & \cdots \\ p_{10}(t) & p_{11}(t) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Plantaremos a continuación el modelo general de cadena de Markov a tiempo continuo, estudiaremos algunas de sus propiedades, y después revisaremos algunos modelos particulares.

5.1. Procesos de saltos

Considere un proceso a tiempo continuo $\{X_t : t \geq 0\}$ que inicia en el estado i_1 al tiempo cero. El proceso permanece en ese estado un tiempo aleatorio T_{i_1} , y después salta a un nuevo estado i_2 distinto del anterior. El sistema permanece ahora en el estado i_2 un tiempo aleatorio T_{i_2} al cabo del cual brinca a otro estado i_3 distinto del inmediato anterior, y así sucesivamente. Esta sucesión de saltos se muestra gráficamente en la Figura 5.1.

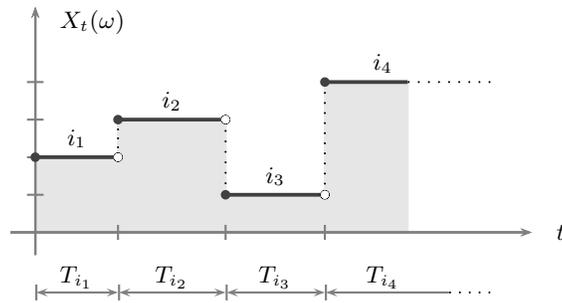


Figura 5.1:

Los tiempos aleatorios T son los tiempos en los que el proceso permanece constante en alguno de sus estados, y se llaman *tiempos de estancia* (*passage times*). Los momentos en donde el proceso tiene saltos son los tiempos $W_n = T_{i_1} + \cdots + T_{i_n}$, para $n \geq 1$. El proceso puede entonces escribirse de la forma siguiente:

$$X_t = \begin{cases} i_1 & \text{si } 0 \leq t < W_1, \\ i_2 & \text{si } W_1 \leq t < W_2, \\ i_3 & \text{si } W_2 \leq t < W_3, \\ \vdots & \end{cases}$$

Un proceso de estas características se llama *proceso de saltos*, y parece ser una buena versión continua de las cadenas de Markov a tiempo discreto. Sin embargo puede comprobarse que un proceso con estas características pudiera no estar definido para todo $t \geq 0$, y tampoco hay garantía de que se cumpla la propiedad de Markov. A fin de evitar situaciones demasiado complicadas, los procesos de saltos que consideraremos estarán restringidos por las siguientes condiciones:

No explosión. Puede suceder que los tiempos de estancia T sean cada vez más pequeños de tal forma que $\lim_{n \rightarrow \infty} W_n < \infty$, es decir, existe la posibilidad de que el proceso efectúe un número infinito de saltos durante un intervalo de tiempo acotado. En tal situación el proceso no estaría bien definido para todo $t \geq 0$, y se dice que el proceso *explota* en un tiempo finito. Para evitar tal comportamiento supondremos que $\lim_{n \rightarrow \infty} W_n = \infty$, y por lo tanto, para cada $t \geq 0$, el valor de X_t es finito.

Probabilidades de transición. Supondremos que el espacio de estados es $\{0, 1, \dots\}$ y que el tiempo de estancia asociado al estado i es la variable aleatoria T_i , la cual supondremos positiva con función de distribución $F_i(t)$. Como en el caso de cadenas a tiempos discreto, se denotará por p_{ij} a la probabilidad de que la cadena salte del estado i al estado j . Adicionalmente impondremos la condición $p_{ii} = 0$, y con ello se inhibe, por ahora, que la cadena salte al mismo estado de partida. Las probabilidades de transición cumplen entonces las siguientes condiciones:

- a) $p_{ij} \geq 0$.
- b) $p_{ii} = 0$.
- c) $\sum_j p_{ij} = 1$.

Independencia. Supondremos que los tiempos de estancia T_{i_1}, T_{i_2}, \dots son independientes entre sí, y también son independientes del mecanismo mediante el cual se escoge el estado j al cual la cadena brinca después de estar en cualquier otro estado i .

Estados absorbentes o no absorbentes. Supondremos que cada variable T_i es finita con probabilidad uno, o bien es infinita con probabilidad uno. En el primer caso se dice que el estado i es *no absorbente*, y en el segundo caso que es *absorbente*. El hecho de que $T_i = \infty$ se interpreta en el sentido de que el proceso deja de saltar y permanece en el estado i el resto del tiempo, es decir, es absorbente. Estamos entonces suponiendo que sólo hay dos tipos de estados: absorbentes o no

absorbentes, en otras palabras, con probabilidad uno el tiempo de estancia es finito o con probabilidad uno es infinito.

Propiedad de Markov. Puede demostrarse que el proceso de saltos descrito arriba satisface la propiedad de Markov si, y sólo si, los tiempos de estancia de estados no absorbentes tienen distribución exponencial. Este es un resultado importante cuya demostración omitiremos y que simplifica el modelo general planteado. Supondremos entonces que el tiempo de estancia en un estado no absorbente i tiene distribución $\exp(\lambda_i)$, con $\lambda_i > 0$, es decir, $F_i(t) = 1 - e^{-\lambda_i t}$, para $t \geq 0$. Cuando $T_i = \infty$ puede considerarse que $\lambda_i = 0$.

Definición. A un proceso de saltos con las características arriba señaladas se le llama *cadena de Markov a tiempo continuo*.

Observe que un proceso de Markov a tiempo continuo queda completamente especificado por los siguientes tres elementos: una distribución de probabilidad inicial, los parámetros λ_i , y las probabilidades p_{ij} .

■ Ejemplo: Cadena de primera ocurrencia. Sea X_t el estado de un sistema tal que en cualquier instante se encuentra en alguno de los dos posibles estados: 0 y 1. Suponga que $X_0 = 0$ y que el proceso cambia al estado 1 después de un tiempo aleatorio con distribución $\exp(\lambda)$, y permanece allí el resto del tiempo. Una posible trayectoria de este proceso se muestra en la Figura 5.2(a). Este proceso modela la situación de espera para la primera ocurrencia de un evento de interés. Las probabilidades de transición son

$$P_t = \begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{pmatrix} = \begin{pmatrix} e^{-\lambda t} & 1 - e^{-\lambda t} \\ 0 & 1 \end{pmatrix}.$$

○

■ Ejemplo: Cadena de dos estados. Considere el proceso X_t con dos posibles estados: 0 y 1, definido por la siguiente dinámica: Cuando el proceso entra al estado 0 permanece en él un tiempo $\exp(\lambda)$, y luego va al estado 1, entonces permanece en el estado 1 un tiempo $\exp(\mu)$ y después regresa a 0, y así sucesivamente. Se postula que los tiempos de estancia en cada estado son variables aleatorias independientes. Una trayectoria de este proceso se muestra en la Figura 5.2(b). Puede demostrarse

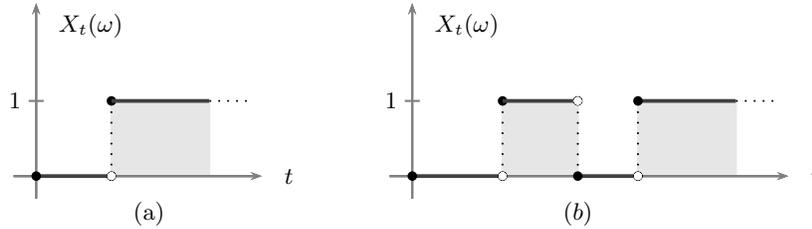


Figura 5.2:

que las probabilidades de transición son

$$P_t = \begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{pmatrix} = \frac{1}{\lambda + \mu} \begin{pmatrix} \mu & \lambda \\ \mu & \lambda \end{pmatrix} + \frac{1}{\lambda + \mu} \begin{pmatrix} \lambda & -\lambda \\ -\mu & \mu \end{pmatrix} e^{-(\lambda + \mu)t}.$$

◦

■ Ejemplo. El proceso de Poisson es una cadena de Markov a tiempo continuo que empieza en cero, los tiempos de estancia son exponenciales de parámetro λ , y las probabilidades de transición de un estado a otro son $p_{i,i+1} = 1$.

◦

5.2. Propiedades generales

Mencionaremos ahora algunos conceptos y propiedades generales de procesos de Markov a tiempo continuo. Varios de estos resultados son similares a los desarrollados antes para el caso de tiempos discretos.

Cadena a tiempo discreto asociada. Toda cadena de Markov a tiempo continuo $\{X_t : t \geq 0\}$ tiene asociada una cadena a tiempo discreto que denotaremos por el mismo símbolo $\{X_n : n = 0, 1, \dots\}$, y que está dada por la primera pero observada en los tiempos en donde se efectúan los saltos. Algunas características de la cadena a tiempo discreto se trasladan a su versión continua. Inversamente, a partir de una cadena a tiempo discreto puede construirse su versión continua en el tiempo tomando tiempos exponenciales independientes entre salto y salto.

Probabilidades de transición. Para una cadena de Markov a tiempo continuo las probabilidades de transición son los números $p_{ij}(t) = P(X_t = j | X_0 = i)$, para cualesquiera estados i y j , y para cualquier tiempo $t \geq 0$. Cuando el espacio de estados es finito, los elementos de cada renglón de esta matriz suman uno, pero para espacios de estados infinito esta suma puede ser estrictamente menor a uno, es decir, en general, $\sum_j p_{ij}(t) \leq 1$. Esta es una diferencia inesperada respecto del modelo a tiempo discreto.

Intentar encontrar $p_{ij}(t)$ para cada par de estados i y j , y para cada $t \geq 0$, es un problema demasiado general, y sólo en algunos casos es posible encontrar explícitamente tales probabilidades. El siguiente resultado nos permitirá obtener algunas conclusiones generales acerca de $p_{ij}(t)$.

Proposición. Sean i y j dos estados. Para cualquier $t \geq 0$,

$$p_{ij}(t) = \delta_{ij} e^{-\lambda_i t} + \lambda_i e^{-\lambda_i t} \int_0^t e^{\lambda_i s} \left(\sum_{k \neq i} p_{ik} p_{kj}(s) \right) ds. \quad (5.1)$$

Demostración. Si i es un estado absorbente, i.e. $\lambda_i = 0$, entonces la fórmula se reduce a $p_{ij}(t) = \delta_{ij}$, lo cual es correcto. Si i es un estado no absorbente, entonces

$$\begin{aligned} p_{ij}(t) &= P(X_t = j | X_0 = i) \\ &= P(X_t = j, T_i > t | X_0 = i) + P(X_t = j, T_i \leq t | X_0 = i) \\ &= \delta_{ij} e^{-\lambda_i t} + \int_0^t f_{X_t, T_i | X_0}(j, u | i) du \\ &= \delta_{ij} e^{-\lambda_i t} + \int_0^t \sum_{k \neq i} f_{X_t, X_u, T_i | X_0}(j, k, u | i) du, \end{aligned}$$

en donde por la propiedad de Markov y la independencia,

$$\begin{aligned} f_{X_t, X_u, T_i | X_0}(j, k, u | i) &= f_{X_t | X_u, T_i, X_0}(j | k, u, i) f_{X_u | T_i, X_0}(k | u, i) f_{T_i | X_0}(u | i) \\ &= p_{kj}(t - u) p_{ik} \lambda_i e^{-\lambda_i u}. \end{aligned}$$

Por lo tanto,

$$p_{ij}(t) = \delta_{ij} e^{-\lambda_i t} + \int_0^t \lambda_i e^{-\lambda_i u} \left(\sum_{k \neq i} p_{ik} p_{kj}(t - u) \right) du.$$

Haciendo el cambio de variable $s(u) = t - u$ en la integral se obtiene el resultado. \square

Las probabilidades de transición satisfacen también una versión continua de la ecuación de Chapman-Kolmogorov, y que en este caso se conoce como la *propiedad de semigrupo*.

Proposición. (Ecuación de Chapman-Kolmogorov). Para cualquiera estados i y j , y para todo $t \geq 0$ y $s \geq 0$,

$$p_{ij}(t + s) = \sum_k p_{ik}(t) p_{kj}(s).$$

En notación matricial, $P_{t+s} = P_t P_s$.

Demostración. Por la propiedad de Markov,

$$\begin{aligned} p_{ij}(t + s) &= P(X_{t+s} = j | X_0 = i) \\ &= \sum_k P(X_{t+s} = j, X_t = k | X_0 = i) \\ &= \sum_k P(X_{t+s} = j | X_t = k) P(X_t = k | X_0 = i) \\ &= \sum_k p_{ik}(t) p_{kj}(s). \end{aligned}$$

\square

La colección $\{P_t : t \geq 0\}$ constituye un semigrupo de matrices estocásticas, esto quiere decir que cumple las siguientes propiedades:

- a) $P_0 = I$, en donde I es la matriz identidad.
- b) P_t es una matriz estocástica.
- c) $P_{t+s} = P_t P_s$, para cualesquiera $t, s \geq 0$.

La ecuación de Chapman-Kolmogorov es interesante pues permite expresar a la probabilidad $p_{ij}(t)$, para cualquier tiempo $t > 0$, en términos de probabilidades

infinitesimales, es decir, probabilidades de transición en intervalos de tiempo de longitud muy pequeña. Por ejemplo, para cualquier n natural, tomando $\Delta t = t/n$,

$$p_{ij}(t) = \sum_{k_1, \dots, k_{n-1}} p_{i, k_1}(\Delta t) p_{k_1, k_2}(\Delta t) \cdots p_{k_{n-1}, j}(\Delta t).$$

Esto quiere decir que es suficiente conocer el comportamiento de $p_{ij}(t)$ en tiempos t pequeños para conocer su comportamiento para cualquier $t > 0$. Especificaremos con detalle este resultado más adelante.

Parámetros infinitesimales. De la fórmula (5.1) es inmediato observar que las probabilidades $p_{ij}(t)$ son funciones continuas en t , y en particular el integrando en (5.1) es una función continua. Esto implica que la integral es diferenciable y por lo tanto también $p_{ij}(t)$. Derivando entonces (5.1) se obtiene el siguiente resultado conocido como el *sistema de ecuaciones hacia atrás* de Kolmogorov.

Proposición. Para cualquier $t > 0$,

$$p'_{ij}(t) = -\lambda_i p_{ij}(t) + \lambda_i \sum_{k \neq i} p_{ik} p_{kj}(t). \quad (5.2)$$

Observe que en particular la derivada $p'_{ij}(t)$ es una función continua del tiempo. Tomando ahora el límite cuando $t \rightarrow 0$, se tiene que

$$\begin{aligned} p'_{ij}(0) &= -\lambda_i \delta_{ij} + \lambda_i \sum_{k \neq i} p_{ik} \delta_{kj} \\ &= -\lambda_i \delta_{ij} + \lambda_i p_{ij} \\ &= \begin{cases} -\lambda_i & \text{si } i = j, \\ \lambda_i p_{ij} & \text{si } i \neq j. \end{cases} \end{aligned}$$

Definición. A las cantidades $p'_{ij}(0)$ se les denota por g_{ij} , y se les conoce con el nombre de *parámetros infinitesimales* del proceso. Es decir, estos parámetros son

$$g_{ij} = \begin{cases} -\lambda_i & \text{si } i = j, \\ \lambda_i p_{ij} & \text{si } i \neq j. \end{cases} \quad (5.3)$$

Variando los índices i y j , estos nuevos parámetros se escriben en notación matricial de la forma siguiente:

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 p_{01} & \lambda_0 p_{02} & \cdots \\ \lambda_1 p_{10} & -\lambda_1 & \lambda_1 p_{12} & \cdots \\ \lambda_2 p_{20} & \lambda_2 p_{21} & -\lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (5.4)$$

Esta matriz determina de manera única el comportamiento de la cadena de Markov a tiempo continuo, y es el concepto equivalente a la matriz de probabilidades de transición en un paso para cadenas a tiempo discreto. Se trata de una matriz con las siguientes propiedades:

- a) $g_{ii} \leq 0$.
- b) $g_{ij} \geq 0$, si $i \neq j$.
- c) $\sum_j g_{ij} = 0$.

Observe que $g_{ii} = 0$ cuando el estado i es absorbente, es decir, cuando $\lambda_i = 0$. Cuando $\lambda_i > 0$ la última propiedad se obtiene a partir del hecho de que $p_{ii} = 0$ pues

$$\sum_j g_{ij} = -\lambda_i + \sum_{j \neq i} \lambda_i p_{ij} = -\lambda_i + \lambda_i(1 - p_{ii}) = 0.$$

Proposición. Los parámetros infinitesimales determinan de manera única a la cadena de Markov a tiempo continuo.

Este resultado es consecuencia de (5.3) pues a partir de g_{ij} se obtiene

$$\begin{aligned} \lambda_i &= -g_{ii}, \\ p_{ij} &= \begin{cases} 0 & \text{si } i = j, \\ -g_{ij}/g_{ii} & \text{si } i \neq j. \end{cases} \end{aligned} \quad (5.5)$$

Probabilidades infinitesimales. Un proceso de Markov a tiempo continuo puede ser también definido a partir del comportamiento de las probabilidades de transición $p_{ij}(t)$ cuando $t \rightarrow 0$. Tales probabilidades se llaman a veces *probabilidades*

infinitesimales, y pueden expresarse en términos de los parámetros infinitesimales como establece el siguiente resultado y del cual omitiremos su demostración.

Proposición. Cuando $t \rightarrow 0$,

1. $p_{ii}(t) = 1 + g_{ii} t + o(t)$.
2. $p_{ij}(t) = g_{ij} t + o(t)$, para $i \neq j$.

Ecuaciones de Kolmogorov. En términos de los parámetros infinitesimales, la ecuación (5.2) puede escribirse en la forma

$$p'_{ij}(t) = \sum_k g_{ik} p_{kj}(t). \quad (5.6)$$

En términos de matrices esta igualdad se escribe como $P'(t) = G P(t)$, es decir,

$$\begin{pmatrix} p_{00}(t) & p_{01}(t) & \cdots \\ p_{10}(t) & p_{11}(t) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} -\lambda_0 & \lambda_0 p_{01} & \cdots \\ \lambda_1 p_{10} & -\lambda_1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} p_{00}(t) & p_{01}(t) & \cdots \\ p_{10}(t) & p_{11}(t) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Como hemos mencionado antes, este sistema de ecuaciones se conoce como las *ecuaciones hacia atrás* de Kolmogorov.

Comunicación. Se dice que el estado j es accesible desde el estado i si $p_{ij}(t) > 0$ para algún $t \geq 0$, y se escribe $i \rightarrow j$. Se dice que los estados i y j se comunican si $i \rightarrow j$ y $j \rightarrow i$, y en tal caso se escribe $i \leftrightarrow j$. Nuevamente puede demostrarse que la comunicación es una relación de equivalencia, y eso lleva a una partición del espacio de estados en clases de comunicación. Se dice que la cadena es irreducible cuando todos los estados pertenecen a la misma clase de comunicación.

Tiempos de primera visita. Para cualesquiera estados i y j , se define

$$T_{ij} = \inf \{t > W_1 : X_t = j\}, \quad \text{cuando } X_0 = i.$$

El tiempo medio de primera visita es entonces $\mu_{ij} = E(T_{ij})$. Cuando los estados i y j coinciden se escribe T_i , y $\mu_i = E(T_i)$ respectivamente. A μ_i se le llama tiempo medio de recurrencia.

Recurrencia y transitoriedad. Se dice que el estado i es transitorio si, partiendo de i , con probabilidad uno el conjunto de tiempos $\{t \geq 0 : X_t = i\}$ es acotado. En cambio, se dice que es recurrente si, partiendo de i , con probabilidad uno el conjunto $\{t \geq 0 : X_t = i\}$ es no acotado. Cuando $E(T_i) = \infty$ se dice que el estado i es recurrente nulo, y cuando $E(T_i) < \infty$ se dice que es recurrente positivo.

Distribuciones estacionarias. Una distribución de probabilidad $\pi = \{\pi_0, \pi_1, \dots\}$ sobre el espacio de estados $\{0, 1, \dots\}$ es estacionaria para la cadena con matriz de probabilidades de transición P_t si para cualquier $t \geq 0$, $\sum_i \pi_i p_{ij}(t) = \pi_j$. En notación matricial, $\pi P_t = \pi$. Si X_0 tiene como distribución inicial una distribución estacionaria π , entonces $P(X_t = j) = \sum_i \pi_i p_{ij}(t) = \pi_j$, es decir, la variable X_t tiene la misma distribución de probabilidad π para cualquier valor de t .

5.3. Procesos de nacimiento y muerte

Un *proceso de nacimiento y muerte* es una cadena de Markov a tiempo continuo con espacio de estados $\{0, 1, \dots\}$ tal que en cada salto la cadena únicamente puede brincar una unidad hacia arriba o una unidad hacia abajo. Un salto hacia arriba se interpreta como un *nacimiento*, mientras que un salto hacia abajo representa una *muerte*. Estas probabilidades de saltos de un estado a otro son

$$p_{ij} = \begin{cases} \frac{\lambda_i}{\lambda_i + \mu_i} & \text{si } j = i + 1, \\ \frac{\mu_i}{\lambda_i + \mu_i} & \text{si } j = i - 1, \\ 0 & \text{otro caso,} \end{cases}$$

en donde $\lambda_0, \lambda_1, \dots$ y μ_1, μ_2, \dots son constantes positivas conocidas como las *tasas instantáneas de nacimiento y muerte*, respectivamente. De acuerdo a las ecuaciones (5.5), el tiempo de estancia en el estado i tiene distribución $\exp(\lambda_i + \mu_i)$, en donde se define $\mu_0 = 0$. Siguiendo la fórmula general del generador infinitesimal (5.4) para un proceso de saltos, tenemos que la matriz de parámetros infinitesimales de este proceso es de la siguiente forma

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Una posible trayectoria de este proceso se muestra en la Figura 5.3 con $X_0 = 0$, los saltos son unitarios, hacia arriba o hacia abajo. La variable X_t puede interpretarse como el número de individuos en una población al tiempo t , en donde pueden presentarse nacimientos y muertes de individuos. Como $\lambda_0 > 0$ y $\mu_0 = 0$, la población puede crecer cuando se encuentre en el estado cero, pero nunca decrecer por abajo de ese nivel.

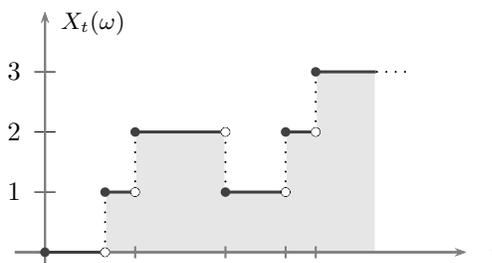


Figura 5.3:

Se pueden tomar las siguientes probabilidades infinitesimales como postulados para definir a este proceso. Para cualquier $t \geq 0$, y cuando $h \rightarrow 0$,

- $P(X_{t+h} - X_t = 1 \mid X_t = k) = \lambda_k h + o(h)$.
- $P(X_{t+h} - X_t = -1 \mid X_t = k) = \mu_k h + o(h)$.
- $P(X_{t+h} - X_t = 0 \mid X_t = k) = 1 - (\lambda_k + \mu_k) h + o(h)$.

Para cualquier $t \geq 0$ y $h > 0$ pequeño consideremos el intervalo $[0, t+h]$ visto como $[0, h] + (h, t+h]$. Haciendo un análisis sobre estos intervalos puede encontrarse nuevamente que las probabilidades $p_{ij}(t)$ satisfacen el sistema de *ecuaciones diferenciales hacia atrás* de Kolmogorov:

$$\begin{aligned} p'_{0j}(t) &= -\lambda_0 p_{0j}(t) + \lambda_0 p_{1j}(t), \\ p'_{ij}(t) &= \mu_i p_{i-1,j}(t) - (\lambda_i + \mu_i) p_{ij}(t) + \lambda_i p_{i+1,j}(t). \end{aligned}$$

De manera análoga, considerando $[0, t+h] = [0, t] + (t, t+h]$, es posible comprobar que se cumple el sistema de *ecuaciones diferenciales hacia adelante* de Kolmogorov:

$$\begin{aligned} p'_{i0}(t) &= -\lambda_0 p_{i0}(t) + \mu_1 p_{i1}(t), \\ p'_{ij}(t) &= \lambda_{j-1} p_{i,j-1}(t) - (\lambda_j + \mu_j) p_{ij}(t) + \mu_{j+1} p_{i,j+1}(t). \end{aligned}$$

■ **Ejemplo.** Un proceso de Poisson es una cadena de nacimiento y muerte en donde las tasas instantáneas de muerte μ_0, μ_1, \dots son cero, y las tasas instantáneas de nacimiento $\lambda_0, \lambda_1, \dots$ son todas ellas iguales a una constante $\lambda > 0$. La matriz de parámetros infinitesimales es entonces de la forma

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}.$$

◦

■ Ejemplo. Las ecuaciones hacia adelante de Kolmogorov del proceso de Poisson son

$$\begin{aligned} p_0'(t) &= -\lambda p_0(t), \\ p_n'(t) &= \lambda p_{n-1}(t) - \lambda p_n(t), \quad \text{para } n \geq 1, \end{aligned}$$

en donde $p_n(t) = p_{0n}(t)$. Usaremos estas ecuaciones para comprobar nuevamente que X_t tiene distribución Poisson(λt). Usando la condición inicial $p_0(0) = 1$, la primera ecuación tiene solución $p_0(t) = e^{-\lambda t}$. Definiendo $q_n(t) = e^{\lambda t} p_n(t)$, la segunda ecuación se transforma en $q_n'(t) = \lambda q_{n-1}(t)$, con condiciones $q_n(0) = \delta_{0n}$ y $q_0(t) = 1$. Esta nueva ecuación se resuelve iterativamente, primero para $q_1(t)$, después para $q_2(t)$, y así sucesivamente. En general, $q_n(t) = (\lambda t)^n / n!$ para $n \geq 0$. De aquí se obtiene $p_n(t) = e^{-\lambda t} (\lambda t)^n / n!$ ◦

■ Ejemplo. Esta es otra derivación mas de la distribución Poisson en el proceso de Poisson, ahora usando las ecuaciones hacia adelante de Kolmogorov y la función generadora de probabilidad. La variable aleatoria X_t puede tomar los valores $0, 1, \dots$ de modo que su función generadora de probabilidad es

$$G_{X_t}(u) = E(u^{X_t}) = \sum_{n=0}^{\infty} p_n(t) u^n,$$

para valores reales de u tales que $|u| < 1$, y en donde $p_n(t) = P(X_t = n)$. Consideraremos a esta función también como función del tiempo t y por comodidad en los siguientes cálculos la llamaremos $G(t, u)$. Derivando respecto de t , para el mismo radio de convergencia $|u| < 1$, y usando las ecuaciones hacia adelante de

Kolmogorov para el proceso de Poisson se tiene que

$$\begin{aligned}
 G'(t, u) &= \sum_{n=0}^{\infty} p'_n(t) u^n \\
 &= \sum_{n=0}^{\infty} (\lambda p_{n-1}(t) - \lambda p_n(t)) u^n \\
 &= \lambda u G(t, u) - \lambda G(t, u) \\
 &= (u - 1) \lambda G(t, u),
 \end{aligned}$$

de donde se obtiene la ecuación diferencial $G'/G = (u - 1)\lambda$. Integrando de 0 a t y usando la condición $G(0, u) = 1$ se llega a

$$G(t, u) = G(0, u) e^{(u-1)\lambda t} = e^{-\lambda t} e^{\lambda t u} = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} u^n.$$

Esta es la función generadora de probabilidad de la distribución Poisson(λt). Por la propiedad de unicidad se obtiene que la variable X_t tiene distribución Poisson(λt). \circ

5.4. Proceso de nacimiento puro

Cuando en un proceso de nacimiento y muerte los parámetros μ_0, μ_1, \dots son todos cero, se obtiene un *proceso de nacimiento puro*. La matriz de parámetros infinitesimales tiene la forma que se muestra en la Figura 5.4, en donde, como antes, los parámetros $\lambda_0, \lambda_1, \dots$ se conocen como las *tasas instantáneas de nacimiento*. En la Figura 5.4 se muestra una trayectoria de este proceso cuando inicia en el estado cero. Por construcción, el tiempo de estancia en el estado i tiene distribución exponencial con parámetro λ_i .

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ 0 & -\lambda_1 & \lambda_1 & 0 & \dots \\ 0 & 0 & -\lambda_2 & \lambda_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix}$$

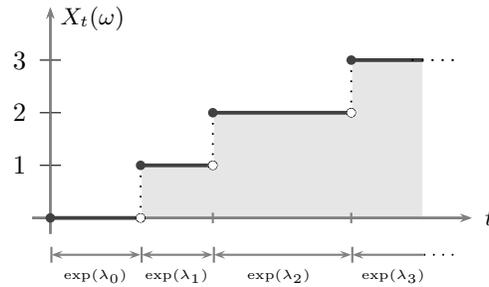


Figura 5.4:

Las probabilidades de saltos son evidentemente $p_{ij} = 1$ cuando $j = i + 1$, y cero en caso contrario. Puede demostrarse que los incrementos de un proceso de nacimiento puro son independientes pero no son necesariamente estacionarios.

Un proceso de nacimiento puro puede también definirse mediante las siguientes probabilidades infinitesimales: Cuando $h \rightarrow 0$,

- a) $P(X_{t+h} - X_t = 1 | X_t = k) = \lambda_k h + o(h)$.
- b) $P(X_{t+h} - X_t = 0 | X_t = k) = 1 - \lambda_k h + o(h)$.

En general no es fácil encontrar una fórmula para las probabilidades de transición $p_{ij}(t)$, sin embargo cuando el estado inicial es el cero se conoce la siguiente fórmula recursiva.

Proposición. Para cualquier $t \geq 0$ y cualquier entero $n \geq 1$,

$$p_{0n}(t) = \lambda_{n-1} e^{-\lambda_n t} \int_0^t e^{\lambda_n s} p_{0,n-1}(s) ds. \quad (5.7)$$

Demostración. El sistema de ecuaciones diferenciales hacia adelante de Kolmogorov para este proceso es

$$p'_{ij}(t) = \lambda_{j-1} p_{i,j-1}(t) - \lambda_j p_{ij}(t).$$

En particular, partiendo de cero,

$$\begin{aligned} p'_{00}(t) &= -\lambda_0 p_{00}(t), \\ p'_{0n}(t) &= \lambda_{n-1} p_{0,n-1}(t) - \lambda_n p_{0n}(t), \quad n \geq 1, \end{aligned}$$

con las condiciones de frontera $p_{00}(0) = 1$ y $p_{0n}(0) = 0$, para $n \geq 1$. La primera ecuación tiene solución $p_{00}(t) = e^{-\lambda_0 t}$, mientras que para el caso $n \geq 1$, multiplicando por el factor integrante $e^{\lambda_n t}$ y resolviendo se obtiene la fórmula enunciada. \square

De manera análoga puede definirse un *proceso de muerte* como un proceso de nacimiento y muerte en donde los parámetros de nacimiento $\lambda_0, \lambda_1, \dots$ son todos cero. El proceso puede iniciar con una población de tamaño $k \geq 1$, y presentar fallecimientos sucesivos hasta una posible extinción completa de la población.

El proceso de Yule. Este es un tipo particular de proceso de nacimiento puro, y para el cual es posible encontrar explícitamente las probabilidades de transición. Este proceso puede definirse a través de los siguientes postulados:

- El estado inicial es $X_0 = k \geq 1$.
- Si $X_t = n$, entonces cada uno de estos elementos puede dar nacimiento a un nuevo elemento durante un periodo de longitud infinitesimal $h > 0$ con probabilidad $\lambda h + o(h)$, en donde $\lambda > 0$, es decir,

$$\begin{aligned} P(X_{t+h} - X_t = 1 | X_t = n) &= \binom{n}{1} [\lambda h + o(h)] [1 - \lambda h + o(h)]^{n-1} \\ &= \lambda n h + o(h). \end{aligned}$$

Es decir, las tasas instantáneas de nacimiento son $\lambda_n = \lambda n$, que crecen de manera lineal conforme la población crece. El tiempo de estancia en el estado n tiene distribución $\exp(\lambda n)$, en consecuencia el tiempo medio de estancia en ese estado es $1/(\lambda n)$, cada vez menor conforme n crece. El sistema de ecuaciones hacia adelante para $p_{kn}(t)$, con $n \geq k$, se reduce a

$$\begin{aligned} p'_{kk}(t) &= -\lambda k p_{kk}(t), \\ p'_{kn}(t) &= \lambda(n-1)p_{k,n-1}(t) - \lambda n p_{kn}(t), \quad n \geq k+1. \end{aligned}$$

La fórmula recursiva (5.7) es, para $n \geq k+1$,

$$p_{kn}(t) = \lambda(n-1) e^{-\lambda n t} \int_0^t e^{\lambda n s} p_{k,n-1}(s) ds. \quad (5.8)$$

Mostraremos a continuación que el incremento $X_t - X_0$ tiene distribución binomial negativa de parámetros (r, p) , con $r = k$ y $p = e^{-\lambda t}$.

Proposición. $p_{kn}(t) = \binom{n-1}{n-k} e^{-\lambda k t} (1 - e^{-\lambda t})^{n-k}$, para $n \geq k$.

Demostración. Usaremos inducción sobre n . Para $n = k$ se tiene la ecuación diferencial $p'_{kk}(t) = -\lambda k p_{kk}(t)$, con condición inicial $p_{kk}(0) = 1$. Esto produce la solución $p_{kk}(t) = e^{-\lambda k t}$, que es de la forma enunciada. Para valores de $n \geq k+1$

usaremos la fórmula recursiva (5.8),

$$\begin{aligned}
p_{kn}(t) &= \lambda(n-1) e^{-\lambda nt} \int_0^t e^{\lambda ns} \binom{n-2}{n-1-k} e^{-\lambda ks} (1 - e^{-\lambda s})^{n-1-k} ds \\
&= \lambda(n-1) \binom{n-2}{n-1-k} e^{-\lambda nt} \int_0^t (e^{\lambda s})^{(n-k)} (1 - e^{-\lambda s})^{n-1-k} ds \\
&= \lambda(n-1) \binom{n-2}{n-1-k} e^{-\lambda nt} \int_0^t (1 - e^{-\lambda s})^{-1} (e^{\lambda s} - 1)^{n-k} ds \\
&= \lambda(n-1) \binom{n-2}{n-1-k} e^{-\lambda nt} \int_0^t e^{\lambda s} (e^{\lambda s} - 1)^{n-1-k} ds \\
&= \lambda(n-1) \binom{n-2}{n-1-k} e^{-\lambda nt} \int_0^t \frac{1}{\lambda(n-k)} \frac{d}{ds} (e^{\lambda s} - 1)^{n-k} ds \\
&= \frac{n-1}{n-k} \binom{n-2}{n-1-k} e^{-\lambda nt} (e^{\lambda t} - 1)^{n-k} \\
&= \binom{n-1}{k-1} e^{-\lambda kt} (1 - e^{-\lambda t})^{n-k}.
\end{aligned}$$

□

Notas y referencias. Una exposición más completa del tema de cadenas de Markov a tiempo continuo puede encontrarse en Basu [1], Karlin y Taylor [19], o Stirzaker [36].

Ejercicios

Ecuaciones de Kolmogorov

103. Sea $\{N_t\}$ un proceso de Poisson de parámetro λ y sean Y_1, Y_2, \dots v.a.i.i.d. con distribución común Bernoulli(p). Encuentre las ecuaciones de Kolmogorov hacia atrás y hacia adelante del proceso

$$X_t = \sum_{j=1}^{N_t} Y_j.$$

Resuelva cualquiera de estos sistemas de ecuaciones y demuestre que para $j \geq i$,

$$p_{ij}(t) = e^{-\lambda pt} \frac{(\lambda pt)^{j-i}}{(j-i)!}.$$

Procesos de nacimiento y muerte

104. Para la cadena de Markov a tiempo continuo de dos estados en donde la estancia en el estado 0 tiene distribución $\exp(\lambda)$, y la estancia en el estado 1 es $\exp(\mu)$, demuestre que
- $p_{00}(t) = \frac{1}{\lambda + \mu} (\mu + \lambda e^{-(\lambda + \mu)t})$.
 - $p_{01}(t) = \frac{1}{\lambda + \mu} (\lambda - \lambda e^{-(\lambda + \mu)t})$.
 - $p_{10}(t) = \frac{1}{\lambda + \mu} (\mu - \mu e^{-(\lambda + \mu)t})$.
 - $p_{11}(t) = \frac{1}{\lambda + \mu} (\lambda + \mu e^{-(\lambda + \mu)t})$.
105. Considere una cadena de Markov a tiempo continuo de dos estados en donde el tiempo de estancia en cada uno de ellos tiene distribución $\exp(\lambda)$. Defina N_t como el número de veces que la cadena de Markov ha efectuado saltos hasta un tiempo $t \geq 0$ cualquiera. Demuestre que $\{N_t\}$ es un proceso de Poisson de parámetro λ .

Procesos de nacimiento puro

106. Sea $\{X_t\}$ un proceso de Yule con estado inicial $X_0 = k \geq 1$. Demuestre que

a) $E(X_t) = k e^{\lambda t}$.

b) $\text{Var}(X_t) = k e^{2\lambda t} (1 - e^{-\lambda t})$.

CAPÍTULO 6

Procesos de renovación y confiabilidad

En este capítulo se presenta otra generalización del proceso de Poisson. Se considera ahora que los tiempos de interarribo no son necesariamente exponenciales. A tales procesos de saltos unitarios se les conoce como *procesos de renovación*, y en general dejan de cumplir la propiedad de Markov. Además de estudiar algunas propiedades generales de los procesos de renovación, estudiaremos también el concepto de confiabilidad para sistemas conformados por varios componentes los cuales son susceptibles de sufrir fallas en tiempos aleatorios.

6.1. Procesos de renovación

Suponga que se pone en operación un cierto componente o artículo cuya duración de vida útil se modela mediante una variable aleatoria T_1 . Una vez que el componente falla, se reemplaza o renueva con otro componente cuyo tiempo de vida es T_2 , y así sucesivamente. La colección de variables aleatorias T_1, T_2, \dots representa la sucesión de tiempos de vida de componentes puestos en operación uno tras otro. Esto se ilustra en la Figura 6.1.

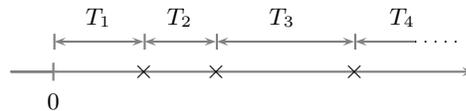


Figura 6.1:

En este contexto es natural suponer que las variables que modelan los tiempos de vida son no negativas, independientes y con la misma distribución de probabilidad. Un proceso de estas características se conoce con el nombre de *proceso de renovación*. Observe que exactamente en los tiempos en los que se efectúan las renovaciones, el proceso reinicia probabilísticamente. Empezaremos por dar un primera definición formal de un proceso de renovación basada en lo recién mencionado.

Definición. Un proceso de renovación es una sucesión infinita de variables aleatorias T_1, T_2, \dots que son no negativas, independientes e idénticamente distribuidas.

Otra forma equivalente de definir a este proceso es a través del registro de los tiempos reales en los que se observan las renovaciones, o bien a través del conteo de renovaciones observadas hasta un tiempo cualquiera. Estos puntos de vista alternativos se definen a continuación.

Definición. Dado un proceso de renovación $\{T_1, T_2, \dots\}$, se definen los tiempos reales de renovación como $W_0 = 0$ y $W_n = T_1 + \dots + T_n$, para $n \geq 1$. El proceso de conteo de renovaciones es

$$N_t = \text{máx} \{n \geq 0 : W_n \leq t\}, \quad \text{para cada } t \geq 0.$$

La variable aleatoria W_n representa el tiempo real en el que se realiza la n -ésima renovación, mientras que N_t indica el número de renovaciones realizadas hasta el tiempo t . En la literatura se le denomina proceso de renovación a cualquiera de los procesos $\{T_n : n = 1, 2, \dots\}$, $\{W_n : n = 0, 1, \dots\}$, o $\{N_t : t \geq 0\}$, pues por construcción existe una correspondencia biunívoca entre cualesquiera dos de estos tres procesos.

Denotaremos por $F(t)$ a la función de distribución de los tiempos de vida. Por lo tanto, la función de distribución de W_n es la convolución de $F(t)$ consigo misma n veces, es decir, $F_{W_n}(t) = P(T_1 + \dots + T_n \leq t) = F^{*n}(t) = (F * \dots * F)(t)$. En particular, cuando $n = 0$ se define

$$F^{*0}(t) = \begin{cases} 0 & \text{si } t < 0, \\ 1 & \text{si } t \geq 0. \end{cases}$$

Una de las primeras preguntas que uno puede hacerse acerca de un proceso de renovación es la de conocer la distribución de la variable N_t . La respuesta no es fácil de encontrar pues la distribución de esta variable depende de la distribución de los tiempos de vida como indica la siguiente fórmula.

Proposición. Para cualquier $n \geq 0$,

$$P(N_t = n) = F^{*n}(t) - F^{*(n+1)}(t).$$

Demostración. Tomando en consideración que las variables W_n y T_{n+1} son independientes, tenemos que

$$\begin{aligned} P(N_t = n) &= P(W_n \leq t, W_{n+1} > t) \\ &= P(W_n \leq t, W_n + T_{n+1} > t) \\ &= P(t - T_{n+1} < W_n \leq t) \\ &= F_{W_n}(t) - F_{W_n}(t - T_{n+1}) \\ &= F^{*n}(t) - \int_0^\infty F_{W_n | T_{n+1}}(t - u | u) dF(u) \\ &= F^{*n}(t) - \int_0^\infty F^{*n}(t - u) dF(u) \\ &= F^{*n}(t) - F^{*(n+1)}(t). \end{aligned}$$

□

Por ejemplo, un proceso de Poisson homogéneo de tasa λ es un proceso de renovación en donde los tiempos de vida tienen distribución $\exp(\lambda)$. En consecuencia, los tiempos reales de renovación W_n tienen distribución $\text{gama}(n, \lambda)$, cuya función de distribución es, para $t > 0$,

$$F_{W_n}(t) = \int_0^t \frac{(\lambda x)^{n-1}}{\Gamma(n)} \lambda e^{-\lambda x} dx = \sum_{k=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

La segunda igualdad puede obtenerse después de aplicar integración por partes varias veces. A partir de esta expresión puede recuperarse la distribución Poisson

pues por la proposición anterior,

$$P(N_t = n) = F^{*n}(t) - F^{*(n+1)}(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Observe la igualdad de eventos $(N_t \geq n) = (W_n \leq t)$, la cual es válida para $t \geq 0$ y para cada valor entero de $n \geq 0$. Esta identidad nos será de utilidad más adelante.

6.2. Función y ecuación de renovación

Otra de las funciones que es natural desear conocer en un proceso de renovación es el número promedio de renovaciones efectuadas hasta un tiempo t cualquiera. A tal función se le llama *función de renovación*, y se le denota por $\Lambda(t)$, es decir, $\Lambda(t) = E(N_t)$. En general no es fácil encontrar una forma explícita para esta función. Sin embargo, se cuenta con la siguiente ecuación integral general.

Proposición. La función de renovación $\Lambda(t)$ satisface la ecuación

$$\Lambda(t) = F(t) + \int_0^t \Lambda(t-s) dF(s). \quad (6.1)$$

Demostración. Condicionando sobre el valor del primer tiempo de vida T_1 se obtiene $\Lambda(t) = \int_0^\infty E(N_t | T_1 = s) dF(s)$, en donde

$$E(N_t | T_1 = s) = \begin{cases} 0 & \text{si } s > t, \\ 1 + \Lambda(t-s) & \text{si } s \leq t. \end{cases}$$

Por lo tanto

$$\Lambda(t) = \int_0^t (1 + \Lambda(t-s)) dF(s) = F(t) + \int_0^t \Lambda(t-s) dF(s).$$

□

Observe que la ecuación (6.1) puede escribirse como $\Lambda(t) = F(t) + (\Lambda * F)(t)$. Debido a que se condiciona sobre el valor del primer tiempo de renovación, a la

técnica de la demostración de este resultado se le llama a veces *análisis del primer paso*.

■ **Ejemplo.** Para el proceso de Poisson, el promedio de renovaciones o saltos al tiempo t es $\Lambda(t) = \lambda t$. Puede comprobarse directamente que esta función satisface (6.1) con $F(t) = 1 - e^{-\lambda t}$. ◻

A una ecuación integral del tipo (6.1) se le llama *ecuación de renovación*, pues algunas funciones de interés en la teoría de la renovación la cumplen. La definición general aparece a continuación.

Definición. Sean $F(t)$, $g(t)$ y $h(t)$ funciones definidas para $t \geq 0$. Suponga que $F(t)$ y $h(t)$ son conocidas, y $g(t)$ es desconocida. Se dice que $g(t)$ satisface una *ecuación de renovación* si cumple la ecuación integral

$$g(t) = h(t) + \int_0^t g(t-s) dF(s). \quad (6.2)$$

Puede demostrarse que si $h(t)$ es una función acotada, entonces existe una única solución $g(t)$ a la ecuación de renovación (6.2) que cumple con la condición de ser acotada sobre intervalos finitos y está dada explícitamente por

$$g(t) = h(t) + \int_0^t h(t-s) d\Lambda(s), \quad (6.3)$$

en donde $\Lambda(s)$ es la función de renovación. La demostración de este resultado general puede ser encontrado en el texto de Karlin y Taylor [19]. Acerca de la función de renovación tenemos la siguiente expresión general.

Proposición. $\Lambda(t) = \sum_{n=1}^{\infty} F^{*n}(t)$, para cualquier $t > 0$.

Demostración.

$$\begin{aligned}
 \Lambda(t) &= \sum_{n=0}^{\infty} n P(N_t = n) \\
 &= P(N_t = 1) + 2P(N_t = 2) + \dots \\
 &= [F^{*1}(t) - F^{*2}(t)] + 2[F^{*2}(t) - F^{*3}(t)] + \dots \\
 &= F^{*1}(t) + F^{*2}(t) + \dots \\
 &= \sum_{n=1}^{\infty} F^{*n}(t).
 \end{aligned}$$

□

Por ejemplo, cuando los tiempos de interarribo tienen distribución $\exp(\lambda)$ se tiene que

$$F^{*n}(t) = \int_0^t \frac{(\lambda x)^{n-1}}{\Gamma(n)} \lambda e^{-\lambda x} dx = \sum_{k=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

Usando esta expresión y la fórmula recién demostrada puede corroborarse que en este caso $\Lambda(t) = \lambda t$.

Es posible demostrar que para un proceso de renovación la variable N_t tiene momentos finitos de cualquier orden, y entonces en particular la suma anterior es siempre convergente. Se puede también demostrar que existe una correspondencia biunívoca entre las funciones $\Lambda(t)$ y $F(t)$, y además el proceso de renovación $\{N_t\}$ queda completamente especificado por la función promedio $\Lambda(t)$. La demostración de estos resultados pueden consultarse en [1].

6.3. Tiempos de vida

Junto a todo proceso de renovación y para cada tiempo fijo, pueden considerarse tres variables aleatorias no negativas, γ_t , δ_t y β_t , cuyo significado geométrico se muestra en la Figura 6.2, y que definiremos con mayor precisión a continuación.

Tiempo de vida restante. Este es el tiempo de vida útil que le resta al elemento que se encuentra en operación al tiempo t , y está dado por la variable

$$\gamma_t = W_{N_t+1} - t.$$

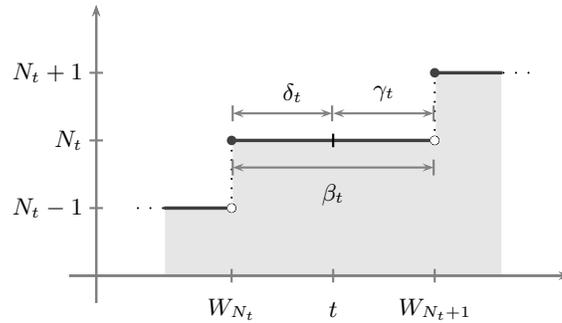


Figura 6.2:

Esta longitud aleatoria se muestra en la Figura 6.2. Usando un argumento de renovación demostraremos que para cada $x > 0$ la función $t \mapsto g(t) = P(\gamma_t > x)$ satisface una ecuación de renovación.

Proposición. La función $g(t) = P(\gamma_t > x)$ satisface la ecuación de renovación

$$g(t) = 1 - F(t + x) + \int_0^t g(t - s) dF(s).$$

Demostración. Se condiciona sobre el valor de T_1 . Se tienen los tres posibles casos que se muestran en la Figura 6.3.

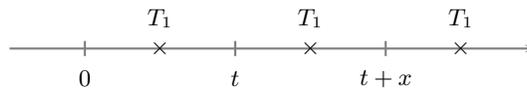


Figura 6.3:

Puesto que un proceso de renovación puede considerarse que inicia nuevamente a

partir del momento en que se observa una renovación, se tiene que

$$P(\gamma_t > x | T_1 = s) = \begin{cases} 1 & \text{si } s > t + x, \\ 0 & \text{si } t < s \leq t + x, \\ P(\gamma_{t-s} > x) & \text{si } 0 < s \leq t. \end{cases}$$

Por lo tanto

$$\begin{aligned} g(t) &= \int_0^\infty P(\gamma_t > x | T_1 = s) dF(s) \\ &= \int_{t+x}^\infty dF(s) + \int_0^t P(\gamma_{t-s} > x) dF(s) \\ &= 1 - F(t+x) + \int_0^t g(t-s) dF(s). \end{aligned}$$

□

En particular, cuando los tiempos de vida son exponenciales de parámetro λ , la única solución de esta ecuación es la función constante $g(t) = e^{-\lambda x}$, que equivale a decir que la variable γ_t tiene distribución $\exp(\lambda)$. Esta es nuevamente la propiedad de *pérdida de memoria* de la distribución exponencial.

Tiempo de vida transcurrido. Este es el tiempo que ha estado en uso el elemento que se encuentra en operación al tiempo t , y está dado por la variable

$$\delta_t = t - W_{N_t}.$$

Véase nuevamente la Figura 6.2. Es claro que esta variable está acotada superiormente por el valor t , pues el elemento en uso al tiempo t no puede tener hasta ese momento un tiempo de uso mayor a t . Sea nuevamente $x > 0$ y defina la función $t \mapsto g(t) = P(\delta_t > x)$. Por las consideraciones anteriores, tenemos que $g(t) = 0$ para $t \in (0, x]$. Usando un argumento de renovación demostraremos que la función $g(t)$ satisface una ecuación de renovación.

Proposición. Para $t \in (x, \infty)$, la función $g(t) = P(\delta_t > x)$ satisface la ecuación de renovación

$$g(t) = 1 - F(t) + \int_0^{t-x} g(t-s) dF(s).$$

Demostración. Nuevamente se condiciona sobre el valor de T_1 . Los tres posibles casos se muestran en la Figura 6.4.

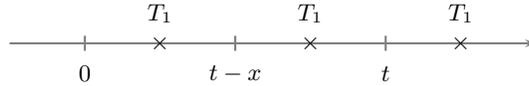


Figura 6.4:

Por lo tanto,

$$P(\delta_t > x | T_1 = s) = \begin{cases} 1 & \text{si } s \geq t, \\ 0 & \text{si } 0 < t - x \leq s < t, \\ P(\delta_{t-s} > x) & \text{si } 0 < s < t - x. \end{cases}$$

Entonces

$$\begin{aligned} g(t) &= \int_0^\infty P(\delta_t > x | T_1 = s) dF(s) \\ &= \int_t^\infty dF(s) + \int_0^{t-x} P(\delta_{t-s} > x) dF(s) \\ &= 1 - F(t) + \int_0^{t-x} g(t-s) dF(s). \end{aligned}$$

En resumen,

$$g(t) = \begin{cases} 0 & \text{si } 0 < t \leq x, \\ 1 - F(t) + \int_0^{t-x} g(t-s) dF(s) & \text{si } t > x. \end{cases}$$

□

Considerando otra vez el caso de tiempos de vida exponencial de parámetro λ , puede demostrarse que el tiempo de vida transcurrido tiene función de distribución

$$P(\delta_t \leq x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } 0 < x < t, \\ 1 & \text{si } x \geq t. \end{cases}$$

Tiempo de vida total. Este es el tiempo completo de vida útil que tiene el elemento que se encuentra en operación al tiempo t , y está dado por la variable

$$\beta_t = \gamma_t + \delta_t = W_{N_t+1} - W_{N_t} = T_{N_t+1}.$$

Nuevamente, para $x > 0$ se define la función $t \mapsto g(t) = P(\beta_t > x)$. Demostraremos que esta función satisface una ecuación de renovación, en la cual aparecerá el término $x \vee y$, que significa $\max\{x, y\}$.

Proposición. La función $g(t) = P(\beta_t > x)$ satisface la ecuación de renovación

$$g(t) = 1 - F(t \vee x) + \int_0^t g(t-s) dF(s),$$

Demostración. Nuevamente condicionaremos sobre un posible valor de T_1 . Los posibles casos se muestran en la Figura 6.5.

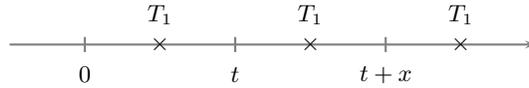


Figura 6.5:

El caso $s \in (t, t+x)$ se debe subdividir en dos casos: $s \leq x$ ó $s > x$. De este modo se obtienen los siguientes resultados

$$P(\beta_t > x | T_1 = s) = \begin{cases} 0 & \text{si } t < s < t+x, \text{ y } s \leq x, \\ 1 & \text{si } t < s < t+x, \text{ y } s > x, \\ 1 & \text{si } s \geq t+x, \\ P(\beta_{t-s} > x) & \text{si } 0 < s \leq t. \end{cases}$$

El segundo y tercer renglón se conjuntan en las condiciones $t \vee x < s < t+x$ ó $s \geq t+x$, lo cual se reduce a la condición $t \vee x < s < \infty$. Por lo tanto,

$$\begin{aligned} g(t) &= \int_0^\infty P(\delta_t > x | T_1 = s) dF(s) \\ &= \int_{t \vee x}^\infty dF(s) + \int_0^t P(\beta_{t-s} > x) dF(s) \\ &= 1 - F(t \vee x) + \int_0^t g(t-s) dF(s). \end{aligned}$$

□

En el caso de tiempos de interarribo exponenciales de parámetro λ , puede demostrarse que el tiempo de vida total tiene función de distribución

$$P(\beta_t \leq x) = \begin{cases} 1 - (1 + \lambda(t \wedge x)) e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{otro caso.} \end{cases}$$

6.4. Teoremas de renovación

En esta sección se estudian algunos resultados sobre el comportamiento límite de un proceso de renovación. Primeramente se demuestra que todo proceso de renovación de conteo crece a infinito con probabilidad uno cuando el tiempo crece a infinito.

Proposición. Para todo proceso de renovación, $\lim_{t \rightarrow \infty} N_t = \infty$ c.s.

Demostración. Recordemos la igualdad de eventos $(N_t \geq n) = (W_n \leq t)$. Entonces para cualquier n natural,

$$\begin{aligned} 1 &= \lim_{t \rightarrow \infty} F_{W_n}(t) \\ &= \lim_{t \rightarrow \infty} P(W_n \leq t) \\ &= \lim_{t \rightarrow \infty} P(N_t \geq n) \\ &= P(\lim_{t \rightarrow \infty} N_t \geq n). \end{aligned}$$

Para la última igualdad se ha utilizado el hecho de que la función $t \rightarrow N_t$ es monótona no decreciente. Como el resultado demostrado vale para cualquier valor de n natural, se tiene que $P(\lim_{t \rightarrow \infty} N_t = \infty) = 1$. \square

Por lo tanto, cuando t tiende a infinito el número de renovaciones N_t también crece a infinito. Por otro lado, el tiempo que le toma al proceso llegar al valor N_t es W_{N_t} y también crece a infinito. El siguiente resultado establece que a largo plazo habrá una renovación cada $\mu = E(T)$ unidades de tiempo.

Proposición. Para un proceso de renovación N_t en donde $E(T) = \mu$, con $0 < \mu < \infty$, se cumple

$$\lim_{t \rightarrow \infty} \frac{W_{N_t}}{N_t} = \mu, \text{ c.s.}$$

Demostración. Para valores enteros de n , por la ley fuerte de los grandes números, $W_n/n \rightarrow \mu$ casi seguramente cuando $n \rightarrow \infty$. Ahora observe la contención de eventos

$$\left(\frac{W_n}{n} \rightarrow \mu\right) \cap (N_t \rightarrow \infty) \subseteq \left(\frac{W_{N_t}}{N_t} \rightarrow \mu\right).$$

Los dos eventos del lado izquierdo tienen probabilidad uno, por lo tanto la intersección también, y ello implica que el lado derecho también tiene probabilidad uno. \square

El siguiente resultado establece la forma en la que N_t crece a infinito. Observe que el cociente N_t/t es una variable aleatoria que cuenta el número de renovaciones por unidad de tiempo efectuadas hasta el tiempo t . Demostraremos a continuación que cuando $t \rightarrow \infty$ la aleatoriedad desaparece y el cociente se vuelve constante.

Teorema elemental de renovación (J. L. Doob, 1948). Para un proceso de renovación en donde $E(T) = \mu$, con $0 < \mu < \infty$, se cumple

$$\lim_{t \rightarrow \infty} \frac{N_t}{t} = \frac{1}{\mu}, \text{ c.s.}$$

Demostración. Para cualquier $t > 0$ se cumple $W_{N_t} \leq t < W_{N_t+1}$. Por lo tanto

$$\frac{W_{N_t}}{N_t} \leq \frac{t}{N_t} < \frac{W_{N_t+1}}{N_t+1} \frac{N_t+1}{N_t}.$$

Cuando t tiende a infinito ambos extremos de estas desigualdades convergen a μ c.s. Por lo tanto $N_t/t \rightarrow 1/\mu$ c.s. \square

Ahora veremos que la esperanza de N_t/t tiene el mismo comportamiento.

Teorema elemental de renovación (W. Feller, 1941). Considere un proceso de renovación en donde $E(T) = \mu$, con $0 < \mu < \infty$. Entonces

$$\lim_{t \rightarrow \infty} \frac{\Lambda(t)}{t} = \frac{1}{\mu}.$$

Demostración. Por la identidad de Wald,

$$E(W_{N_t+1}) = E(N_t + 1) E(T) = (\Lambda(t) + 1)\mu.$$

Despejando $\Lambda(t)$, dividiendo entre t y arreglando términos adecuadamente se obtiene

$$\frac{\Lambda(t)}{t} = \frac{1}{\mu t} E(W_{N_t+1} - t) + \frac{1}{\mu} - \frac{1}{t}.$$

Como $W_{N_t+1} > t$, se tiene que $E(W_{N_t+1} - t) \geq 0$. Por lo tanto

$$\liminf_{t \rightarrow \infty} \frac{\Lambda(t)}{t} \geq \frac{1}{\mu}.$$

Ahora estimaremos el límite superior del mismo cociente. Como $W_{N_t+1} - t \leq W_{N_t+1} - W_{N_t} = T_{N_t+1}$, se tiene que $E(W_{N_t+1} - t) \leq E(T_{N_t+1})$ y por lo tanto

$$\frac{\Lambda(t)}{t} \leq \frac{1}{\mu} + \frac{1}{\mu t} E(T_{N_t+1}).$$

Consideremos primero el caso particular cuando las variables T son uniformemente acotadas, es decir, supongamos que existe un entero $k > 0$ tal que para cada $n \geq 1$, $P(T_n \leq k) = 1$. Entonces $E(T_{N_t+1}) \leq k$ y por lo tanto

$$\limsup_{t \rightarrow \infty} \frac{\Lambda(t)}{t} \leq \frac{1}{\mu}.$$

Esto concluye la demostración en el caso particular mencionado. Ahora consideremos que las variables T no son necesariamente acotadas de la forma anterior. En este caso se definen los tiempos recortados

$$T_n^k = \begin{cases} T_n & \text{si } T_n \leq k, \\ k & \text{si } T_n > k. \end{cases}$$

Observe que $T_n^k \nearrow T_n$ cuando $k \rightarrow \infty$. A partir de estos tiempos se considera un nuevo proceso de renovación N_t^k con función de renovación Λ_t^k . Se cumple que $\Lambda(t) \leq \Lambda_t^k$ y por lo tanto

$$\limsup_{t \rightarrow \infty} \frac{\Lambda(t)}{t} \leq \frac{1}{E(T^k)}.$$

Ahora se hace k tender a infinito. Por el teorema de convergencia monótona, $E(T^k) \nearrow E(T)$, y ello prueba la desigualdad faltante. \square

El cociente $\Lambda(t)/t$ es el número de renovaciones promedio por unidad de tiempo. El resultado recién demostrado establece que a largo plazo habrá en promedio $1/E(T)$ renovaciones por unidad de tiempo. Por ejemplo, para el proceso Poisson, $E(T) = 1/\lambda$ y $\Lambda(t) = \lambda t$. Por lo tanto $\Lambda(t)/t = \lambda$.

Los siguientes dos resultados se enuncian sin demostración.

Teorema de renovación. (D. Blackwell, 1948). Si $F(t)$ es no aritmética, entonces para cualquier $h > 0$,

$$\lim_{t \rightarrow \infty} \Lambda(t+h) - \Lambda(t) = \frac{h}{\mu}.$$

Si $F(t)$ es aritmética con base d , entonces para cualquier natural n ,

$$\lim_{t \rightarrow \infty} \Lambda(t+nd) - \Lambda(t) = \frac{nd}{\mu}.$$

Por ejemplo, para el proceso de Poisson se tiene que $\Lambda(t+h) - \Lambda(t) = \lambda(t+h) - \lambda t = \lambda h = h/(1/\lambda) = h/E(T)$.

Teorema clave de renovación. (W. L. Smith, 1953). Sea $A(t)$ la solución a la ecuación de renovación

$$A(t) = H(t) + \int_0^t A(t-s) dF(s),$$

en donde $H : [0, \infty) \rightarrow \mathbb{R}$ es una función directamente Riemann integrable. Si $F(t)$ es no aritmética,

$$\lim_{t \rightarrow \infty} A(t) = \frac{1}{\mu} \int_0^{\infty} H(t) dt.$$

Si $F(t)$ es aritmética con base d , entonces

$$\lim_{t \rightarrow \infty} A(t + nd) = \frac{d}{\mu} \sum_{k=0}^{\infty} H(t + kd).$$

Puede demostrarse que los teoremas de Smith y Blackwell son equivalentes, y que cualquiera de ellos implica el teorema elemental de Feller.

6.5. Confiabilidad

Suponga que $T > 0$ es una variable aleatoria continua que registra el tiempo de falla de un cierto componente que se encuentra en operación al tiempo cero. Nuevamente denotaremos por $F(t)$ y $f(t)$ a la función de distribución y de densidad de esta variable aleatoria. En la teoría de la confiabilidad interesa conocer la probabilidad de que el componente funcione correctamente por lo menos hasta un tiempo t cualquiera, o la probabilidad de que deje de funcionar en el siguiente instante de tiempo, dado que el componente ha funcionado bien hasta un cierto momento. Es por ello que se definen las siguientes dos funciones.

Definición. La función de confiabilidad se define como $R(t) = P(T > t)$, mientras que la función de tasa de falla es $r(t) = f(t)/(1 - F(t))$.

A la función de confiabilidad se le llama también *función de supervivencia*, y a la función de tasa de falla se le conoce también como *función hazard*, y se le llama así pues un componente que ha sobrevivido al tiempo t , fallará en el intervalo $(t, t + \Delta t]$ con probabilidad condicional $r(t) \Delta t + o(\Delta t)$, en efecto,

$$\begin{aligned} P(t < T \leq t + \Delta t | T > t) &= \frac{P(t < T \leq t + \Delta t)}{P(T > t)} \\ &= \frac{f(t)}{1 - F(t)} \Delta t + o(\Delta t) \\ &= r(t) \Delta t + o(\Delta t). \end{aligned}$$

Después de algunos cálculos sencillos es inmediato comprobar que estas dos funciones están relacionadas de la siguiente forma.

$$\begin{aligned} r(t) &= \frac{f(t)}{R(t)} = -\frac{d}{dt} \ln R(t). \\ R(t) &= \exp\left(-\int_0^t r(s) ds\right). \end{aligned}$$

Además, el tiempo promedio de falla $E(T)$ puede escribirse en términos de la función $R(t)$ como

$$E(T) = \int_0^{\infty} R(t) dt.$$

■ Ejemplo. Suponga que el tiempo de falla T de un componente tiene una distribución exponencial de parámetro λ . La función de confiabilidad es $R(t) = e^{-\lambda t}$, y la función de tasa de falla es constante $r(t) = \lambda$. El hecho de que la función de tasa de falla sea constante significa que la probabilidad de falla en un intervalo de tiempo pequeño Δt , dado que el componente se encuentra operando al tiempo t , es aproximadamente $\lambda \Delta t$, independiente del valor de t . Esta es otra manifestación de la propiedad de pérdida de memoria de la distribución exponencial, y es la única distribución continua con tasa de falla constante. Por otro lado, el tiempo medio de falla es $E(T) = 1/\lambda$. ◻

■ Ejercicio. Demuestre que la función $R(t)$ es decreciente, es decir, si $t_1 \leq t_2$, entonces $R(t_1) \geq R(t_2)$. ¿Puede decirse algo sobre la monotonía de $r(t)$? ▀

Confiabilidad para sistemas en serie.

Considere un sistema de n componentes puestos en serie como se muestra en la Figura 6.6. Supondremos que cada uno de estos componentes funciona de manera independiente uno del otro.



Figura 6.6:

Es intuitivamente claro que tal sistema

en su conjunto funciona si todos y cada uno de los componentes se encuentra en buen estado. Nos interesa encontrar las funciones de confiabilidad y de tasa de falla de este tipo de sistemas.

Observe que el tiempo de falla T de un sistema de este tipo es el tiempo de falla más pequeño de cada uno de los componentes, es decir, $T = \min\{T_1, \dots, T_n\}$. Esquemas físicos de este tipo ayudan a justificar el plantearse la pregunta de encontrar la distribución de probabilidad del mínimo de n variables aleatorias independientes.

■ Ejercicio. Suponga que los tiempos de falla T_1, \dots, T_n de los componentes del sistema en serie de la Figura 6.6 son idénticamente distribuidos con función de distribución $F(t)$, y función de densidad $f(t)$. Demuestre que el tiempo de falla T del sistema tiene función de distribución y densidad dada por las siguientes expresiones:

- a) $F_T(t) = 1 - (1 - F(t))^n$.
- b) $f_T(t) = n(1 - F(t))^{n-1} f(t)$. .

Calcularemos a continuación las funciones de confiabilidad y de tasa de falla para sistemas en serie. Denotaremos por $R_1(t), \dots, R_n(t)$ las funciones de confiabilidad de cada componente en el sistema, y las funciones de tasa de falla serán $r_1(t), \dots, r_n(t)$.

Proposición. Las funciones de confiabilidad y de tasa de falla del sistema en serie de la Figura 6.6 son

1. $R(t) = R_1(t) \cdots R_n(t)$.
2. $r(t) = r_1(t) + \cdots + r_n(t)$.

Demostración.

$$1. R(t) = P(T_1 > t, \dots, T_n > t) = P(T_1 > t) \cdots P(T_n > t) = R_1(t) \cdots R_n(t).$$

$$2. r(t) = -\frac{d}{dt} \ln R(t) = -\frac{d}{dt} \sum_{k=1}^n \ln R_k(t) = \sum_{k=1}^n r_k(t).$$

□

■ **Ejemplo.** La función de confiabilidad y la función de tasa de falla de un sistema de n componentes puestos en serie en donde cada uno de ellos tiene un tiempo de falla exponencial de parámetro λ son, respectivamente, $R(t) = e^{-\lambda nt}$, y $r(t) = \lambda nt$. La distribución del tiempo de falla T es exponencial de parámetro λn . El tiempo medio de falla es $E(T) = 1/(\lambda n)$, naturalmente decreciente conforme mayor es el número de componentes en el sistema, es decir, mientras mas componentes haya en el sistema en serie, menor es el tiempo de vida operacional del sistema en su conjunto pues una falla de cualesquiera de los componentes provoca que el sistema completo deje de funcionar. ◦

■ **Ejercicio.** En el contexto de sistemas de n componentes en serie, demuestre que la función de confiabilidad $R(t) = R_1(t) \cdots R_n(t)$ es una función decreciente de n , y en consecuencia el tiempo medio de falla también es una función decreciente de n . ■

Confiabilidad para sistemas en paralelo.

Considere ahora sistemas de n componentes puestos en paralelo como se muestra en la Figura 6.7, en donde cada uno de estos componentes funciona de manera independiente uno del otro. Tales tipos de sistemas funcionan en su conjunto si por lo menos uno de los componentes se encuentra en buen estado. Sistemas de este tipo se utilizan cuando se desea mantener un nivel alto de confiabilidad de funcionamiento, como por ejemplo los sistemas de vuelo de los aviones, o los sistemas de manejo de reactores nucleares.

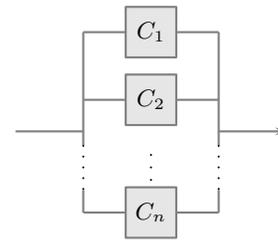


Figura 6.7:

El tiempo de falla T de un sistema de este tipo es el tiempo de falla más grande de los componentes, es decir, $T = \max\{T_1, \dots, T_n\}$, y que el evento $(T \leq t)$ es equivalente a que todos los componentes fallen antes del tiempo t .

Proposición. La función de confiabilidad del sistema en paralelo de la Figura 6.7 es $R(t) = 1 - (1 - R_1(t)) \cdots (1 - R_n(t))$.

Demostración. Primeramente se tiene que

$$F(t) = P(T \leq t) = P(T_1 \leq t, \dots, T_n \leq t) = F_1(t) \cdots F_n(t).$$

Por lo tanto

$$R(t) = 1 - F(t) = 1 - F_1(t) \cdots F_n(t) = 1 - (1 - R_1(t)) \cdots (1 - R_n(t)).$$

□

■ Ejercicio. Demuestre que la función de confiabilidad $R(t)$ para sistemas en paralelo recién encontrada es una función creciente de n . Es decir, mientras mas componentes haya en el sistema más seguro es éste. En consecuencia, demuestre que el tiempo medio de falla es una función creciente de n . ■

■ Ejercicio. Demuestre que las funciones de distribución y de densidad del tiempo de falla T del sistema en paralelo de la Figura 6.7 son

a) $F_T(t) = F^n(t)$.

b) $f_T(t) = n F^{n-1}(t) f(t)$. ■

Puede calcularse la función de tasa de falla del sistema $r(t)$ en términos de las funciones $r_1(t), \dots, r_n(t)$ pero la expresión no resulta ser simple.

■ Ejemplo. Para un sistema de n componentes puestos en paralelo en donde cada uno de ellos tiene un tiempo de falla exponencial de parámetro λ se tiene que

$$R(t) = 1 - (1 - e^{-\lambda t})^n.$$

$$r(t) = \lambda n (1 - e^{-\lambda t})^{n-1} e^{-\lambda t} / (1 - (1 - e^{-\lambda t})^n). \quad \circ$$

Confiabilidad para sistemas mezcla serie/paralelo. Pueden también considerarse sistemas que son una combinación de sistemas en serie y en paralelo como el que se muestra en la parte izquierda de la Figura 6.8. Sistemas de este tipo pueden

ser representados de manera equivalente como se muestra en la parte derecha de la misma figura.

■ Ejercicio. Demuestre que la función de confiabilidad del sistema de la Figura 6.8 es $R(t) = R_1(t)(1 - (1 - R_2(t))(1 - R_3(t)))$. ■

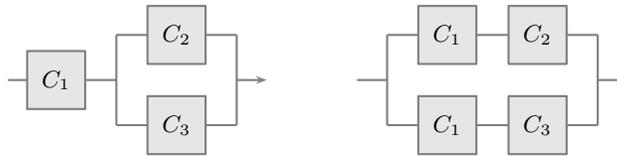


Figura 6.8:

■ Ejercicio. Encuentre la función de confiabilidad para cada uno de los sistemas de componentes que aparecen en la Figura 6.9. ■

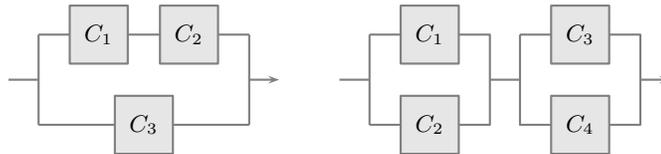


Figura 6.9:

Notas y referencias

En los textos de Karlin y Taylor [19], Basu [1], y Lawler [23] el lector puede encontrar algunos otros resultados y aplicaciones de los procesos de renovación.

Ejercicios

Procesos de renovación

107. Diga falso o verdadero para cada una de las siguientes igualdades de eventos. Explique en cada caso.
- $(N_t > n) = (W_n < t)$
 - $(N_t \geq n) = (W_n < t)$
 - $(N_t = n) = (W_n \leq t)$
 - $(N_t \leq n) = (W_n \geq t)$
 - $(N_t < n) = (W_n \geq t)$
108. Demuestre que $F^{*n}(t) \geq F^{*m}(t)$, para $n \leq m$.
109. Respecto de la definición de proceso de renovación, demuestre que los procesos $\{T_n : n = 1, 2, \dots\}$, $\{W_n : n = 0, 1, \dots\}$ y $\{N_t : t \geq 0\}$ pueden escribirse cada uno en términos de cualquier otro.
110. ¿Es la suma de dos procesos de renovación independientes nuevamente un proceso de renovación? Los siguientes ejercicios dan respuesta a esta pregunta.
- Demuestre que si la suma de dos procesos de renovación independientes resulta ser un proceso de Poisson, entonces ambos sumandos son procesos de Poisson.
 - Demuestre que si la suma de dos procesos de renovación independientes y con la misma función de distribución de interarribo es un proceso de renovación, entonces la suma y cada uno de los sumandos es un proceso de Poisson.
 - Mediante un contraejemplo demuestre que, en general, la suma de dos procesos de renovación independientes y con la misma función de distribución de interarribo no es necesariamente un proceso de renovación.
111. Sea $\{N_t\}$ un proceso de renovación. Demuestre que para cada $k \geq 1$,

$$E(N_t^k) = \sum_{n=0}^{\infty} [(n+1)^k - n^k] F^{*(n+1)}(t).$$

112. Sea $\{N_t\}$ un proceso de renovación con función de renovación $\Lambda(t) = E(N_t)$. Demuestre que para cualquier $t \geq 0$,

$$0 \leq \Lambda(t) \leq \frac{1}{1 - F(t)}.$$

113. Sea $\{N_t\}$ un proceso de renovación. Demuestre que la función $A(t) = E(W_{N_t+1})$ satisface la ecuación de renovación

$$A(t) = E(T) + \int_0^t A(t-s) dF(s).$$

114. Demuestre que el tiempo de vida restante γ_t y el tiempo de vida transcurrido δ_t , para un proceso de renovación, tienen distribución conjunta dada por la siguiente expresión: Para $0 < y < t$,

$$P(\gamma_t > x, \delta_t \leq y) = \int_{t-y}^t (1 - F(t+x-u)) d\Lambda(u).$$

115. Considere el proceso de Poisson visto como un proceso de renovación, es decir, los tiempos de vida tienen distribución exponencial con parámetro λ . Demuestre que la función de renovación $\Lambda(t) = \lambda t$ satisface la ecuación de renovación

$$\Lambda(t) = 1 - e^{-\lambda t} + \int_0^t \Lambda(t-s) \lambda e^{-\lambda s} ds.$$

116. Sea $\{N_t\}$ un proceso de Poisson de parámetro $\lambda > 0$. Demuestre que el tiempo de vida total δ_t tiene función de distribución la que aparece abajo. Demuestre además que $E(\delta_t) = \frac{1}{\lambda}(2 - e^{-\lambda t})$.

$$P(\delta_t \leq x) = \begin{cases} 1 - (1 + \lambda(t \wedge x)) e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{otro caso.} \end{cases}$$

117. Sea $\{N_t\}$ el proceso de conteo de un proceso de renovación. Defina $\Lambda(t) = E(N_t)$, y $\Lambda_2(t) = E(N_t^2)$. Demuestre que

$$a) \quad \Lambda_2(t) = \sum_{n=1}^{\infty} (2n-1) F^{*n}(t).$$

$$b) \quad \Lambda_2(t) = \Lambda(t) + 2 \int_0^t \Lambda(t-s) d\Lambda(s).$$

Encuentre además $\Lambda_2(t)$ cuando $\{N_t\}$ es un proceso de Poisson.

118. Encuentre la distribución de probabilidad del tiempo de vida restante, del tiempo de vida transcurrido y del tiempo de vida total para un proceso de Poisson. Encuentre además la esperanza de estas variables aleatorias.
119. Sea $\{N_t\}$ un proceso de renovación. Demuestre directamente que para cada n fijo,

$$\lim_{t \rightarrow \infty} P(N_t = n) = 0.$$

Confiabilidad

120. Sean T_1, \dots, T_n los tiempos de falla de cada uno de los n componentes puestos en serie como en la Figura 6.6. Suponga que las correspondientes funciones de distribución y de densidad son $F_1(t), \dots, F_n(t)$, y $f_1(t), \dots, f_n(t)$, respectivamente. Demuestre que las funciones de distribución y de densidad del tiempo de falla del sistema dado por $T = \min\{T_1, \dots, T_n\}$ son
- a) $F(t) = 1 - (1 - F_1(t)) \cdots (1 - F_n(t))$.
- b) $f(t) = \sum_{k=1}^n f_k(t) (1 - F_1(t) 1_{(k \neq 1)}) \cdots (1 - F_n(t) 1_{(k \neq n)})$.
121. Sean T_1, \dots, T_n los tiempos de falla de cada uno de los n componentes puestos en paralelo como en la Figura 6.7. Suponga que las correspondientes funciones de distribución y de densidad son $F_1(t), \dots, F_n(t)$, y $f_1(t), \dots, f_n(t)$, respectivamente. Demuestre que las funciones de distribución y de densidad del tiempo de falla del sistema dado por $T = \max\{T_1, \dots, T_n\}$ son
- a) $F(t) = F_1(t) \cdots F_n(t)$.
- b) $f(t) = \sum_{k=1}^n f_k(t) (1 - (1 - F_1(t)) 1_{(k \neq 1)}) \cdots (1 - (1 - F_n(t)) 1_{(k \neq n)})$.

CAPÍTULO 7

Martingalas

Existen varias acepciones para el término *martingala*. Una de ellas hace referencia a un tipo de proceso estocástico que básicamente cumple la identidad $E(X_{n+1} | X_0 = x_0, \dots, X_n = x_n) = x_n$. Hemos mencionado antes que este tipo de modelo corresponde a un proceso que representa la evolución del capital de un jugador que realiza una sucesión de apuestas justas. Parte de la motivación para el estudio de este tipo de procesos fue el buscar demostrar la inexistencia de estrategias ventajosas en este juego de apuestas. El concepto de martingala fue incorporado a la teoría de la probabilidad por Paul Lèvy, y buena parte de su desarrollo inicial fue realizado por Joseph Doob.

En este capítulo se presenta una introducción a la teoría de martingalas a tiempo discreto, aunque también se mencionan algunas definiciones generales a tiempo continuo.

7.1. Filtraciones

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad para modelar un cierto experimento aleatorio. La σ -álgebra \mathcal{F} es la estructura que agrupa a los eventos del experimento aleatorio a los cuales se les puede calcular su probabilidad. Suponga ahora que se consideran dos sub σ -álgebras \mathcal{F}_1 y \mathcal{F}_2 tales que $\mathcal{F}_1 \subseteq \mathcal{F}_2$. Entonces \mathcal{F}_2 contiene más información que \mathcal{F}_1 en el sentido de que, en general, un mayor número de conjuntos son considerados como eventos. Más generalmente, puede considerarse

una sucesión no decreciente de sub σ -álgebras $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$. Este tipo de estructuras surgen de manera natural, por ejemplo, para un proceso estocástico a tiempo discreto $\{X_n\}$, pues puede construirse la sucesión $\{\mathcal{F}_n\}$ de la siguiente forma

$$\mathcal{F}_n = \sigma\{X_1, \dots, X_n\},$$

en donde efectivamente se cumple que $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$. En este caso la σ -álgebra \mathcal{F}_n contiene los eventos para los cuales puede determinarse su ocurrencia o no ocurrencia, sólo con la información del proceso disponible hasta el tiempo n .

■ **Ejemplo.** Si las variables $\{X_n\}$ representan los resultados de lanzar sucesivamente un dado, y si se define el evento A como “La suma de los dos primeros resultados es mayor a 3”, entonces es claro que $A \notin \mathcal{F}_1$, sin embargo $A \in \mathcal{F}_2$. Por supuesto, si sucede por ejemplo que $X_1 = 4$, entonces sabemos que el evento A ocurre, pero sin esta información adicional la ocurrencia o no ocurrencia del evento A no puede determinarse sino hasta el segundo lanzamiento del dado. ◻

Estas consideraciones sobre sucesiones no decrecientes de σ -álgebras llevan a la definición de filtración.

Definición. Una *filtración* es una colección de σ -álgebras $\{\mathcal{F}_n\}_{n \geq 1}$ tal que $\mathcal{F}_n \subseteq \mathcal{F}_m$, cuando $n \leq m$. En particular, la *filtración natural* o *canónica* de un proceso $\{X_n\}$ es aquella sucesión de σ -álgebras definidas por $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$, $n \geq 1$.

A tiempo continuo las definiciones de estos conceptos son análogas: Una *filtración* es una colección no numerable de sub σ -álgebras $\{\mathcal{F}_t\}_{t \geq 0}$ tal que $\mathcal{F}_s \subseteq \mathcal{F}_t$, cuando $0 \leq s \leq t$. La *filtración natural* o *canónica* de un proceso a tiempo continuo $\{X_t : t \geq 0\}$ es la colección de σ -álgebras $\{\mathcal{F}_t\}_{t \geq 0}$ dadas por $\mathcal{F}_t = \sigma\{X_s : 0 \leq s \leq t\}$, esto es, \mathcal{F}_t es la mínima σ -álgebra que hace que cada una de las variables X_s , para valores de s en el intervalo $[0, t]$, sean medibles. A la σ -álgebra \mathcal{F}_t se le denomina la *historia* del proceso al tiempo t . Regresemos al caso de tiempo discreto.

Definición. Se dice que un proceso $\{X_n\}$ es *adaptado* a una filtración $\{\mathcal{F}_n\}$ si la variable X_n es \mathcal{F}_n -medible, para cada $n \geq 1$.

Sabemos que cada X_n es \mathcal{F} -medible por ser variable aleatoria. La condición de adaptabilidad requiere que X_n sea también una variable aleatoria respecto de la

sub σ -álgebra \mathcal{F}_n . Naturalmente todo proceso es adaptado a su filtración natural, y puede demostrarse que la filtración natural es la filtración más pequeña respecto de la cual el proceso es adaptado. El siguiente caso particular es necesario de considerar.

Definición. Se dice que el proceso $\{X_n : n \geq 1\}$ es *predecible* respecto de la filtración $\{\mathcal{F}_n\}_{n \geq 0}$ si para cada $n \geq 1$, la variable X_n es \mathcal{F}_{n-1} -medible.

Evidentemente todo proceso predecible es adaptado. La definición de adaptabilidad es la misma en el caso de procesos a tiempo continuo, y se pueden definir además las siguientes dos σ -álgebras:

$$\mathcal{F}_\infty = \sigma\left(\bigcup_{t \geq 0} \mathcal{F}_t\right),$$

$$\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s.$$

Se dice entonces que una filtración es *continua por la derecha* si $\mathcal{F}_{t+} = \mathcal{F}_t$. Si \mathcal{F}_t es una filtración continua por la derecha y la σ -álgebra inicial \mathcal{F}_0 contiene a todos los subconjuntos insignificantes de \mathcal{F} , entonces la filtración se llama *estándar*, o bien que *satisface las condiciones usuales*. Algunos cálculos requieren suponer estas hipótesis.

7.2. Tiempos de paro

Sea $\{X_n\}$ un proceso adaptado a una filtración $\{\mathcal{F}_n\}$. Un *tiempo de paro* es una variable aleatoria con valores en el espacio parametral de este proceso, que registra el tiempo en el que ocurre un cierto evento del proceso de tal forma que puede determinarse si ha ocurrido o no ha ocurrido tal evento al tiempo n con la información de la σ -álgebra \mathcal{F}_n . Esta variable aleatoria puede tomar el valor infinito cuando el evento de interés nunca ocurre.

Definición. Una variable aleatoria τ con valores en $\{1, 2, \dots\} \cup \{\infty\}$ es un *tiempo de paro* respecto de una filtración $\{\mathcal{F}_n\}$ si para cada $n \geq 1$ se cumple que $(\tau \leq n) \in \mathcal{F}_n$.

Bajo la interpretación de que τ es un tiempo aleatorio en el cual ocurre un cierto evento de interés, la condición $(\tau \leq n) \in \mathcal{F}_n$ puede interpretarse del siguiente modo: la pregunta de si el evento de interés ha ocurrido al tiempo n o antes, debe poder ser respondida con la información dada por la filtración al tiempo n . No es difícil comprobar que la condición que aparece en la definición es equivalente a la condición $(\tau = n) \in \mathcal{F}_n$, para cada $n \geq 1$.

■ Ejemplo: *Tiempo de primer arribo*. Sea X_1, X_2, \dots una sucesión de variables aleatorias adaptada a la filtración \mathcal{F}_n . Sea A un conjunto de Borel de \mathbb{R} , y defina $\tau = \min\{n \geq 1 : X_n \in A\}$, en donde conviene definir $\min \emptyset = \infty$. Es decir, τ es el primer momento en el que la sucesión toma un valor dentro del conjunto A , si acaso ello sucede. La variable aleatoria τ es un tiempo de paro pues para cualquier valor entero de n ,

$$(\tau = n) = (X_1 \notin A) \cap \dots \cap (X_{n-1} \notin A) \cap (X_n \in A) \in \mathcal{F}_n.$$

En particular, y recordando el problema de la ruina del jugador, tomando A como el conjunto $\{0, N\}$, el primer momento en el que la variable $X_n = \xi_1 + \dots + \xi_n$ toma uno de los dos valores del conjunto A es un tiempo de paro, y es el tiempo aleatorio en el cual el juego finaliza, ya sea porque el jugador se arruina o porque gana todo el capital. \circ

■ Ejemplo: *Tiempo de segundo arribo*. Sea nuevamente X_1, X_2, \dots una sucesión de variables aleatorias adaptada a la filtración \mathcal{F}_n , y sea A un conjunto de Borel de \mathbb{R} . Dado un tiempo de paro cualquiera τ_1 , defina ahora un segundo tiempo de paro de la siguiente forma $\tau_2 = \min\{n > \tau_1 : X_n \in A\}$. Es decir τ_2 es el primer momento, después de τ_1 , en el que el proceso toma un valor dentro del conjunto A . Naturalmente se cumple $\tau_1 \leq \tau_2$. Esta nueva variable resulta también ser un tiempo de paro pues el siguiente conjunto es un elemento de \mathcal{F}_n .

$$(\tau_2 = n) = \bigcup_{k=1}^{n-1} (\tau_1 = k) \cap (X_{k+1} \notin A) \cap \dots \cap (X_{n-1} \notin A) \cap (X_n \in A).$$

\circ

■ Ejercicio. Sea n un entero positivo. Demuestre que la variable aleatoria constante $\tau = n$ es un tiempo de paro. Este es un tiempo de paro determinista. Compruebe además que $\tau = \infty$ es también un tiempo de paro. \bullet

■ Ejercicio. Sea τ un tiempo de paro con valores en $\{1, 2, \dots\} \cup \{\infty\}$, y sea n un entero positivo. Demuestre que las siguientes variables también son tiempos de paro.

- a) $\tau \wedge n$.
- b) $\tau \vee n$.
- c) $\tau + n$.

En el caso de tiempos continuos, la definición de tiempo de paro es la siguiente. Una variable aleatoria $\tau : \Omega \rightarrow [0, \infty) \cup \{\infty\}$ es un *tiempo de paro* respecto de una filtración $\{\mathcal{F}_t\}$ si se cumple que $(\tau \leq t) \in \mathcal{F}_t$, para cada $t \geq 0$.

7.3. Martingalas

Definición. Se dice que un proceso $\{X_n\}$ es una *martingala* respecto de una filtración $\{\mathcal{F}_n\}$ si cumple las siguiente tres condiciones:

- a) Es integrable.
- b) Es adaptado a la filtración.
- c) Para cualesquiera $n \leq m$,

$$E(X_m | \mathcal{F}_n) = X_n, \text{ c.s.} \quad (7.1)$$

Cuando en lugar de (7.1) se cumple la desigualdad $E(X_m | \mathcal{F}_n) \geq X_n$, entonces el proceso es una *submartingala*, y si $E(X_m | \mathcal{F}_n) \leq X_n$, entonces es una *supermartingala*.

Las martingalas tienen una interpretación sencilla en términos de juegos justos que ya hemos mencionado antes: Si X_n denota el capital de un jugador al tiempo n , entonces la igualdad (7.1) establece que la fortuna promedio al tiempo futuro m , dado que se conoce la historia del juego hasta el tiempo n , es su capital al tiempo n , es decir, el juego es justo pues en promedio el jugador no pierde ni gana. En este sentido, la desigualdad $E(X_m | \mathcal{F}_n) \geq X_n$, correspondiente a la definición de submartingala, equivale a un juego favorable al jugador. La desigualdad contraria, el caso de supermartingala, corresponde a un juego desfavorable al jugador.

Puede comprobarse que la condición (7.1) es equivalente a la igualdad aparentemente más débil $E(X_{n+1} | \mathcal{F}_n) = X_n$. Esta última condición es la que a menudo usaremos para verificar la propiedad de martingala de un proceso a tiempo discreto. Además, cuando la filtración es la natural, es decir, cuando $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$, la condición de martingala puede escribirse en la forma $E(X_{n+1} | X_1, \dots, X_n) = X_n$.

Observe que toda martingala es al mismo tiempo una submartingala y una supermartingala, y que si $\{X_n\}$ es una submartingala, entonces $\{-X_n\}$ es una supermartingala. Por lo tanto, toda propiedad para submartingalas puede ser escrita también para supermartingalas bajo este cambio de signo. Por otro lado, tomando esperanza en (7.1) con $n = 1$ se obtiene la igualdad $E(X_m) = E(X_1)$, para cada $m \geq 1$, esto quiere decir que todas las variables aleatorias que conforman una martingala tienen la misma esperanza. Análogamente, tomando esperanza ahora en la condición de submartingala se obtiene que para cualesquiera tiempos $1 \leq n \leq m$, $E(X_m) \geq E(X_n)$, esto puede interpretarse en el sentido de que, en promedio, las submartingalas son procesos que *tienden a crecer*. Más adelante demostraremos que cuando la submartingala es acotada superiormente, es convergente. Este interesante resultado es el análogo estocástico al hecho de que toda sucesión de números reales creciente y acotada es convergente.

Para procesos a tiempos continuos, la condición de martingala se escribe $E(X_t | \mathcal{F}_s) = X_s$, para cualesquiera tiempos s y t tales que $0 \leq s \leq t$, sin olvidar la condiciones de adaptabilidad e integrabilidad para poder llamar a tal proceso una martingala.

7.4. Ejemplos

Veremos a continuación algunos ejemplos de martingalas.

Martingala del juego de apuestas. Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes idénticamente distribuidas y con esperanza finita. Defina $X_n = \xi_1 + \dots + \xi_n$, y considere la filtración $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$. La variable aleatoria X_n puede interpretarse como la fortuna de un jugador después de n apuestas sucesivas en donde la ganancia o pérdida en la k -ésima apuesta es ξ_k .

Claramente el proceso X_n es integrable y es adaptado. Además

$$\begin{aligned} E(X_{n+1} | \mathcal{F}_n) &= E(X_n + \xi_{n+1} | \mathcal{F}_n) \\ &= X_n + E(\xi_{n+1} | \mathcal{F}_n) \\ &= X_n + E(\xi_{n+1}). \end{aligned}$$

Cuando el resultado promedio de cualquiera de las apuestas es cero, el segundo sumando se anula y se tiene un ejemplo de martingala, es decir, un juego justo. En particular, una caminata aleatoria simétrica es una martingala. Si el resultado promedio de las apuestas es mayor o igual a cero, entonces $E(X_{n+1} | \mathcal{F}_n) \geq X_n$, y por lo tanto el proceso es una submartingala, un juego favorable al jugador. Finalmente cuando el resultado promedio de cualquier apuesta es menor o igual a cero, el proceso es una supermartingala pues $E(X_{n+1} | \mathcal{F}_n) \leq X_n$, correspondiente a un juego desfavorable al jugador.

■ **Ejercicio.** Para la martingala del juego de apuestas justas $X_n = \xi_1 + \dots + \xi_n$, demuestre que el proceso $X_n^2 - n$ es una martingala si, y sólo si, $\text{Var}(\xi) = 1$. ■

Martingala del proceso de Poisson centrado. Sea X_t un proceso de Poisson de parámetro λ . Entonces el proceso centrado $X_t - \lambda t$ es una martingala. En efecto, este nuevo proceso es integrable y adaptado. Además, para cada $0 \leq s \leq t$,

$$\begin{aligned} E(X_t - \lambda t | \mathcal{F}_s) &= E(X_t | \mathcal{F}_s) - \lambda t \\ &= E(X_t - X_s + X_s | \mathcal{F}_s) - \lambda t \\ &= E(X_t - X_s | \mathcal{F}_s) + X_s - \lambda t \\ &= E(X_t - X_s) + X_s - \lambda t \\ &= \lambda(t - s) + X_s - \lambda t \\ &= X_s - \lambda s. \end{aligned}$$

Este ejemplo es un caso particular del siguiente resultado que el lector puede fácilmente verificar siguiendo el cálculo anterior: Si un proceso integrable $\{X_t\}$ tiene incrementos independientes, entonces el proceso centrado $\{X_t - E(X_t)\}$ es una martingala.

Martingala de la esperanza condicional. Sea X una variable aleatoria integrable, y sean \mathcal{F}_1 y \mathcal{F}_2 dos sub σ -álgebras tales que $\mathcal{F}_1 \subseteq \mathcal{F}_2$. No es difícil comprobar la siguiente propiedad de la esperanza condicional: $E(E(X | \mathcal{F}_2) | \mathcal{F}_1) =$

$E(X | \mathcal{F}_1)$. Sea \mathcal{F}_n una filtración dada. Usando la identidad anterior demostraremos que la sucesión de variables aleatorias dada por $X_n = E(X | \mathcal{F}_n)$ es una martingala. Es claro que cada una de estas variables es integrable y por definición de esperanza condicional el proceso es adaptado a la filtración. Además

$$E(X_{n+1} | \mathcal{F}_n) = E(E(X | \mathcal{F}_{n+1}) | \mathcal{F}_n) = E(X | \mathcal{F}_n) = X_n.$$

Martingala de la urna de Polya. Suponga que una urna contiene inicialmente una bola blanca y una bola negra. Un experimento consiste en escoger una bola al azar y regresarla a la urna junto con otra bola del mismo color. Este experimento se repite varias veces. Sea X_n el número de bolas blancas en la urna después del n -ésimo ensayo. Es claro que $X_0 = 1$, y que después del n -ésimo ensayo hay $n+2$ bolas en la urna. Además $1 \leq X_n \leq n+1$, y por lo tanto $E|X_n| < \infty$. Las probabilidades de transición son las siguientes

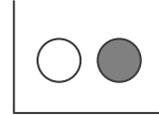


Figura 7.1:

$$P(X_{n+1} = k + 1 | X_n = k) = \frac{k}{n + 2},$$

$$P(X_{n+1} = k | X_n = k) = \frac{n + 2 - k}{n + 2}.$$

Observe que se puede escribir $X_{n+1} = X_n + Y_{n+1}$, en donde Y_{n+1} es una variable aleatoria que toma el valor $+1$ cuando se escoge una bola blanca en la extracción $n + 1$, y el valor 0 cuando la bola escogida es negra, por lo tanto

$$E(X_{n+1} | X_n) = X_n + (0) \frac{n + 2 - X_n}{n + 2} + (1) \frac{X_n}{n + 2} = \frac{n + 3}{n + 2} X_n.$$

Este cálculo demuestra que el proceso $\{X_n\}$ no es una martingala. Sin embargo, si se define $M_n = X_n/(n + 2)$, lo cual representa la fracción de bolas blancas en la urna después de la n -ésima extracción, entonces se tiene que

$$E(M_{n+1} | \mathcal{F}_n) = E\left(\frac{X_{n+1}}{n + 3} | \mathcal{F}_n\right) = \frac{X_n}{n + 2} = M_n.$$

Es decir, hemos comprobado que el proceso $\{M_n\}$ es una martingala. ¿Se modifica el resultado si la configuración inicial de la urna es distinta a la considerada? ¿Se modifica el resultado si en lugar de agregar una bola adicional se agregan r bolas del mismo color?

7.5. Procesos detenidos

Sea X_n un proceso estocástico adaptado a una filtración \mathcal{F}_n , y sea τ un tiempo de paro respecto de la misma filtración. En ocasiones es necesario considerar un proceso de la forma $X_{\tau \wedge n}$, en donde $\tau \wedge n = \min\{\tau, n\}$. A este tipo de procesos se les llama *procesos detenidos*, por ejemplo, suponga que $\tau = k$, entonces

$$X_{\tau \wedge n} = \begin{cases} X_n & \text{si } n \leq k, \\ X_k & \text{si } n > k. \end{cases}$$

Es decir, como función del parámetro n el proceso se vuelve constante a partir del tiempo aleatorio τ .

Medibilidad y adaptabilidad de $\{X_{\tau \wedge n}\}$. Las variables del proceso detenido son efectivamente variables aleatorias y el proceso mismo es adaptado a la misma filtración pues para cualquier número real x , y para cada número natural n ,

$$(X_{\tau \wedge n} \leq x) = \left[\bigcup_{k=1}^n (\tau = k) \cap (X_k \leq x) \right] \cup [(\tau > n) \cap (X_n \leq x)].$$

La expresión del lado derecho es claramente un elemento de \mathcal{F}_n . Si el proceso original es integrable, entonces el proceso detenido es también integrable pues

$$\begin{aligned} E|X_{\tau \wedge n}| &= E(|X_\tau| 1_{(\tau \leq n)}) + E(|X_n| 1_{(\tau > n)}) \\ &\leq \sum_{k=1}^n E(|X_k| 1_{(\tau=k)}) + E(|X_n| 1_{(\tau > n)}) \\ &\leq \sum_{k=1}^n E|X_k| + E|X_n| < \infty. \end{aligned}$$

Por ejemplo, considere que $\{X_n\}$ es la martingala del juego de apuestas. Suponga que el jugador decide dejar de jugar cuando pierda todo su capital o cuando consigue ganar el doble de su capital inicial. El momento aleatorio en el que ocurre alguno de estos dos eventos es un tiempo de paro τ . El capital del jugador en cualquier tiempo n puede expresarse como $X_{\tau \wedge n}$.

■ Ejercicio. Sea $\{X_n\}$ un proceso adaptado a la filtración $\{\mathcal{F}_n\}$, y sea τ un tiempo de paro a tiempo discreto que además es finito, es decir, $P(\tau < \infty) = 1$. Demuestre que X_τ es una variable aleatoria. ■

Demostremos a continuación que una martingala detenida sigue siendo martingala.

Proposición. Si $\{X_n\}$ es una martingala, submartingala o supermartingala, y τ es un tiempo de paro respecto de la misma filtración, entonces el proceso detenido $\{X_{\tau \wedge n}\}$ también lo es.

Demostración. Hemos demostrado antes que el proceso detenido es adaptado e integrable. Resta demostrar la propiedad de martingala. El caso submartingala y supermartingala se obtienen modificando adecuadamente la penúltima igualdad en el siguiente análisis.

$$\begin{aligned}
 E(X_{\tau \wedge (n+1)} | \mathcal{F}_n) &= \sum_{k=1}^n E(X_k \mathbf{1}_{(\tau=k)} | \mathcal{F}_n) + E(X_{n+1} \mathbf{1}_{(\tau>n)} | \mathcal{F}_n) \\
 &= \sum_{k=1}^n X_k \mathbf{1}_{(\tau=k)} + E(X_{n+1} | \mathcal{F}_n) \mathbf{1}_{(\tau>n)} \\
 &= \sum_{k=1}^n X_k \mathbf{1}_{(\tau=k)} + X_n \mathbf{1}_{(\tau>n)} \\
 &= X_{\tau \wedge n}
 \end{aligned}$$

□

7.6. Una aplicación: Estrategias de juego

Considere nuevamente la sucesión de variables aleatorias independientes idénticamente distribuidas $\{\xi_n\}$ tal que $P(\xi = +1) = 1/2$ y $P(\xi = -1) = 1/2$, y con filtración natural $\{\mathcal{F}_n\}$. Considere la suma $X_n = \xi_1 + \dots + \xi_n$, es decir, X_n es una martingala que representa el total de ganancias en una serie de n apuestas justas de una unidad monetaria. Suponga ahora que el monto de cada apuesta no es uno, sino una cantidad a_n para la n -ésima apuesta. Supondremos que a_n es una variable aleatoria que el jugador determina a partir de la información de las $n-1$ apuestas previas, y por lo tanto es una variable \mathcal{F}_{n-1} -medible, es decir, se trata de un proceso predecible. A la colección de variables aleatorias $\{a_n\}$ con esta característica se le llama una *estrategia de juego*. Bajo una de estas estrategias, el capital del

jugador después de la n -ésima apuesta es ahora $A_n = a_1 \xi_1 + \cdots + a_n \xi_n$, que es \mathcal{F}_n -medible. Bajo la hipótesis de que la estrategia de juego consta de variables aleatorias acotadas, se cumple que el proceso $\{A_n\}$ es integrable y cumple además la propiedad de martingala pues

$$\begin{aligned}
 E(A_{n+1} | \mathcal{F}_n) &= E(A_n + a_{n+1} \xi_{n+1} | \mathcal{F}_n) \\
 &= A_n + a_{n+1} E(\xi_{n+1} | \mathcal{F}_n) \\
 &= A_n + a_{n+1} E(X_{n+1} - X_n | \mathcal{F}_n) \\
 &= A_n + a_{n+1} (E(X_{n+1} | \mathcal{F}_n) - X_n) \\
 &= A_n.
 \end{aligned}$$

Esto quiere decir que bajo cualquier estrategia de juego, el proceso de ganancias $\{A_n\}$ es una martingala siempre y cuando el proceso original $\{X_n\}$ lo sea. Este resultado es importante saber para los apostadores pues quiere decir que no existe una estrategia de juego que convierta un juego justo en un juego favorable o desfavorable al jugador. El mismo análisis demuestra que si el proceso original $\{X_n\}$ es una submartingala o supermartingala y la estrategia de juego consta de variables aleatorias no negativas y acotadas, entonces el proceso $\{A_n\}$ sigue siendo una submartingala o supermartingala.

Uno de los varios significados del término martingala, y que parece ser el original, establece que una martingala es una estrategia de juego en la cual un jugador apuesta sucesivamente en cada lanzamiento de una moneda honesta del siguiente modo: Inicialmente apuesta una unidad monetaria. En caso de perder, dobla el monto de la apuesta para el siguiente lanzamiento. En caso de ganar, vuelve a apostar una unidad monetaria en el siguiente lanzamiento. En la Figura 7.2 se muestra una tabla con algunos resultados siguiendo esta estrategia de juego.

Monto de la apuesta	1	2	4	8	1	1	2	1
Resultado del lanzamiento	x	x	x	✓	✓	x	✓	x
Ganancia	-1	-3	-7	1	2	1	3	2

Figura 7.2:

Bajo esta estrategia de juego resulta que, cada vez que el jugador acierta, se recupera de las pérdidas anteriores e incluso incrementa su fortuna en una unidad monetaria. En efecto, si el jugador pierde las primeras n apuestas y gana la apuesta

$n + 1$, entonces su capital al tiempo $n + 1$ es

$$\begin{aligned}
 \underbrace{-1 - 2^1 - 2^2 - \dots - 2^{n-1}}_{n \text{ apuestas perdidas}} + \underbrace{2^n}_{\text{apuesta } n+1 \text{ ganada}} &= -\sum_{k=0}^{n-1} 2^k + 2^n \\
 &= -\frac{1-2^n}{1-2} + 2^n \\
 &= (1-2^n) + 2^n \\
 &= 1.
 \end{aligned}$$

De esta manera si el jugador pudiera mantener esta estrategia de juego tendría una unidad ganada por cada acierto que haya conseguido. Parece ser una estrategia segura de ganar, sin embargo, veamos cuánto, en promedio, podría ser su déficit justo antes de recuperarse, es decir, calcularemos $E(X_{\tau-1})$, en donde τ es el tiempo de paro en el que el jugador acierta por primera vez. Puede comprobarse que este tiempo de paro es finito casi seguramente, es decir, $P(\tau < \infty) = 1$. De hecho, con probabilidad uno, el jugador tendría eventualmente un éxito aún cuando sus probabilidades de acertar fueran pequeñas. Como hemos visto, después de n apuestas consecutivas perdidas el capital empeñado es $-1 - 2^1 - 2^2 - \dots - 2^{n-1} = 1 - 2^n$, y la probabilidad de perder n apuestas sucesivas y ganar la apuesta $n + 1$ es $1/2^{n+1}$. Por lo tanto

$$\begin{aligned}
 E(X_{\tau-1}) &= \sum_{n=1}^{\infty} E(X_{\tau-1} | \tau = n) P(\tau = n) \\
 &= \sum_{n=1}^{\infty} E(X_{n-1}) P(\tau = n) \\
 &= \sum_{n=1}^{\infty} (1 - 2^{n-1}) \frac{1}{2^n} \\
 &= -\infty.
 \end{aligned}$$

Es decir, se requeriría de un capital promedio infinito y poder apostar una infinidad de veces para llevar a cabo con éxito esta estrategia, algo no muy factible en términos de dinero y tiempo.

7.7. Teorema de paro opcional y aplicaciones

Hemos observado antes que para una martingala X_n se cumple que $E(X_n) = E(X_1)$, para cualquier valor de n . Si además se tiene un tiempo de paro finito τ , no necesariamente es cierto que $E(X_\tau) = E(X_1)$, e incluso el número $E(X_\tau)$ podría no ser finito como en la estrategia de juego llamada martingala analizada en la sección anterior. El siguiente resultado establece condiciones bajo las cuales la variable X_τ tiene la misma esperanza que la martingala.

Teorema de paro opcional. Sea $\{X_n\}$ una martingala y sea τ un tiempo de paro finito, ambos respecto de una filtración $\{\mathcal{F}_n\}$, tales que

a) X_τ es integrable.

b) $\lim_{n \rightarrow \infty} E(X_n 1_{(\tau > n)}) = 0$.

Entonces $E(X_\tau) = E(X_n)$, para cualquier $n \geq 1$.

Demostración. La observación importante es que para cualquier $n \geq 1$,

$$X_\tau = X_{\tau \wedge n} + (X_\tau - X_n) 1_{(\tau > n)}.$$

Tomando esperanza y usando el hecho de que $X_{\tau \wedge n}$ es una martingala,

$$\begin{aligned} E(X_\tau) &= E(X_{\tau \wedge n}) + E((X_\tau - X_n) 1_{(\tau > n)}) \\ &= E(X_1) + E(X_\tau 1_{(\tau > n)}) - E(X_n 1_{(\tau > n)}). \end{aligned}$$

Como el proceso original es una martingala, el primer sumando es $E(X_n)$. Haciendo $n \rightarrow \infty$, el tercer sumando se anula por hipótesis. Usando la hipótesis de integrabilidad de X_τ y el teorema de convergencia dominada, el segundo sumando converge también a cero pues es la cola de la serie convergente

$$E(X_\tau) = E\left(\sum_{k=1}^{\infty} X_k 1_{(\tau=k)}\right) = \sum_{k=1}^{\infty} E(X_k 1_{(\tau=k)}).$$

□

Como una aplicación del teorema de paro opcional calcularemos algunos tiempos medios de arriba en caminatas aleatorias.

Caso caminata simétrica, barrera simétrica

Sea $\{X_n\}$ una caminata aleatoria simétrica simple sobre \mathbb{Z} que inicia en cero, y sea $b \geq 1$ un entero cualquiera. Defina el tiempo de paro

$$\tau = \min \{n \geq 1 : |X_n| = b\},$$

es decir, τ es el primer momento en el que la caminata alcanza, en valor absoluto, el nivel b , véase la Figura 7.3. Nos interesa encontrar $E(\tau)$, esto es, el número promedio de pasos que le toma a la caminata llegar al nivel $\pm b$. Sabemos que tanto el proceso $\{X_n\}$ como $\{X_n^2 - n\}$ son martingalas. Suponiendo de manera preliminar que las condiciones del teorema de paro opcional se cumplen, se tiene que $E(X_\tau^2 - \tau) = E(X_1^2 - 1) = 0$. Por lo tanto, $E(\tau) = E(X_\tau^2) = b^2$. La última igualdad se obtiene al observar que $|X_\tau| = b$. En palabras, este resultado dice que la caminata aleatoria simétrica simple que inicia en cero necesita, en promedio, b^2 pasos para alcanzar, en valor absoluto, el nivel b .

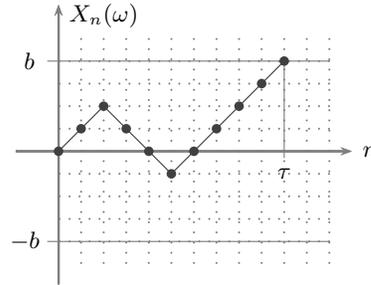


Figura 7.3:

Caso caminata simétrica, barrera asimétrica

Generalizaremos ahora el cálculo anterior para el caso en el que se tiene una barrera inferior $-a$ y una barrera superior b , con $a, b \in \mathbb{N}$, no necesariamente idénticos. La idea es aplicar nuevamente el teorema de paro opcional aunque los cálculos no son tan inmediatos. Supongamos entonces que $\{X_n\}$ es una caminata aleatoria simétrica simple que inicia en cero y defina el tiempo de paro

$$\tau = \min \{n \geq 0 : X_n = b \text{ ó } X_n = -a\},$$

en donde $a, b \in \mathbb{N}$. Nuevamente el proceso centrado $\{X_n^2 - n\}$ es una martingala y por lo tanto $E(X_\tau^2 - \tau) = E(X_1^2 - 1) = 0$, de donde se obtiene $E(\tau) = E(X_\tau^2)$. Ahora no es tan sencillo calcular esta esperanza pues X_τ^2 puede tomar dos valores, b^2 o a^2 . Entonces

$$E(\tau) = E(X_\tau^2) = b^2 P(X_\tau = b) + a^2 P(X_\tau = -a).$$

Defina $u_k = P(X_\tau = b | X_0 = k)$. Usando análisis del primer paso puede comprobarse que u_k cumple la ecuación en diferencias $2u_k = u_{k+1} + u_{k-1}$, con condiciones de frontera $u_b = 1$ y $u_{-a} = 0$, y cuya solución es $u_k = (a + k)/(a + b)$.

Análogamente, definiendo $v_k = P(X_\tau = -a | X_0 = k)$, se encuentra que v_k cumple la ecuación en diferencias $2v_k = v_{k+1} + v_{k-1}$, con condiciones de frontera $v_b = 0$ y $v_{-a} = 1$, y cuya solución es $v_k = (b - k)/(a + b)$.

Por lo tanto,

$$\begin{aligned} E(\tau) &= E(X_\tau^2) \\ &= b^2 P(X_\tau = b) + a^2 P(X_\tau = -a) \\ &= b^2 u_0 + a^2 v_0 \\ &= b^2 \frac{a}{a+b} + a^2 \frac{b}{a+b} \\ &= ab. \end{aligned}$$

Caso caminata asimétrica, barrera asimétrica

En este caso $\{X_n\}$ no es una martingala pero debido a la propiedad de incrementos independientes, el proceso centrado $\{X_n - n(p - q)\}$ lo es. Suponiendo nuevamente que las condiciones del teorema de paro opcional se cumplen, se tiene que $E(X_\tau - \tau(p - q)) = E(X_1 - (p - q)) = 0$. De donde se obtiene $E(\tau) = E(X_\tau)/(p - q)$. El problema es nuevamente encontrar $E(X_\tau)$. Tenemos que

$$E(X_\tau) = b P(X_\tau = b) - a P(X_\tau = -a).$$

Defina nuevamente $u_k = P(X_\tau = b | X_0 = k)$. Usando análisis del primer paso puede comprobarse que u_k cumple la ecuación en diferencias $u_k = p u_{k+1} + q u_{k-1}$, con condiciones de frontera $u_b = 1$ y $u_{-a} = 0$, y cuya solución es

$$u_k = \frac{(p/q)^{b-k} - (p/q)^{a+b}}{1 - (p/q)^{a+b}}.$$

Análogamente, definiendo $v_k = P(X_\tau = -a | X_0 = k)$, se encuentra que v_k cumple la ecuación en diferencias $v_k = p v_{k+1} + q v_{k-1}$, con condiciones de frontera $v_b = 0$ y $v_{-a} = 1$, y cuya solución es

$$v_k = \frac{(q/p)^{a+k} - (q/p)^{a+b}}{1 - (q/p)^{a+b}}.$$

Por lo tanto,

$$\begin{aligned}
 E(X_\tau) &= b u_0 - a v_0 \\
 &= b \frac{(p/q)^b - (p/q)^{a+b}}{1 - (p/q)^{a+b}} - a \frac{(q/p)^a - (q/p)^{a+b}}{1 - (q/p)^{a+b}} \\
 &= b - (a+b) \frac{1 - (p/q)^b}{1 - (p/q)^{a+b}}.
 \end{aligned}$$

Entonces

$$E(\tau) = \frac{b}{p-q} - \frac{a+b}{p-q} \cdot \frac{1 - (p/q)^b}{1 - (p/q)^{2b}}.$$

Verificaremos a continuación la validez de las tres condiciones del teorema de paro opcional para esta aplicación cuando la caminata y las barreras son simétricas.

a) Demostraremos que $\tau < \infty$ c.s. La estrategia consiste en considerar bloques sucesivos de $2b$ pasos en la caminata aleatoria. Observe que el evento $(\tau = \infty)$ es el límite de la sucesión decreciente de eventos $(\tau > 2bk)$, para $k = 1, 2, \dots$, y que el evento $(\tau > 2bk)$ está contenido en el evento “Ninguno de los primeros k bloques contiene únicamente valores $+1$ ”. La utilidad de este evento mayor radica en que es fácil calcular su probabilidad como veremos a continuación. Tenemos que

$$\begin{aligned}
 P(\tau = \infty) &= \lim_{k \rightarrow \infty} P(\tau > 2bk) \\
 &\leq \lim_{k \rightarrow \infty} P(\text{“Ninguno de los primeros } k \text{ bloques} \\
 &\quad \text{contiene únicamente valores } +1\text{”}) \\
 &= \lim_{k \rightarrow \infty} (1 - (1/2)^{2b})^k \\
 &= 0.
 \end{aligned}$$

b) Demostraremos ahora que $E(|X_\tau^2 - \tau|) < \infty$. Como $X_\tau^2 = b^2$, se tiene que

$$\begin{aligned}
 E(|X_\tau^2 - \tau|) &\leq b^2 + E(\tau) \\
 &= b^2 + \sum_{k=0}^{\infty} k P(\tau = k) \\
 &= b^2 + \sum_{k=0}^{\infty} \sum_{j=1}^{2b} (2bk + j) P(\tau = 2bk + j) \\
 &\leq b^2 + \sum_{k=0}^{\infty} \sum_{j=1}^{2b} (2b(k+1)) P(\tau > 2bk) \\
 &\leq b^2 + (2b)^2 \sum_{k=0}^{\infty} (k+1) (1 - (1/2)^{2b})^k \\
 &< \infty.
 \end{aligned}$$

c) Finalmente verificaremos que $E((X_n^2 - n)1_{(\tau > n)}) \rightarrow 0$, cuando $n \rightarrow \infty$.

$$\begin{aligned}
 E(|X_n^2 - n|1_{(\tau > n)}) &\leq E(X_n^2 1_{(\tau > n)}) + E(n 1_{(\tau > n)}) \\
 &\leq N^2 P(\tau > n) + E(\tau 1_{(\tau > n)}).
 \end{aligned}$$

La sucesión de eventos $(\tau > n)$ es monótona decreciente y por lo tanto convergente, su límite es el evento $(\tau = \infty)$ que tiene probabilidad cero. El primer sumando por lo tanto se hace pequeño cuando n crece. Como $E(\tau) < \infty$, la sucesión de variables $\tau 1_{(\tau > n)}$ es también decreciente y consta de variables aleatorias integrables cuyo límite es cero c.s. El segundo sumando por tanto también converge a cero.

7.8. Algunas desigualdades

Proposición. (Desigualdad maximal de Doob.) Sea X_n una submartingala no negativa y defina $X_n^* = \max\{X_1, \dots, X_n\}$. Entonces para cualquier $\lambda > 0$,

$$\lambda P(X_n^* \geq \lambda) \leq E(X_n 1_{(X_n^* \geq \lambda)}).$$

Demostración. Para cada n natural defina el tiempo de paro

$$\tau = n \wedge \min\{1 \leq k \leq n : X_k \geq \lambda\}.$$

Es decir, τ es el primer momento hasta n en el que el proceso alcanza o rebasa el valor λ . Si tal evento nunca ocurre, entonces $\tau = n$. Como X_n es una submartingala y $1 \leq \tau \leq n$, se tiene que $E(X_n) \geq E(X_\tau)$. Observe que si ocurre el evento $(X_n^* \geq \lambda)$, entonces $X_\tau \geq \lambda$, y si ocurre $(X_n^* < \lambda)$, entonces $\tau = n$. Por lo tanto

$$\begin{aligned} E(X_n) &\geq E(X_\tau) \\ &= E(X_\tau 1_{(X_n^* \geq \lambda)}) + E(X_\tau 1_{(X_n^* < \lambda)}) \\ &\geq \lambda P(X_n^* \geq \lambda) + E(X_n 1_{(X_n^* < \lambda)}). \end{aligned}$$

Es decir,

$$\begin{aligned} \lambda P(X_n^* \geq \lambda) &\leq E(X_n) - E(X_n 1_{(X_n^* < \lambda)}) \\ &= E(X_n 1_{(X_n^* \geq \lambda)}). \end{aligned}$$

□

Proposición. (Desigualdad maximal de Doob en L^2 .) Sea X_n una submartingala no negativa y cuadrado integrable. Para $X_n^* = \max\{X_1, \dots, X_n\}$ se tiene que

$$E(|X_n^*|^2) \leq 4 E(X_n^2).$$

Demostración. El segundo momento de una variable aleatoria no negativa X puede expresarse, cuando existe, de la siguiente forma

$$E(X^2) = 2 \int_0^\infty x P(X > x) dx.$$

Usando esta expresión, la primera desigualdad maximal, el teorema de Fubini y

después la desigualdad de Cauchy-Schwarz se tiene que

$$\begin{aligned}
 E(|X_n^*|^2) &= 2 \int_0^\infty x P(X_n^* > x) dx \\
 &\leq 2 \int_0^\infty E(X_n \mathbf{1}_{(X_n^* \geq x)}) dx. \\
 &= 2 \int_0^\infty \left(\int_{(X_n^* \geq x)} X_n dP \right) dx \\
 &= 2 \int_\Omega X_n \int_0^{X_n^*} dx dP \\
 &= 2 \int_\Omega X_n X_n^* dP \\
 &= 2E(X_n X_n^*) \\
 &\leq 2\sqrt{E|X_n|^2} \sqrt{E|X_n^*|^2}.
 \end{aligned}$$

Por lo tanto,

$$\frac{E(|X_n^*|^2)}{\sqrt{E|X_n^*|^2}} \leq 2\sqrt{E|X_n|^2}.$$

Elevando al cuadrado se obtiene el resultado. \square

7.9. Convergencia de martingalas

Proposición. Sea $\{X_n\}$ una submartingala y sean τ_1 y τ_2 dos tiempos de paro acotados tales que $0 \leq \tau_1 \leq \tau_2 \leq N$, con $N \in \mathbb{N}$ fijo. Entonces

$$E(X_{\tau_1}) \leq E(X_{\tau_2}).$$

Demostración. Sea k fijo tal que $k \leq n \leq N$. Entonces

$$\begin{aligned}
 E(X_{n+1} \mathbf{1}_{(\tau_1=k)} \mathbf{1}_{(\tau_2>n)}) &= E(E(X_{n+1} \mathbf{1}_{(\tau_1=k)} \mathbf{1}_{(\tau_2>n)} | \mathcal{F}_n)) \\
 &= E(E(X_{n+1} | \mathcal{F}_n) \mathbf{1}_{(\tau_1=k)} \mathbf{1}_{(\tau_2>n)}) \\
 &\geq E(X_n \mathbf{1}_{(\tau_1=k)} \mathbf{1}_{(\tau_2>n)}).
 \end{aligned}$$

Por lo tanto

$$\begin{aligned} E(X_{\tau_2 \wedge n} 1_{(\tau_1=k)}) &= E(X_{\tau_2} 1_{(\tau_1=k)} 1_{(\tau_2 \leq n)}) + E(X_n 1_{(\tau_1=k)} 1_{(\tau_2 > n)}) \\ &\leq E(X_{\tau_2} 1_{(\tau_1=k)} 1_{(\tau_2 \leq n)}) + E(X_{n+1} 1_{(\tau_1=k)} 1_{(\tau_2 > n)}) \\ &= E(X_{\tau_2 \wedge (n+1)} 1_{(\tau_1=k)}). \end{aligned}$$

Esto quiere decir que la función $n \mapsto E(X_{\tau_2 \wedge n} 1_{(\tau_1=k)})$ es monótona creciente. Evaluando esta función en $n = k$ y después en $n = N$ se obtiene la desigualdad

$$E(X_k 1_{(\tau_1=k)}) \leq E(X_{\tau_2} 1_{(\tau_1=k)}).$$

Entonces

$$\begin{aligned} E(X_{\tau_1}) &= \sum_{k=0}^N E(X_{\tau_1} 1_{(\tau_1=k)}) \\ &= \sum_{k=0}^N E(X_k 1_{(\tau_1=k)}) \\ &\leq \sum_{k=0}^N E(X_{\tau_2} 1_{(\tau_1=k)}) \\ &= E(X_{\tau_2}). \end{aligned}$$

□

Sea $\{X_k\}$ un proceso adaptado a una filtración $\{\mathcal{F}_k\}$, y sean $a < b$ dos números reales. Consideraremos que n es un número natural cualquiera. Defina la sucesión creciente de tiempos de paro

$$\begin{aligned} \tau_1 &= \text{mín} \{k \geq 1 : X_k > b\}, \\ \tau_2 &= \text{mín} \{k > \tau_1 : X_k < a\}, \\ \tau_3 &= \text{mín} \{k > \tau_2 : X_k > b\}, \\ \tau_4 &= \text{mín} \{k > \tau_3 : X_k < a\}, \\ &\vdots \end{aligned}$$

Si alguno de los conjuntos señalados es vacío o bien cuando $\tau_k \geq n$, para alguna k , se define $\tau_k = n$. De esta forma se tiene la sucesión $\tau_1 \leq \tau_2 \leq \dots \leq n$. Una de tales sucesiones se muestra en la Figura 7.4, en donde se presenta una trayectoria del proceso con sus valores

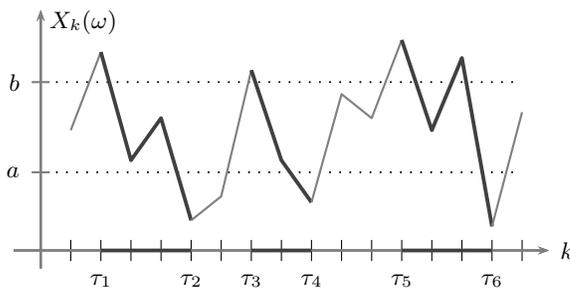


Figura 7.4:

unidos por una línea continua para una mejor visualización. Nos interesa contar el número de *cruces* de arriba hacia abajo, que las trayectorias del proceso realizan sobre el intervalo $[a, b]$. En la gráfica se muestran tres cruces de este tipo, los cuales se encuentran remarcados en la trayectoria.

Observe que antes del valor n y para valores impares en el subíndice de τ , el proceso se encuentra arriba de b en ese momento, mientras que para valores pares del subíndice, el proceso se encuentra por abajo del nivel a , es decir, entre los tiempos τ_{2k-1} y τ_{2k} el proceso realiza un cruce de arriba hacia abajo.

El número de cruces completos, de arriba hacia abajo, que el proceso realiza sobre el intervalo $[a, b]$ antes del tiempo n es el máximo entero k tal que $\tau_{2k} < n$, y se denota por $D_n[a, b]$, es decir,

$$D_n[a, b] = \text{máx} \{k \geq 1 : \tau_{2k} < n\}.$$

Si el conjunto de cruces es vacío, se define $D_n[a, b] = 0$. La letra D usada para denotar este número proviene del término en inglés *Downcrossing*. Observe que para cada n , la función $D_n[a, b] : \Omega \rightarrow \{0, 1, \dots\}$ es una variable aleatoria pues para cada entero k , el conjunto $(D_n[a, b] \geq k) = (\tau_{2k} < n)$ es medible. De esta manera se tiene la sucesión monótona de variables aleatorias $D_1[a, b] \leq D_2[a, b] \leq \dots$ que es convergente al número total de cruces

$$D[a, b] = \sup \{k \geq 1 : \tau_{2k} < \infty\}.$$

En la demostración que se presenta a continuación sobre la convergencia de submartingalas se hace uso de este número de cruces. Empezaremos estimando la esperanza de $D_n[a, b]$.

Proposición. Sea $\{X_n\}$ una submartingala. Entonces

$$E(D_n[a, b]) \leq \frac{1}{b-a} E(X_n - b)^+ \leq \frac{1}{b-a} (\sup_n E|X_n| + |b|).$$

Demostración. Defina la sucesión de eventos $A_k = (\tau_k < n)$ para $k = 1, 2, \dots$. Por la monotonía de los tiempos de paro definidos antes se tiene que $A_1 \supseteq A_2 \supseteq \dots$. Eventualmente los elementos de esta sucesión son el conjunto vacío pues no pueden efectuarse demasiados cruces en un tiempo limitado n . Observe además que cuando ocurre el evento A_{2k-1} el proceso al tiempo τ_{2k-1} se encuentra por arriba de b , mientras que cuando ocurre A_{2k} , el proceso al tiempo τ_{2k} se encuentra por abajo de a . A continuación usaremos la propiedad de submartingala aplicada a los tiempos de paro acotados $\tau_{2k-1} \leq \tau_{2k} \leq n$. Tenemos entonces que $E(X_{\tau_{2k-1}}) \leq E(X_{\tau_{2k}})$, es decir,

$$\int_{\Omega} X_{\tau_{2k-1}} dP \leq \int_{\Omega} X_{\tau_{2k}} dP.$$

Separando ambas regiones de integración como $\Omega = A_{2k-1} \cup A_{2k-1}^c$, se tiene que

$$\int_{A_{2k-1}} X_{\tau_{2k-1}} dP + \int_{A_{2k-1}^c} X_{\tau_{2k-1}} dP \leq \int_{A_{2k-1}} X_{\tau_{2k}} dP + \int_{A_{2k-1}^c} X_{\tau_{2k}} dP.$$

Ahora observe que sobre el conjunto A_{2k-1}^c , se cumple que $\tau_{2k-1} = \tau_{2k} = n$. Por lo tanto, la segunda y cuarta integral coinciden y podemos omitirlas de esta desigualdad. Añadiendo la constante b , se llega al siguiente resultado

$$0 \leq \int_{A_{2k-1}} (X_{\tau_{2k-1}} - b) dP \leq \int_{A_{2k-1}} (X_{\tau_{2k}} - b) dP.$$

Entonces

$$\begin{aligned}
0 &\leq \int_{A_{2k-1}} (X_{\tau_{2k}} - b) dP \\
&= \int_{A_{2k}} (X_{\tau_{2k}} - b) dP + \int_{A_{2k-1} \setminus A_{2k}} (X_{\tau_{2k}} - b) dP \\
&\leq (a - b)P(A_{2k}) + \int_{A_{2k-1} \setminus A_{2k}} (X_n - b) dP.
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
P(A_{2k}) &\leq \frac{1}{b-a} \int_{A_{2k-1} \setminus A_{2k}} (X_n - b) dP \\
&\leq \frac{1}{b-a} \int_{A_{2k-1} \setminus A_{2k}} (X_n - b)^+ dP.
\end{aligned}$$

Como $P(A_{2k}) = P(\tau_{2k} < n) = P(D_n[a, b] \geq k)$,

$$\begin{aligned}
E(D_n[a, b]) &= \sum_{k=1}^n P(D_n[a, b] \geq k) \\
&= \sum_{k=1}^n P(A_{2k}) \\
&\leq \frac{1}{b-a} \sum_{k=1}^n \int_{A_{2k-1} \setminus A_{2k}} (X_n - b)^+ dP \\
&\leq \frac{1}{b-a} E(X_n - b)^+.
\end{aligned}$$

Para la última desigualdad se hace uso del hecho de que las diferencias $A_{2k-1} \setminus A_{2k}$ son conjuntos ajenos. Esto concluye la demostración de la primera estimación. La segunda desigualdad del enunciado se sigue de las siguientes estimaciones

$$(X_n - b)^+ \leq |X_n - b| \leq |X_n| + |b|.$$

□

Ahora estamos en condiciones de probar que toda submartingala acotada en media es convergente.

Teorema de convergencia de submartingalas de Doob. Sea $\{X_n\}$ una submartingala tal que $\sup_n E|X_n| < \infty$. Entonces existe una variable aleatoria integrable X tal que

$$\lim_{n \rightarrow \infty} X_n = X \quad c.s.$$

Demostración. Verificaremos primero que la convergencia de tal sucesión de variables aleatorias es equivalente a la condición: $D[a, b] < \infty$ casi seguramente, para cualesquiera números $a < b$. Sea ω en Ω y considere la sucesión numérica $X_1(\omega), X_2(\omega), \dots$ cuyo número de cruces es $D[a, b](\omega)$. Demostraremos que la sucesión $\{X_n(\omega)\}$ es convergente si, y sólo si, $D[a, b](\omega) < \infty$, para cualesquiera $a < b$.

(\Rightarrow) Suponga que la sucesión es convergente pero que $D[a, b](\omega) = \infty$ para algún par de números a y b tales que $a < b$. Entonces

$$\liminf_{n \rightarrow \infty} X_n(\omega) \leq a < b \leq \limsup_{n \rightarrow \infty} X_n(\omega).$$

Esto contradice la hipótesis de que la sucesión es convergente.

(\Leftarrow) Suponga ahora que $D[a, b](\omega) < \infty$ para cualesquiera $a < b$. Suponga que la sucesión no es convergente. Entonces existen $a_1 < b_1$ tales que

$$\liminf_{n \rightarrow \infty} X_n(\omega) \leq a_1 < b_1 \leq \limsup_{n \rightarrow \infty} X_n(\omega).$$

Entonces, en particular, para este par de números reales se tiene que $D[a_1, b_1](\omega) = \infty$, lo cual contradice la hipótesis inicial.

Por lo tanto es suficiente demostrar que con probabilidad uno, $D[a, b] < \infty$. Para llegar a esta conclusión demostraremos que $E(D[a, b]) < \infty$, pero ello es consecuencia del teorema de convergencia monótona y la última proposición pues

$$\begin{aligned} E(D[a, b]) &= \lim_{n \rightarrow \infty} E(D_n[a, b]) \\ &\leq \frac{1}{b-a} (\sup_n E|X_n| + |b|) < \infty. \end{aligned}$$

La integrabilidad del límite X se sigue del lema de Fatou pues

$$E|X| = E(\liminf_n |X_n|) \leq \liminf_n E|X_n| \leq \sup_n E|X_n| < \infty.$$

□

Como toda martingala es una submartingala, y toda supermartingala se convierte en una submartingala a través de un cambio de signo, se tiene que el teorema anterior es válido en cualquiera de los tres casos. Es decir, toda martingala, submartingala o supermartingala acotada en la forma en la que indica el enunciado del teorema, es convergente casi seguramente, y su límite es una variable aleatoria integrable.

La demostración que se ha presentado aquí sobre la convergencia de submartingalas es la prueba original de Doob de 1940. En [14] pueden encontrarse algunos otros métodos alternativos de demostración. Como hemos mencionado antes, las submartingalas son procesos que en promedio tienden a crecer, de modo que en este caso hemos encontrado una cota superior para el número promedio de cruces hacia abajo. En algunos textos, por ejemplo [4], se enuncia y prueba el mismo resultado para supermartingalas, procesos que, en promedio, tienden a decrecer.

7.10. Representación de martingalas

En esta sección demostraremos que toda martingala que cumple la condición de ser uniformemente integrable puede escribirse en términos de una esperanza condicional. Antes de enunciar el resultado explicaremos la condición de integrabilidad uniforme para un proceso.

Integrabilidad uniforme. Puede comprobarse que una variable aleatoria X es integrable si, y sólo si, para cada $\epsilon > 0$ puede encontrarse un número $M > 0$ tal que

$$\int_{(|X|>M)} |X| dP < \epsilon.$$

Considere ahora una sucesión infinita de variables aleatorias integrables X_1, X_2, \dots . Para cada $\epsilon > 0$ puede encontrarse entonces una sucesión de números reales $M_n > 0$ tales que

$$\int_{(|X_n|>M_n)} |X_n| dP < \epsilon.$$

Cuando la sucesión M_n no depende de n , es decir, cuando sea una sucesión constante, se dice que la sucesión de variables aleatorias es *uniformemente integrable*. Es

evidente que la integrabilidad uniforme es más fuerte que la simple integrabilidad de las variables de un proceso. Tenemos entonces la siguiente definición, la cual ilustraremos después con un par de ejemplos.

Definición. Se dice que una sucesión de variables aleatorias integrables X_1, X_2, \dots es *uniformemente integrable* si para cada $\epsilon > 0$ existe un número $M > 0$ tal que para toda $n \geq 1$,

$$\int_{(|X_n|>M)} |X_n| dP < \epsilon.$$

■ Ejemplo. Considere el espacio muestral $\Omega = (0, 1)$ con la σ -álgebra los subconjuntos de Borel de $(0, 1)$, y como medida de probabilidad la medida de Lebesgue sobre dicho intervalo. La sucesión de variables aleatorias dada por $X_n = n 1_{(0, 1/n)}$ no es uniformemente integrable pues para cualquier $M > 0$,

$$\int_{(|X_n|>M)} |X_n| dP = \begin{cases} 1 & \text{si } n > M, \\ 0 & \text{si } n \leq M. \end{cases}$$

◦

■ Ejemplo. Sea X una variable aleatoria integrable y sea \mathcal{F}_n un filtración. Demostraremos que la martingala $X_n = E(X | \mathcal{F}_n)$ es uniformemente integrable. Como $|X|$ es integrable, tenemos que para cada $\epsilon > 0$ existe $\delta > 0$ tal que si $P(A) < \delta$, entonces $\int_A |X| dP < \epsilon$. Además, como $|X_n| \leq E(|X| | \mathcal{F}_n)$, tomando esperanzas y para cualquier $M > 0$ se tiene que

$$\begin{aligned} E|X| &\geq E|X_n| \\ &\geq \int_{(|X_n|>M)} |X_n| dP \\ &\geq M P(|X_n| > M). \end{aligned}$$

De modo que si se toma $M > E(|X|)/\delta$, con $\delta > 0$ arbitrario, entonces $P(|X_n| >$

$M) \leq E(|X|)/M < \delta$. Por lo tanto

$$\begin{aligned} \int_{(|X_n|>M)} |X_n| dP &\leq \int_{(|X_n|>M)} E(|X| | \mathcal{F}_n) dP \\ &= \int_{(|X_n|>M)} |X| dP \\ &< \epsilon. \end{aligned}$$

◦

La siguiente proposición establece que la convergencia en media es una condición suficiente para que se cumpla la propiedad de integrabilidad uniforme en un proceso.

Proposición. Todo proceso que consta de variables aleatorias integrables y que es convergente en media, es uniformemente integrable.

Demostración. Suponga que X_1, X_2, \dots es una sucesión de variables aleatorias integrables convergente en media a la variable aleatoria integrable X , es decir, $E|X_n - X| \rightarrow 0$. Esto es, para cada $\epsilon > 0$ existe un natural N tal que para cualquier $n > N$, $E|X_n - X| < \epsilon/2$. Dado que el límite X es integrable, para cada $\epsilon > 0$ existe $\delta > 0$ tal que si $P(A) < \delta$, entonces $\int_A |X| dP < \epsilon/2$. Tomando un valor de $\delta > 0$ más pequeño si es necesario se tiene además que $\int_A |X_n| dP < \epsilon$, para cada $n = 1, \dots, N$, cuando $P(A) < \delta$. Por otro lado, para cualquier $n \geq 1$,

$$E(|X_n|) \geq \int_{(|X_n|>M)} |X_n| dP \geq MP(|X_n| > M).$$

Es decir, $P(|X_n| > M) \leq \frac{1}{M} E(|X_n|) \leq \delta$ si se toma $M = \frac{1}{\delta} \sup_n E(|X_n|)$. Tal valor de M es finito pues como la sucesión converge en media, es acotada en media. En particular y con el valor de M mencionado, lo anterior demuestra la integrabilidad uniforme de las variables X_n para valores de n menores o iguales a N . Para el caso $n > N$ se tiene que

$$\begin{aligned} \int_{(|X_n|>M)} |X_n| dP &\leq \int_{(|X_n|>M)} |X| dP + \int_{(|X_n|>M)} |X_n - X| dP \\ &\leq \int_{(|X_n|>M)} |X| dP + E(|X_n - X|) \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

□

El siguiente resultado es un recíproco de la proposición anterior, sólo que hay que añadir la condición de que el proceso sea una submartingala, en particular, una martingala.

Proposición. Toda submartingala uniformemente integrable es convergente en media.

Demostración. Sea $\{X_n : n \geq 1\}$ una submartingala uniformemente integrable. Hemos demostrado antes que en tales condiciones la submartingala es necesariamente acotada en L^1 , y por lo tanto satisface las condiciones del teorema de convergencia de martingalas de Doob. Existe entonces una variable aleatoria integrable X tal que $X_n \rightarrow X$ c.s. Demostraremos que $X_n \rightarrow X$ en media, es decir, que $E|X_n - X| \rightarrow 0$. Sea $\epsilon > 0$ arbitrario. Debido a la hipótesis de integrabilidad uniforme, existe $M > 0$ tal que para toda $n \geq 1$,

$$\int_{(|X_n - X| > M)} |X_n - X| dP < \frac{\epsilon}{3}.$$

Por otro lado, como la convergencia casi segura implica la convergencia en probabilidad se tiene que $P(|X_n - X| > \epsilon/3) \rightarrow 0$, cuando $n \rightarrow \infty$, es decir, existe un entero N tal que para $n > N$,

$$P(|X_n - X| > \epsilon/3) < \frac{\epsilon}{3M},$$

en donde $M > 0$ es la constante de la estimación anterior. Entonces

$$\begin{aligned} E(|X_n - X|) &= \int_{(|X_n - X| > M)} |X_n - X| dP \\ &\quad + \int_{(\epsilon/3 < |X_n - X| \leq M)} |X_n - X| dP \\ &\quad + \int_{(|X_n - X| \leq \epsilon/3)} |X_n - X| dP \\ &< \frac{\epsilon}{3} + MP(|X_n - X| > \epsilon/3) + \frac{\epsilon}{3}P(|X_n - X| \leq \epsilon/3) \\ &< \epsilon. \end{aligned}$$

□

Finalmente se presenta el resultado que establece que toda martingala uniformemente integrable se puede escribir como una esperanza condicional.

Teorema de representación de martingalas. Sea X_n una martingala uniformemente integrable, con filtración natural \mathcal{F}_n , y teniendo a la variable X como su límite en media. Entonces

$$X_n = E(X | \mathcal{F}_n).$$

Demostración. Para cualquier $m > n$, $E(X_m | \mathcal{F}_n) = X_n$. Esto quiere decir que para cualquier evento A en \mathcal{F}_n ,

$$\int_A X_m dP = \int_A X_n dP.$$

Entonces

$$\begin{aligned} \left| \int_A (X_n - X) dP \right| &= \left| \int_A (X_m - X) dP \right| \\ &\leq \int_A |X_m - X| dP \\ &\leq \int_{\Omega} |X_m - X| dP. \end{aligned}$$

Por lo demostrado antes, el último término converge a cero cuando $m \rightarrow \infty$. Es decir, para cualquier A en \mathcal{F}_n , $\int_A X_n dP = \int_A X dP$. Esto significa que $X_n = E(X | \mathcal{F}_n)$ c.s. \square

La variable $X_n = E(X | \mathcal{F}_n)$ puede interpretarse como una aproximación de la variable desconocida X cuando se cuenta con la información dada por la σ -álgebra \mathcal{F}_n . Conforme n crece, la información acerca de X aumenta a través de la filtración, y en el límite se obtiene o reconstruye X .

Notas y referencias. El lector puede encontrar otras exposiciones introductorias al tema de martingalas a tiempo discreto en Karlin y Taylor [19], y Lawler [23]. Para lecturas más avanzadas pueden consultarse Revuz y Yor [29] y Tudor [38].



J. L. Doob

Joseph L. Doob (USA 1910–2004).

Doob empezó a tener interés por la ciencia desde pequeño cuando cursaba la escuela secundaria. Estuvo muy interesado en el funcionamiento del radio e incluso construyó él mismo su propio equipo. Este interés suyo por la electrónica y las comunicaciones se incrementó durante la preparatoria obteniendo incluso una licencia para llevar a cabo transmisiones por radio. Dado este interés en la electrónica, Doob pensó que la física era el área que debía estudiar al ingresar a la universidad. Y así lo hizo al ingresar a la universidad de Harvard en 1926. Sin embargo, después de un año de estudios, se convenció que el curso que verdaderamente disfrutó fue el de cálculo y sus aplicaciones. Para el segundo año se registró en cursos de matemáticas. En 1930 obtuvo el grado de licenciatura de la universidad de Harvard, y en 1931 obtuvo el grado de maestría bajo la supervisión de J. L. Walsh en la misma universidad. En Junio de ese mismo año se casó con Elsie Field con quien tuvo tres niños. Al año siguiente, en 1932, obtuvo el grado de doctorado con un trabajo de tesis sobre las funciones analíticas y sus valores de frontera. Tiempo después obtuvo un beca para trabajar en la teoría de la probabilidad con H. Hotelling en la universidad de Columbia de 1934 a 1935. Al final de ese periodo fue contratado como profesor asociado en la universidad de Illinois en 1935, allí fue donde desarrolló su carrera como profesor hasta el año de 1978 cuando se jubiló. Dentro de sus estudiantes de doctorado figuran, entre otros, Paul Halmos (1938), David Blackwell (1941), J. Laurie Snell (1951) y John Walsh (1966). El trabajo de Doob se centra principalmente en la teoría de la medida, la teoría de la probabilidad y las relaciones de ésta última con la teoría del potencial. En particular, y profundizando parte del trabajo de Paul Lévy, durante los años cuarentas y cincuentas Doob desarrolló la teoría básica de las martingalas y algunas de sus aplicaciones. En 1953 publicó su libro clásico *Stochastic Processes*, el cual fue reeditado en 1990. En 1984 publicó *Classical Potential Theory and its Probabilistic Counterparts*, reimpresso en 2001. En 1994, a la edad de 84 años, publicó su último texto titulado *Measure Theory*. Doob fue miembro de las Academias de Ciencias de los Estados Unidos y de Francia, presidió el Instituto de Estadística Matemática (IMS) en 1950, y la Sociedad Matemática Norteamericana (AMS) de 1963 a 1964. Recibió varios premios prestigiosos de su país por la trascendencia y profundidad de sus trabajos. Durante muchos años de su vida Doob mantuvo la costumbre y el gusto por organizar y realizar las famosas caminatas de los sábados por la tarde junto a profesores universitarios, colegas y amigos, de la universidad de Illinois. A petición propia, las cenizas de los restos mortales de Doob fueron esparcidas por

sus compañeros en uno de sus sitios favoritos donde acostumbraba caminar junto a sus amigos. Para una información más amplia sobre la vida y el trabajo de Doob véanse los trabajos de Bingham [2], Burkholder y Protter [5], y Snell [32], [33]. Fuente: Archivo MacTutor, Universidad de St. Andrews.

Ejercicios

Filtraciones

122. Sea $\{X_t\}$ un proceso adaptado a la filtración $\{\mathcal{F}_t\}$, y sea $\{\mathcal{G}_t\}$ la filtración natural del proceso. Demuestre que para cada $t \geq 0$ se cumple $\mathcal{G}_t \subseteq \mathcal{F}_t$. Esto significa que la filtración natural es la filtración más pequeña respecto de la cual un proceso es adaptado.

Tiempos de paro

123. Sea $\{X_n\}$ la sucesión de resultados que se obtienen al lanzar sucesivamente un dado equilibrado. Sea $\{\mathcal{F}_n\}$ la filtración natural de este proceso y defina la variable τ como el primer tiempo n tal que $X_1 + \dots + X_n \geq 10$. Determine si τ es un tiempo de paro respecto de la filtración dada.
124. Demuestre que la condición en la definición de tiempo de paro a tiempo discreto " $(\tau \leq n) \in \mathcal{F}_n$ " es equivalente a la condición " $(\tau = n) \in \mathcal{F}_n$ ".
125. Sea n un entero positivo. Demuestre que la variable aleatoria constante $\tau = n$ es un tiempo de paro. Compruebe además que $\tau = \infty$ es también un tiempo de paro.
126. Sea τ un tiempo de paro con valores en $\{1, 2, \dots, \infty\}$ y sea n un número natural. Demuestre que tanto $\tau \wedge n$ como $\tau \vee n$ son también tiempos de paro respecto de la misma filtración.
127. Sean τ_1 y τ_2 dos tiempos de paro respecto de la misma filtración. Demuestre que las siguientes variables aleatorias también son tiempos de paro.
a) $\tau_1 \wedge \tau_2$ b) $\tau_1 \vee \tau_2$.
128. Sean τ_1 y τ_2 dos tiempos de paro respecto de la misma filtración. Demuestre que $\tau_1 + \tau_2$ es también un tiempo de paro.
129. Sea X_n un proceso a tiempo discreto, y sea x un número real cualquiera dentro del espacio de estados del proceso. Defina las variables aleatorias: τ_1 como el primer momento en el que el proceso toma el valor x , τ_2 como el segundo momento en el que el proceso toma el valor x , etc. Demuestre que $\tau_1 \leq \tau_2 \leq \dots$ son tiempos de paro.

130. Sea τ un tiempo de paro con valores en $[0, \infty)$ y sea $c \geq 1$ una constante. Demuestre que $c\tau$ es tiempo de paro.
131. Sea τ_1, τ_2, \dots una sucesión infinita de tiempos de paro. Demuestre que tanto $\sup_n \tau_n$ como $\inf_n \tau_n$ son tiempos de paro.

Martingalas

132. Sea $\{X_n\}$ una martingala a tiempo discreto. Demuestre que la propiedad de martingala “ $E(X_m | \mathcal{F}_n) = X_n$, para cualquier $m \geq n$ ”, es equivalente a la condición “ $E(X_{n+1} | \mathcal{F}_n) = X_n$, para cada $n \geq 1$ ”. Enuncie y demuestre la equivalencia análoga al caso submartingala y supermartingala.
133. Demuestre que
- toda martingala es una submartingala y una supermartingala.
 - todo proceso que es al mismo tiempo una submartingala y una supermartingala es una martingala.
134. Demuestre que para cualquier $t > s \geq 0$,
- $E(X_t) = E(X_s)$ cuando $\{X_t\}$ es una martingala.
 - $E(X_t) \geq E(X_s)$ cuando $\{X_t\}$ es una submartingala.
 - $E(X_t) \leq E(X_s)$ cuando $\{X_t\}$ es una supermartingala.
135. Sea X una variable aleatoria integrable, y sea $\{\mathcal{F}_t\}_{t \geq 0}$ una filtración a tiempo continuo. Demuestre que $X_t = E(X | \mathcal{F}_t)$ es una martingala.
136. Sea $\{X_n : n \geq 1\}$ un proceso adaptado a la filtración $\{\mathcal{F}_n\}_{n \geq 1}$. Demuestre que si A es un evento \mathcal{F}_1 -medible, entonces el proceso $\{X_n 1_A\}$ es una martingala, submartingala o supermartingala cuando $\{X_n\}$ lo es.
137. Sea M una constante. Demuestre que
- si $\{X_n\}$ es una submartingala, entonces $\{X_n \vee M\}$ es una submartingala.
 - si $\{X_n\}$ es una supermartingala, entonces $\{X_n \wedge M\}$ es una supermartingala.
138. Sean $\{X_t\}$ and $\{Y_t\}$ dos martingalas, submartingalas o supermartingalas respecto de la misma filtración. Demuestre que el proceso $\{aX_t + bY_t + c\}$ es también una martingala, submartingala o supermartingala, respectivamente,

en donde a , b y c son constantes. Para el caso submartingala y supermartingala se necesita suponer además que a y b son no negativos. En particular esto demuestra que la suma de dos martingalas es una martingala, y que el conjunto de martingalas respecto de la misma filtración y definidas en un mismo espacio de probabilidad es un espacio vectorial.

139. Sea $\{X_n\}$ un proceso integrable. Demuestre que el proceso dado por $X_n^* = \max\{X_1, \dots, X_n\}$ es una submartingala.
140. Sean $\{X_t\}$ y $\{Y_t\}$ dos martingalas o submartingalas respecto de la misma filtración. Demuestre que el proceso $\{X_t \vee Y_t\}$ es una submartingala.
141. Sean $\{X_t\}$ y $\{Y_t\}$ dos martingalas o supermartingalas respecto de la misma filtración. Demuestre que el proceso $\{X_t \wedge Y_t\}$ es una supermartingala.
142. *Martingala de de Moivre.* Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes cada una de ellas con la misma distribución dada por $P(\xi = +1) = p$ y $P(\xi = -1) = q = 1 - p$. Sea $X_n = \xi_1 + \dots + \xi_n$. Demuestre que $Y_n = (q/p)^{X_n}$ es una martingala respecto de la filtración generada por el proceso X_n .
143. Sea $\{X_t\}$ una martingala respecto de una filtración $\{\mathcal{F}_t\}$. Demuestre que el proceso también es una martingala respecto de su filtración natural. En general el recíproco es falso.
144. Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes con esperanza cero. Defina $X_0 = 0$ y $X_n = \xi_1 + \dots + \xi_n$, para $n \geq 1$. Se ha demostrado que el proceso X_n es una martingala respecto de la filtración $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$. Suponga ahora que las variables ξ tienen segundo momento finito. Demuestre que el proceso $X_n^2 - n$ es también una martingala respecto de la misma filtración si, y sólo si, las variables ξ tienen varianza uno.
145. *Martingala producto.* Sean ξ_1, ξ_2, \dots variables aleatorias independientes con esperanza unitaria. Demuestre que el proceso $X_n = \xi_1 \cdots \xi_n$ es una martingala respecto de su filtración natural.
146. Sea $\{X_n\}$ una submartingala. Demuestre que los siguientes procesos también son submartingalas.
- $X_n^+ = \max\{X_n, 0\}$.
 - $X_n^- = -\min\{X_n, 0\}$.
 - $|X_n|$. Suponga, en este caso, que X_n es una martingala.

147. *Martingala del juego de apuestas.* Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas. Defina $X_n = \xi_1 + \dots + \xi_n$. Demuestre que el proceso
- $\{X_n - nE(\xi)\}$ es una martingala, suponiendo $E(\xi) < \infty$.
 - $\{X_n^2 - nE(\xi^2)\}$ es una martingala, suponiendo $E(\xi^2) < \infty$.
148. Sea $\{X_t\}$ una martingala cuadrado integrable. Demuestre que $\{X_t^2\}$ es una submartingala respecto de la misma filtración. Vea el siguiente ejercicio para una generalización de este resultado.
149. Sea $\{X_t\}$ una martingala tal que $E|X_t|^p < \infty$ para cada $t \geq 0$, con $p \geq 1$. Demuestre que $\{|X_t|^p\}$ es también una submartingala. Sugerencia: Use la desigualdad de Jensen. En el siguiente ejercicio se generaliza este resultado.
150. Sea $\{X_t\}$ un proceso integrable, y sea g una función convexa tal que $g(X_t)$ es integrable. Demuestre que bajo cualquiera de las siguientes condiciones se cumple que el proceso $g(X_t)$ es una submartingala.
- cuando $\{X_t\}$ es una martingala.
 - cuando $\{X_t\}$ es una submartingala y g es una función no decreciente.
151. Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias independientes tales que $E(\xi) = \mu_k$, $\text{Var}(\xi) = \sigma_k^2 < \infty$, y con filtración natural \mathcal{F}_n . Demuestre que el siguiente proceso es una martingala respecto de esta filtración.

$$X_n^2 = \left(\sum_{k=1}^n (\xi_k - \mu_k) \right)^2 - \sum_{k=1}^n \sigma_k^2.$$

152. Sea $\{\xi_n : n \geq 0\}$ un proceso integrable y adaptado a la filtración $\{\mathcal{F}_n\}$. Demuestre que el siguiente proceso es una martingala.

$$X_n = \sum_{k=1}^n (\xi_k - E(\xi_k | \mathcal{F}_{k-1})).$$

153. Sea $\{X_t\}$ un proceso de Poisson de parámetro $\lambda > 0$. Demuestre que los siguientes procesos son martingalas.
- $Y_t = (X_t - \lambda t)^2 - \lambda t$.
 - $Y_t = \exp(-\theta X_t + \lambda t(1 - e^{-\theta}))$, en donde $\theta \in \mathbb{R}$.

154. Sea $\{X_t\}$ un proceso integrable con incrementos independientes. Demuestre que el proceso centrado $\{X_t - E(X_t)\}$ es una martingala.
155. Considere la martingala de la urna de Polya con una configuración inicial de una bola blanca y una negra. Suponga ahora que cuando se extrae una bola de un color se regresa a la urna junto con k bolas del mismo color. Defina $M_n = X_n/(2 + nk)$. ¿Es $\{M_n\}$ una martingala?
156. Demuestre que el proceso $\{X_n\}$ de la estrategia de juego llamada martingala, cuando las apuestas son justas, es una martingala.
157. Sea $X_n = \xi_1 + \dots + \xi_n$ una caminata aleatoria simple sobre \mathbb{Z} y que inicia en el origen. Suponga $P(\xi = +1) = p$ y $P(\xi = -1) = q = 1 - p$. Sean a y b dos enteros positivos fijos. Defina el tiempo de paro

$$\tau = \min\{n \geq 1 : X_n = -a \text{ ó } X_n = b\}.$$

Demuestre que

$$E(\tau) = \begin{cases} ab & \text{si } p = q, \\ \frac{b}{p-q} - \frac{a+b}{p-q} \cdot \frac{1 - (p/q)^b}{1 - (p/q)^{a+b}} & \text{si } p \neq q. \end{cases}$$

158. Sea $\{X_n\}$ una martingala. Demuestre que para $n_1 \leq n_2 < n_3$,

$$E((X_{n_3} - X_{n_2})X_{n_1}) = 0.$$

159. Sea ξ_1, ξ_2, \dots una sucesión de variables aleatorias tal que el proceso $X_n = \xi_1 + \dots + \xi_n$ es una martingala. Demuestre que $E(\xi_i \xi_j) = 0$ para $i \neq j$.
160. Sea $\{X_n\}$ una cadena de Markov con espacio de estados $\{0, 1, \dots\}$. Suponga que las probabilidades de transición son $p_{00} = 1$ y $p_{ij} = e^{-i} i^j / j!$ en los otros casos. Demuestre que $\{X_n\}$ es una martingala.

Teorema de paro opcional

161. Sea $X_n = \xi_1 + \dots + \xi_n$ una caminata aleatoria simple sobre \mathbb{Z} y que inicia en el origen. Suponga $P(\xi = +1) = p$ y $P(\xi = -1) = q = 1 - p$. Sean a y b

dos enteros positivos fijos. Defina el tiempo de paro $\tau = \min\{n \geq 1 : X_n = -a \text{ ó } X_n = b\}$. Demuestre que

$$E(\tau) = \begin{cases} ab & \text{si } p = q, \\ \frac{b}{p-q} - \frac{a+b}{p-q} \cdot \frac{1-(p/q)^b}{1-(p/q)^{a+b}} & \text{si } p \neq q. \end{cases}$$

Integrabilidad uniforme

162. Demuestre que una variable aleatoria X es integrable si, y sólo si, para cada $\epsilon > 0$ existe $M > 0$ tal que

$$\int_{(|X|>M)} |X| dP < \epsilon.$$

CAPÍTULO 8

Movimiento Browniano

El fenómeno natural que ahora se conoce como movimiento Browniano tiene una interesante historia. El primer registro, aunque no así la primera observación de él, data de 1828 cuando el botánico Robert Brown reportó en una revista científica que granos de polen suspendidos en una cierta substancia y vistos a través de un microscopio, realizaban un movimiento irregular e inexplicable [3]. Este extraño movimiento fue objeto de mucha discusión y muy diversas controversias se suscitaron a partir de su divulgación en la comunidad científica de aquella época. Con la intención de dar una explicación satisfactoria del extraño fenómeno observado, se llevaron a cabo diversos experimentos y se formularon muy diversas hipótesis [10]. Hoy en día este movimiento es entendido y explicado a través de las múltiples colisiones aleatorias de las moléculas del líquido con los granos de polen. Llegar a tal aseveración tomó muchos años pues debió aceptarse la teoría cinético molecular de la materia, y el seminal trabajo de Einstein de 1905 sobre este fenómeno contribuyó decididamente a tal tarea [10].



R. Brown

En este capítulo se presenta una introducción al modelo matemático para el movimiento Browniano. Se trata del ejemplo más importante de un proceso de Markov a tiempo continuo y con espacio de estados continuo.

8.1. Definición

Las observaciones reales del movimiento de granos de polen a través del microscopio sugieren que las trayectorias son continuas y que los desplazamientos son independientes en intervalos de tiempo disjuntos. Además, debido al gran número de colisiones del grano de polen con las moléculas circundantes en longitudes de tiempo no pequeños, y teniendo en cuenta el teorema central del límite, los incrementos pueden modelarse como variables aleatorias Gaussianas. La estructura matemática de un proceso estocástico a tiempo continuo, denotado en este caso por $\{B_t : t \geq 0\}$, ha resultado adecuada para modelar este tipo de fenómenos. La variable B_t puede representar la posición de la partícula al tiempo t . La definición matemática, en el caso unidimensional, es la siguiente.

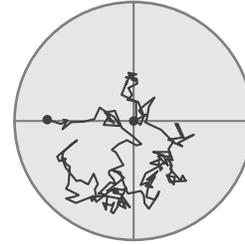


Figura 8.1: Movimiento Browniano

Definición. (I) Un *movimiento Browniano* unidimensional de parámetro σ^2 es un proceso estocástico $\{B_t : t \geq 0\}$ con valores en \mathbb{R} que cumple las siguientes propiedades.

1. $B_0 = 0$ c.s.
2. Las trayectorias son continuas.
3. El proceso tiene incrementos independientes.
4. La variable $B_t - B_s$ tiene distribución $N(0, \sigma^2(t - s))$, para cualesquiera tiempos $0 \leq s < t$.

Las condiciones que aparecen en esta definición son consecuencia directa de las observaciones del fenómeno físico, pero eso no garantiza que tal objeto matemático exista. En 1923 el matemático norteamericano Norbert Wiener demostró la existencia y unicidad de un proceso con tales condiciones. Es por esta razón que a menudo a este proceso también se le llama *proceso de Wiener*, y se le denota también por $\{W_t : t \geq 0\}$. En sentido estricto, el movimiento Browniano es el fenómeno físico, mientras que su modelo matemático es el proceso de Wiener, aunque es común

llamar a ambas cosas por el mismo nombre: movimiento Browniano. Demostraremos a continuación que las condiciones (3) y (4) de la definición anterior son equivalentes a solicitar que las distribuciones finito dimensionales del proceso sean las siguientes.

Definición. (II) Un *movimiento Browniano* unidimensional de parámetro σ^2 es un proceso estocástico $\{B_t : t \geq 0\}$ con valores en \mathbb{R} que cumple las siguientes propiedades.

1. $B_0 = 0$ c.s.
2. Las trayectorias $t \mapsto B_t$ son continuas.
3. Para cualesquiera tiempos $0 < t_1 < \dots < t_n$, y para cualesquiera conjuntos de Borel A_1, \dots, A_n de \mathbb{R} , se cumple que

$$P(B_{t_1} \in A_1, \dots, B_{t_n} \in A_n) = \int_{A_1} \dots \int_{A_n} p(t_1, 0, x_1) \cdot p(t_2 - t_1, x_1, x_2) \dots p(t_n - t_{n-1}, x_{n-1}, x_n) dx_n \dots dx_1,$$

en donde

$$p(t, x, y) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-(y-x)^2/2\sigma^2 t}. \quad (8.1)$$

Observe que la tercera propiedad de la última definición establece que la función de densidad del vector $(B_{t_1}, \dots, B_{t_n})$ evaluada en el punto (x_1, \dots, x_n) es

$$p(t_1, 0, x_1)p(t_2 - t_1, x_1, x_2) \dots p(t_n - t_{n-1}, x_{n-1}, x_n).$$

En particular, la variable B_t tiene distribución $N(0, \sigma^2 t)$. Demostraremos a continuación que las dos definiciones anteriores son equivalentes.

(I) \Rightarrow (II). Se hace uso de la independencia de los incrementos y de la hipótesis

de que éstos tienen distribución normal.

$$\begin{aligned}
f_{B_{t_1}, B_{t_2}, \dots, B_{t_n}}(x_1, x_2, \dots, x_n) &= f_{B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}}(x_1, x_2 - x_1, \dots, x_n - x_{n-1}) \\
&= f_{B_{t_1}}(x_1) f_{B_{t_2} - B_{t_1}}(x_2 - x_1) \cdots f_{B_{t_n} - B_{t_{n-1}}}(x_n - x_{n-1}) \\
&= p(t_1, 0, x_1) p(t_2 - t_1, 0, x_2 - x_1) \cdots p(t_n - t_{n-1}, 0, x_n - x_{n-1}) \\
&= p(t_1, 0, x_1) p(t_2 - t_1, x_1, x_2) \cdots p(t_n - t_{n-1}, x_{n-1}, x_n).
\end{aligned}$$

(II) \Rightarrow (I). De acuerdo al tercer postulado, para $0 \leq s < t$, la función de densidad del vector (B_s, B_t) es $f_{B_s, B_t}(x, y) = p(s, 0, x) p(t - s, x, y)$. Aplicando la fórmula general para la función de densidad de la diferencia de dos variables aleatorias, $f_{Y-X}(u) = \int f_{X,Y}(x, u+x) dx$, se obtiene

$$\begin{aligned}
f_{B_t - B_s}(u) &= \int_{-\infty}^{\infty} p(s, 0, x) p(t - s, x, u+x) dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 s}} e^{-x^2/2s} \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} e^{-u^2/2\sigma^2(t-s)} dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} e^{-u^2/2\sigma^2(t-s)} \\
&= p(t-s, 0, u),
\end{aligned}$$

es decir, $B_t - B_s$ tiene distribución $N(0, t-s)$. Entonces

$$\begin{aligned}
f_{B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}}(x_1, x_2, \dots, x_n) &= f_{B_{t_1}, B_{t_2}, \dots, B_{t_n}}(x_1, x_2 + x_1, \dots, x_n + x_{n-1} + \cdots + x_1) \\
&= p(t_1, 0, x_1) p(t_2 - t_1, x_1, x_2 + x_1) \cdots \\
&\quad \cdots p(t_n - t_{n-1}, x_1 + \cdots + x_{n-1}, x_1 + \cdots + x_n) \\
&= p(t_1, 0, x_1) p(t_2 - t_1, 0, x_2) \cdots p(t_n - t_{n-1}, 0, x_n) \\
&= f_{B_{t_1}}(x_1) f_{B_{t_2} - B_{t_1}}(x_2) \cdots f_{B_{t_n} - B_{t_{n-1}}}(x_n).
\end{aligned}$$

Esto demuestra que las variables $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ son independientes cada una de ellas con distribución normal con los parámetros mencionados.

Se dice que un movimiento Browniano es *estándar* cuando $\sigma^2 = 1$. A través del cambio de variable $\tau = \sigma^2 t$ un movimiento Browniano no estándar puede convertirse en uno estándar. Entonces, a menos que se especifique lo contrario y sin pérdida

de generalidad, a partir de ahora supondremos que el movimiento Browniano es estándar, es decir, $B_t - B_s$ tiene distribución $N(0, t - s)$.

■ Ejercicio. Sea $\{B_t\}$ un movimiento Browniano. Demuestre que los siguientes procesos también son movimientos Brownianos.

- a) $W_t = -B_t$.
- b) $W_t = \frac{1}{c}B_{c^2t}$, con c constante positiva.
- c) $W_t = tB_{1/t}$, con $W_0 = 0$.
- d) $W_t = B_{t+t_0} - B_{t_0}$, con $t_0 \geq 0$ fijo. .

8.2. Función de probabilidad de transición

A la función $p(t, x, y)$ definida por (8.1) se le llama *función de probabilidad de transición* del movimiento Browniano de parámetro σ^2 . En particular, la probabilidad de que un movimiento Browniano que inicia en x se encuentre en un conjunto $A \subseteq \mathbb{R}$ (apropiadamente medible) después de t unidades de tiempo es

$$p(t, x, A) = \int_A p(t, x, y) dy.$$

Hemos hecho énfasis en la propiedad que aparece en el inciso (3) de la definición II del movimiento Browniano, pues ésta tiene la ventaja de que proporciona una expresión explícita para la probabilidad del conjunto de trayectorias Brownianas que cumplen las condiciones de encontrarse en el conjunto A_1 al tiempo t_1 , estar en el conjunto A_2 al tiempo posterior t_2 , etcétera.

La condición de que el movimiento Browniano inicie en el origen no es absolutamente necesaria. Pueden considerarse trayectorias Brownianas que inicien en un punto x cualquiera a través del proceso $x + B_t$, el cual se denota a veces por B_t^x para recordar la posición de origen.

Integrando directamente puede comprobarse que la probabilidad de transición $p(t, x, y)$ cumple la *ecuación de Chapman-Kolmogorov*,

$$p(t + s, x, y) = \int_{-\infty}^{\infty} p(t, x, u) p(s, u, y) du,$$

y también cumple la *ecuación de difusión*, o *ecuación de calor*,

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}.$$

8.3. Propiedades básicas

Tenemos entonces que para el movimiento Browniano estándar cada variable aleatoria B_t tiene distribución $N(0, t)$ y por lo tanto $E(B_t) = 0$ y $\text{Var}(B_t) = E(B_t^2) = t$. En particular $E(B_t - B_s)^2 = t - s$, para $0 \leq s < t$. El movimiento Browniano físico real se presenta en tres dimensiones y es completamente errático.

En la Figura 8.2 puede apreciarse una posible trayectoria Browniana cuando ésta se proyecta sobre una de sus coordenadas. Esta gráfica fue generada en computadora y es sólo una aproximación del modelo matemático. En la gráfica pueden apreciarse algunas pequeñas partes en donde aparentemente el comportamiento es lineal y suave, ello no sucede en el modelo teórico.

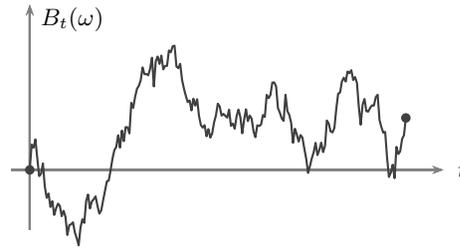


Figura 8.2:

Usando esta probabilidad de transición demostraremos a continuación que el movimiento Browniano cumple la propiedad de Markov en el sentido débil.

Proposición. El movimiento Browniano es un proceso de Markov.

Demostración. Para cualesquiera tiempos $0 \leq t_1 < t_2 < \dots < t_n < t_{n+1}$, y para

cualquier evento A en \mathbb{R} ,

$$\begin{aligned}
 & P(B_{t_{n+1}} \in A \mid B_{t_1} = x_1, \dots, B_{t_n} = x_n) \\
 &= \frac{p(t_1, 0, x_1) p(t_2 - t_1, x_1, x_2) \cdots p(t_n - t_{n-1}, x_{n-1}, x_n) p(t_{n+1} - t_n, x_n, A)}{p(t_1, 0, x_1) p(t_2 - t_1, x_1, x_2) \cdots p(t_n - t_{n-1}, x_{n-1}, x_n)} \\
 &= p(t_{n+1} - t_n, x_n, A) \\
 &= p(t_{n+1} - t_n, x_n, A) \frac{p(t_n, 0, x_n)}{p(t_n, 0, x_n)} \\
 &= P(B_{t_{n+1}} \in A \mid B_{t_n} = x_n).
 \end{aligned}$$

□

Puede demostrarse que cumple además la propiedad fuerte de Markov: Si τ es un tiempo de paro respecto de la filtración del movimiento Browniano, entonces el proceso $B_{t+\tau} - B_\tau$ es también un movimiento Browniano y es independiente de la σ -álgebra $\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap (\tau \leq t) \in \mathcal{F}_t \text{ para cada } t \geq 0\}$. En particular, cuando τ es constante t_0 , el proceso $B_{t+t_0} - B_{t_0}$ es un movimiento Browniano.

Como una consecuencia del hecho de que los incrementos de este proceso son independientes, demostraremos a continuación la propiedad de martingala.

Proposición. El movimiento Browniano es una martingala continua.

Demostración. Claramente el proceso es adaptado a su filtración natural y cada variable aleatoria del proceso es integrable. Por otro lado, para cualesquiera tiempos s y t tales que $0 \leq s < t$,

$$\begin{aligned}
 E(B_t \mid \mathcal{F}_s) &= E(B_t - B_s + B_s \mid \mathcal{F}_s) \\
 &= E(B_t - B_s \mid \mathcal{F}_s) + E(B_s \mid \mathcal{F}_s) \\
 &= E(B_t - B_s) + B_s \\
 &= B_s.
 \end{aligned}$$

□

El siguiente resultado no trivial se debe a Paul Lèvy y establece condiciones que caracterizan de manera única al movimiento Browniano en términos de la propiedad de martingala.

Teorema de caracterización de Paul Lèvy. Un proceso $\{X_t : t \geq 0\}$ es un movimiento Browniano si, y sólo si, tiene trayectorias continuas, empieza en cero, y tanto $\{X_t\}$ como $\{X_t^2 - t\}$ son martingalas.

El movimiento Browniano $\{B_t\}$ cumple claramente cada una de las condiciones mencionadas, aunque posiblemente no sea tan evidente que el proceso $\{B_t^2 - t\}$ sea también una martingala, ello no es difícil de verificar y se deja como ejercicio. La parte fuerte de esta caracterización radica en que estas condiciones definen a un movimiento Browniano.

■ Ejercicio. Use la caracterización de Paul Lèvy para demostrar nuevamente que los siguientes procesos son movimientos Brownianos.

- a) $W_t = -B_t$.
- b) $W_t = \frac{1}{c}B_{c^2t}$, con $c > 0$ constante.
- c) $W_t = tB_{1/t}$, con $W_0 = 0$.
- d) $W_t = B_{t+t_0} - B_{t_0}$, con $t_0 \geq 0$ fijo.

El movimiento Browniano como límite de una caminata aleatoria. Considere una caminata aleatoria simétrica simple sobre \mathbb{Z} que inicia en el origen, es decir, sea $X_n = \xi_1 + \dots + \xi_n$, en donde ξ_1, ξ_2, \dots son variables aleatorias independientes e idénticamente distribuidas tales que $P(\xi = +1) = P(\xi = -1) = 1/2$. Es decir, $E(\xi) = 0$ y $\text{Var}(\xi) = E(\xi^2) = 1$. Suponga que la unidad en la variable tiempo es ahora de longitud $\Delta t = 1/N$, con N entero. El objetivo es hacer Δt cada vez más pequeño. Para lograr una versión discreta del movimiento Browniano es necesario hacer también un cambio en la escala en el tamaño de los saltos, ahora no serán unitarios sino de longitud $\sqrt{\Delta t}$, más adelante explicaremos las razones de esta elección. Para cada número natural n defina ahora la caminata aleatoria

$$W_{n\Delta t} = \sqrt{\Delta t} \xi_1 + \dots + \sqrt{\Delta t} \xi_n,$$

una de cuyas trayectorias aparece en la Figura 8.3.

Dada la simetría de la caminata, sigue cumpliéndose que $E(W_{n\Delta t}) = 0$. La razón por la que se ha tomado esa nueva escala en el tamaño de los saltos es que con ello se logra similitud con el movimiento Browniano estándar al cumplirse también que para cualquier valor de n , $\text{Var}(W_{n\Delta t}) = n\Delta t \text{Var}(\xi) = n\Delta t$. Puede demostrarse que cuando $\Delta t \rightarrow 0$ esta caminata tiende a un proceso a tiempo continuo con trayectorias continuas. El proceso límite resultante es el movimiento Browniano estándar.

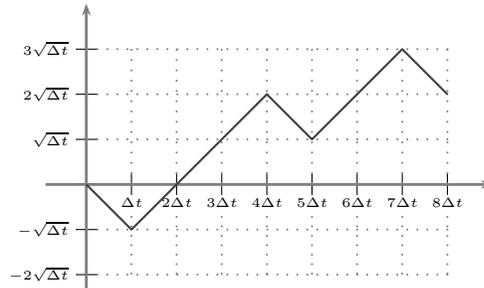


Figura 8.3:

Esta aproximación del movimiento Browniano como límite de una caminata aleatoria sugiere un mecanismo para simular trayectorias Brownianas por computadora: Se escoge Δt pequeño y N el número de puntos que se deseen graficar. Se generan entonces N valores independientes de la variable ξ con distribución uniforme en el conjunto $\{-1, +1\}$, y se grafica la sucesión de puntos $(k\Delta t, W_{k\Delta t})$, para $k = 0, 1, \dots, N$. En la práctica suelen generarse valores continuos para ξ con distribución normal estándar. Este es el mecanismo seguido para generar la gráfica de la Figura 8.2 y las otras trayectorias Brownianas que aparecen en este capítulo.

Difusión. Suponga que se coloca una partícula al azar en la recta real de acuerdo a una densidad de probabilidad $f(y)$. Suponga ahora que esta partícula se mueve siguiendo un movimiento Browniano estándar unidimensional. Entonces la densidad de probabilidad de la posición de la partícula después de t unidades de tiempo es la función $f(t, x)$ dada por

$$f(t, x) = \int_{-\infty}^{\infty} f(y) p(t, y, x) dx = \int_{-\infty}^{\infty} f(y) p(t, x, y) dx,$$

en donde para la segunda igualdad se ha hecho uso de la identidad $p(t, y, x) = p(t, x, y)$, pero ahora esta última expresión adquiere una interpretación interesante pues corresponde a la esperanza de la variable $f(B_t)$ para un movimiento Browniano que inicia en x , es decir, $f(t, x) = E(f(B_t^x))$. A esta función también se le denota por $E^x(f(B_t))$, y puede demostrarse que satisface la ecuación

$$\frac{\partial}{\partial t} f(t, x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, x).$$

8.4. Propiedades de las trayectorias

Antes de establecer las siguientes propiedades, recordemos la definición de variación de una función. Sea $a = t_0 < t_1 < \dots < t_n = b$ una partición del intervalo $[a, b]$, y defina $\Delta t = \max\{|t_{i+1} - t_i| : i = 0, \dots, n-1\}$. La *variación* de una función $g : [a, b] \rightarrow \mathbb{R}$ es el número

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)|.$$

Cuando este número es finito se dice que la función tiene *variación finita* en dicho intervalo. Análogamente, la *variación cuadrática* es

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)|^2.$$

Demostremos a continuación que sobre un intervalo de tiempo acotado $[a, b]$, casi todas las trayectorias del movimiento Browniano tienen variación no acotada, esto es,

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=1}^{n-1} |B_{t_{i+1}} - B_{t_i}| = \infty, \quad \text{c.s.}$$

Esta propiedad es particularmente importante pues tiene como consecuencia el hecho de que no se pueden usar las trayectorias Brownianas como funciones integradoras en el sentido de Riemann-Stieltjes. El hecho de que se desee definir algún tipo de integral respecto del movimiento Browniano será claro en el siguiente capítulo cuando se estudian ecuaciones diferenciales estocásticas. Por otro lado demostraremos también que la variación cuadrática del movimiento Browniano sobre $[a, b]$ es finita, de hecho, es la longitud del intervalo en cuestión.

Proposición. La variación cuadrática de una trayectoria del movimiento Browniano sobre el intervalo $[a, b]$ es la longitud del intervalo, es decir,

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=1}^{n-1} |B_{t_{i+1}} - B_{t_i}|^2 = b - a, \quad \text{en el sentido } L^2(P).$$

Demostración. Sea $\{\mathcal{P}_n\}$ una sucesión de particiones finitas del intervalo $[a, b]$. Denote por Δt_i el incremento $t_{i+1} - t_i$, y sea ΔB_i la diferencia $B_{t_{i+1}} - B_{t_i}$. Entonces

$$\begin{aligned}
& E \left| \sum_i (\Delta B_i)^2 - (b-a) \right|^2 \\
&= E \sum_{i,j} (\Delta B_i)^2 (\Delta B_j)^2 - 2(b-a) E \sum_i (\Delta B_i)^2 + (b-a)^2 \\
&= \sum_i E (\Delta B_i)^4 + \sum_{i \neq j} E (\Delta B_i)^2 E (\Delta B_j)^2 - 2(b-a) \sum_i (\Delta t_i) + (b-a)^2 \\
&= \sum_i 3(\Delta t_i)^2 + \sum_{i \neq j} \Delta t_i \Delta t_j - (b-a)^2 \\
&= \sum_i 2(\Delta t_i)^2 + \left(\sum_i \Delta t_i \right)^2 - (b-a)^2 \\
&= \sum_i 2(\Delta t_i)^2 \\
&\leq 2(b-a) \max_{0 \leq i < n} \Delta t_i \rightarrow 0.
\end{aligned}$$

□

Recordemos ahora el resultado que establece que toda toda sucesión convergente en el sentido $L^2(P)$ tiene una subsucesión convergente casi seguramente. Por lo tanto existe una subsucesión de particiones $\{\mathcal{P}_{n_k}\}$ del intervalo $[a, b]$ tal que

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=1}^{n-1} |B_{t_{i+1}} - B_{t_i}|^2 = b-a, \quad \text{c.s.}$$

Proposición. (Variación del movimiento Browniano). La variación de una trayectoria del movimiento Browniano sobre el intervalo $[a, b]$ es infinita, casi seguramente, es decir,

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=1}^{n-1} |B_{t_{i+1}} - B_{t_i}| = \infty, \quad \text{c.s.}$$

Demostración. Para cada n natural sea \mathcal{P}_n la partición uniforme del intervalo $[a, b]$ en n subintervalos, es decir cada incremento $\Delta t_i = t_{i+1} - t_i$ tiene longitud $(b-a)/n$. Entonces se tiene la estimación

$$\sum_{i=0}^{n-1} |\Delta B_i|^2 \leq \left(\max_{0 \leq i < n} |\Delta B_i| \right) \sum_{i=0}^{n-1} |\Delta B_i|. \quad (8.2)$$

Sea $\{\mathcal{P}_{n_k}\}$ subsucesión de particiones uniformes tal que

$$\lim_{k \rightarrow \infty} \sum_{i=0}^{n_k-1} |\Delta B_i|^2 = b - a, \quad \text{c.s.}$$

Por otro lado, como las trayectorias del movimiento Browniano son continuas casi seguramente, se tiene que

$$\lim_{k \rightarrow \infty} \left(\max_{0 \leq i < n_k} |\Delta B_i| \right) = 0, \quad \text{c.s.}$$

Substituyendo los últimos dos resultados en (8.2) se obtiene que, respecto de la subsucesión de particiones,

$$\lim_{k \rightarrow \infty} \sum_{i=0}^{n_k-1} |\Delta B_i| = \infty, \quad \text{c.s.}$$

De donde se sigue que el límite superior es infinito casi seguramente. \square

No diferenciabilidad. Demostraremos a continuación que para cualquier tiempo $t_0 \geq 0$ fijo, con probabilidad uno la trayectoria $t \mapsto B_t$ no es diferenciable en t_0 . Más adelante se enuncia sin demostración un resultado más fuerte acerca de esta no diferenciabilidad.

Proposición. Sea $t_0 \geq 0$ fijo. Con probabilidad uno, el movimiento Browniano $\{B_t\}$ no es diferenciable en t_0 .

Demostración. Debido a que $B_{t+t_0} - B_{t_0}$ es también un movimiento Browniano, es suficiente demostrar la no diferenciabilidad de B_t en $t = 0$. Demostraremos que con probabilidad uno, para cada número natural n existe t en el intervalo $[0, 1/n^2]$

tal que $|\frac{1}{t}B_t| > n$. Esta propiedad implica que B_t no es diferenciable en $t = 0$. Para cada número natural n defina el evento

$$A_n = \left\{ \left| \frac{1}{t} B_t \right| > n, \text{ para algún } t \in [0, 1/n^2] \right\},$$

y observe que $A_1 \supseteq A_2 \supseteq \dots$. Entonces

$$\begin{aligned} P(A_n) &\geq P\left(\left| \frac{1}{1/n^4} B_{1/n^4} \right| > n \right) \\ &= P\left(|B_{1/n^4}| > \frac{1}{n^3} \right) \\ &= P\left(|n^2 B_{(1/n^4)(1)}| > \frac{1}{n} \right) \\ &= P\left(|B_1| > \frac{1}{n} \right) \rightarrow 1 \text{ cuando } n \rightarrow \infty. \end{aligned}$$

Hemos usado el hecho de que $\frac{1}{c}B_{c^2t}$ es también un movimiento Browniano para cualquier $c > 0$ constante. Por lo tanto $P(A_1) \geq P(A_2) \geq \dots \geq 1$. Es decir, $P(A_n) = 1$ para cualquier $n \geq 1$. \square

Es decir, para cada $t_0 \geq 0$, el conjunto de trayectorias $t \mapsto B_t$ que no son diferenciables en t_0 tiene probabilidad uno. Este conjunto de trayectorias puede cambiar para cada valor de t_0 , aunque cada una de ellas tenga probabilidad uno. El siguiente resultado, más fuerte y que se enuncia sin demostración, asegura que con probabilidad uno no hay diferenciabilidad en ningún punto. Observe que el conjunto de tiempos $t_0 \geq 0$ no es numerable y por lo tanto la afirmación no se sigue de manera obvia del resultado anterior.

Proposición. Con probabilidad uno, el movimiento Browniano $\{B_t\}$ no es diferenciable en ningún $t \geq 0$.

Las trayectorias Brownianas son entonces ejemplos de funciones, otrora consideradas extrañas, que son continuas pero no diferenciables en ningún punto. La gráfica de la Figura 8.2 muestra una de tales trayectorias, el zigzagueo incesante del movimiento de la partícula no permite la diferenciabilidad de su trayectoria en ningún punto. Este tipo de resultados son los que dan la pauta para buscar desarrollar una teoría de la diferenciabilidad de funciones un poco más amplia que la proporcionada por el cálculo diferencial tradicional.

8.5. Movimiento Browniano multidimensional

El movimiento Browniano que hemos estudiado con valores en \mathbb{R} puede extenderse a un proceso con valores en \mathbb{R}^n de la siguiente forma.

Definición. Sean $B_1(t), \dots, B_n(t)$ movimientos Brownianos independientes unidimensionales. El movimiento Browniano en \mathbb{R}^n es el proceso

$$B(t) = (B_1(t), \dots, B_n(t)).$$

En la Figura 8.4 puede apreciarse una simulación de una trayectoria Browniana en \mathbb{R}^2 . En completa analogía con el caso unidimensional, este proceso puede definirse de manera alternativa mediante los siguientes postulados. Primeramente se pide que $B(0) = (0, \dots, 0)$ casi seguramente. Se presupone además que las trayectorias $t \mapsto B(t)$ son continuas, y que el proceso tiene incrementos independientes. Finalmente, para cualesquiera tiempos $0 \leq s < t$, el vector $B(t) - B(s)$ tiene una distribución normal multivariada con media el vector de ceros $(0, \dots, 0)$, y matriz de covarianzas la matriz diagonal

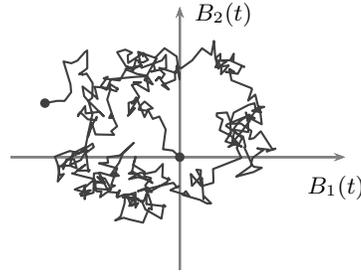


Figura 8.4:

$$\text{Var}(B(t) - B(s)) = \begin{pmatrix} (t-s)\sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & (t-s)\sigma_n^2 \end{pmatrix}.$$

Es decir, la función de densidad del vector $B(t) - B(s)$ es

$$f(x_1, \dots, x_n) = \frac{1}{\sqrt{2\pi(t-s)\sigma_1^2}} e^{-x_1^2/2(t-s)\sigma_1^2} \dots \\ \dots \frac{1}{\sqrt{2\pi(t-s)\sigma_n^2}} e^{-x_n^2/2(t-s)\sigma_n^2}.$$

Cuando los valores de los parámetros σ son todos uno, se dice nuevamente que

el movimiento Browniano es *estándar*, y la función de densidad de $B(t) - B(s)$ adquiere la expresión compacta

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi(t-s))^{n/2}} e^{-\|x\|^2/2(t-s)},$$

en donde $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$. Como en el caso unidimensional, puede considerarse un movimiento Browniano que inicie en $x \in \mathbb{R}^n$, y entonces para cualquier $t > 0$ la probabilidad de transición o densidad de $B(t)$ es

$$p(t, x, y) = \frac{1}{(2\pi t)^{n/2}} e^{-\|y-x\|^2/2t},$$

que nuevamente cumple al ecuación de Chapman-Kolmogorov

$$p(t+s, x, y) = \int_{\mathbb{R}^n} p(t, x, u) p(s, u, y) du.$$

El proceso de Bessel. Sea $B(t)$ un movimiento Browniano en \mathbb{R}^n . El proceso de Bessel es el proceso dado por $R(t) = \|B(t)\| = (B_1^2(t) + \dots + B_n^2(t))^{1/2}$. Es decir, $R(t)$ es la distancia Euclideana que guarda un movimiento Browniano n -dimensional respecto al origen al tiempo t , y por eso se le llama a veces *movimiento Browniano radial*. Se trata pues de un proceso con valores en $[0, \infty)$ que evidentemente tiene trayectorias continuas. Puede demostrarse (véase [1]) que este proceso cumple la propiedad de Markov y que la función de probabilidades de transición $p(t, x, y)$ puede expresarse en términos de las funciones de Bessel, y de allí es de donde adquiere este nombre alternativo.

Ecuación de calor en dominios acotados. Vamos a enunciar sin demostración un resultado que nos llevará a una aplicación del movimiento Browniano. Considere una región abierta y acotada D de \mathbb{R}^n con frontera ∂D , y suponga que la función $u(t, x)$ representa la temperatura en el punto $x \in D$ al tiempo $t \geq 0$. La evolución en el tiempo de esta función está dada por la ecuación de calor

$$\frac{\partial}{\partial t} u(t, x) = \frac{d}{2} \Delta u(t, x),$$

en donde Δ es el operador Laplaciano, y d es una constante positiva. Suponga además que se tiene una temperatura inicial $u(0, x) = f(x)$ para $x \in D$, y la condición de frontera $u(t, x) = g(x)$ para $x \in \partial D$. Sea x un punto cualquiera en D y defina el tiempo $\tau = \inf \{t > 0 : B_t^x \in \partial D\}$, en donde $\{B_t^x\}$ es un

movimiento Browniano de parámetro $\sigma^2 = d$, y que inicia en x . La solución a esta ecuación de calor con las condiciones mencionadas se puede expresar en términos del movimiento Browniano de la siguiente forma

$$u(t, x) = E(f(B_t^x)1_{(\tau > t)} + g(B_\tau^x)1_{(\tau \leq t)}). \quad (8.3)$$

Conforme $t \rightarrow \infty$ la estructura de la ecuación de calor hace que la solución $u(t, x)$ se aproxime a una función $u(x)$, la solución estacionaria de la ecuación de calor, que satisface

$$\Delta u(x) = 0,$$

para $x \in D$, y conservando la condición de frontera $u(x) = g(x)$ para $x \in \partial D$, es decir,

$$u(x) = \lim_{t \rightarrow \infty} u(t, x) = \begin{cases} E(g(B_\tau^x)) & \text{si } x \in D, \\ g(x) & \text{si } x \in \partial D. \end{cases}$$

■ Ejemplo: El problema de la ruina del jugador con trayectorias Brownianas. Suponga que un movimiento Browniano unidimensional inicia en el punto x dentro del intervalo (a, b) , con $0 \leq a < b < \infty$. ¿Cuál es la probabilidad de que el proceso tome el valor a antes que b ? Una trayectoria Browniana que cumple tal condición se muestra en la Figura 8.5(a). Este es el problema de la ruina del jugador estudiado antes sólo que ahora el capital del jugador cambia continuamente siguiendo un movimiento Browniano. Llegar primero al valor a se interpreta como ruina, y el juego es justo pues los saltos del movimiento Browniano tienen esperanza nula. Defina nuevamente el tiempo de paro $\tau = \inf\{t > 0 : B_t^x = a \text{ ó } B_t^x = b\}$. Nos interesa encontrar

$$u(x) = P(B_\tau^x = a) = E(1_{\{a\}}(B_\tau^x)).$$

Por lo anterior, esta función cumple la ecuación $u''(x) = 0$, para $a < x < b$, con condiciones de frontera $u(a) = 1$ y $u(b) = 0$. La solución es $u(x) = (b - x)/(b - a)$, cuya gráfica se muestra en la Figura 8.5(b).

o

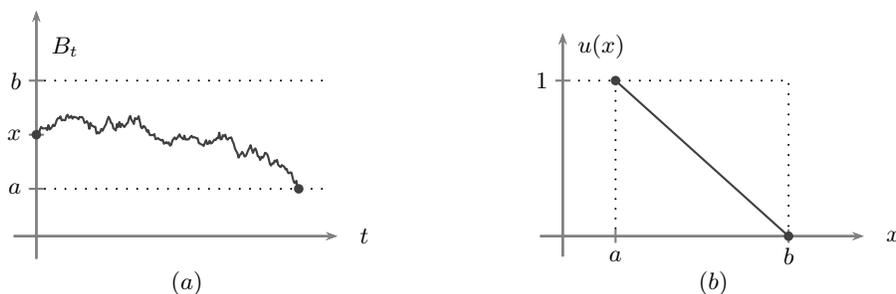


Figura 8.5:

8.6. El principio de reflexión

Este resultado tiene una interpretación geométrica fácil de entender. Considere un movimiento Browniano que inicia en a , y sea b otro número tal que $b > a$. El conjunto de trayectorias que tocan la línea horizontal $y = b$ en algún tiempo $s \in [0, t]$, se descompone en dos conjuntos ajenos que son simétricos uno del otro y tienen idéntica probabilidad: aquel conjunto de trayectorias que finalizan en algún punto

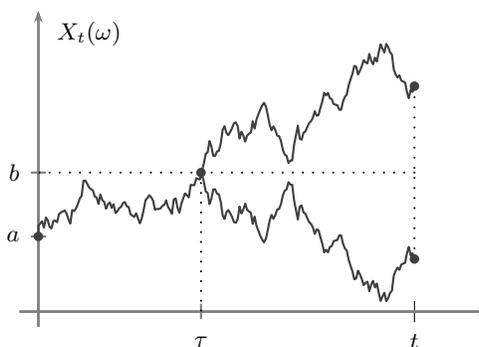


Figura 8.6:

arriba de b , y el conjunto de trayectorias que terminan por abajo de b . En la Figura 8.6 se ilustra esta situación. Una vez que una trayectoria toca el punto b , es igualmente probable que finalice al tiempo t arriba de b o abajo de b . Este resultado adquiere su nombre debido a esta propiedad de reflexión y facilita el cálculo de algunas probabilidades, en particular, lo usaremos para demostrar la propiedad de recurrencia puntual del movimiento Browniano unidimensional.

Proposición. (Principio de reflexión). Sea B_t^a un movimiento Browniano que empieza en a , y sea $b > a$. Para cualquier $t > 0$,

$$P(B_s^a \geq b \text{ para algún } s \in [0, t]) = 2P(B_t^a \geq b). \quad (8.4)$$

Demostación. Sea τ el primer momento en el que el movimiento Browniano es igual a b , es decir, sea $\tau = \inf\{t \geq 0 : B_t^a = b\}$. Esta variable aleatoria puede tomar el valor infinito si el evento mencionado nunca ocurre. Entonces

$$\begin{aligned} P(B_s^a \geq b \text{ para algún } s \in [0, t]) &= P(\tau \leq t) \\ &= P(\tau < t) + P(\tau = t) \\ &= P(\tau < t). \end{aligned}$$

La última igualdad se debe a que $P(\tau = t) \leq P(B_t^a = b) = 0$, por ser B_t^a una variable aleatoria continua. Por otro lado,

$$\begin{aligned} P(B_t^a \geq b) &= P(B_t^a \geq b | \tau \leq t) P(\tau \leq t) \\ &= P(B_t^a - b \geq 0 | \tau < t) P(\tau < t), \end{aligned} \quad (8.5)$$

en donde, por la propiedad de Markov, y condicionada a la ocurrencia del evento $(\tau < t)$, la variable $B_t^a - b = B_t^a - B_\tau^a$ tiene distribución $N(0, (t - \tau)\sigma^2)$. Por lo tanto $P(B_t^a - b \geq 0 | \tau < t) = 1/2$. Substituyendo en (8.5) se obtiene

$$P(\tau < t) = 2P(B_t^a \geq b).$$

□

8.7. Recurrencia y transitoriedad

En esta sección encontraremos la probabilidad de que el movimiento Browniano eventualmente regrese a su posición de origen. Veremos que la respuesta depende de la dimensión del proceso. Empezaremos estudiando una propiedad general en el caso unidimensional.

Proposición. Sea $\{B_t\}$ un movimiento Browniano unidimensional que inicia en cero y considere dos tiempos t_1 y t_2 tales que $0 < t_1 < t_2$. Entonces

$$P(B_t = 0 \text{ para algún } t \in [t_1, t_2]) = 1 - \frac{2}{\pi} \arctan \frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}}.$$

Demostración. Para cualquier $u > 0$, mediante argumentos de traslación y por el principio de reflexión se tiene que

$$\begin{aligned} P(B_t = 0 \text{ para algún } t \in [t_1, t_2] \mid B_{t_1} = u) &= P(B_t \leq -u \text{ para algún } t \in [0, t_2 - t_1]) \\ &= P(B_t \geq u \text{ para algún } t \in [0, t_2 - t_1]) \\ &= 2P(B_{t_2-t_1} \geq u). \end{aligned}$$

Por simetría se tiene la misma probabilidad para el caso $u < 0$. Entonces

$$\begin{aligned} P(B_t = 0 \text{ para algún } t \in [t_1, t_2]) &= \int_{-\infty}^{\infty} P(B_t = 0 \text{ para algún } t \in [t_1, t_2] \mid B_{t_1} = u) p(t_1, 0, u) du \\ &= \int_{-\infty}^{\infty} 2P(B_{t_2-t_1} \geq u) p(t_1, 0, u) du \\ &= 4 \int_0^{\infty} P(B_{t_2-t_1} \geq u) p(t_1, 0, u) du \\ &= 4 \int_0^{\infty} \left(\int_u^{\infty} p(t_2 - t_1, 0, v) dv \right) p(t_1, 0, u) du \\ &= 4 \int_0^{\infty} \int_u^{\infty} \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-v^2/2(t_2-t_1)} \frac{1}{\sqrt{2\pi t_1}} e^{-u^2/2t_1} dv du. \end{aligned}$$

Haciendo el cambio de variable $(x, y) = (u/\sqrt{t_1}, v/\sqrt{t_2 - t_1})$ se obtiene la expresión equivalente

$$4 \int_0^{\infty} \int_{x\sqrt{t_1}/\sqrt{t_2-t_1}}^{\infty} \frac{1}{2\pi} e^{-(x^2+y^2)/2} dy dx.$$

Ahora se resuelve esta integral usando coordenadas polares. La región de integración $\{(x, y) : x \geq 0 \text{ y } y \geq x\sqrt{t_1}/\sqrt{t_2 - t_1}\}$ que se muestra en la Figura 8.7,

corresponde a la región polar $\{(r, \theta) : r > 0 \text{ y } \theta \in (\arctan \frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}}, \pi/2)\}$. Por lo tanto la probabilidad buscada es

$$\begin{aligned} & 4 \int_{\arctan \frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}}}^{\pi/2} \int_0^{\infty} \frac{1}{2\pi} e^{-r^2/2} r \, dr \, d\theta \\ &= 4 \left(\frac{\pi}{2} - \arctan \frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}} \right) \frac{1}{2\pi} \int_0^{\infty} e^{-r^2/2} r \, dr \\ &= 1 - \frac{2}{\pi} \arctan \frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}}. \end{aligned}$$

□

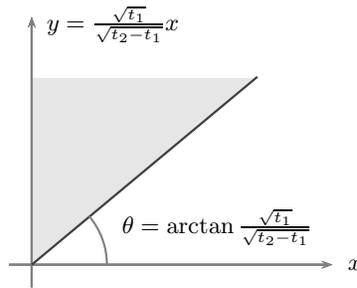


Figura 8.7:

Con ayuda de este resultado demostramos ahora la recurrencia puntual del movimiento Browniano unidimensional.

Proposición. (Recurrencia puntual del mov. Browniano unidimensional). Con probabilidad uno el movimiento Browniano unidimensional es recurrente puntual, es decir, regresa a su posición de origen una infinidad de veces casi seguramente.

Demostración. Haciendo t_2 tender a infinito en la fórmula recién demostrada e intercambiando el límite con la probabilidad, lo cual es válido pues la sucesión de

eventos es creciente, se obtiene que para cualquier valor positivo de t_1 ,

$$P(B_t = 0 \text{ para algún } t \in [t_1, \infty)) = \lim_{t_2 \rightarrow \infty} \left(1 - \frac{2}{\pi} \arctan \frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}}\right) = 1.$$

Esto es, la probabilidad de que el movimiento Browniano unidimensional regrese al cero para algún tiempo t dentro del intervalo $[t_1, \infty)$ es uno, sin importar la magnitud de t_1 . Ahora, una vez que regresa a cero, por la propiedad fuerte de Markov, inicia en ese momento otro movimiento Browniano que eventualmente regresará nuevamente a cero con probabilidad uno. De esta manera regresará a cero una infinidad de veces con probabilidad uno. \square

Esta propiedad de recurrencia significa que una trayectoria como la de la Figura 8.2 cruzará una infinidad de veces el eje horizontal, casi seguramente. Puede uno también considerar que el movimiento inicia en un punto cualquiera x obteniéndose la misma propiedad de recurrencia al punto x . La siguiente conclusión es también muy interesante: Hemos mencionado antes que el proceso $W_t = tB_{1/t}$, con $W_0 = 0$, es también un movimiento Browniano. Ahora, dado que B_t es cero para una infinidad de valores de t en cualquier intervalo de la forma $[t_1, \infty)$, se tiene entonces que W_t es cero una infinidad de veces dentro del intervalo $(0, 1/t_1)$. Es decir, en cualquier vecindad de $t = 0$, las trayectorias del movimiento Browniano cruzan el eje horizontal una infinidad de veces, casi seguramente. Esta misma conclusión puede obtenerse directamente de la fórmula recién demostrada tomando $t_2 > 0$ fijo y haciendo $t_1 \rightarrow 0$, es decir,

$$P(B_t = 0 \text{ para algún } t \in (0, t_2]) = \lim_{t_1 \rightarrow 0} \left(1 - \frac{2}{\pi} \arctan \frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}}\right) = 1.$$

En el caso de dimensiones mayores la situación es distinta.

Proposición. Sea $B(t)$ un movimiento Browniano n -dimensional que inicia en el origen, y sea el disco $D = \{x \in \mathbb{R}^n : \|x\| < r\}$, para algún $r > 0$.

1. Cuando $n = 2$, con probabilidad uno el movimiento Browniano visita la vecindad D en una infinidad de tiempos no acotados. Esto es, el proceso es recurrente por vecindades. Sin embargo no es recurrente puntual pues la probabilidad de que regrese exactamente al punto de partida es cero.
2. Cuando $n \geq 3$, el proceso es transitorio por vecindades, es decir, existe una probabilidad positiva de que el proceso nunca regrese a la vecindad del punto de partida.

Demostración. Sean r_1 y r_2 dos radios tales que $0 < r_1 < r_2$, y defina la región $A = \{x \in \mathbb{R}^n : r_1 < \|x\| < r_2\}$ como se muestra en la Figura 8.8. La frontera de A es $\partial A = \{x \in \mathbb{R}^n : \|x\| = r_1 \text{ ó } \|x\| = r_2\}$, y $\|x\| = \sqrt{x_1^2 + x_2^2}$.

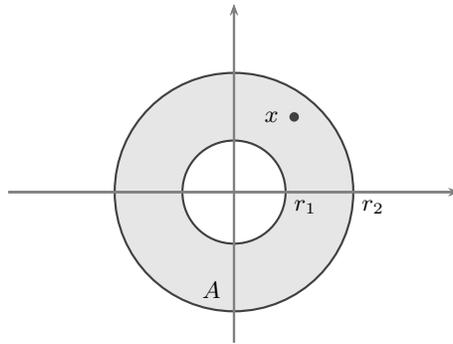


Figura 8.8:

Suponga que el movimiento Browniano inicia en el origen y que en algún tiempo posterior se encuentra en un punto x dentro de la región A . Defina la función $f(x)$ como la probabilidad de que el movimiento Browniano que parte de x llegue a la circunferencia exterior $\{x \in \mathbb{R}^n : \|x\| = r_2\}$ antes que a la circunferencia interior $\{x \in \mathbb{R}^n : \|x\| = r_1\}$. Es decir, si se define el tiempo de paro

$$\tau = \inf \{t > 0 : B(t) \in \partial D\},$$

entonces $f(x) = P(\|B(\tau)\| = r_2 \mid B(0) = x)$. Esta función puede también escribirse como

$$f(x) = E(g(B(\tau)) \mid B(0) = x),$$

en donde $g(x) : \partial A \rightarrow \mathbb{R}$ es la función indicadora

$$g(y) = \begin{cases} 1 & \text{si } \|y\| = r_2, \\ 0 & \text{si } \|y\| = r_1. \end{cases}$$

La función $f(x)$ satisface

$$\Delta f(x) = 0, \tag{8.6}$$

con condiciones de frontera

$$f(y) = g(y) = \begin{cases} 1 & \text{si } \|y\| = r_2, \\ 0 & \text{si } \|y\| = r_1. \end{cases}$$

Dada la simetría del movimiento Browniano, la función $f(x)$ depende de x sólo a través de la magnitud $\|x\|$. Sea entonces $f(x) = \phi(\|x\|)$ para alguna función $\phi(r)$ con $r = \|x\|$. En este caso particular, conviene escribir al operador de Laplace en coordenadas esféricas adquiriendo la expresión

$$\begin{aligned} \Delta \phi &= \frac{1}{r^{n-1}} \frac{d}{dr} (r^{n-1} \frac{d}{dr} \phi) \\ &= \frac{d^2}{dr^2} \phi + \frac{n-1}{r} \frac{d}{dr} \phi. \end{aligned}$$

Por lo tanto, la ecuación (8.6) se escribe

$$\phi''(r) + \frac{n-1}{r} \phi'(r) = 0, \quad \text{para } r \in (r_1, r_2),$$

y las nuevas condiciones de frontera son $\phi(r_2) = 1$ y $\phi(r_1) = 0$. La solución general de esta ecuación es

$$\phi(r) = \begin{cases} c_1 \ln r + c_2 & \text{si } n = 2, \\ c_1 r^{2-n} + c_2 & \text{si } n \geq 3, \end{cases}$$

con c_1 y c_2 constantes. Usando ahora las condiciones de frontera se obtiene

$$\phi(\|x\|) = \frac{\ln \|x\| - \ln r_1}{\ln r_2 - \ln r_1} \quad \text{si } n = 2, \tag{8.7}$$

$$\phi(\|x\|) = \frac{r_1^{2-n} - \|x\|^{2-n}}{r_1^{2-n} - r_2^{2-n}} \quad \text{si } n \geq 3. \tag{8.8}$$

Estos resultados nos permitirán encontrar las probabilidades buscadas tomando algunos límites sobre los radios r_1 y r_2 . Para el caso $n = 2$, la probabilidad de que el movimiento Browniano que inicia en x nunca visite la bola de radio r_1 alrededor del cero es, por (8.7),

$$\begin{aligned} & \lim_{r_2 \rightarrow \infty} P(\|B(\tau)\| = r_2 \text{ antes que } \|B(\tau)\| = r_1 \mid B(0) = x) \\ &= \lim_{r_2 \rightarrow \infty} \frac{\ln \|x\| - \ln r_1}{\ln r_2 - \ln r_1} \\ &= 0. \end{aligned}$$

Es decir, la probabilidad de que el movimiento Browniano en \mathbb{R}^2 visite el disco de radio r_1 alrededor del origen es uno. Y regresará a dicho disco en una infinidad de tiempos no acotados. Este comportamiento se conoce con el nombre de *recurrencia por vecindades*. Sin embargo, no se presenta la recurrencia puntual, es decir, la probabilidad de que el movimiento Browniano en \mathbb{R}^2 regrese exactamente al punto de origen es cero. Esto es consecuencia nuevamente de (8.7) al tomar el límite cuando $r_1 \rightarrow 0$. Es decir, la probabilidad de que el proceso tome el valor $(0, 0)$ antes de que toque el círculo de radio r_2 es

$$\begin{aligned} & \lim_{r_1 \rightarrow 0} P(\|B(\tau)\| = r_1 \text{ antes que } \|B(\tau)\| = r_2 \mid B(0) = x) \\ &= \lim_{r_1 \rightarrow 0} \left(1 - \frac{\ln \|x\| - \ln r_1}{\ln r_2 - \ln r_1} \right) \\ &= 0. \end{aligned}$$

Ahora consideremos el caso $n \geq 3$. La probabilidad de que el proceso que inicia en x nunca visite el disco de radio r_1 alrededor del cero es, por (8.8),

$$\begin{aligned} & \lim_{r_2 \rightarrow \infty} P(\|B(\tau)\| = r_2 \text{ antes que } \|B(\tau)\| = r_1 \mid B(0) = x) \\ &= \lim_{r_2 \rightarrow \infty} \frac{r_1^{2-n} - \|x\|^{2-n}}{r_1^{2-n} - r_2^{2-n}} \\ &= 1 - \left(\frac{r_1}{\|x\|} \right)^{n-2} \\ &> 0. \end{aligned}$$

Es decir, existe una probabilidad estrictamente positiva de que el proceso nunca visite el disco de radio r_1 alrededor del origen, cuando inicia en x . Este es el comportamiento transitorio del movimiento Browniano para dimensiones $n \geq 3$. Explícitamente, la probabilidad de un retorno exacto al origen es cero pues por (8.8), esta

probabilidad es

$$\begin{aligned} & \lim_{r_1 \rightarrow 0} P(\|B(\tau)\| = r_1 \text{ antes que } \|B(\tau)\| = r_2 \mid B(0) = x) \\ &= \lim_{r_1 \rightarrow 0} \left(1 - \frac{r_1^{2-n} - \|x\|^{2-n}}{r_1^{2-n} - r_2^{2-n}}\right) \\ &= 0. \end{aligned}$$

□

Notas y referencias

El lector puede encontrar una muy interesante y motivadora exposición histórica sobre el descubrimiento del movimiento Browniano en el excelente libro de Edward Nelson [24]. Este libro se encuentra disponible en formato electrónico en la página web del autor. Para otras primeras lecturas y mas resultados sobre el movimiento Browniano pueden consultarse, por ejemplo, los textos de Karlin y Taylor [19], Lawler [23], Nelson [24], y Tudor [38].



N. Wiener

Norbert Wiener (USA 1894–1964).

Norbert Wiener fue ante todo un niño prodigio. En 1903 a la edad de 9 años ingresó a la preparatoria Ayer en Massachusetts, la cual concluyó en 1906 para ingresar después al colegio Tufts en donde estudió matemáticas con el apoyo y asesoría de su padre. En 1909 se graduó del colegio Tufts a la edad 14 años e ingresó a la universidad de Harvard para realizar estudios de posgrado en Zoología. Con ayuda de una beca ingresó después a la universidad de Cornell en 1910 en donde tomó cursos de posgrado en matemáticas y filosofía, pero su desempeño allí no fue del todo satisfactorio y su padre lo regresó a Harvard para continuar sus estudios de filosofía. Se graduó en Harvard a la edad de

18 años con una tesis sobre lógica matemática bajo la dirección de Karl Schmidt. Después de Harvard viajó a Cambridge, Inglaterra, para estudiar junto a Bertrand Russell y en donde asistió a algunos cursos dictados por G. H. Hardy. En 1914 viajó a Gottingen para estudiar ecuaciones diferenciales con David Hilbert. Regresó a los Estados Unidos dos días antes del inicio de la primera guerra mundial. Murió en Estocolmo a la edad de 69 años después de sufrir un segundo ataque al corazón. Una universidad en el Perú lleva ahora su nombre. Fuente: Archivo MacTutor, Universidad de St. Andrews.



P. P. Lèvy

Paul Pierre Lèvy (Francia 1886–1971).

Paul Lèvy nació dentro de una familia con tradición matemática. Su abuelo fue profesor de matemáticas y su padre estudió geometría y trabajó en la *École Polytechnique*. De pequeño, Lèvy asistió al *Lycée Saint Louis* en París, en donde mostró ser un excelente estudiante, sobresaliendo y ganando premios no sólo en matemáticas sino también en Griego, física y química. Ingresó después a la *École Polytechnique* y siendo allí aún estudiante publicó su primer artículo en matemáticas en 1905, el cual tocaba temas de series convergentes. Después de un año de servicio militar, ingresó en 1907 a la *École des Mines*, y asistió a cursos de matemáticas en la *Sorbonne*. En 1910 concluyó sus estudios en la *École des Mines*, y realizó a partir de entonces trabajos de investigación en análisis funcional que le llevaron a obtener el grado de *Docteur és Sciences* en 1912 bajo el escrutinio de E. Picard, H. Poincaré y J. Hadamard. Trabajó como profesor en la *École des Mines de Saint-Etienne* en París de 1910 a 1913, y en la *École Nationale Supérieure de Mines* de 1914 a 1951. También impartió clases en la *École Polytechnique* de 1920 a 1959, año en el que se jubiló. En 1963 fue nombrado miembro honorario de la *London Mathematical Society*, y en 1964 fue elegido miembro de la *Académie des Sciences*. Paul Lèvy realizó contribuciones importantes en la teoría de la probabilidad, el análisis funcional y las ecuaciones diferenciales parciales. Dentro de sus textos publicados se encuentran *Leçons d'analyse fonctionnelle* (1922), *Calcul des probabilités* (1925), *Theorie de l'addition des variables aléatoires* (1937), y *Processus stochastiques et mouvement Brownien* (1948). Paul Lèvy fue uno de los más grandes matemáticos de su tiempo. Como científico fue un ejemplo de individualismo absoluto, era un investigador solitario que sólo se preocupaba por plantearse problemas matemáticos de su interés y buscaba su solución a través de la reflexión interior. No participaba mayormente en los congresos internacionales excepto posiblemente hacia el final de su vida. A menudo encontraba por sí mismo resultados ya conocidos, y otras veces descubría resultados importantes nuevos sin darles mucha publicidad pues los creía ya conocidos. Paul Lèvy fue un ejemplo de un hombre de pensamiento de originalidad profunda, indiferente a las escuelas y métodos establecidos, y quien no dudaba en lanzarse sobre nuevos caminos de pensamiento pues no temía de ninguna manera a la soledad intelectual. Fuente: Archivo MacTutor, Universidad de St. Andrews.

Ejercicios

Movimiento Browniano unidimensional

163. Sea $\{B_t\}$ un movimiento Browniano de parámetro σ^2 . Demuestre que $W_t = B_t/\sigma^2$ es un movimiento Browniano estándar.
164. Sea $\sigma \neq 0$ una constante. Demuestre que si $\{B_t\}$ es un movimiento Browniano estándar, entonces los siguientes procesos son movimientos Brownianos de parámetro σ^2 .
- $W_t = \sigma B_t$.
 - $W_t = B_{\sigma^2 t}$.
165. Sea B_t un movimiento Browniano de parámetro σ^2 . A partir de la definición, demuestre que el proceso $W_t = tB_{1/t}$, con $W_0 = 0$, es un movimiento Browniano de parámetro σ^2 .
166. Demuestre que la función de covarianza del movimiento Browniano estándar es $\text{Cov}(B_s, B_t) = s \wedge t$.
167. Sea $t_0 > 0$ fijo y sea B_t es un movimiento Browniano estándar. Demuestre que el proceso $\{B_{t_0} - B_{t_0-t} : 0 \leq t \leq t_0\}$ es un movimiento Browniano en el intervalo $[0, t_0]$. Este es un movimiento Browniano con tiempo invertido.
168. Sea $\{B_t\}$ un movimiento Browniano de parámetro σ^2 . Demuestre que para cualesquiera tiempos $t \neq s$, $P(B_t > B_s) = 1/2$.
169. Demuestre que la probabilidad de transición $p(t, x, y)$ del movimiento Browniano unidimensional de parámetro σ^2 satisface la ecuación de difusión (también llamada ecuación de calor):

$$\frac{\partial p}{\partial t} = \frac{\sigma}{2} \frac{\partial^2 p}{\partial y^2}.$$

170. Demuestre que la probabilidad de transición $p(t, x, y)$ del movimiento Browniano unidimensional de parámetro σ^2 cumple la ecuación de Chapman-Kolmogorov

$$p(t+s, x, y) = \int_{-\infty}^{\infty} p(t, x, u) p(s, u, y) du.$$

171. Sea $\{B_t\}$ un movimiento Browniano estándar. Demuestre que

- a) $E|B_t - B_s| = \sqrt{\frac{2|t-s|}{\pi}}$.
- b) $E(B_t - B_s)^2 = |t - s|$.
- c) $E(B_s B_t) = \min\{s, t\}$.
- d) $\rho(B_t, B_s) = \sqrt{s/t}$, para $0 \leq s \leq t$.

172. Demuestre que el movimiento Browniano es una martingala respecto de su filtración natural.

173. Sea $\{B_t\}$ un movimiento Browniano. Demuestre que el proceso $\{B_t^2 - t\}$ es también una martingala respecto de la misma filtración.

174. Sea $\{B_t\}$ un movimiento Browniano de parámetro σ^2 que inicia en cero. Sea a una constante y defina el tiempo $\tau = \inf\{t \geq 0 : B_t = a\}$. Encuentre la función de densidad de τ .

175. *Movimiento Browniano reflejado en el origen.* Sea B_t un movimiento Browniano estándar. El proceso $X_t = |B_t|$ corresponde a un movimiento Browniano que sólo toma valores positivos pues las trayectorias son reflejadas respecto del origen hacia la parte positiva del eje. Demuestre que X_t es un proceso de Markov con trayectorias continuas y que tiene función de probabilidad de transición $q(t, x, y)$ dada por

$$q(t, x, y) = p(t, x, y) + p(t, x, -y),$$

para cualesquiera valores $x, y \geq 0$, en donde $p(t, x, y)$ es la correspondiente función de probabilidad de transición del movimiento Browniano. Compruebe además que

- a) $E(X_t) = \sqrt{2t/\pi}$.
- b) $\text{Var}(X_t) = (1 - 2/\pi)t$.

176. *Movimiento Browniano con absorción en el origen.* Sea $\{B_t^x\}$ un movimiento Browniano estándar que inicia en $x > 0$. Defina el tiempo $\tau = \inf\{t \geq 0 : B_t^x = 0\}$ y el proceso

$$X_t = \begin{cases} B_t^x & \text{si } 0 \leq t \leq \tau, \\ 0 & \text{si } t > \tau, \end{cases}$$

el cual representa un movimiento Browniano que inicia en x con la característica de que una vez que llega al cero permanece en ese estado el resto del tiempo.

- a) Demuestre que $\{X_t\}$ es un proceso de Markov con trayectorias continuas.
- b) Observe que la variable X_t es mixta, es decir, no es discreta ni continua. Demuestre que la función de probabilidad de transición del proceso $\{X_t\}$ es, en su parte continua, $q(t, x, y) = p(t, 0, y - x) - p(t, 0, y + x)$, para $x, y > 0$, en donde $p(t, x, y)$ es la correspondiente función para B_t , un movimiento Browniano que inicia en cero. Demuestre además que, para la parte discreta, $q(t, x, 0) = 2(1 - F(x))$, en donde $F(x)$ es la función de distribución de B_t .
- c) Calcule $E(X_t)$ y $\text{Var}(X_t)$.

177. *La martingala exponencial.* El proceso que se obtiene al tomar la exponencial de un movimiento Browniano no es, en general, una martingala, sin embargo, añadiendo un término extra este nuevo proceso puede convertirse en una martingala y se le llama *martingala exponencial*. Más específicamente, sea B_t un movimiento Browniano estándar y sea c una constante. Defina el proceso $X_t = \exp(cB_t - c^2t/2)$. Compruebe que X_t es efectivamente una martingala respecto de la filtración natural del movimiento Browniano y verifique además que $E(X_t) = 1$ y $\text{Var}(X_t) = e^{c^2t} - 1$.
178. *El puente Browniano en $[0, 1]$.* Sea $\{B_t\}$ un movimiento Browniano unidimensional estándar. Demuestre que la distribución condicional de la variable B_t , con $t \in (0, 1)$, dado que $B_0 = 0$ y $B_1 = 0$, es

$$f(x) = \frac{1}{\sqrt{2\pi t(1-t)}} e^{-x^2/2t(1-t)},$$

para $-\infty < x < \infty$, es decir, B_t tiene distribución condicional $N(0, t(1-t))$. A este proceso condicionado se le conoce con el nombre de *puente Browniano* en el intervalo unitario $(0, 1)$. El siguiente ejercicio generaliza este resultado.

179. *El puente Browniano en $[t_1, t_2]$.* Sea $\{B_t\}$ un movimiento Browniano unidimensional estándar, y sean t_1 y t_2 dos tiempos fijos tales que $0 \leq t_1 < t_2$. Demuestre que la distribución condicional de la variable B_t , con $t \in (t_1, t_2)$, dado que $B_{t_1} = a$ y $B_{t_2} = b$, es $N(\mu, \sigma^2)$ con

$$\begin{aligned} \mu &= a + \frac{t - t_1}{t_2 - t_1}(b - a), \\ \sigma^2 &= \frac{(t_2 - t)(t - t_1)}{t_2 - t_1}. \end{aligned}$$

Movimiento Browniano multidimensional

180. Sea $B(t) = (B_1(t), \dots, B_n(t))$ un movimiento Browniano estándar en \mathbb{R}^n , y sea $r > 0$. Sea $\|x\| = (x_1^2 + \dots + x_n^2)^{1/2}$ la norma Euclideana en \mathbb{R}^n . Demuestre que la función de densidad de $\|B(t)\|$ es, para $x > 0$,

$$f(x) = \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} \left(\frac{x^2}{t}\right)^{n/2-1} \frac{2x}{t} e^{-x^2/2t}.$$

En particular, demuestre que para $n = 2$,

$$P(\|B(t)\| \leq x) = 1 - e^{-x^2/2t}.$$

181. Demuestre que el operator Laplaciano en \mathbb{R}^2 , $\Delta f(x, y) = \partial^2 f / \partial x^2 + \partial^2 f / \partial y^2$ adquiere la siguiente expresión en coordenadas polares

$$\Delta f(r, \theta) = \frac{\partial^2}{\partial r^2} f + \frac{1}{r} \frac{\partial}{\partial r} f + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} f.$$

Compruebe que cuando f depende de x y de y únicamente a través de $r = \sqrt{x^2 + y^2}$, el Laplaciano se reduce a la expresión

$$\Delta f(r, \theta) = \frac{d^2}{dr^2} f + \frac{1}{r} \frac{d}{dr} f.$$

Este resultado fue usado en la demostración de la propiedad de recurrencia por vecindades del movimiento Browniano en \mathbb{R}^2 .

182. Demuestre que el operator Laplaciano en \mathbb{R}^3 , $\Delta f(x, y, z) = \partial^2 f / \partial x^2 + \partial^2 f / \partial y^2 + \partial^2 f / \partial z^2$ adquiere la siguiente expresión en coordenadas esféricas

$$\Delta f(r, \theta, \phi) = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \frac{\partial}{\partial r}) f + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \frac{\partial}{\partial \theta}) f + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} f.$$

Compruebe que cuando f depende de x , y y z únicamente a través de $r = \sqrt{x^2 + y^2 + z^2}$, el Laplaciano se reduce a la expresión

$$\Delta f(r, \theta) = \frac{d^2}{dr^2} f + \frac{2}{r} \frac{d}{dr} f.$$

Este resultado fue usado en la demostración de la propiedad de transitoriedad del movimiento Browniano en \mathbb{R}^3 .

CAPÍTULO 9

Cálculo estocástico

En este último capítulo se presenta una breve introducción al cálculo estocástico de Itô. Vamos a definir la integral de Itô de un proceso estocástico respecto del movimiento Browniano. Mostraremos además el uso y aplicación de la fórmula de Itô, y resolveremos algunos modelos sencillos de ecuaciones estocásticas.

9.1. Integración estocástica

El objetivo de esta sección es presentar la definición de la integral de Itô de un proceso X_t respecto del movimiento Browniano, es decir, una integral de la forma

$$\int_0^T X_t dB_t. \quad (9.1)$$

Este tipo de integrales no pueden definirse trayectoria por trayectoria, es decir, como si fuera una integral de Riemann-Stieltjes de una función respecto de otra función, pues en este caso la función integradora es una trayectoria del movimiento Browniano que, como hemos visto antes, no tiene variación finita. La justificación para desear definir este tipo de integrales será evidente más adelante cuando estudiemos ecuaciones estocásticas basadas en este tipo de integrales. Definir la integral de Itô a un nivel elemental nos conducirá necesariamente a dejar sin demostración algunos resultados técnicos. Daremos la definición de integral estocástica en varios pasos, primero para procesos simples y después, por aproximación, para procesos más generales.

Primeras hipótesis. Consideraremos como elementos iniciales un espacio de probabilidad (Ω, \mathcal{F}, P) , y un movimiento Browniano estándar unidimensional $\{B_t\}$, junto con su filtración natural $\{\mathcal{F}_t\}$. Supondremos dado un proceso $\{X_t\}$ con espacio parametral el intervalo $[0, T]$, con $T > 0$ fijo, y que visto como función $X : \Omega \times [0, T] \rightarrow \mathbb{R}$, es $\mathcal{F}_T \otimes \mathcal{B}[0, T]$ -medible, en donde el término $\mathcal{F}_T \otimes \mathcal{B}[0, T]$ corresponde a la mínima σ -álgebra generada por el espacio producto $\mathcal{F}_T \times \mathcal{B}[0, T]$. Supondremos además que el proceso es adaptado, es decir, para cada t en el intervalo $[0, T]$, la variable aleatoria X_t es medible respecto de \mathcal{F}_t .

El espacio $L^2(P)$. Denotaremos por $L^2(P)$ al espacio vectorial de variables aleatorias X que son cuadrado integrables, es decir, que cumplen la condición

$$\|X\|_{L^2(P)} = (E|X|^2)^{1/2} < \infty.$$

La función $\|\cdot\|_{L^2(P)}$ define una norma en $L^2(P)$, es decir, es una función real definida sobre este espacio lineal que cumple las siguientes cuatro condiciones:

- a) $\|X\| \geq 0$.
- b) $\|X\| = 0 \Leftrightarrow X = 0$ c.s.
- c) $\|X + Y\| \leq \|X\| + \|Y\|$.
- d) $\|\alpha X\| = |\alpha| \|X\|$, α constante.

Se puede verificar que el espacio lineal $L^2(P)$ es completo respecto de esta norma, es decir, es un espacio de Banach. Esto quiere decir que toda sucesión de Cauchy en este espacio tiene límite en él. A la convergencia usando esta norma se le llama *convergencia en $L^2(P)$* , o también *convergencia en media cuadrática*. Por ejemplo, la variable aleatoria B_t del movimiento Browniano pertenece a este espacio pues

$$\|B_t\|_{L^2(P)} = (E|B_t|^2)^{1/2} = \sqrt{t} < \infty.$$

El espacio $L^2(P \times dt)$. Denotaremos también por $L^2(P \times dt)$ al espacio lineal de todos los procesos $X = \{X_t : 0 \leq t \leq T\}$, que cumplen la condición

$$\|X\|_{L^2(P \times dt)} = (E \int_0^T |X_t|^2 dt)^{1/2} < \infty.$$

Puede demostrarse que la función $\|\cdot\|_{L^2(P \times dt)}$ es efectivamente una norma y que este espacio es completo respecto de esta norma, es decir, es un espacio de Banach.

Por ejemplo, el movimiento Browniano $B = \{B_t : 0 \leq t \leq T\}$ pertenece a este espacio pues

$$\begin{aligned} \|B\|_{L^2(P \times dt)} &= (E \int_0^T |B_t|^2 dt)^{1/2} \\ &= (\int_0^T E |B_t|^2 dt)^{1/2} \\ &= (\int_0^T t dt)^{1/2} \\ &= \frac{T}{\sqrt{2}} < \infty. \end{aligned}$$

El espacio \mathcal{H}_0^2 de procesos simples. Definiremos primero la integral de Itô para procesos que tienen la forma indicada a continuación.

Definición. Sea $0 = t_0 < t_1 < \dots < t_n = T$ una partición finita del intervalo $[0, T]$. Un proceso estocástico *simple* es un proceso de la forma

$$X_t = \sum_{k=0}^{n-1} X^{(k)} 1_{[t_k, t_{k+1})}(t), \tag{9.2}$$

en donde $X^{(0)}, \dots, X^{(n-1)}$ es una colección de variables aleatorias adaptadas a la filtración $\{\mathcal{F}_{t_k}\}_{k=0}^{n-1}$, y que son cuadrado integrables.

La expresión $1_{[a,b)}(t)$ denota a la función indicadora del intervalo $[a, b)$. Un proceso simple es entonces un proceso constante por pedazos con trayectorias càdlàg (continuas por la derecha, con límite por la izquierda), y las condiciones solicitadas garantizan que el proceso es adaptado y tiene trayectorias cuadrado integrables. Una trayectoria de este tipo de procesos se muestra en la Figura 9.1. Denotaremos por \mathcal{H}_0^2 al espacio de todos los procesos simples.

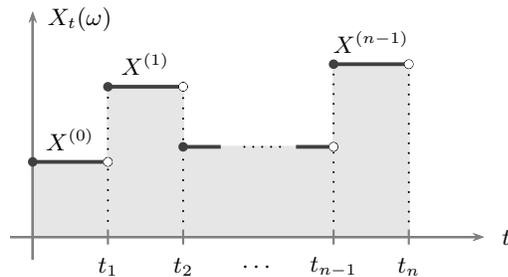


Figura 9.1:

Haciendo posiblemente algunos refinamientos en las particiones, dos procesos sim-

ples pueden siempre expresarse en términos de una misma partición común. De modo que la suma de dos procesos simples tiene sentido y resultará ser también un proceso simple. El espacio \mathcal{H}_0^2 es efectivamente un espacio vectorial.

Integral para procesos simples. Esta es la definición intuitiva de integral y establece simplemente que si el integrando es constante en algún subintervalo, entonces la integral debe ser esa constante multiplicada por el incremento del movimiento Browniano en dicho subintervalo.

Definición. La integral estocástica de Itô de un proceso simple X de la forma (9.2), respecto del movimiento Browniano, denotada por $I(X)$, se define como la variable aleatoria

$$I(X) = \int_0^T X_s dB_s = \sum_{k=0}^{n-1} X^{(k)} (B_{t_{k+1}} - B_{t_k}).$$

Veamos algunas propiedades de esta variable aleatoria.

- a) Es integrable pues siendo las variables $X^{(k)}$ y $B_{t_{k+1}} - B_{t_k}$ independientes, cada sumando tiene esperanza cero, y por lo tanto la esperanza de la integral es cero.
- b) La integral es además cuadrado integrable y de hecho se cumple la siguiente igualdad fundamental llamada *Isometría de Itô*:

$$\|I(X)\|_{L^2(P)} = \|X\|_{L^2(P \times dt)}. \quad (9.3)$$

Para comprobar esta identidad vamos a denotar nuevamente por ΔB_k a la diferencia $B_{t_{k+1}} - B_{t_k}$, y sea $\Delta t_k = t_{k+1} - t_k$. Nuevamente por la independencia

de $X^{(k)}$ y ΔB_k se tiene que

$$\begin{aligned}
 \|I(X)\|_{L^2(P)}^2 &= E\left(\left|\sum_{k=0}^{n-1} X^{(k)} (B_{t_{k+1}} - B_{t_k})\right|^2\right) \\
 &= \sum_{j,k=0}^{n-1} E(X^{(j)} X^{(k)} \Delta B_j \Delta B_k) \\
 &= \sum_{k=0}^{n-1} E((X^{(k)})^2 (\Delta B_k)^2) \\
 &= \sum_{k=0}^{n-1} E(X^{(k)})^2 \Delta t_k \\
 &= E\left(\int_0^T |X_t|^2 dt\right) \\
 &= \|X\|_{L^2(P \times dt)}^2.
 \end{aligned}$$

Esta identidad establece que tanto el proceso simple X como la variable aleatoria $I(X)$ tienen la misma norma en sus respectivos espacios. Como se verá más adelante, esta igualdad juega un papel primordial en la definición general de integral estocástica. La integral estocástica asigna entonces a cada elemento del espacio \mathcal{H}_0^2 una variable aleatoria en el espacio $L^2(P)$. De esta forma se tiene la transformación lineal $I : \mathcal{H}_0^2 \rightarrow L^2(P)$, que resulta ser continua por la isometría de Itô.

Observe que se puede tomar como ejemplo de proceso simple el movimiento Browniano discretizado, es decir, se puede tomar como proceso simple $X^{(k)} = B_{t_k}$, y de esta forma tener la integral estocástica discreta del movimiento Browniano respecto de sí mismo,

$$\sum_{k=0}^{n-1} B_{t_k} (B_{t_{k+1}} - B_{t_k}).$$

Extensión por aproximación. Ahora extenderemos la integral estocástica a procesos un poco más generales. Sea \mathcal{H}^2 el espacio de todos los procesos X_t medibles y adaptados, tales que

$$E \int_0^T |X_t|^2 dt < \infty.$$

El espacio \mathcal{H}^2 resulta ser un subespacio lineal cerrado de $L^2(P \times dt)$. Observe que la única diferencia entre estos dos espacios es que a los elementos de \mathcal{H}^2 se les pide

además que sean medibles y adaptados. En particular, todo proceso simple es un elemento de \mathcal{H}^2 . Tenemos entonces la contención de espacios $\mathcal{H}_0^2 \subset \mathcal{H}^2 \subset L^2(P \times dt)$, en donde puede probarse que \mathcal{H}_0^2 es denso en \mathcal{H}^2 respecto de la norma en $L^2(P \times dt)$. Esto significa que para cualquier proceso X en \mathcal{H}^2 existe una sucesión de procesos X^k en \mathcal{H}_0^2 tales que

$$\lim_{k \rightarrow \infty} \|X - X^k\|_{L^2(P \times dt)} = 0. \quad (9.4)$$

Este procedimiento de aproximación puede llevarse a cabo de la siguiente forma: Mediante la técnica de truncación todo proceso en \mathcal{H}^2 puede ser aproximado por un proceso acotado. A su vez todo proceso en \mathcal{H}^2 que es acotado se puede aproximar por procesos acotados y continuos. Y éstos a su vez se aproximan por procesos simples de la forma (9.2). Los detalles completos de esta sucesión de aproximaciones pueden encontrarse en [26]. Usando la isometría de Itô es sencillo comprobar que la sucesión $I(X^k)$ es una sucesión de Cauchy en el espacio $L^2(P)$, en efecto,

$$\begin{aligned} \|I(X^k) - I(X^l)\|_{L^2(P)} &= \|I(X^k - X^l)\|_{L^2(P)} \\ &= \|X^k - X^l\|_{L^2(P \times dt)} \\ &\leq \|X - X^k\|_{L^2(P \times dt)} + \|X - X^l\|_{L^2(P \times dt)}. \end{aligned}$$

Debido a (9.4) la última expresión puede hacerse tan pequeña como se desee, tomando índices k y l suficientemente grandes.

Definición. Sea X un proceso en \mathcal{H}^2 , y sea X^k una sucesión de procesos en \mathcal{H}_0^2 aproximante a X . Se define la integral estocástica de X como

$$I(X) = \lim_{k \rightarrow \infty} I(X^k),$$

en donde el límite debe entenderse dentro del espacio $L^2(P)$, es decir, se trata de la convergencia en media cuadrática de una sucesión de variables aleatorias.

Esto significa que la variable aleatoria $I(X)$ es un elemento de $L^2(P)$ y es tal que $\lim_{n \rightarrow \infty} \|I(X) - I(X^k)\|_{L^2(P)} = 0$. No es difícil verificar que tal definición es correcta en el sentido de que el límite no depende de la sucesión aproximante. En este punto empieza a perderse nuestra concepción tradicional de integral pues ahora ésta se encuentra definida a través de una sucesión aproximante del proceso a integrar. De manera gráfica esta extensión se ilustra en la Figura 9.2.

La isometría de Itô se cumple también para procesos en \mathcal{H}^2 como se muestra a continuación.

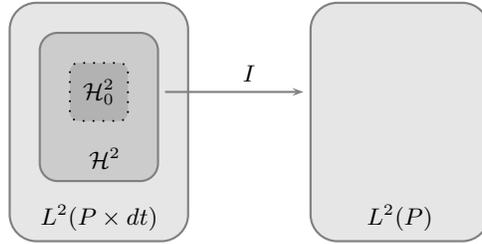


Figura 9.2:

Proposición. (Isometría de Itô). Para cualquier proceso X en \mathcal{H}^2 se cumple

$$\|I(X)\|_{L^2(P)} = \|X\|_{L^2(P \times dt)}. \tag{9.5}$$

Demostración. Sea X en \mathcal{H}^2 y sea X_n en \mathcal{H}_0^2 tal que $\|X - X_n\|_{L^2(P \times dt)} \rightarrow 0$. Esta convergencia y la desigualdad $|||a| - |b|| \leq \|a - b\|$ implican que $\|X_n\|_{L^2(P \times dt)} \rightarrow \|X\|_{L^2(P \times dt)}$. Análogamente, como $\|I(X) - I(X_n)\|_{L^2(P)} \rightarrow 0$ se tiene que $\|I(X_n)\|_{L^2(P)} \rightarrow \|I(X)\|_{L^2(P)}$. El resultado buscado se obtiene al tomar el límite en la isometría de Itô como se muestra en el siguiente diagrama:

$$\begin{array}{ccc} \|I(X_n)\|_{L^2(P)} & = & \|X_n\|_{L^2(P \times dt)} \\ \downarrow & & \downarrow \\ \|I(X)\|_{L^2(P)} & & \|X\|_{L^2(P \times dt)}. \end{array}$$

□

La propiedad de esperanza nula se cumple también para procesos en \mathcal{H}^2 , pues usando probabilidad elemental, o por la desigualdad de Jensen,

$$0 \leq E^2(I(X) - I(X^k)) \leq E(I(X) - I(X^k))^2 \rightarrow 0.$$

De donde se obtiene $E(I(X) - I(X^k)) \rightarrow 0$, es decir,

$$E(I(X)) = \lim_{k \rightarrow \infty} E(I(X^k)) = 0.$$

De esta forma se tiene ahora la transformación lineal y continua $I : \mathcal{H}^2 \rightarrow L^2(P)$. Observe nuevamente que el movimiento Browniano B_t es un ejemplo de un proceso en el espacio \mathcal{H}^2 , y es posible demostrar que tal proceso puede ser aproximado en el sentido de la norma del espacio $L^2(P \times dt)$ por el proceso simple

$$X_t = \sum_{k=0}^{n-1} B_{t_k} 1_{[t_k, t_{k+1})}(t),$$

en donde $0 = t_0 < t_1 < \dots < t_n = T$ es una partición de $[0, T]$. Se tiene entonces la siguiente integral estocástica particular, y su aproximación como límite en media cuadrática

$$\int_0^T B_t dB_t = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} B_{t_k} (B_{t_{k+1}} - B_{t_k}),$$

en donde el límite debe entenderse en el sentido de que la distancia máxima entre dos puntos sucesivos de la partición tiende a cero. Más adelante calcularemos esta integral estocástica de dos maneras, primero usando esta representación como límite en el espacio $L^2(P)$, y después usando la fórmula de Itô.

La integral como un proceso. Haremos ahora una pequeña extensión. Para cada t en $[0, T]$ y para cualquier X en \mathcal{H}^2 se define el proceso

$$I_t(X) = \int_0^t X_s 1_{[0, t]}(s) dB_s = \int_0^t X_s dB_s.$$

Este pequeño artificio permite ver a la integral estocástica no como una variable aleatoria sino como un proceso. Es claro que tal proceso no es necesariamente continuo, sin embargo puede demostrarse que existe una versión continua de él, y que esa versión es una martingala respecto de la filtración natural del movimiento Browniano. Denotaremos por el mismo símbolo a tal martingala continua.

Extensión por localización. Mediante un procedimiento llamado de localización es posible extender la definición de integral de Itô a procesos medibles y adaptados que cumplen la condición menos restrictiva

$$P\left(\int_0^T |X_t|^2 dt < \infty\right) = 1. \quad (9.6)$$

Denotaremos por \mathcal{L}_{loc}^2 el espacio de todos estos procesos. Este nuevo espacio contiene a \mathcal{H}^2 y es tal que para cada proceso X en \mathcal{L}_{loc}^2 existe una sucesión creciente de tiempos de paro $0 \leq \tau_1 \leq \tau_2 \leq \dots$ tales que $\tau_n \nearrow T$ cuando $n \rightarrow \infty$, y para

cada $n \geq 1$ el proceso $X_t \cdot 1_{(\tau_n \geq t)}$ pertenece al espacio \mathcal{H}^2 . Se define entonces la integral estocástica como el siguiente límite en el espacio $L^2(P)$,

$$\int_0^t X_s dB_s = \lim_{n \rightarrow \infty} \int_0^T X_s \cdot 1_{(\tau_n \geq t)}(\omega) \cdot 1_{[0,t]}(s) dB_s.$$

Nuevamente es posible demostrar que tal límite existe, que existe una versión continua de él, y que es independiente de la sucesión de tiempos de paro localizante. En este caso la integral ya no es una martingala sino una *martingala local*, esto quiere decir que el proceso detenido $I_{t \wedge \tau_n}(X)$ es una martingala para cada natural n . En general, la isometría de Itô ya no se cumple cuando la integral estocástica tiene como dominio de definición el espacio \mathcal{L}_{loc}^2 .

■ Ejemplo. Para el movimiento Browniano unidimensional B_t y para cualquier función continua f , el proceso $f(B_t)$ tiene trayectorias continuas y acotadas, por lo tanto se cumplen las condiciones de adaptabilidad y medibilidad y se cumple también (9.6), por lo tanto este proceso es un elemento de \mathcal{L}_{loc}^2 , y tiene sentido la expresión

$$\int_0^t f(B_s) dB_s.$$

Se puede demostrar que esta integral puede ser calculada mediante el siguiente límite en el espacio $L^2(P)$, aunque también se verifica la convergencia en probabilidad:

$$\int_0^t f(B_s) dB_s = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} f(B_{t_k}) (B_{t_{k+1}} - B_{t_k}), \quad (9.7)$$

en donde $0 = t_0 < t_1 < \dots < t_n = t$ es una partición de $[0, t]$, y nuevamente el límite debe entenderse en el sentido de que la distancia máxima entre dos puntos sucesivos de la partición tiende a cero. \circ

Con esto concluimos la serie de ideas generales bajo las cuales puede construirse la integral de Itô. Las demostraciones de algunos detalles técnicos que hemos simplemente mencionado se pueden encontrar en [35]. El esquema simplificado del procedimiento seguido para definir la integral estocástica se ilustra la Figura 9.3.

■ Ejemplo. Con la ayuda de (9.7) calcularemos la integral estocástica

$$\int_0^t B_s dB_s.$$

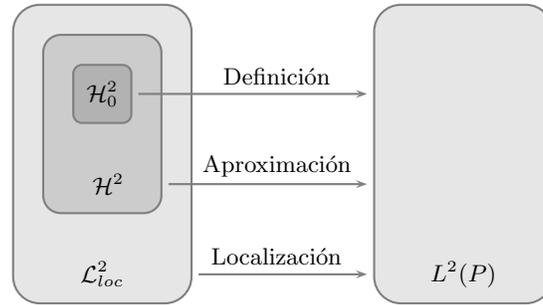


Figura 9.3:

Sea $0 = t_0 < t_1 < \dots < t_n = t$ una partición uniforme de $[0, t]$, es decir $t_{i+1} - t_i = 1/n$. Usando la identidad

$$a(b - a) = \frac{1}{2}(b^2 - a^2) - \frac{1}{2}(a - b)^2$$

se obtiene

$$\begin{aligned} \int_0^t B_s dB_s &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} B_{t_k} (B_{t_{k+1}} - B_{t_k}) \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} \left[\frac{1}{2}(B_{t_{k+1}}^2 - B_{t_k}^2) - \frac{1}{2}(B_{t_{k+1}} - B_{t_k})^2 \right] \\ &= \frac{1}{2}B_t^2 - \frac{1}{2}t. \end{aligned}$$

La primera suma es telescópica mientras que la segunda suma corresponde a la variación cuadrática del movimiento Browniano. Los límites indicados son válidos en el sentido de media cuadrática. Observe que aparece el término $\frac{1}{2}B_t^2$ como si se siguieran las reglas de integración usual, pero aparece también el término $-\frac{1}{2}t$, conocido como la *corrección de Itô*.

Ahora veamos el cambio en la solución de la integral cuando se modifica ligeramente la forma de calcular la integral. El cambio consiste en evaluar el integrando en el extremo derecho de cada subintervalo. Observe que en este caso el proceso a integrar ya no es adaptado y por lo tanto queda fuera de la teoría desarrollada antes. Usando

la identidad

$$b(b-a) = \frac{1}{2}(b^2 - a^2) + \frac{1}{2}(a-b)^2$$

se obtiene, nuevamente en el sentido de media cuadrática,

$$\begin{aligned} \int_0^t B_s dB_s &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} B_{t_{k+1}} (B_{t_{k+1}} - B_{t_k}) \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} \left[\frac{1}{2}(B_{t_{k+1}}^2 - B_{t_k}^2) + \frac{1}{2}(B_{t_{k+1}} - B_{t_k})^2 \right] \\ &= \frac{1}{2}B_t^2 + \frac{1}{2}t. \end{aligned}$$

El signo del segundo término cambió de negativo a positivo. Esto muestra que, a diferencia de la integral de Riemann, la integral estocástica es sensible al punto donde se evalúa el integrando. Al considerar el promedio de las dos evaluaciones en los extremos se obtiene la así llamada *integral de Stratonovich*, denotada de la forma siguiente

$$\int_0^t B_s \circ dB_s = \frac{1}{2}B_t^2.$$

Observe que el término adicional de la integral de Itô ha desaparecido. La integral de Stratonovich tiene algunas ventajas operacionales pues sigue algunas reglas usuales del cálculo integral, pero vista como proceso deja de ser una martingala. ◦

■ Ejemplo. calcularemos ahora la esperanza y varianza del proceso

$$\int_0^t B_s dB_s.$$

La esperanza es cero pues la integral estocástica en este caso es una martingala.

Para la varianza se tiene que

$$\begin{aligned}
 \text{Var}\left(\int_0^t B_s dB_s\right) &= E\left(\int_0^t B_s dB_s\right)^2 \\
 &= E\left(\int_0^t B_s^2 ds\right) \\
 &= \int_0^t E(B_s^2) ds \\
 &= \int_0^t s ds \\
 &= \frac{1}{2}t^2.
 \end{aligned}$$

Alternativamente, como $\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$, las cantidades anteriores pueden calcularse usando el lado derecho de esta igualdad. Claramente $E(\frac{1}{2}B_t^2 - \frac{1}{2}t) = 0$. Además

$$\begin{aligned}
 \text{Var}\left(\frac{1}{2}B_t^2 - \frac{1}{2}t\right) &= \frac{1}{4}\text{Var}(B_t^2) \\
 &= \frac{1}{4}(E(B_t^4) - E^2(B_t^2)) \\
 &= \frac{1}{4}(3t^2 - t^2) \\
 &= \frac{1}{2}t^2.
 \end{aligned}$$

◦

■ Ejemplo. Sean X_t y Y_t dos procesos en \mathcal{H}^2 . Entonces

$$E\left(\int_0^t X_s dB_s \int_0^t Y_s dB_s\right) = \int_0^t E(X_s Y_s) ds.$$

Esta fórmula es fácil de comprobar usando la isometría de Itô y la igualdad $ab =$

$\frac{1}{2}(a+b)^2 - \frac{1}{2}(a^2 + b^2)$. En efecto,

$$\begin{aligned} E\left(\int_0^t X_s dB_s \int_0^t Y_s dB_s\right) &= E\left(\frac{1}{2}\left|\int_0^t (X_s + Y_s) dB_s\right|^2\right) - \frac{1}{2}\left(E\left|\int_0^t X_s dB_s\right|^2 + E\left|\int_0^t Y_s dB_s\right|^2\right) \\ &= \frac{1}{2}\int_0^t E|X_s + Y_s|^2 ds - \frac{1}{2}\left(\int_0^t E|X_s|^2 ds + \int_0^t E|Y_s|^2 ds\right) \\ &= \int_0^t E(X_s Y_s) ds. \end{aligned}$$

◦

Propiedades de la integral. La integral estocástica de Itô cumple varias propiedades aunque sólo mencionaremos algunas de ellas aquí, a manera de resumen de las características señaladas antes.

- a) La integral $I_t : \mathcal{L}_{loc}^2 \rightarrow L^2(P)$ es *lineal*, es decir, para c constante y para cualesquiera procesos X_s y Y_s en \mathcal{L}_{loc}^2 , se cumple que

$$\int_0^t (cX_s + Y_s) dB_s = c \int_0^t X_s dB_s + \int_0^t Y_s dB_s, \quad c.s.$$

- b) Tiene esperanza es cero, es decir, para cualquier proceso X_s en \mathcal{L}_{loc}^2 ,

$$E\left(\int_0^t X_s dB_s\right) = 0, \quad c.s.$$

- c) Cuando la integral se restringe al espacio \mathcal{H}^2 , se cumple la *isometría de Itô*, es decir,

$$E\left|\int_0^t X_s dB_s\right|^2 = E\int_0^t |X_s|^2 ds.$$

- d) Nuevamente restringida al espacio \mathcal{H}^2 , la integral es una martingala, es decir, es integrable, adaptada y para $0 \leq s \leq t$, se cumple

$$E\left(\int_0^t X_u dB_u \mid \mathcal{F}_s\right) = \int_0^s X_u dB_u.$$

En general, para procesos en \mathcal{L}_{loc}^2 , la integral ya no es una martingala sino una martingala local.

- e) Existe una versión continua de la integral estocástica.

9.2. Fórmula de Itô

Usualmente una integral de Riemann no se calcula a partir la definición, en lugar de ello existen fórmulas bien conocidas para calcular integrales. La misma situación se presenta para integrales estocásticas: en muy raros casos se calculan éstas a través de su definición. La famosa fórmula de Itô es la herramienta fundamental para este tipo de integrales. Se enuncia a continuación este resultado en una versión simple y se ejemplifica su uso. Más adelante se presenta una versión un poco más general.

En lo sucesivo haremos referencia a las siguientes espacios de funciones. Una función real de variable real es de *clase* C^1 , cuando es diferenciable y su derivada es continua. Análogamente, una función es de *clase* C^2 , si es dos veces diferenciable y su segunda derivada es una función continua.

Teorema. (Fórmula de Itô I). Sea $f(x)$ es un función de clase C^2 . Entonces

$$f(B_t) - f(B_0) = \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds.$$

Explicaremos una forma de obtener este resultado usando el teorema de Taylor pero sin dar una justificación rigurosa de la prueba. Para una función $f(x)$ suficientemente suave, se tiene que

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + R(x),$$

en donde el residuo $R(x)$ puede escribirse como sigue

$$\begin{aligned} R(x) &= \int_{x_0}^x f''(t)(x - t) dt \\ &= \int_0^1 (1 - \theta) f''(x_0 + \theta(x - x_0))(x - x_0)^2 d\theta. \end{aligned}$$

La segunda igualdad se obtiene después de un evidente cambio de variable. Por lo

tanto, si $0 = t_0 < t_1 < \dots < t_n = t$ es una partición de $[0, t]$, entonces

$$\begin{aligned} f(B_t) - f(B_0) &= \sum_{k=1}^n [f(B_{t_k}) - f(B_{t_{k-1}})] \\ &= \sum_{k=1}^n f'(B_{t_{k-1}}) \Delta B_k \\ &\quad + \int_0^1 (1 - \theta) \sum_{k=1}^n f''(B_{t_{k-1}} + \theta \Delta B_k) (\Delta B_k)^2 d\theta. \end{aligned}$$

Puede comprobarse que al tomar el límite cuando $n \rightarrow \infty$ las sumas convergen casi seguramente y entonces se obtiene la igualdad

$$\begin{aligned} f(B_t) - f(B_0) &= \int_0^t f'(B_s) dB_s + \int_0^1 (1 - \theta) \int_0^t f''(B_s) ds d\theta \\ &= \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds. \end{aligned}$$

Esta fórmula es una versión estocástica de la regla de la cadena del cálculo diferencial usual, y es común escribirla en su forma diferencial del siguiente modo

$$df(B_t) = f'(B_t) dB_t + \frac{1}{2} f''(B_t) dt.$$

Esta expresión debe entenderse en el sentido de su forma integral arriba enunciado. Ilustraremos su uso mediante algunos ejemplos.

■ Ejemplo. Sea $f(x) = \frac{1}{2}x^2$. Entonces la fórmula de Itô establece que

$$\frac{1}{2}B_t^2 - \frac{1}{2}B_0^2 = \int_0^t B_s dB_s + \frac{1}{2} \int_0^t 1 ds.$$

Es decir,

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t.$$

Este resultado había sido encontrado antes, ahora lo hemos obtenido de manera inmediata de la fórmula de Itô. De manera análoga, para la función $f(x) = \frac{1}{3}x^3$ se obtiene

$$\int_0^t B_s^2 dB_s = \frac{1}{3}B_t^3 - \int_0^t B_s ds.$$

Más generalmente, para $f(x) = \frac{1}{n+1}x^{n+1}$ se obtiene

$$\int_0^t B_s^n dB_s = \frac{1}{n+1} B_t^{n+1} - \frac{1}{2} \int_0^t n B_s^{n-1} ds.$$

◦

■ Ejemplo. Usaremos la fórmula de Itô para encontrar una expresión de los momentos pares de una distribución normal centrada. Demostraremos que

$$E(B_t^{2n}) = \frac{(2n)!}{2^n n!} t^n.$$

Los momentos impares de dicha distribución se anulan pues en tal caso el integrando resulta ser una función impar. Consideremos entonces la función $f(x) = \frac{1}{2n}x^{2n}$, para cualquier entero natural n . De la fórmula de Itô se sigue que

$$\frac{1}{2n} B_t^{2n} - \frac{1}{2n} B_0^{2n} = \int_0^t B_s^{2n-1} dB_s + \frac{1}{2} \int_0^t (2n-1) B_s^{2n-2} ds.$$

Tomando esperanza y resolviendo de manera iterada se obtiene

$$\begin{aligned} E(B_t^{2n}) &= \frac{2n(2n-1)}{2} \int_0^t E(B_s^{2n-2}) ds \\ &= \frac{2n(2n-1)}{2} \frac{(2n-2)(2n-3)}{2} \int_0^t \int_0^{t_1} E(B_s^{2n-4}) ds dt_1 \\ &\vdots \\ &= \frac{(2n)!}{2^n} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} 1 ds dt_{n-1} \cdots dt_1. \end{aligned}$$

No es difícil verificar que los resultados sucesivos de estas integrales son: t_{n-1} , $t_{n-2}^2/2!$, $t_{n-3}^3/3!$, ..., $t^n/n!$. De esta forma se obtiene la fórmula enunciada. ◦

9.3. Ecuaciones diferenciales estocásticas

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad, y sea $\{B_t\}$ un movimiento Browniano unidimensional adaptado a la filtración $\{\mathcal{F}_t\}$.

Definición. Sean $b(t, x)$ y $\sigma(t, x)$ dos funciones de $[0, T] \times \mathbb{R}$ en \mathbb{R} . Una *ecuación estocástica* es una ecuación de la forma

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t, \quad (9.8)$$

definida para valores de t en el intervalo $[0, T]$, y con condición inicial la variable aleatoria X_0 que se presupone \mathcal{F}_0 -medible e independiente del movimiento Browniano. La ecuación (9.8) se interpreta como la ecuación integral

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s, \quad (9.9)$$

en donde la primera es una integral de Riemann, mientras que la segunda es una integral estocástica de Itô. Al proceso X_t se le llama *proceso de Itô*.

Los elementos conocidos de esta ecuación son los coeficientes $b(t, x)$ y $\sigma(t, x)$, y la variable aleatoria inicial X_0 . La incógnita es el proceso X_t . A la función $b(t, x)$ se le conoce como *coeficiente de tendencia* (*drift* en inglés o también *deriva* en español).

A la función $\sigma(t, x)$ se le llama *coeficiente de difusión*. El proceso solución puede interpretarse como el estado de un sistema que evoluciona de manera determinista gobernado por la parte no aleatoria de la ecuación (la tendencia), pero alterado por un ruido aditivo dado por la integral estocástica (la difusión).

Para que una ecuación estocástica tenga solución se deben pedir condiciones en los coeficientes. De manera análoga al caso de ecuaciones diferenciales deterministas, existen teoremas básicos de existencia y unicidad para ecuaciones estocásticas que establecen condiciones de regularidad para los coeficientes $b(t, x)$ y $\sigma(t, x)$, bajo las cuales la ecuación (9.8) tiene solución única. El siguiente es uno de tales resultados.

Teorema de existencia y unicidad. Si los coeficientes $b(t, x)$ y $\sigma(t, x)$ de la ecuación (9.8) satisfacen la condición de Lipschitz en la variable x ,

$$|b(t, x) - b(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leq K|x - y|^2,$$

y la condición de crecimiento en x ,

$$|b(t, x)|^2 + |\sigma(t, x)|^2 \leq K(1 + |x|^2),$$

para alguna constante $K > 0$, entonces existe un proceso estocástico $\{X_t\}$ solución de (9.8) que es adaptado a la filtración, tiene trayectorias continuas, es uniformemente acotado en $L^2(P)$, es decir, $\sup_{0 \leq t \leq T} E(X_t^2) < \infty$, y además es único en el sentido de indistinguibilidad.

En este caso a tal solución se le llama *solución fuerte* de la ecuación estocástica. No presentaremos la demostración de este resultado, simplemente comentaremos algunos aspectos de los que consta la prueba completa. La demostración es semejante al caso determinista, y hace uso del método de iteraciones de Picard. Mediante este método se define la sucesión de procesos

$$\begin{aligned} X_t^{(0)} &= X_0, \\ X_t^{(n+1)} &= X_0 + \int_0^t b(s, X_s^{(n)}) ds + \int_0^t \sigma(s, X_s^{(n)}) dB_s. \end{aligned}$$

Para que estas iteraciones tengan sentido es necesario verificar que los integrandos involucrados son efectivamente susceptibles de ser integrados respecto de la diferencial respectiva. Para comprobar que tal sucesión de procesos es convergente se demuestra que, con probabilidad uno, esta sucesión constituye una sucesión de Cauchy en el espacio de funciones continuas $C[0, T]$, respecto de la norma uniforme $\|X\| = \sup_{0 \leq t \leq T} |X_t|$. Dado lo anterior, existe entonces un proceso continuo X , tal que con probabilidad uno, $X_t^{(n)}$ converge a X_t de manera uniforme en el intervalo $[0, T]$. Adicionalmente puede demostrarse que el proceso límite es L^2 -acotado en $[0, T]$, y que la convergencia $X_t^{(n)} \rightarrow X_t$ también es válida en $L^2(P)$. También debe demostrarse que el proceso límite es efectivamente solución de la ecuación estocástica. Para ello se toma el límite en la ecuación que define las iteraciones, y se verifica la convergencia uniforme en $[0, T]$ con probabilidad uno, término a término. Los detalles de esta demostración pueden encontrarse por ejemplo en [35].

Observe que el teorema anterior no establece la forma de encontrar la solución a

una ecuación estocástica dada, sino que asegura únicamente la existencia de dicha solución. La siguiente versión de la fórmula de Itô es un resultado bastante útil para resolver algunas ecuaciones estocásticas y generaliza la versión anteriormente enunciada.

Teorema.(Fórmula de Itô II). Si $\{X_t\}$ es un proceso de Itô dado por (9.8) y $f(t, x)$ es un función de clase C^1 en t y de clase C^2 en x , entonces el proceso $Y_t = f(t, X_t)$ es también un proceso de Itô y satisface la ecuación estocástica

$$dY_t = f_t(t, X_t)dt + f_x(t, X_t)dX_t + \frac{1}{2}f_{xx}(t, X_t)(dX_t)^2. \quad (9.10)$$

Los subíndices indican derivada y (9.8) se substituye en (9.10) usando la siguiente tabla de multiplicación de McKean que se muestra en la Figura 9.4. Observe que como las derivadas involucradas son funciones continuas, las integrales estocásticas resultantes están bien definidas. La demostración de este resultado sigue las mismas líneas que la versión más simple. Ilustraremos a continuación el uso de esta fórmula con varios ejemplos.

\times	dt	dB_t
dt	0	0
dB_t	0	dt

Figura 9.4:

■ Ejemplo. Demostraremos que $\int_0^t s dB_s = tB_t - \int_0^t B_s ds$.

Para verificar esta fórmula puede tomarse el proceso $X_t = B_t$ y la función $f(t, x) = tx$. Entonces

$$\begin{aligned} d(f(t, B_t)) &= f_t(t, B_t) dt + f_x(t, B_t) dB_t + \frac{1}{2}f_{xx}(t, B_t) (dB_t)^2 \\ d(tB_t) &= B_t dt + t dB_t. \end{aligned}$$

Esta es la forma diferencial de la fórmula enunciada. ◦

■ Ejemplo. Considere la función $f(x) = e^x$. Por la fórmula de Itô,

$$e^{B_t} - e^{B_0} = \int_0^t e^{B_s} dB_s + \frac{1}{2} \int_0^t e^{B_s} ds,$$

es decir, el proceso $X_t = e^{B_t}$ satisface la ecuación diferencial

$$dX_t = X_t dB_t + \frac{1}{2}X_t dt,$$

con condición inicial $X_0 = 1$. A este proceso se le llama *movimiento Browniano geométrico*. \circ

■ Ejemplo. Demostraremos que el proceso $X_t = B_t/(1+t)$ es solución de la ecuación estocástica

$$dX_t = -\frac{X_t}{1+t} dt + \frac{1}{1+t} dB_t,$$

con condición inicial $X_0 = 0$. Sea $f(t, x) = x/(1+t)$. El proceso $X_t = f(t, B_t)$ cumple la condición inicial y por la fórmula de Itô satisface la ecuación

$$\begin{aligned} dX_t &= f_t(t, B_t)dt + f_x(t, B_t)dB_t + \frac{1}{2}f_{xx}(t, B_t)dt \\ &= -\frac{B_t}{(1+t)^2}dt + \frac{1}{1+t}dB_t \\ &= -\frac{X_t}{1+t}dt + \frac{1}{1+t}dB_t. \end{aligned}$$

\circ

■ Ejemplo. Usando el método de igualación de coeficientes resolveremos la ecuación

$$dX_t = -X_t dt + e^{-t} dB_t,$$

con condición inicial $X_0 = 0$. Se busca un función $f(t, x)$ tal que el proceso solución pueda escribirse como $X_t = f(t, B_t)$. Igualando los coeficientes de esta ecuación con los de la fórmula de Itô

$$dX_t = f_t(t, B_t)dt + f_x(t, B_t)dB_t + \frac{1}{2}f_{xx}(t, B_t)dt,$$

se obtiene el sistema de ecuaciones

$$\begin{aligned} f_x(t, x) &= e^{-t} \\ f_t(t, x) + \frac{1}{2}f_{xx}(t, x) &= -f(t, x). \end{aligned}$$

De la primera ecuación se obtiene $f(t, x) = e^{-t}x + c(t)$. Sustituyendo en la segunda ecuación y simplificando se obtiene $c'(t) = -c(t)$, cuya solución es $c(t) = ce^{-t}$, en donde c es una constante. Por lo tanto $f(t, x) = e^{-t}(x + c)$. Para que el proceso $X_t = f(t, B_t) = e^{-t}(B_t + c)$ cumpla la condición inicial $X_0 = 0$ forzosamente la constante c debe ser cero. De esta forma la función buscada es $f(t, x) = e^{-t}x$. En tal caso la fórmula de Itô asegura que efectivamente

$$\begin{aligned} dX_t &= -e^{-t}B_t dt + e^{-t} dB_t \\ &= -X_t dt + e^{-t} dB_t. \end{aligned}$$

◦

9.4. Simulación

Una ecuación estocástica ha resultado muy útil para modelar sistemas que presentan algún tipo de ruido o perturbación aleatoria. Aplicaciones de tales modelos se estudian en ingeniería, finanzas y física entre muchas otras áreas del conocimiento. Debido a la imposibilidad de encontrar soluciones explícitas a ciertas ecuaciones de interés, los métodos numéricos del caso determinista se han extendido al caso estocástico. En las siguientes secciones presentaremos algunos modelos particulares de ecuaciones estocásticas, y explicaremos un mecanismo para simular las trayectorias del proceso solución.

Una trayectoria de un proceso X_t que sigue la ley de movimiento de una ecuación estocástica de la forma

$$\begin{aligned} dX_t &= b(t, X_t) dt + \sigma(t, X_t) dB_t, \\ X_0 &= x_0, \end{aligned}$$

puede obtenerse mediante el *método de discretización de Euler-Maruyama*. En este procedimiento se divide el intervalo $[0, t]$ de manera uniforme en n subintervalos de idéntica longitud $\Delta t = t/n$, y se define $t_j = j\Delta t$ para $j = 0, 1, 2, \dots, N$. Suponiendo que Y_j es un valor al azar de la distribución normal estándar, se definen los valores sucesivos de la trayectoria solución de la ecuación estocástica como sigue

$$\begin{aligned} X_0 &= x_0, \\ X_{t_{j+1}} &= X_{t_j} + b(t_j, X_{t_j})\Delta t + \sigma(t_j, X_{t_j}) \sqrt{\Delta t} Y_j. \end{aligned}$$

Más adelante se presentará una implementación de este procedimiento en MATLAB para simular trayectorias de soluciones de ecuaciones particulares.

9.5. Algunos modelos particulares

Movimiento Browniano geométrico. Este modelo es de amplio uso en finanzas y sirve para representar el precio de algunos bienes que fluctúan siguiendo los vaivenes de los mercados financieros

Definición. Sean μ y $\sigma > 0$ dos constantes, y $x_0 > 0$. El *movimiento Browniano geométrico* es el proceso solución de la ecuación estocástica

$$\begin{aligned} dX_t &= \mu X_t dt + \sigma X_t dB_t, \\ X_0 &= x_0. \end{aligned} \tag{9.11}$$

Esta ecuación puede interpretarse de la siguiente forma. En ausencia del término estocástico, la ecuación se reduce a $dX_t = \mu X_t dt$, cuya solución es $X_t = x_0 e^{\mu t}$. Esta función representa el comportamiento en el tiempo de un capital inicial positivo x_0 que crece de manera continua y determinista a una tasa efectiva del $100\mu\%$, suponiendo $\mu > 0$. Por otro lado la parte estocástica corresponde a la volatilidad de una inversión con riesgo sujeta a las fluctuaciones de los mercados financieros. El modelo supone que dicha variabilidad es proporcional al valor de la inversión. Observe que los coeficientes de esta ecuación satisfacen las condiciones para la existencia y unicidad de la solución. Resolveremos esta ecuación usando el método de igualación de coeficientes.

Proposición. La solución a la ecuación (9.11) es

$$X_t = x_0 \exp \left[\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right]. \tag{9.12}$$

Demostración. Encontraremos una función $f(t, x)$ tal que al aplicar la fórmula de

Itô al proceso $X_t = f(t, B_t)$ se obtenga la ecuación (9.11). Comparando entonces los coeficientes de la fórmula general

$$dX_t = f_t(t, X_t) dt + f_x(t, X_t) dB_t + \frac{1}{2} f_{xx}(t, X_t) dt$$

con los de (9.11), se obtienen las igualdades

$$\begin{aligned} \mu f(t, x) &= f_t(t, x) + \frac{1}{2} f_{xx}(t, x), \\ \sigma f(t, x) &= f_x(t, x). \end{aligned}$$

De la segunda ecuación se obtiene que $f(t, x) = \exp[\sigma x + g(t)]$, para alguna función $g(t)$. Substituyendo en la primera ecuación se obtiene $g'(t) = \mu - \frac{1}{2}\sigma^2$, cuya solución es $g(t) = (\mu - \frac{1}{2}\sigma^2)t$. De donde $X_t = f(t, B_t)$ adquiere la expresión indicada. \square

A este proceso se le llama también se le conoce con el nombre de *movimiento Browniano exponencial*. En la Figura 9.5 puede apreciarse una trayectoria de este proceso con una inversión inicial x_0 de una unidad monetaria, y con parámetros $\mu = 1$, y $\sigma = 1/3$. La curva creciente corresponde al crecimiento determinista de la inversión cuando no hay aleatoriedad, es decir, cuando el coeficiente de difusión es cero. El valor de σ en la simulación es pequeño y por esa razón la trayectoria aleatoria mostrada se mantiene cerca de la trayectoria determinista.

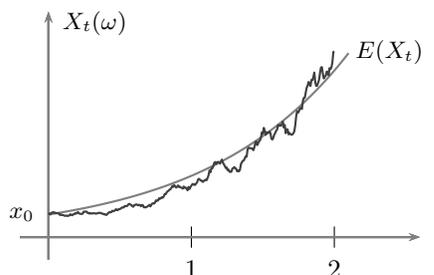


Figura 9.5:

Cuando se incrementa el valor de σ las trayectorias pueden diferir considerablemente.

En el programa de computadora de la Figura 9.6 se muestra una manera de simular trayectorias de este proceso. El código es una traducción a MATLAB de la discretización de la ecuación estocástica, y es una adaptación del código que aparece en [15]. Este programa puede ser encontrado en la página web de Desmond J. Higham, junto con la implementación de otros modelos y otras técnicas de discretización. La función `randn` produce un valor al azar de la distribución normal estándar.

Demostraremos ahora algunas características numéricas de este proceso.

```

randn('state',100)
T=2; N=300; dt=T/N; xcero=1; mu=1; sigma=1/3;
dW=zeros(1,N); MBG=zeros(1,N);
dW(1)=sqrt(dt)*randn;
MBG(1)=xcero+mu*dt+sigma*xcero*dW(1);
for j=2:N
    dW(j)=sqrt(dt)*randn
    MBG(j)=MBG(j-1)+mu*MBG(j-1)*dt+sigma*MBG(j-1)*dW(j)
end
plot([0:dt:T],[xcero,MBG],'r-')

```

Figura 9.6:

Proposición. Para el movimiento Browniano geométrico se cumple lo siguiente.

1. $E(X_t) = x_0 e^{\mu t}$.
2. $\text{Var}(X_t) = x_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1)$.
3. $\text{Cov}(X_t, X_s) = x_0^2 e^{\mu(s+t)} (e^{\sigma^2 s} - 1)$, para $0 \leq s \leq t$.

Demostración. Usaremos el hecho de que la función generadora de momentos de la distribución $N(\mu, \sigma^2)$ es $M(s) = \exp(\mu s + \frac{1}{2}\sigma^2 s^2)$.

1. Para la esperanza se tiene que

$$\begin{aligned}
 E(X_t) &= E(x_0 \exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t]) \\
 &= x_0 \exp[(\mu - \frac{1}{2}\sigma^2)t] E(\exp[\sigma B_t]) \\
 &= x_0 \exp[(\mu - \frac{1}{2}\sigma^2)t] \exp[\frac{1}{2}t\sigma^2] \\
 &= x_0 e^{\mu t}.
 \end{aligned}$$

2. Ahora calcularemos la varianza.

$$\begin{aligned}
 \text{Var}(X_t) &= \text{Var}(x_0 \exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t]) \\
 &= x_0^2 \exp[2(\mu - \frac{1}{2}\sigma^2)t] \text{Var}(\exp[\sigma B_t]) \\
 &= x_0^2 \exp[2(\mu - \frac{1}{2}\sigma^2)t] (E(\exp[2\sigma B_t]) - E^2(\exp[\sigma B_t])) \\
 &= x_0^2 \exp[2(\mu - \frac{1}{2}\sigma^2)t] (\exp[\frac{1}{2}t(2\sigma)^2] - \exp[2(\frac{1}{2}t\sigma^2)]) \\
 &= x_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1).
 \end{aligned}$$

3. Calcularemos primero $E(X_t X_s)$. Observe que $B_t + B_s$ se puede escribir como $2B_s + (B_t - B_s)$, siendo estos sumandos independientes. Entonces

$$\begin{aligned}
 E(X_t X_s) &= E(x_0^2 \exp[(\mu - \frac{1}{2}\sigma^2)(t+s) + \sigma(B_t + B_s)]) \\
 &= x_0^2 \exp[(\mu - \frac{1}{2}\sigma^2)(t+s)] E(\exp[\sigma(B_t + B_s)]) \\
 &= x_0^2 \exp[(\mu - \frac{1}{2}\sigma^2)(t+s)] E(e^{2\sigma B_s}) E(e^{\sigma(B_t - B_s)}) \\
 &= x_0^2 \exp[(\mu - \frac{1}{2}\sigma^2)(t+s)] e^{2s\sigma^2} e^{\frac{1}{2}(t-s)\sigma^2} \\
 &= x_0^2 \exp[\mu(t+s) + s\sigma^2].
 \end{aligned}$$

Por lo tanto,

$$\begin{aligned}
 \text{Cov}(X_t, X_s) &= E(X_t X_s) - E(X_t)E(X_s) \\
 &= x_0^2 e^{\mu(t+s) + s\sigma^2} - x_0^2 e^{\mu(t+s)} \\
 &= x_0^2 e^{\mu(t+s)} (e^{s\sigma^2} - 1)
 \end{aligned}$$

□

Proceso de Ornstein-Uhlenbeck. Este modelo fue propuesto por Ornstein y Uhlenbeck para modelar la velocidad del movimiento difuso de una partícula en intervalos de tiempo pequeños.

Definición. Sean α y σ dos constantes positivas. El *proceso de Ornstein-Uhlenbeck* es aquel proceso solución de la ecuación estocástica

$$\begin{aligned} dX_t &= -\alpha X_t dt + \sigma dB_t, \\ X_0 &= x_0. \end{aligned} \quad (9.13)$$

La variable X_t se interpreta como la velocidad de la partícula al tiempo t . La parte determinista $-\alpha X_t$ corresponde a la fuerza de fricción, y el sumando σdB_t es un ruido aleatorio. Encontraremos la solución de esta ecuación.

Proposición. La solución a la ecuación (9.13) está dada por

$$X_t = x_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dB_s. \quad (9.14)$$

Demostración. Considere una solución de la forma

$$X_t = a(t) \left[x_0 + \int_0^t b(s) dB_s \right], \quad (9.15)$$

en donde $a(t)$ y $b(t)$ son funciones diferenciables. Derivando (9.15) y usando la fórmula de Itô (9.10) se obtiene

$$\begin{aligned} dX_t &= a'(t) \left[x_0 + \int_0^t b(s) dB_s \right] dt + a(t) b(t) dB_t \\ &= \frac{a'(t)}{a(t)} X_t dt + a(t) b(t) dB_t. \end{aligned}$$

Comparando con (9.13), las funciones $a(t)$ y $b(t)$ deben cumplir las ecuaciones

$$\frac{a'(t)}{a(t)} = -\alpha, \quad a(t) b(t) = \sigma.$$

Suponiendo $a(0) = 1$ se obtiene $a(t) = \exp(-\alpha t)$, y $b(t) = \sigma \exp(\alpha t)$. Substituyendo en (9.15) se obtiene (9.14). \square

En la Figura 9.7 se muestra la simulación de una trayectoria de este proceso y se compara con $E(X_t)$. El proceso muestra un decaimiento conforme el tiempo avanza y ello es debido al factor de fricción del modelo. Calcularemos a continuación la esperanza y varianza de este proceso.

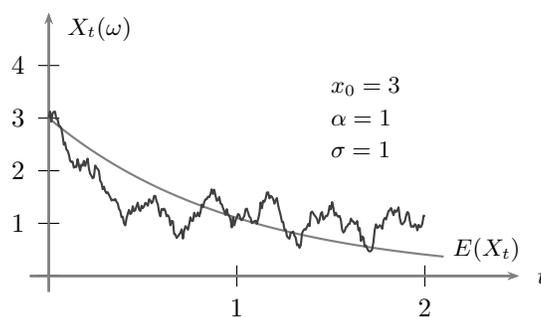


Figura 9.7:

Proposición. Para el proceso de Ornstein-Uhlenbeck se cumple lo siguiente.

1. $E(X_t) = x_0 e^{-\alpha t}$.
2. $\text{Var}(X_t) = \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha t})$.
3. $\text{Cov}(X_t, X_s) = \frac{\sigma^2}{2\alpha} (e^{-\alpha(t-s)} - e^{-\alpha(t+s)})$.

Demostración.

1. Este resultado se obtiene al tomar esperanza en (9.14), y observar que la integral estocástica es una martingala que inicia en cero.

2. El cálculo de la varianza de X_t es una aplicación de la isometría de Itô,

$$\begin{aligned}
 \text{Var}(X_t) &= \sigma^2 \text{Var}\left(\int_0^t e^{-\alpha(t-s)} dB_s\right) \\
 &= \sigma^2 E\left(\int_0^t e^{-\alpha(t-s)} dB_s\right)^2 \\
 &= \sigma^2 \left(\int_0^t e^{-2\alpha(t-s)} ds\right) \\
 &= \sigma^2 e^{-2\alpha t} \left[\frac{1}{2\alpha} e^{2\alpha s}\right]_0^t \\
 &= \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha t}).
 \end{aligned}$$

2. Nuevamente usaremos la isometría de Itô. Para $0 \leq s \leq t$,

$$\begin{aligned}
 \text{Cov}(X_t, X_s) &= E(X_t X_s) - E(X_t)E(X_s) \\
 &= E\left[\left(x_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-u)} dB_u\right)\right. \\
 &\quad \left.(x_0 e^{-\alpha s} + \sigma \int_0^s e^{-\alpha(s-u)} dB_u\right)] - x_0^2 e^{-\alpha(t+s)} \\
 &= \sigma^2 E\left(\int_0^t e^{-\alpha(t-u)} dB_u \int_0^s e^{-\alpha(s-u)} dB_u\right).
 \end{aligned}$$

La primera integral puede descomponerse en la suma de dos integrales, una sobre el intervalo $[0, s]$ y otra sobre $(s, t]$. Dada la propiedad de incrementos independientes del movimiento Browniano, el segundo sumando desaparece. De modo que, por la isometría de Itô,

$$\begin{aligned}
 \text{Cov}(X_t, X_s) &= \sigma^2 e^{-\alpha(t+s)} E\left(\int_0^s e^{\alpha u} dB_u\right)^2 \\
 &= \sigma^2 e^{-\alpha(t+s)} \int_0^s e^{2\alpha u} du \\
 &= \frac{\sigma^2}{2\alpha} (e^{-\alpha(t-s)} - e^{-\alpha(t+s)}).
 \end{aligned}$$

□

Puente Browniano.

El puente Browniano sobre el intervalo unitario $[0, 1]$ es un movimiento Browniano con espacio parametral dicho intervalo y es tal que en los extremos de este intervalo el proceso se hace cero. Una trayectoria de tal proceso se muestra en la Figura 9.8. Existen varias formas equivalentes de definir a este proceso.

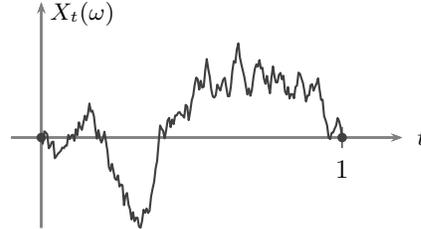


Figura 9.8:

Definición. El puente Browniano en el intervalo $[0, 1]$ es aquel proceso $\{X_t\}$ solución de la ecuación estocástica

$$\begin{aligned} dX_t &= -\frac{X_t}{1-t} dt + dB_t, & t \in [0, 1], & (9.16) \\ X_0 &= 0. \end{aligned}$$

Proposición. La solución a la ecuación (9.16) es

$$X_t = (1-t) \int_0^t \frac{1}{1-s} dB_s. \quad (9.17)$$

Demostración. Puede resolverse (9.16) proponiendo nuevamente una solución de la forma

$$X_t = a(t) \left[x_0 + \int_0^t b(s) dB_s \right], \quad (9.18)$$

en donde $a(t)$ y $b(t)$ son dos funciones diferenciables, y $x_0 = 0$. Derivando se obtiene

nuevamente

$$\begin{aligned} dX_t &= a'(t) \left[x_0 + \int_0^t b(s) dB_s \right] dt + a(t) b(t) dB_t \\ &= \frac{a'(t)}{a(t)} X_t dt + a(t) b(t) dB_t. \end{aligned}$$

Igualando coeficientes se obtienen las ecuaciones $a'(t)/a(t) = -1/(1-t)$, y $a(t)b(t) = 1$. Suponiendo $a(0) = 1$ se obtiene $a(t) = 1 - t$, y por lo tanto $b(t) = 1/(1 - t)$. Substituyendo en (9.18) se obtiene (9.17). \square

■ Ejercicio. Sea X_t un puente Browniano en $[0, 1]$. Demuestre que efectivamente $\lim_{t \rightarrow 1^-} X_t = 0$. \cdot

Vamos a calcular a continuación la esperanza, varianza y covarianza de este proceso.

Proposición. Para el puente Browniano (9.17) se cumple

1. $E(X_t) = 0$.
2. $\text{Var}(X_t) = t(1 - t)$.
3. $\text{Cov}(X_t, X_s) = s(1 - t)$, para $0 \leq s \leq t$.

Demostración.

1. La integral es una martingala continua que inicia en cero, por lo tanto $E(X_t) = 0$.

2. Por la isometría de Itô,

$$\begin{aligned}
 \text{Var}(X_t) &= (1-t)^2 \text{Var}\left(\int_0^t \frac{1}{1-s} dB_s\right) \\
 &= (1-t)^2 E\left(\int_0^t \frac{1}{1-s} dB_s\right)^2 \\
 &= (1-t)^2 E \int_0^t \left(\frac{1}{1-s}\right)^2 ds \\
 &= (1-t)^2 \left[\frac{1}{1-s}\right]_0^t \\
 &= t(1-t).
 \end{aligned}$$

3. Para $0 \leq s \leq t$,

$$\begin{aligned}
 \text{Cov}(X_s, X_t) &= E(X_s X_t) \\
 &= (1-s)(1-t) E\left(\int_0^s \frac{1}{1-u} dB_u \int_0^t \frac{1}{1-u} dB_u\right).
 \end{aligned}$$

Nuevamente la segunda integral puede descomponerse en la suma de dos integrales, una sobre el intervalo $[0, s]$ y otra sobre $(s, t]$. Dada la propiedad de incrementos independientes del movimiento Browniano, el segundo sumando desaparece. De modo que, por la isometría de Itô,

$$\begin{aligned}
 \text{Cov}(X_s, X_t) &= (1-s)(1-t) E\left(\int_0^s \frac{1}{1-u} dB_u\right)^2 \\
 &= (1-s)(1-t) \int_0^s \left(\frac{1}{1-u}\right)^2 du \\
 &= (1-s)(1-t) \left[\frac{1}{1-u}\right]_0^s \\
 &= s(1-t).
 \end{aligned}$$

□

Como era de esperarse, la varianza se anula en los extremos del intervalo pues allí el proceso es cero con probabilidad uno. Observe además que la varianza se hace máxima exactamente en la mitad de dicho intervalo. El puente Browniano en $[0, 1]$ puede representarse de varias formas como se muestra en el siguiente resultado.

■ Ejercicio. Demuestre que los siguientes procesos son puentes Brownianos.

a) $X_t = B_t - tB_1$, para $t \in [0, 1]$.

b) $X_t = B_{1-t} - (1-t)B_1$, para $t \in [0, 1]$. ■

Notas y referencias

Para el desarrollo de la definición de integral estocástica respecto del movimiento Browniano hemos seguido el lineamiento general presentado en el texto de Steele [35]. Allí pueden encontrarse las demostraciones completas de varios de los resultados que hemos solamente enunciado. Otros trabajos en donde pueden encontrarse exposiciones elementales sobre integración estocástica son Øksendal [26], Kuo [22], o Klebaner [20]. Para una exposición más completa y general puede consultarse por ejemplo Protter [27], o Revuz y Yor [29]. En los trabajos de D. J Higham como [15] pueden encontrarse algunas primeras lecturas sobre los métodos para simular ecuaciones estocásticas. En el texto de Kloeden y Platen [21] se expone la teoría general, varios métodos numéricos, y diversas aplicaciones de ecuaciones estocásticas.

Ejercicios

Integral de Itô

183. A partir de la definición de integral estocástica demuestre que

$$a) \int_0^t s dB_s = tB_t - \int_0^t B_s ds.$$

$$b) \int_0^t B_s^2 dB_s = \frac{1}{3}B_t^3 - \int_0^t B_s ds.$$

Para el segundo inciso use la identidad

$$x^2(y-x) + x(y-x)^2 = \frac{1}{3}(y^3 - x^3 - (y-x)^3).$$

Fórmula de Itô

184. Use la fórmula de Itô para demostrar que el proceso X_t es una solución de la ecuación estocástica indicada.

$$a) X_t = B_t^2, \quad dX_t = dt + 2B_t dB_t.$$

$$b) X_t = B_t^3, \quad dX_t = 3X_t^{1/3} dt + 3X_t^{2/3} dB_t.$$

$$c) X_t = tB_t, \quad dX_t = \frac{X_t}{t} dt + t dB_t.$$

185. Use la fórmula de Itô para demostrar que

$$\int_0^t B_s^2 ds = \frac{1}{3}B_t^3 - \int_0^t B_s ds.$$

APÉNDICE A

Conceptos y resultados varios

Igualdad de procesos. Se dice que dos procesos $\{X_t : t \geq 0\}$ y $\{Y_t : t \geq 0\}$ son *equivalentes*, o también se dice que uno es una *versión* o *modificación* del otro, si para cada valor de $t \geq 0$ fijo se cumple que $P(X_t = Y_t) = 1$, es decir, si las variables X_t y Y_t son iguales c.s. Un tipo de igualdad más fuerte establece que los procesos son *indistinguibles* si $P(X_t = Y_t \text{ para cada } t \geq 0) = 1$. Esto significa que con probabilidad uno las trayectorias de los dos procesos son idénticas. Claramente la indistinguibilidad es más fuerte que la equivalencia. Sin embargo, puede demostrarse que cuando los procesos son continuos, es decir, cuando sus trayectorias son funciones continuas del parámetro, ambas nociones de igualdad coinciden. Cuando el parámetro es discreto, nuevamente los dos tipos de igualdad son equivalentes.

Distribuciones finito dimensionales. Las distribuciones finito dimensionales de un proceso $\{X_t : t \geq 0\}$ es la colección de todas las funciones de distribución conjuntas $F_{X_{t_1}, \dots, X_{t_n}}(x_1, \dots, x_n)$, para cualesquiera tiempos $0 \leq t_1 < \dots < t_n$, y cualquier n natural.

Independencia de procesos. Se dice que una variable aleatoria X es *independiente* de un proceso $\{X_t : t \geq 0\}$ si para cualesquiera tiempos $0 \leq t_1 < t_2 < \dots < t_n$, $n \in \mathbb{N}$, la distribución conjunta de la variable X y el vector $(X_{t_1}, \dots, X_{t_n})$ es el producto de las distribuciones marginales, es decir,

$$F_{X, X_{t_1}, \dots, X_{t_n}}(x, x_1, \dots, x_n) = F_X(x) F_{X_{t_1}, \dots, X_{t_n}}(x_1, \dots, x_n),$$

o en términos de conjuntos de Borel A, A_1, \dots, A_n ,

$$P(X \in A, X_{t_1} \in A_1, \dots, X_{t_n} \in A_n) = P(X \in A) \\ P(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n).$$

Más generalmente, dos procesos estocásticos $\{X_t : t \geq 0\}$ y $\{Y_t : t \geq 0\}$ son *independientes* si para cualesquiera dos enteros naturales n y m , y tiempos $0 \leq t_1 < t_2 < \dots < t_n$ y $0 \leq s_1 < s_2 < \dots < s_m$, se cumple que

$$F_{X_{t_1}, \dots, X_{t_n}, Y_{s_1}, \dots, Y_{s_m}}(x_1, \dots, x_n, y_1, \dots, y_m) = F_{X_{t_1}, \dots, X_{t_n}}(x_1, \dots, x_n) \\ F_{Y_{s_1}, \dots, Y_{s_m}}(y_1, \dots, y_m).$$

En palabras, esta condición significa que las distribuciones finito dimensionales conjuntas son el producto de las distribuciones finito dimensionales marginales.

Lema de Abel

- a. Si la serie $\sum_{k=0}^{\infty} a_k$ es convergente, entonces $\lim_{t \nearrow 1} \sum_{k=0}^{\infty} a_k t^k = \sum_{k=0}^{\infty} a_k$.
 b. Inversamente, si $a_k \geq 0$ y $\lim_{t \nearrow 1} \sum_{k=0}^{\infty} a_k t^k \leq \infty$, entonces $\sum_{k=0}^{\infty} a_k = \lim_{t \nearrow 1} \sum_{k=0}^{\infty} a_k t^k$.

En Karlin y Taylor [19] puede encontrarse una demostración de estos resultados.

Lema de Fatou

- a. Sea $\{a_{nm} : n, m \in \mathbb{N}\}$ una colección de números reales. Entonces

$$\sum_m \liminf_{n \rightarrow \infty} a_{nm} \leq \liminf_{n \rightarrow \infty} \sum_m a_{nm}.$$

- b. Además, si $a_{nm} \leq b_m$ con $\sum_m b_m < \infty$, entonces

$$\limsup_{n \rightarrow \infty} \sum_m a_{nm} \leq \sum_m \limsup_{n \rightarrow \infty} a_{nm}.$$

Teorema de convergencia monótona. Sea X_1, X_2, \dots una sucesión de variables aleatorias tales que $0 \leq X_1 \leq X_2 \leq \dots$ y $\lim_{n \rightarrow \infty} X_n = X$ casi seguramente. Entonces

$$\lim_{n \rightarrow \infty} E(X_n) = E(\lim_{n \rightarrow \infty} X_n).$$

Teorema de convergencia dominada

a. Sea X_1, X_2, \dots una sucesión de variables aleatorias tales que $\lim_{n \rightarrow \infty} X_n = X$ casi seguramente, y para cada valor de n , $|X_n| \leq Y$, para alguna variable Y con $E|Y| < \infty$. Entonces

$$\lim_{n \rightarrow \infty} E(X_n) = E(\lim_{n \rightarrow \infty} X_n).$$

b. Sea $\{a_{nm} : n, m \in \mathbb{N}\}$ una colección de números reales tales que $\lim_{n \rightarrow \infty} a_{nm}$ existe para cada m , $|a_{nm}| \leq b_m$, independiente de n , y $\sum_{m=0}^{\infty} b_m < \infty$. Entonces

$$\lim_{n \rightarrow \infty} \sum_{m=0}^{\infty} a_{nm} = \sum_{m=0}^{\infty} \lim_{n \rightarrow \infty} a_{nm}.$$

Ecuación de Wald. Sea X_1, X_2, \dots una sucesión de variables aleatorias independientes con la misma distribución que X y con esperanza finita. Sea N una variable aleatoria con valores en el conjunto $\{1, 2, \dots\}$, con esperanza finita e independiente de la sucesión. Entonces

$$E\left(\sum_{k=1}^N X_k\right) = E(X) E(N).$$

Distribución tipo reticular. Se dice que una variable aleatoria discreta X , o su distribución de probabilidad, es de tipo *reticular* o que es de tipo *lattice* si $P(X = c + nd) > 0$, en donde c y $d > 0$ son constantes reales y $n = \pm 1, \pm 2, \dots$

Notación o pequeña. Sean $f(t)$ y $g(t)$ dos funciones que se encuentran definidas y son positivas para valores de t suficientemente grandes. Se dice que $f(t)$ es de *orden más pequeño* que $g(t)$ cuando $t \rightarrow \infty$, y se escribe $f(t) = o(g(t))$ cuando $t \rightarrow \infty$, si se cumple

$$\lim_{t \rightarrow \infty} \frac{f(t)}{g(t)} = 0.$$

En particular, se escribe $f(t) = o(1)$ cuando $f(t)$ converge a cero cuando $t \rightarrow \infty$. Se usa la misma notación cuando t tiende a algún valor finito particular y las funciones se encuentran definidas y son positivas en una vecindad no trivial alrededor de ese valor. En este texto se usa la notación *o* pequeña para establecer el comportamiento de probabilidades dependientes del tiempo $p(t)$, cuando t se aproxima a cero a través de valores positivos.

Fórmula de Stirling. $n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}$.

Esperanza condicional. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Sea X una variable aleatoria con esperanza finita, y sea \mathcal{G} una sub σ -álgebra de \mathcal{F} . La esperanza condicional de X dado \mathcal{G} , es una variable aleatoria denotada por $E(X | \mathcal{G})$, que cumple las siguientes tres propiedades:

- a) Es \mathcal{G} -medible.
- b) Tiene esperanza finita.
- c) Para cualquier evento G en \mathcal{G} ,

$$\int_G E(X | \mathcal{G}) dP = \int_G X dP. \quad (\text{A.1})$$

Es importante enfatizar que la esperanza condicional es una variable aleatoria. Usando el teorema de Radon-Nikodym (véase por ejemplo [9]), puede demostrarse que esta variable aleatoria existe y es única casi seguramente. Esto significa que si existe otra variable aleatoria con las tres propiedades anteriores, entonces con probabilidad uno coincide con $E(X | \mathcal{G})$. Cuando la σ -álgebra \mathcal{G} es generada por una variable aleatoria Y , es decir, cuando $\mathcal{G} = \sigma(Y)$, la esperanza condicional se escribe simplemente como $E(X | Y)$. Mencionaremos a continuación algunas propiedades de esta variable aleatoria, en estas expresiones se postula de manera implícita que la variable aleatoria a la que se le aplica la esperanza condicional es integrable.

1. Si c es constante, entonces $E(c | \mathcal{G}) = c$.
2. $E(X | \{\emptyset, \Omega\}) = E(X)$.
3. Si A es un evento, entonces $E(1_A | \{\emptyset, \Omega\}) = P(A)$.
4. Si A y B son eventos con $0 < P(B) < 1$, entonces

$$E(1_A | \{\emptyset, B, B^c, \Omega\}) = P(A | B) 1_B + P(A | B^c) 1_{B^c}.$$

5. Si A es un evento y B_1, \dots, B_n es una partición de Ω tal que $P(B_i) > 0$ para $i = 1, \dots, n$, entonces

$$E(1_A | \sigma\{B_1, \dots, B_n\}) = \sum_{i=1}^n P(A | B_i) 1_{B_i}.$$

6. Si Y es una variable aleatoria discreta con valores $0, 1, \dots$, entonces

$$E(X | Y) = \sum_{n=0}^{\infty} E(X | Y = n) 1_{(Y=n)}.$$

7. $E(E(X | \mathcal{G})) = E(X)$.
8. $|E(X | \mathcal{G})| \leq E(|X| | \mathcal{G})$. Este es un caso particular de la desigualdad de Jensen que se enunciará más adelante. En particular, tomando esperanza se tiene el siguiente resultado.
9. $E|E(X | \mathcal{G})| \leq E(|X|)$.
10. Si c es constante, entonces $E(cX + Y | \mathcal{G}) = cE(X | \mathcal{G}) + E(Y | \mathcal{G})$.
11. Si X es \mathcal{G} -medible, entonces $E(X | \mathcal{G}) = X$ c.s.
12. Si $X \geq 0$, entonces $E(X | \mathcal{G}) \geq 0$.
13. Si $X \leq Y$, entonces $E(X | \mathcal{G}) \leq E(Y | \mathcal{G})$.
14. Si $\mathcal{G}_1 \subseteq \mathcal{G}_2$, entonces

$$E(E(X | \mathcal{G}_1) | \mathcal{G}_2) = E(E(X | \mathcal{G}_2) | \mathcal{G}_1) = E(X | \mathcal{G}_1).$$

15. Si X es independiente de \mathcal{G} , entonces $E(X | \mathcal{G}) = E(X)$.
16. Si \mathcal{G}_1 y \mathcal{G}_2 son independientes, entonces
- $$E(X | \sigma(\mathcal{G}_1 \cup \mathcal{G}_2)) = E(X | \mathcal{G}_1) + E(X | \mathcal{G}_2) - E(X).$$

17. Si X es independiente de \mathcal{G}_2 , entonces

$$E(X | \sigma(\mathcal{G}_1 \cup \mathcal{G}_2)) = E(X | \mathcal{G}_1).$$

18. Si $X_n \xrightarrow{m} X$, entonces $E(X_n | \mathcal{G}) \xrightarrow{m} E(X | \mathcal{G})$.
19. *Teorema de convergencia monótona.* Si $X_n \geq 0$ y $X_n \nearrow X$ c.s., entonces $E(X_n | \mathcal{G}) \nearrow E(X | \mathcal{G})$ c.s.
20. Si XY es integrable y X es \mathcal{G} -medible, entonces $E(XY | \mathcal{G}) = X E(Y | \mathcal{G})$.
21. X es independiente de \mathcal{G} si, y sólo si, $E(f(X) | \mathcal{G}) = E(f(X))$ para cualquier función Lebesgue medible f tal que $f(X)$ es integrable.

22. *Desigualdad de Jensen.* Si u es convexa y $u(X)$ es integrable, entonces $u(E(X | \mathcal{G})) \leq E(u(X) | \mathcal{G})$.

Funciones directamente Riemann integrables. Sea $H : [0, \infty) \rightarrow [0, \infty)$ una función Borel medible. Sea $h > 0$, y para cada n natural defina las funciones

$$\begin{aligned}\alpha_n(h) &= \inf \{ H(t) : (n-1)h \leq t < nh \}, \\ \beta_n(h) &= \sup \{ H(t) : (n-1)h \leq t < nh \}.\end{aligned}$$

Suponga que las siguientes funciones son absolutamente convergentes

$$\begin{aligned}\alpha(h) &= h \sum_{n=1}^{\infty} \alpha_n(h), \\ \beta(h) &= h \sum_{n=1}^{\infty} \beta_n(h),\end{aligned}$$

y que además $\lim_{h \rightarrow 0} \alpha(h) = \lim_{h \rightarrow 0} \beta(h)$. Entonces se dice que la función H es *directamente Riemann integrable*. Esta condición de integrabilidad es más fuerte que la integrabilidad usual de Riemann, es decir, toda función directamente Riemann integrable es Riemann integrable. por ejemplo,

a) $H(t) = 1_{[0,a]}(t)$ es d. R. integrable.

b) $H : [0, \infty) \rightarrow [0, \infty)$ no creciente y tal que $\int_0^{\infty} H(t) dt < \infty$, es d. R. integrable.

Bibliografía

- [1] Basu A. K. *Introduction to stochastic processes*. Alpha Science International Ltd., 2003.
- [2] Bingham N. H. Doob: a half century on. *Journal of Applied Probability*, Vol. **42**, 257–266, 2005.
- [3] Brown R. *A brief account of Microscopical Observations made in the Months of June, July, and August, 1827, on the Particles contained in the Pollen of Plants; and on the general Existence of active Molecules in Organic and Inorganic Bodies*, Philosophical Magazine N. S. 4, 161-173, 1828.
- [4] Brzeźniak Z. y Zastawniak T. *Basic stochastic processes*. Springer, 1999.
- [5] Burkholder D. y Protter P. *Joseph Leo Doob, 1910-2004*. Stochastic Processes and their Applications, Vol. **115**, 1061-1072, 2005.
- [6] Caballero M. E. et al. *Cadenas de Markov: un enfoque elemental*. Aportaciones Matemáticas, Serie Textos, **29**. Sociedad Matemática Mexicana, 2004.
- [7] Chung K. L. *Markov chains with stationary transition probabilities (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen Band 104)*. Springer, 1960.
- [8] Chung K. L. y Williams R. J. *Introduction to stochastic integration*. Birkhäuser, 1983.
- [9] Cohn D. L. *Measure theory*. Birkhäuser, 1980.
- [10] Einstein A. *Investigations on the theory of the Brownian movement*. Dover, 1956.

- [11] Feller W. *Introducción a la teoría de probabilidades y sus aplicaciones*. Limusa, 1973.
- [12] Gard T. C. *Introduction to stochastic differential equations*. Marcel Dekker, 1988.
- [13] Grimmett G. R. y Stirzaker D. R. *Probability and random processes*. Clarendon Press, Oxford, 1982.
- [14] Gut A. *Probability: a graduate course*. Springer, 2005.
- [15] Higham D. J. *An algorithmic introduction to numerical simulation of stochastic differential equations*. SIAM Review Vol. **43**, No. 3, pp. 525-546, 2001.
- [16] Hoel P. G., Port S. C. y Stone C. J. *Introduction to stochastic processes*. Houghton Mifflin Company, 1972.
- [17] Jones P. W. y Smith P. *Stochastic processes: an introduction*. Arnold, 2001.
- [18] Jordan D. W. y Smith P. *Mathematical techniques*. Oxford University Press, 1997.
- [19] Karlin S. y Taylor H. M. *A first course in stochastic processes*. Academic Press, 1975.
- [20] Klebaner F. C. *Introduction to stochastic calculus with applications*. Imperial College Press, 1998.
- [21] Kloeden P. E. y Platen E. *Numerical solution of stochastic differential equations*. Springer-Verlag, 1999.
- [22] Kuo H. *Introduction to stochastic integration*. Springer, 2006.
- [23] Lawler G. F. *Introduction to stochastic processes*. Chapman & Hall / CRC, 2006.
- [24] Nelson E. *Dynamical theories of Brownian motion*. Princeton University Press, 1967.
- [25] Norris J. *Markov chains*. Cambridge University Press, 1997.
- [26] Øksendal B. *Stochastic differential equations: an introduction with applications*. Springer-Verlag, 1992.
- [27] Protter P. H. *Stochastic integration and differential equations*. Springer, 1990.

- [28] Resnick S. *Adventures in stochastic processes*. Birkhäuser, 1992.
- [29] Revuz D. y Yor M. *Continuous martingales and Brownian motion*. Springer-Verlag, 1991.
- [30] Rincón L. *Sobre el problema del mono que escribe caracteres al azar*. Miscelánea Matemática **42**, Sociedad Matemática Mexicana, 79–90, 2006.
- [31] Ross S. *A first course in probability - 4th ed.* Macmillan Publishing Company, 1994.
- [32] Snell J. L. *A Conversation with Joe Doob*. Statistical Science, Vol. **12**, No. 4, 301–311, 1997.
- [33] Snell J. L. *Obituary: Joseph Leonard Doob*. Journal of Applied Probability, Vol. **42**, 247–256, 2005.
- [34] Spitzer F. *Principles of random walk. 2nd. ed.* Springer, 2001.
- [35] Steele J. M. *Stochastic calculus and financial applications*. Springer-Verlag, 2001.
- [36] Stirzaker D. *Stochastic processes and models*. Oxford University Press, 2005.
- [37] Taylor H. M. y Karlin S. *An introduction to stochastic modeling*. Academic Press, 1994.
- [38] Tudor C. *Procesos estocásticos*. Aportaciones Matemáticas, Serie Textos 2, Sociedad Matemática Mexicana, 1994.

Índice

- Abel
 - lema, 272
- Cadena de Markov, 23
 - a tiempo continuo, 126
 - accesibilidad de edos., 38
 - comunicación de edos., 38
 - de dos estados, 27
 - de Ehrenfest, 31
 - de inventarios, 33
 - de la caminata aleatoria, 30
 - de la fila de espera, 33
 - de rachas de éxitos, 29
 - de ramificación, 32
 - de v.a.s independientes, 28
 - del jugador, 31
 - distribución inicial, 26
 - ergódica, 40
 - estacionaria, 24
 - existencia, 26
 - finita, 24
 - irreducible, 40
 - recurrente, 49
 - regular, 76
 - reversible, 79
 - transitoria, 49
- Caminata aleatoria, 5
 - asimétrica, 7
 - del jugador, 14
 - simétrica, 7
- Chapman-Kolmogorov, 34, 209
- Clase
 - aperiódica, 41
 - cerrada, 52
 - de comunicación, 38
 - periódica, 41
- Coficiente
 - de difusión, 253
 - de tendencia, 253
- Comunicación, 37
- Confiabilidad, 157
- Corrección de Itô, 246
- Coseno hiperbólico, 117
- Cox
 - proceso de, 111
- Delta de Kronecker, 27
- Deriva, 253
- Desigualdad
 - de Jensen, 276
- Distribución
 - estacionaria, 65
 - invariante, 65
 - reticulada, 273
 - tipo lattice, 273
- Distribuciones
 - finito dimensionales, 271
- Doob, J. L., 196
- Downcrossing, 187
- Drift, 253

- Ecuación
 - de balance detallado, 79
 - de calor, 210
 - de Chapman-Kolmogorov, 34, 209
 - de difusión, 210
 - de renovación, 146, 147
 - de Wald, 273
 - estocástica, 253
 - solución fuerte, 254
- Ecuaciones
 - de Kolmogorov, 132, 134
- Ecuaciones de Kolmogorov
 - hacia atrás, 130, 132
- Espacio
 - C^1 , 250
 - C^2 , 250
 - $L^2(P)$, 238
 - $L^2(P \times dt)$, 238
 - \mathcal{H}^2 , 241
 - \mathcal{H}_0^2 , 239
 - \mathcal{L}_{loc}^2 , 244
 - de estados, 1
 - parametral, 1
- Esperanza
 - condicional, 274
- Estado
 - absorbente, 39
 - aperiódico, 40
 - periódico, 40
 - recurrente, 45
 - recurrente nulo, 60
 - recurrente positivo, 60
 - transitorio, 45
- Estrategia de juego, 176
- Euler-Maruyama, 257
- Fórmula
 - de Stirling, 273
- Fatou
 - lema, 272
- Filtración, 168
 - canónica, 168
 - continua por la derecha, 169
 - estándar, 169
 - natural, 168
- Función
 - coseno hiperbólico, 117
 - de confiabilidad, 157
 - de intensidad, 111
 - de renovación, 146
 - de supervivencia, 158
 - de tasa de falla, 157
 - de valor medio, 111
 - delta de Kronecker, 27
 - dir. Riemann integrable, 276
 - hazard, 158
 - seno hiperbólico, 117
 - variación cuadrática de una, 214
 - variación de una, 214
- Igualdad de procesos, 271
- Independencia
 - de procesos, 271
- Integrabilidad uniforme, 191
- Integral estocástica, 237
 - como un proceso, 244
 - de Itô, 240, 242, 244, 245
 - de Stratonovich, 247
 - extensión por aproximación, 241
 - extensión por localización, 244
 - para procesos simples, 240
 - propiedades, 249
- Itô
 - corrección de, 246
 - fórmula de, 250, 255
 - integral de, 240, 242, 244, 245
 - isometría de, 243
 - proceso de, 253

- Kronecker, 27
- Lèvy, P. P., 231
- Lema
 - de Abel, 272
 - de Fatou, 272
- Método de discretización, 257
- Markov, A. A., 83
- Martingala, 4, 171
 - de de Moivre, 200
 - detenida, 176
 - estrategia de juego, 176, 177
 - exponencial, 234
 - producto, 200, 201
- Martingalas
 - teorema de convergencia, 190
 - teorema de representación, 195
- Matriz
 - de prob. de transición, 24
 - doblemente estocástica, 25
 - estocástica, 25
 - regular, 76
- McKean
 - Tabla de multiplicación, 255
- Movimiento Browniano, 205
 - estándar, 206–208
 - exponencial, 259
 - geométrico, 258
 - martingala, 211
 - multidimensional, 218
 - probabilidad de transición, 209
 - radial, 219
 - recurrencia
 - por vecindades, 228
 - puntual, 224
 - reflejado en el origen, 233
 - unidimensional, 206, 207
 - variación, 215
 - variación cuadrática, 214
- Número de cruces, 186
- Notación *o* pequeña, 273
- Parámetros infinitesimales, 130
- Periodo, 40
- Probabilidades
 - de transición, 24, 27
 - de transición en n pasos, 27
 - de transición en un paso, 24
 - infinitesimales, 106, 131
- Problema
 - de la ruina del jugador, 14
- Proceso, 1
 - a tiempo continuo, 1
 - a tiempo discreto, 1
 - adaptado, 168
 - con incrementos estacionarios, 3
 - con incrementos indep, 3
 - de Bessel, 219
 - de Cox, 111
 - de ensayos independientes, 3
 - de Itô, 253
 - de Lèvy, 4
 - de Markov, 3
 - de muerte, 138
 - de nacimiento puro, 137
 - de nacimiento y muerte, 133
 - de Ornstein-Uhlenbeck, 261
 - de Poisson, 98, 106, 107
 - compuesto, 112
 - generalizado, 113
 - homogéneo, 98
 - no homogéneo, 109
 - simulaciones, 104
 - de renovación, 144
 - de saltos, 124
 - de Wiener, 206
 - de Yule, 138
 - detenido, 175

- distribuciones fin. dim., 271
- equivalencia de —s, 271
- espacio de estados, 1
- espacio parametral, 1
- estacionario, 3
- filtración de un, 168
- Gausiano, 4
- historia de un —, 168
- igualdad de —s, 271
- independencia, 271
- indistinguibilidad, 271
- modificación de un —, 271
- predecible, 169
- realización de un —, 2
- simple, 239
- trayectoria de un —, 2
- versión de un —, 271
- Propiedad
 - de Markov, 3
 - de pérdida de memoria, 99, 150
 - de semigrupo, 129
- Puente Browniano, 234, 265
 - densidad, 234
- Racha de éxitos, 29
- Recurrencia, 45
 - nula, 60
 - positiva, 60
- Renovación
 - ecuación de, 147
 - función de, 146
- Seno hiperbólico, 117
- Stirling, 273
- Stratonovich, 247
- Submartingala, 171
- Supermartingala, 171
- Técnica de acople, 74
- Tabla de multiplicación de McKean, 255
- Teorema
 - de conv. de mart. de Doob, 190
 - de conv. dominada, 273
 - de conv. monótona, 272
 - de convergencia monótona, 275
 - de paro opcional, 179
 - de representación de martingalas, 195
 - ergódico para cadenas, 58
- Tiempo
 - s de estancia, 98, 124
 - s de interarribo, 98
 - s de paro, 169
 - de primera visita, 43, 51, 59
 - de vida restante, 148
 - medio de recurrencia, 51, 60
- Tiempo de paro, 171
- Transitoriedad, 45
- Variación, 214
 - cuadrática, 214
- Wald
 - ecuación de, 273
- Wiener, N., 230