

# Procesos Estocásticos

Fernando Baltazar-Larios

20 de febrero de 2023



# Índice general

|  |           |
|--|-----------|
| <b>1. Introducción a los Procesos Estocásticos</b>                                       | <b>5</b>  |
| 1.1. Definiciones elementales . . . . .  | 5         |
| 1.2. Clasificación general. . . . .  | 6         |
| 1.3. Condiciones de regularidad: una primera tipología de procesos estocásticos. . . . . | 6         |
| 1.4. Características principales de los procesos estocásticos . . . . .                  | 7         |
| 1.4.1. Esperanza de un proceso estocástico . . . . .                                     | 7         |
| 1.4.2. Autocorrelación de un proceso estocástico . . . . .                               | 7         |
| 1.4.3. Autocovarianza de un proceso estocástico . . . . .                                | 7         |
| 1.4.4. Correlación cruzada de procesos estocásticos . . . . .                            | 7         |
| 1.4.5. Covarianza cruzada de procesos estocásticos . . . . .                             | 8         |
| 1.5. Ejemplos . . . . .  | 8         |
| <b>2. Cadenas de Markov a tiempo discreto</b>  | <b>11</b> |
| 2.1. Introducción . . . . .  | 11        |
| 2.2. Conceptos básicos . . . . .   | 11        |
| 2.3. Ejemplos . . . . .  | 12        |
| 2.4. Resultados Importantes . . . . .  | 12        |
| 2.5. Clases de comunicación . . . . .  | 13        |
| 2.6. Clasificación de estados . . . . .  | 13        |
| 2.6.1. Tiempos de llegada y tiempos de absorción . . . . .                               | 15        |
| 2.6.2. Criterios de clasificación de estados . . . . .                                   | 16        |
| 2.6.3. Recurrencia positiva y nula. . . . .  | 17        |

|  |           |
|--|-----------|
| 2.7. Distribuciones invariantes y distribuciones límites . . . . . | 18        |
| 2.7.1. Cadenas regulares, absorbentes y convergentes. . . . .      | 19        |
| 2.8. Simulación de cadenas de Markov . . . . .                     | 21        |
| 2.9. Inferencia para Cadenas de Markov . . . . .                   | 23        |
| 2.10. Ejemplos . . . . .   | 24        |
| 2.11. Cadenas de Markov a tiempo discreto en Seguros . . . . .     | 25        |
| 2.11.1. El sistema Bonus-Malus . . . . .                           | 25        |
| <b>3. Proceso de Poisson</b>                                       | <b>27</b> |
| 3.1. Definiciones . . . . .  | 27        |
| 3.1.1. Propiedades principales . . . . .                           | 29        |
| 3.1.2. El proceso Poisson compuesto . . . . .                      | 30        |
| 3.1.3. Proceso de Poisson no homogéneo . . . . .                   | 32        |
| 3.2. Simulación de un proceso de Poisson . . . . .                 | 32        |
| 3.2.1. Simulación de un proceso de Poisson homogéneo . . . . .     | 32        |
| 3.2.2. Simulación de Proceso de Poisson no homogéneo . . . . .     | 33        |
| 3.2.3. Simulación de Proceso de Poisson compuesto . . . . .        | 34        |
| <b>4. Martingalas</b>  | <b>37</b> |
| <b>5. Procesos con valores continuos</b>                           | <b>39</b> |
| 5.1. El Movimiento Browniano y sus transformaciones . . . . .      | 39        |
| 5.2. Ecuaciones diferenciales estocásticas . . . . .               | 40        |

# Capítulo 1

## Introducción a los Procesos Estocásticos

En este capítulo se presenta una breve introducción al concepto de proceso estocástico, se da una descripción general de sus características principales, se hace una clasificación básica de los procesos estocásticos y se dan algunos ejemplos.

### 1.1. Definiciones elementales

Consideremos un sistema cuya evolución en el tiempo entre diferentes estados (determinados previamente) está modelada por una ley de movimiento. Si  $X_t$  representa el estado del sistema en el tiempo  $t$  y suponemos que la evolución de éste no es determinista, sino está regida por algún componente aleatorio, resulta natural considerar a  $X_t$  como una variable aleatoria para cada valor  $t$ . Hasta aquí, tenemos una colección de variables aleatorias indexadas por el tiempo, a esta colección se le conoce como un proceso estocástico. Resulta de gran relevancia estudiar las relaciones entre las variables aleatorias de dicha colección que permiten clasificar a los procesos estocásticos.

Comenzaremos dando una definición formal de un proceso estocástico.

**Definición 1.1.** *Un proceso estocástico  $\mathbf{X} = \{X_t\}_{t \in T}$  es una colección de variables aleatorias (definidas en el mismo espacio de probabilidad) que toman valores en un conjunto  $E$  con parámetro  $T$ , es decir, para cada  $t \in T$ ,  $X_t$  es una variable aleatoria.*

En esta definición a  $T$  se le conoce como el espacio parametral y es interpretado como el tiempo, de esta forma a  $X_t$  se le entenderá como el estado del proceso al tiempo  $t$ . Además se considera que las variables aleatorias  $X_t$  para toda  $t \in T$ , están definidas sobre el mismo espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ . Un proceso estocástico puede ser interpretado como la evolución en el tiempo de algún fenómeno cuya dinámica se rige por el azar.

Siguiendo la definición de proceso estocástico, se le puede ver como una función de dos variables

$$\mathbf{X} : T \times \Omega \rightarrow E,$$

tal que a cada pareja  $(t, \omega) \in (T, \Omega)$  se le asocia  $X(t, \omega) = X_t(\omega)$  y de esta manera para cada  $t \in T$  la

función de

$$\omega \rightarrow X_t(\omega)$$

es una variable aleatoria y para cada  $\omega \in \Omega$ , la función

$$t \rightarrow X_t(\omega)$$

es una trayectoria o realización del proceso.

## 1.2. Clasificación general.

Según las características del espacio parametral se puede hacer la primera distinción entre procesos estocásticos.

**Definición 1.2. Proceso estocástico a tiempo discreto** Si  $T$  es un conjunto a lo más numerable, se dice que  $\mathbf{X}$  es un proceso estocástico a tiempo discreto, para nosotros será  $T = M$ , donde  $M \subset \mathbb{N}$  y en general se denota por  $\mathbf{X} = \{X_n\}_{n \geq 0}$ .

**Definición 1.3. Proceso estocástico a tiempo continuo** Cuando  $T$  sea un conjunto infinito no numerable, se dice que  $\mathbf{X}$  es un proceso estocástico a tiempo continuo, para este caso  $T = [0, \tau]$  ó  $[0, \infty)$ . y en será denotado por  $\mathbf{X} = \{X_t\}_{t \geq 0}$ .

En cuanto al conjunto donde las variables aleatorias toman valores ( $E$ ), se le conoce como espacio de estados del proceso y de manera análoga este también puede ser discreto o continuo.

## 1.3. Condiciones de regularidad: una primera tipología de procesos estocásticos.

En esta sección hablaremos sobre características que pueden tener los procesos estocásticos que además sirven para hacer una clasificación de los mismos.

**Definición 1.4. Procesos de variables independientes.** Sea  $\mathbf{X} = \{X_n\}_{n \geq 0}$  se dice que es un proceso estocástico de variables aleatorias independientes si  $X_m$  y  $X_n$  son independientes para toda  $n, m \geq 0$  con  $m \neq n$ .

**Definición 1.5. Procesos con incrementos independientes.** Un proceso estocástico  $\mathbf{X} = \{X_t\}_{t \in T}$ , se dice que tiene incrementos independientes si para cualesquiera  $t_0 < t_1 < \dots < t_n$ ,  $t_i \in T$ , las variables aleatorias

$$X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$$

son independientes.

**Definición 1.6. Procesos con incrementos estacionarios.** Un proceso estocástico  $\mathbf{X} = \{X_t\}_{t \in T}$ , se dice que es estacionario si para cualesquiera  $t_0 < t_1 < \dots < t_n$ ,  $t_i \in T$ , la distribución del vector aleatorio  $(X_{t_0}, X_{t_1}, \dots, X_{t_n})$  es la misma que la del vector aleatorio  $(X_{t_0+h}, X_{t_1+h}, \dots, X_{t_n+h})$  para toda  $h > 0$  y entonces tiene incrementos estacionarios si para cada  $s < t$  y  $h > 0$  se tiene que  $X_{t+h} - X_{s+h}$  y  $X_t - X_s$  tienen la misma distribución.

**Definición 1.7. Procesos Gaussianos.** Se dice que un proceso estocástico  $\mathbf{X} = \{X_t\}_{t \in T}$  es un proceso de Gaussiano si para cualquier colección finita de tiempos  $t_1, \dots, t_n$  el vector  $(X_{t_1}, \dots, X_{t_n})$  tiene distribución Gaussiana multivariada.

**Definición 1.8. Procesos de Markov.** Se dice que un proceso estocástico  $\mathbf{X} = \{X_t\}_{t \in T}$  es un proceso de Markov si satisface la propiedad de Markov, es decir, si

$$P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, \dots, X_{t_0} = x_0) = P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1})$$

con  $t_0 < t_1 < \dots < t_n$ ,  $t_i \in T$  y  $x_i \in E$  para  $i = 0, 1, \dots, n$ . Este tipo de procesos son aquellos cuyo evolución futura depende solamente del estado presente del proceso.

## 1.4. Características principales de los procesos estocásticos

### 1.4.1. Esperanza de un proceso estocástico

Para un proceso estocástico  $\mathbf{X} = \{X_t\}_{t \in T}$  se define la esperanza de  $X_t$  como el valor esperado de la variable aleatoria observando el proceso en algún tiempo  $t$ , es decir

$$E(X_t) = \int_E x dF_{X_t}(x),$$

donde  $F_{X_t}$  es la función de distribución de  $X_t$ .

### 1.4.2. Autocorrelación de un proceso estocástico

La función de autocorrelación del proceso  $\mathbf{X} = \{X_t\}_{t \in T}$  se define como el valor esperado del producto de las variables aleatorias  $X_{t_1}$  y  $X_{t_2}$  en los momentos  $t_1$  y  $t_2$  respectivamente, es decir,

$$E(X_{t_1} X_{t_2}) = \int_E \int_E x_1 x_2 dF_{X_{t_1}, X_{t_2}}(x_1, x_2)$$

donde  $F_{X_{t_1}, X_{t_2}}$  es la función de densidad conjunta de las variables aleatorias  $X_{t_1}$  y  $X_{t_2}$ .

### 1.4.3. Autocovarianza de un proceso estocástico

La función de autocovarianza del proceso  $\mathbf{X} = \{X_t\}_{t \in T}$  se define por:

$$R_X(t_1, t_2) = E[(X_{t_1} - E(X_{t_1}))(X_{t_2} - E(X_{t_2}))].$$

### 1.4.4. Correlación cruzada de procesos estocásticos

Sean  $\mathbf{X} = \{X_t\}_{t \in T}$  e  $\mathbf{Y} = \{Y_t\}_{t \in T}$  dos procesos estocásticos se define la correlación cruzada entre estos procesos por:

$$R_{XY}(t_1, t_2) = E(X_{t_1} Y_{t_2})$$

o

$$R_{YX}(t_1, t_2) = E(Y_{t_1} X_{t_2})$$

para cualesquiera tiempos  $t_1$  y  $t_2$ .

#### 1.4.5. Covarianza cruzada de procesos estocásticos

Sean  $\mathbf{X} = \{X_t\}_{t \in T}$  e  $\mathbf{Y} = \{Y_t\}_{t \in T}$  dos procesos estocásticos se define la covarianza cruzada entre estos procesos por:

$$E[(X_{t_1} - E(X_{t_1}))(Y_{t_2} - E(Y_{t_2}))]$$

para cualesquiera tiempos  $t_1$  y  $t_2$ .

### 1.5. Ejemplos

Presentaremos algunos ejemplos de procesos estocásticos.

**Ejemplo 1.5.1.** (*Caminatas aleatorias simples y el problema de la ruina.*) Consideremos la siguiente situación: teniendo un capital de  $k$  pesos al tiempo cero, a cada instante de tiempo se apuesta un peso en un volado, ganando si cae águila. Si  $X_n$  denota el capital después del  $n$ -ésimo volado entonces  $\mathbf{X} = \{X_n\}_{n \geq 0}$  es un proceso estocástico que modela la evolución del capital en el tiempo.

Para estudiar matemáticamente el capital en el tiempo, consideremos a la variable aleatoria que nos indica el instante que se llega a la ruina (podría ser infinita). El modelo matemático puede ser descrito como sigue: sean  $U_1, U_2, \dots$  variables aleatorias uniformes en  $(0, 1)$  e independientes. Considerando la variable aleatoria  $\mathbb{I}_{\{U_i \leq 1/2\}}$  que puede ser interpretada como la indicadora de que el volado tenga como resultado águila. Por último se considera a la variable aleatoria  $Y_i = 2\mathbb{I}_{\{U_i \leq 1/2\}} - 1$  que toma los valores 1 si cae águila y -1 si cae sol. De esta forma tenemos que

$$X_n = \sum_{i=0}^n Y_i.$$

**Ejemplo 1.5.2.** Considere una partícula que se mueve en un conjunto de  $m + 1$  nodos, etiquetados por  $0, 1, \dots, m$ , distribuidos en un círculo. Sea  $X_n$  la posición de la partícula después de  $n$ -pasos y supongamos que se mueve a los nodos vecinos con igual probabilidad.

**Ejemplo 1.5.3. Tiempos de espera** Consideremos la fila en un banco y supongamos que los clientes van llegando a tiempos aleatorios y cada uno requiere un servicio que es también una variable aleatoria. Supongamos que un cliente llega a un tiempo  $t$ , nos interesa saber ¿cuánto tiempo tarda en salir del banco?.

**Ejemplo 1.5.4. (Movimiento Browniano Geométrico).** Supongamos que tenemos un proceso estocástico a tiempo continuo  $\mathbf{X} = \{X_t\}_{t \geq 0}$  donde  $X_t$  representa el precio de una acción al tiempo  $t$ .



Estudiando los elementos de este proceso podemos conocer características del comportamiento futuro del precio del activo.

**Ejemplo 1.5.5. *El modelo clásico de Cramér-Lundberg.***

Este modelo tiene sus orígenes en la tesis doctoral de Filip Lundberg defendida en el año de 1903. En este trabajo Lundberg analiza el reaseguro de riesgos colectivos y presenta el proceso de Poisson compuesto. En 1930 Harald Cramér retoma las ideas originales de Lundberg y las pone en el contexto de los procesos estocástico. El modelo ha sido muy estudiado y se han propuesto varias formas de generalizarlo.

El modelo que estudiaremos es el proceso a tiempo continuo  $\{C_t\}_{t \geq 0}$  dado por

$$C_t = u + ct - \sum_{j=1}^{N_t} Y_j$$

donde:

- $u$  es el capital inicial de la compañía aseguradora
- $ct$  es la entrada por primas hasta el tiempo  $t$  con  $c$  una constante positiva
- $Y_j$  es el monto de la  $j$ -ésima reclamación, y
- $N = \{N_t\}$  es un proceso de Poisson de parámetro  $\lambda$ .

$C_t$  representa el balance más sencillo de ingresos menos egresos de una compañía aseguradora. Al proceso  $C$  se le llama proceso de riesgo o proceso de superávit.

La variable aleatoria  $C_t$  se puede interpretar como el capital de la compañía aseguradora al tiempo  $t$  y por razones naturales y legales es importante que  $C_t$  permanezca por arriba de cierto nivel mínimo. Cuando  $C_t < 0$  se dice que hay ruina.

La ruina casi nunca sucede en la práctica, es solamente un término técnico que produce alguna toma de decisión. Por ejemplo, si el capital de una compañía aseguradora asignado a una cartera decrece en forma significativa, ésta automáticamente puede tomar ciertas medidas para subsanar esta situación y no se trata de un evento insalvable.



# Capítulo 2

## Cadenas de Markov a tiempo discreto

### 2.1. Introducción

Una propiedad de especial importancia que poseen algunos procesos estocásticos a tiempo discreto, es que sus valores en el  $n$ -ésimo momento tienen toda la información para determinar el valor del  $n + 1$ -ésimo momento. Esta propiedad conocida como propiedad de Markov es de gran importancia en el estudio, en general, de la teoría de procesos estocásticos, y por ello prestamos especial dedicación en este capítulo. En este capítulo se supondrá que trabajamos con procesos a tiempo discreto y que toman valores en un conjunto a lo más numerable.

### 2.2. Conceptos básicos

Consideremos una colección de variables aleatorias  $X_0, X_1, \dots$ , interpretando a  $X_n$  como el estado de un sistema al tiempo  $n$  y supongamos que el conjunto de posibles valores de  $X_n$  es a lo más numerable, enumerándolos, tenemos que  $E = \{1, 2, \dots, N\}$  para el caso finito y  $E = \{1, 2, \dots\}$  para el caso numerable,  $E$  es llamado espacio de estados. Si este proceso satisface la **propiedad de Markov**

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

con  $x_0, \dots, x_{n+1} \in E$ , decimos que  $\{X_n\}_{n \geq 0}$  es **una cadena de Markov**.

A la distribución de  $X_0$  se le llama distribución inicial y la denotaremos por  $\pi = (\pi_1, \dots, \pi_N)$  o  $\pi = (\pi_1, \dots)$  para el caso finito o infinito numerable respectivamente donde  $\pi_i = P(X_0 = i)$  para toda  $i \in E$ .

Algunas observaciones importantes:

1. Si  $X_0 = x_0, \dots, X_n = x_n$  entonces decimos que se ha presentado la trayectoria  $x_0, \dots, x_n$ .
2. A la familia  $P(X_{n+1} = j | X_n = i)$  se le conoce como familia de probabilidades de transición de la cadena de Markov y describe la evolución de la cadena en el tiempo.

3. En caso de que exista independencia del tiempo se le llama homogénea y

$$P(X_{n+1} = j | X_n = i) = p_{ij}$$

son las probabilidades de transición y para  $m \geq 1$

$$p_{ij}^{(m)} = P(X_{m+n} = j | X_n = i),$$

que representa la probabilidad de pasar del estado  $i$  al estado  $j$  en  $m$  unidades de tiempo.

**Definición 2.1.** *Matriz de transición.* Es la matriz cuadrada de orden  $N$  formada por las probabilidades de transición

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{pmatrix}$$

con entradas no negativas y la suma de cada renglón es uno (matriz estocástica).

## 2.3. Ejemplos

Veamos algunos ejemplos de cadenas de Markov

**Ejemplo 2.3.1.** *Caminata aleatoria simple.* Un individuo está parado e inicia una caminata en el cual sólo se mueve hacia adelante con probabilidad  $p$  o hacia atrás con probabilidad  $1 - p$  ( $p \in (0, 1)$ ). Definiendo a  $X_n$  como la posición del individuo después de  $n$ -pasos, se tiene entonces que  $E = \mathbb{Z}$  y las probabilidades de transición para cualquier  $n \geq 0$  e  $i, j \in E$  están dadas por:

$$p_{ij} = \begin{cases} p & j = i + 1 \\ 1 - p & j = i - 1 \\ 0 & e.o.c. \end{cases}$$

**Ejemplo 2.3.2.** *Problema del jugador.* Suponga que un jugador  $A$  apuesta \$1 sucesivamente contra otro jugador  $B$ .  $A$  inicia con  $k$  unidades y  $B$  con  $N - k$ . En cada apuesta,  $A$  gana con probabilidad  $p$  y pierde con probabilidad  $1 - p$  ( $p \in (0, 1)$ ). Si definimos  $X_n$  como la fortuna de  $A$  al tiempo  $n$ , entonces  $\mathbf{X} = \{X_n\}_{n \geq 0}$  define una cadena de Markov.

**Ejemplo 2.3.3.** *Cadena de Ehrenfest.* Se tienen dos urnas, entre ambas hay un total de  $N$  bolas. En cada paso se elige una de las bolas al azar y se cambia de urna. Si  $X_n =$  número de bolas en la urna  $A$  después de  $n$  ensayos, entonces  $\mathbf{X} = \{X_n\}_{n \geq 0}$  define una cadena de Markov.

## 2.4. Resultados Importantes

A continuación presentaremos algunos resultados importantes.

**Proposición 2.1.** *Ecuación de Chapman-Kolmogorov.* Sea  $\mathbf{X} = \{X_n\}_{n \geq 0}$  una cadena de Markov con matriz de transición  $P = \{p_{ij}\}_{ij \in E}$  entonces

$$p_{ij}^{(n+m)} = \sum_{k \in E} p_{ik}^{(n)} p_{kj}^{(m)}.$$

Este resultado nos dice que  $p_{ij}^{(n)}$  es la entrada  $ij$  de la  $n$ -ésima potencia de  $P$ . En general, no es posible calcular explícitamente las potencias de la matriz de transición. Sin embargo, nos podemos auxiliar de paquetes computacionales para esta tarea.

**Corolario 2.1.** *Sea  $\mathbf{X} = \{X_n\}_{n \geq 0}$  una cadena de Markov con matriz de transición  $P = \{p_{ij}\}_{ij \in E}$  y distribución inicial  $\pi$  entonces*

$$P[X_n = i] = \pi P_{*i}^n,$$

para toda  $n \in \mathbb{N}$  e  $i \in E$ , donde  $P_{*i}^n$  es la  $i$ -ésima columna de la  $n$ -ésima potencia de  $P$ .

**Proposición 2.2.** *Propiedad de corrimiento.* Sean  $n, k \in \mathbb{N}$  fijos y  $x_0, x_1, \dots, x_n, \dots, x_{n+k} \in E$  entonces

$$P(X_{n+k} = x_{n+k}, \dots, X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0) = P(X_k = x_{n+k}, \dots, X_1 = x_{n+1} | X_0 = x_n).$$

## 2.5. Clases de comunicación

Sean  $i$  y  $j$  dos estados de  $E$ , se dice que el estado  $j$  es accesible desde el estado  $i$  si existe  $n \in \mathbb{N}$ , tal que  $p_{ij}^{(n)} > 0$  ( $i \rightarrow j$ ). Si  $i \rightarrow j$  e  $j \rightarrow i$  entonces,  $i$  y  $j$  se **comunican**, ( $i \leftrightarrow j$ ), en particular, se tiene el siguiente resultado sobre la comunicación.

**Proposición 2.3.** *La relación  $i \leftrightarrow j$  (comunicación) es una relación de equivalencia, por lo tanto, induce a una partición en el espacio de estados.*

A las clases de equivalencia inducidas por esta relación les llamaremos clases de comunicación. En particular diremos que una cadena de Markov es irreducible si tiene una sola clase de comunicación.

Sea  $i \in E$ , a su clase de comunicación la denotaremos por  $C_i$ , es decir

$$C_i = \{j \in E : i \leftrightarrow j\}.$$

En cuanto a la partición inducida por la comunicación de estados decimos que, tomando  $C \subset E$ , es un conjunto cerrado si para toda  $j \in E - C$ ,  $i \in C$  no puede acceder a  $j$ , en caso contrario decimos que es una clase abierta.

## 2.6. Clasificación de estados

De acuerdo a las propiedades que tiene cada estado en una cadena de Markov en cuanto a la probabilidad de regresar a él, los estados pueden ser clasificados de la siguiente manera.

**Definición 2.2.** Se dice que un estado  $i$  es **recurrente** si la probabilidad de (eventualmente) regresar a  $i$ , es uno, es decir, si

$$P(X_n = i | X_0 = i) = 1,$$

para alguna  $n \geq 1$ . Un estado que no es recurrente se llama **transitorio**.

Tratar de clasificar los estado en recurrentes y transitorios partiendo de la definición no resulta, en general, tarea fácil. Una manera distinta de enunciar estos conceptos es en términos de las siguientes probabilidades.

**Definición 2.3.** Para cada  $n \geq 1$ , denotamos por  $f_{ij}^{(n)}$  a 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,

$$f_{ij}^{(n)} = P(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j, X_0 = i).$$

Además definimos

$$f_{ij}^{(0)} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{e.o.c} \end{cases}$$

**Definición 2.4.** Definimos a

$$f_i = \sum_{n=1}^{\infty} f_{ii}^{(n)},$$

que es la probabilidad de un eventual regreso al estado  $i$ .

**Nota 2.1.** Decimos que:

- El estado  $i$  es **recurrente** si  $f_i = 1$ , en particular si  $f_{ii}^{(1)} = 1$  el estado se le llama **absorbente**.
- El estado  $x$  es **transitorio** si  $f_i < 1$ .

Un resultado importante es el que permite descomponer la probabilidad de visitar el estado  $j$  iniciando en  $i$  en términos de todas las posibilidades de tiempos que pueden ser la primera visita.

**Proposición 2.4.** Para  $n \geq 1$  se cumple que

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

para tod  $i, k \in E$ .

Podemos ya formular el siguiente importante teorema.

**Teorema 2.1. (Teorema de descomposición de las Cadenas de Markov)** El espacio de estados  $E$  de una cadena de Markov, tiene la siguiente partición única:

$$E = T \cup C_1 \cup C_2 \cup \dots$$

donde  $T$  es un conjunto de estados transitorios, y  $C_i$  son clases cerradas e irreducibles de estados recurrente.

**Nota 2.2.** El teorema de descomposición nos muestra las posibilidades que pueden darse en una cadena de Markov. Esto es, si  $x_0 \in C_k$ , entonces la cadena nunca abandonará la clase  $C_k$  y entonces, podemos considerar el espacio de estados  $E = C_k$ . Por otra parte, si  $x_0 \in T$  entonces, o la cadena permanece por siempre en  $T$  o se mueve a una clase  $C_k$  y permanece ahí por siempre. Así, o la cadena siempre toma valores en el conjunto de estados transitorios o acaba en un conjunto cerrado persistente de estados donde permanecerá por siempre. Veamos ahora que en el caso en el que  $E$  sea finito la primera situación no puede darse.

**Definición 2.5.** Definamos al número de visitas al estado  $i$  como la variable aleatoria

$$V_i := \sum_{n=0}^{\infty} \mathbb{I}_{\{X_n=i\}}.$$

**Nota 2.3.** La variable aleatoria anterior toma valor infinito con probabilidad cero o uno.

Decimos que si  $V_i$  es infinita con probabilidad 1, partiendo de  $i$ , entonces  $i$  es estado recurrente, en caso contrario es un estado transitorio.

Motivados en este criterio tenemos el siguiente resultado

**Proposición 2.5.**

1. Un estado  $i$  es recurrente si y sólo si  $\sum_n p_{ii}^{(n)} = \infty$ .
2. Un estado  $i$  es transitorio si y sólo si  $\sum_n p_{ii}^{(n)} < \infty$ .

Este resultado nos proporciona una equivalencia, en términos de las matrices de transición, para que un estado sea recurrente.

### 2.6.1. Tiempos de llegada y tiempos de absorción

**Definición 2.6.** Un *tiempo aleatorio* es una variable aleatoria

$$T : \Omega \rightarrow \mathbb{N} \cup \{\infty\}.$$

**Definición 2.7.** Un tiempo aleatorio  $T$  es un *tiempo de paro* si para  $n \in \mathbb{N}$  existe  $A_n \subset E^{n+1}$  tal que

$$\{T = n\} = \{(X_0, \dots, X_n) \in A_n\}.$$

A la variable aleatoria tiempo de paro la podemos interpretar como el tiempo que se obtiene de observar una trayectoria hasta que se cumpla cierta condición. Veamos algunos ejemplos de tiempo de paro.

**Ejemplo 2.6.1.** Sea  $T$  el tiempo en el que una cadena de Markov regresa a su estado inicial, es decir

$$T = \begin{cases} \infty & X_n \neq X_0 \text{ para toda } n \\ \min\{n \geq 1 | X_n = X_0\} & \text{e.o.c} \end{cases}$$

**Ejemplo 2.6.2.** Sea  $T_n$  el tiempo en el ocurre la  $n$ ésima visita al estado inicial es un tiempo de paro.

En el estudio de las cadenas de Markov resulta de particular interés el estudio de los tiempos aleatorios de la ocurrencia de eventos relacionados con la llegada a un estado o un conjunto de estados del espacio de la cadena. En esta sección estudiaremos este tipo de variables aleatorias.

**Definición 2.8.** Sea  $T_A$  primera vez que la cadena accede a  $A$  un subconjunto del espacio de estados, es decir

$$T_A = \begin{cases} \infty & X_n \in E - A \text{ para toda } n \\ \min\{n \geq 1 | X_n \in A\} & \text{e.o.c} \end{cases}$$

**Nota 2.4.** En general, no resulta tarea fácil encontrar la distribución de este tipo de variables aleatorias.

En el caso que el conjunto  $A$  consta de un solo estado ( $j$ ) y la cadena empieza en el estado  $i$  denotaremos por  $T_{ij}$  la primera vez que la cadena visita el estado  $j$  iniciando en  $i$  si  $i \neq j$ . Si  $i = j$  lo denotamos por  $T_i$ .

Veamos un ejemplo donde podemos ilustrar esta definición.

**Ejemplo 2.6.3. Racha de éxitos.** Sea  $E_1, E_2, \dots$ , una sucesión de ensayos Bernoulli con probabilidad de éxito  $p$  y probabilidad de fracaso  $q = 1 - p$ . Si  $X_n$  es el número de éxitos consecutivos previos al tiempo  $n$  (incluido el tiempo  $n$ ). Se dice que una cadena de  $r$  éxitos ocurre al tiempo  $n$  si en el ensayo  $n - r$  ocurre un fracaso y los resultados de los ensayos  $n - r + 1$  al  $n$  son éxitos.

Para la cadena de Markov de racha de éxitos tenemos (por ejemplo)

1.  $f_{01}^{(n)} = q^{n-1}p$ .
2.  $f_{00}^{(n)} = p^{n-1}q$ .

## 2.6.2. Criterios de clasificación de estados

A continuación presentamos una serie de resultados respecto a la clasificación de estados y la comunicación.

**Proposición 2.6.** La recurrencia y transitividad son propiedades de clase.

Un criterio fácil para saber si una clase es recurrente es el siguiente.

**Proposición 2.7.** Si el espacio de estados es finito, una clase es recurrente si y sólo si es cerrada.

**Corolario 2.2.** En una cadena de Markov irreducible con espacio de estados finito todos sus estados son recurrente.

**Corolario 2.3.** Toda cadena de Markov con espacio de estados finito tiene por lo menos un estado recurrente.

**Corolario 2.4.** Sea  $C \subset E$  una clase abierta. Entonces  $C$  es transitoria.

**Nota 2.5.** En general, el análisis de recurrencia y transitoriedad, en cadenas con espacio de estados infinito, es más complicado.



**Definición 2.9.** Sea  $P$  la matriz de transición, definimos al periodo de un estado  $i$  como el máximo común divisor del conjunto  $\{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}$ , es decir,

$$d(i) := M.C.D.\{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}.$$

**Proposición 2.8.** El periodo es una propiedad de clase de comunicación.

### 2.6.3. Recurrencia positiva y nula.

Siguiendo la clasificación de estados, resulta importante conocer el número esperado de pasos para regresar a  $i$ , es decir,

$$\mu_i := \sum_{n=1}^{\infty} n f_{ii}^{(n)},$$

a la cual llamaremos **tiempo medio de recurrencia**.

**Nota 2.6.** Puede darse el caso de que siendo  $i$  un estado recurrente el tiempo medio de recurrencia,  $\mu_i$ , sea infinito; siendo éste el caso en el que aunque el regreso a  $i$  es cierto se necesita un tiempo infinito. Así, realizamos la siguiente distinción entre los estados recurrentes obteniendo una subclasificación de estos.

**Definición 2.10.** Decimos que

1. El estado  $i$  es **recurrente positivo** si  $\mu_i < \infty$ .
2. El estado  $i$  es **recurrente nulo** si  $\mu_i = \infty$ .

**Definición 2.11.** Sean  $i, j \in E$ , definimos a la variable aleatoria  $V_{ij}(n)$  como el número de visitas al estado  $j$  empezando en  $i$  en una trayectoria de longitud  $n$ , es decir,

$$V_{ij}(n) = \sum_{k=1}^n \mathbb{I}_{X_k=j},$$

dado que  $X_0 = i$ . Además  $V_{ij} = \lim_{n \rightarrow \infty} V_{ij}(n)$ .

Se tiene directamente el siguiente resultado.

**Proposición 2.9.** Para cualesquiera  $i, j \in E$  se tiene que

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

$$5. P(V_{ij} < \infty) = \begin{cases} 1 & \text{si } j \text{ es transitorio} \\ 1 - f_{ij} & \text{si } j \text{ es recurrente} \end{cases}$$

**Teorema 2.2. Propiedad Fuerte de Markov.** Sea  $A \in E^{n+1}$ ,  $T$  un tiempo de paro tales que  $P(A, X_n = j, T = n) > 0$ . Entonces, condicionado a  $A \cap \{T = n, X_n = j\}$ , el proceso  $\{X_{n+m}\}_{m \geq 0}$  es una cadena de Markov que comienza en  $j$  y tiene matriz de  $P$ .

**Teorema 2.3. Teorema ergódico para cadenas de Markov** Sean  $i, j$  cualesquiera estados de una cadena de Markov irreducible se cumple que

$$\lim_{n \rightarrow \infty} \frac{V_{ij}(n)}{n} = \frac{1}{\mu_j}$$

casi seguramente.

**Corolario 2.5.**

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

**Proposición 2.10.** La recurrencia nula y positiva es una propiedad de clase

**Proposición 2.11.** En una cadena irreducible con espacio de estados finito, todos los estados son recurrentes positivos, es decir, o existen estados recurrentes nulos en cadenas de Markov finitas.

**Ejemplo 2.6.4.** Veamos que la cadena de Markov de racha de éxitos es una cadena recurrente positiva.

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

**Ejemplo 2.6.5.** Considere una cadena de Markov con espacio de estados  $\{0, \dots, 5\}$  y matriz de transición

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{8} & 0 & \frac{7}{8} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{5} & \frac{2}{5} \end{pmatrix}$$

Determine que estados son transitorios y cuales recurrentes. Calcular y estimar tiempos medios de recurrencia y número esperado de visitas a cada estado.

## 2.7. Distribuciones invariantes y distribuciones límites

**Definición 2.12.** Sea  $\nu \in R^{|E|}$ , decimos que  $\nu$  es una **distribución invariante o estacionaria** de la cadena de Markov con matriz de transición  $P$ , si satisface

1. Tiene todas sus entradas no negativas.
2.  $\sum_{i \in E} \nu_i = 1$ .

$$3. \sum_{j \in E} \nu_i p_{ij} = \nu_j.$$

En términos matriciales, la distribución de probabilidad  $\nu$  es invariante si

$$\nu P = \nu.$$

**Nota 2.7.** *La interpretación de una distribución invariante es que una variable aleatoria con dicha distribución mantendrá su distribución en la evolución de la cadena de Markov en el tiempo.*

Desde que encontrar una distribución estacionaria en una cadena de Markov consiste en resolver un sistema de  $|E| + 1$  ecuaciones y  $|E|$  incógnitas podemos no tener solución, una infinidad de soluciones o una única solución.

Los siguientes resultados son de gran utilidad en el cálculo de la distribución estacionaria de una cadena de Markov.

**Proposición 2.12.** *Sea  $\nu$  una distribución estacionaria para una cadena de Markov. Si  $i$  es un estado transitorio o recurrente nulo, entonces  $\nu_i = 0$ .*

Una propiedad importante de una distribución invariante es que si  $X_0$  tiene distribución  $\nu$ , entonces  $X_n$  tendrá la misma distribución para toda  $n$ .

**Proposición 2.13.** *Toda cadena de Markov que es irreducible y recurrente positiva tiene una única distribución estacionaria dada por*

$$\nu_i = \frac{1}{\mu_i} > 0.$$

**Teorema 2.4.** *Para una cadena irreducible las siguientes condiciones son equivalentes.*

1. *Todos los estados son recurrentes positivos.*
2. *Algún estado es recurrente positivo.*
3. *La cadena tiene una única distribución invariante*

**Ejemplo 2.7.1.**

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

$$\nu = (1/3, 2/3).$$

**Ejemplo 2.7.2. Cadena de Ehrenfest.** *Como esta cadena de Markov es irreducible y con espacio de estados finito se tiene que es recurrente positiva y por los resultados anteriores tiene una única distribución invariante  $\nu$ .*

### 2.7.1. Cadenas regulares, absorbentes y convergentes.

En esta sección presentaremos resultados del comportamiento de una cadena de Markov recurrente cuando la trayectoria es lo suficientemente grande. Otro concepto importante en el estudio de las cadenas de Markov es el relacionado con la posible periodicidad con la que se visita a cada estado.

**Definición 2.13.** Sea  $P$  la matriz de transición, definimos al periodo de un estado  $i$  como el máximo común divisor del conjunto  $\{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}$ , es decir,

$$d(i) := M.C.D.\{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}.$$

**Definición 2.14.** Decimos que una cadena de Markov  $\{X_n\}$  con espacio de estados finito es **regular** si existe una potencia de la matriz de transición con todas las entradas positivas.

**Proposición 2.14.** El periodo es una propiedad de clase.

**Proposición 2.15.** Dado un estado  $i$ , existe  $k \in \mathbb{N}$  tal que para toda  $n \geq k$ , se tiene que  $p_{ii}(nd(i)) > 0$ .

**Corolario 2.6.** Si para  $m \in \mathbb{N}$  se tiene que  $p_{ij}^{(m)} > 0$  entonces existe  $k \in \mathbb{N}$  tal que para toda  $n \geq k$  se tiene que  $p_{ij}^{(m+nd(j))} > 0$ .

Ahora combinando esta definición con el concepto de recurrencia positiva tenemos que un estado es **ergódico** si es recurrente positivo y tiene periodo uno (aperiódico).

**Proposición 2.16.** Una cadena de Markov es regular si y sólo si es finita, aperiódica e irreducible.

**Definición 2.15.** Si los límites

$$\mu_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)}$$

existen para cada estado  $j$  en el espacio de estados entonces  $\pi$ , es la distribución límite.

**Nota 2.8.** Las siguientes observaciones son consecuencia de la definición anterior.

1. La distribución límite es inducida por la matriz de transición  $P$ .
2. La distribución límite no depende de la distribución inicial.
3. El límite de las potencias de  $P$  es una matriz estocástica para la cual todos sus renglones son la distribución límite.
4. La distribución límite es una distribución estacionaria.

Ahora presentamos resultados fundamentales en el estudio de la convergencia de las cadenas de Markov.

**Teorema 2.5.** (Existencia de la distribución límite.) Si la cadena de Markov es irreducible, aperiódica y con distribución invariante  $\nu$ , entonces

$$\nu_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)}$$

**Teorema 2.6.** Si la cadena de Markov es irreducible, aperiódica y recurrente positiva existen las probabilidades límite y son la única distribución invariante.

**Lema 2.1.** Sea  $P = \{p_{ij}\}_{i,j \in E}$  una matriz de transición de una cadena de Markov con espacio de estados  $E$ , tal que  $|E| = r$  y  $d := \min\{p_{ij}\}_{i,j \in E}$ . Dado  $Y \in \mathbb{R}^r$  un vector columna de con todas sus entradas no negativas, si  $M_0$  y  $m_0$  denotan el máximo y mínimo de  $Y$  y  $M_1$  y  $m_1$  el máximo y mínimo de  $PY$ , entonces

$$M_1 - m_1 \leq (1 - 2d)(M_0 - m_0),$$

y

$$m_0 \leq m_1 \leq M_1 \leq M_0.$$

**Teorema 2.7.** *Sea  $P$  la matriz de transición de una cadena de Markov regular con espacio de estados finito  $E$ , se cumple que:*

1. *Si  $n \rightarrow \infty$ , las potencias  $P^n$  se aproximan a una matriz  $W$  tal que todos sus renglones son iguales a un mismo vector de probabilidad  $w$  con componentes estrictamente positivos.*
2. *Se tiene que  $w$  es el vector de probabilidad invariante y cualquier otro vector  $v$  tal que  $vP = v$  es un escalar multiplicado por  $w$  (como vector de probabilidad es único).*
- 3.

$$\lim_{n \rightarrow \infty} P(X_n = i) = w_i$$

para todo  $i \in E$  y

$$\lim_{n \rightarrow \infty} P(X_n = j | X_0 = i) = w_j$$

para todo  $i, j \in E$ .

Por último estudiaremos una clase particular de cadenas de Markov, las cadenas de Markov absorbentes.

**Definición 2.16.** *Una cadena absorbente es una cadena de Markov con espacio de estados finita, con al menos un estado absorbente y que desde cualquier estado se puede llegar a algún estado absorbente.*

Para una cadena de Markov absorbente es importante estudiar conceptos como el tiempo medio de absorción o ¿cual será el estado más probable de absorción?

## 2.8. Simulación de cadenas de Markov

Consideremos una cadena de Markov  $\mathbf{b}X = \{X_n\}_{n \geq 0}$  a tiempo discreto, con espacio de estados  $E = \{1, 2, \dots, N\}$  y homogénea en el tiempo. Supongamos además que su distribución inicial está dada por  $\pi = \{\pi_1, \dots, \pi_N\}$  y con matriz de probabilidades de transición:

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1N} \\ p_{21} & p_{22} & \dots & p_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ p_{N1} & p_{N2} & \dots & p_{NN} \end{pmatrix}.$$

Sabemos que, dada una cadena de Markov homogénea, podemos obtener su matriz de transición y su distribución inicial. Ahora nos preguntamos si dada una matriz estocástica y un vector de distribución de probabilidades, es posible asociar una cadena de Markov que tenga a tal matriz como matriz de transición y a tal vector como distribución inicial. El siguiente resultado asegura que sí se puede y su demostración constructiva provee un algoritmo para la simulación de una cadena de Markov.

**Teorema 2.8.** *Dada una matriz estocástica de  $P \in \mathcal{M}_{N \times N}$ , una distribución inicial  $\pi$  sobre el espacio de estados  $E = \{1, 2, \dots, N\}$ , existe un espacio de probabilidad en el que están definidas una sucesión de variables aleatorias  $\{X_n\}_{n \in \mathbb{N}}$  con soporte en  $E$ , que conforman una cadena Markov homogénea con espacio de estados  $E$ , matriz de transición  $P$  y distribución inicial  $\pi$ .*

*Demostración.* Consideremos un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$  en el cual está definida una sucesión  $\{U_n\}_{n \in \mathbb{N}}$  de variables aleatorias independientes uniformes en el intervalo  $(0,1)$ . Definimos  $\phi_i(y)$  como la

función de cuantiles sobre  $E$  asociada a la distribución  $\{p_{ij}\}_{i,j \in E}$  y a  $\phi_0(y)$  como la función de cuantiles sobre  $E$  asociada a la distribución  $\pi$ , con  $y$  en  $(0, 1)$ .

Construimos a continuación el proceso deseado de forma iterativa. Definamos

$$X_0 = \phi_0(U_1)$$

y, habiendo definido a  $X_n$ , definimos  $X_{n+1} = \phi_{X_n}(U_{n+1})$  para  $n \geq 0$ .

Analicemos las probabilidades de transición del proceso  $\{X_n\}_{n \geq 0}$ :

$$\mathbb{P}[X_n = x_n | X_{n-1} = x_{n-1}] = \mathbb{P}[\phi_{X_{n-1}}(U_n) = x_n | X_{n-1} = x_{n-1}].$$

Notemos que  $X_{n-1}$  depende exclusivamente de  $\{U_k\}_{k \leq n-1}$ , además de que las variables aleatorias  $\{U_k\}_{k \in \mathbb{N}}$  son independientes. Esto implica que

$$\mathbb{P}[\phi_{X_{n-1}}(U_n) = x_n | X_{n-1} = x_{n-1}] = \mathbb{P}[\phi_{x_{n-1}}(U_n) = x_n] = p_{x_{n-1}x_n},$$

por definición de la función de cuantiles.

Mostraremos ahora que el proceso  $\{X_n\}_{n \geq 0}$  propuesto satisface la propiedad de Markov.

$$\begin{aligned} \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] &= \mathbb{P}[\phi_0(U_1) = x_0, \phi_{X_0}(U_1) = x_1, \dots, \phi_{X_{n-1}}(U_n) = x_n] \\ &= \mathbb{P}[\phi_0(U_1) = x_0, \phi_{x_0}(U_1) = x_1, \dots, \phi_{x_{n-1}}(U_n) = x_n], \\ &= \mathbb{P}[\phi_0(U_1) = x_0] \mathbb{P}[\phi_{x_0}(U_1) = x_1] \dots \mathbb{P}[\phi_{x_{n-1}}(U_n) = x_n] \\ &= \mathbb{P}[X_0 = x_0] \mathbb{P}[X_1 = x_1 | X_0 = x_0] \dots \mathbb{P}[X_n = x_n | X_{n-1} = x_{n-1}] \\ &= \mathbb{P}[X_0 = x_0] \prod_{i=0}^{n-1} p_{x_i x_{i+1}}. \end{aligned}$$

De lo cual se sigue directamente la propiedad de Markov. Por lo tanto  $\{X_n\}_{n \in \mathbb{N}}$  es una cadena Markov, con distribución inicial  $\pi$  y matriz de transición  $P$ .  $\square$

El resultado anterior nos proporciona una justificación de la existencia de cadenas de Markov además de indicarnos un procedimiento para simular trayectorias de éstas. A continuación mostramos tal procedimiento explícitamente, en forma de algoritmo.

El Algoritmo 1 genera trayectorias de una cadena de Markov homogénea con distribución inicial  $\pi$  y matriz de transición  $P$  hasta el horizonte de tiempo  $m$ .

---

**Algorithm 1** Simulación de trayectorias de cadenas de Markov

---

- 1: Generar  $X_0$  de la distribución inicial  $\pi$ .
  - 2: Inicializamos  $n=1$ .
  - 3: Generamos  $X_n$  de la distribución correspondiente al renglón de la matriz de transición asociado al estado  $X_{n-1}$ .
  - 4: Si  $n$  es igual al horizonte de tiempo  $m$ , entonces detenemos el proceso, en otro caso actualizamos  $n = n + 1$  y volvemos al paso 3.
- 

**Ejemplo 2.8.1.** *Se simuló una trayectoria de una cadena de Markov usando el Algoritmo 1 con las siguientes características:*

Matriz de transición

$$P = \begin{pmatrix} 0,2 & 0,1 & 0,5 & 0,2 \\ 0,5 & 0,2 & 0,2 & 0,1 \\ 0,2 & 0,5 & 0,1 & 0,2 \\ 0,3 & 0,5 & 0,1 & 0,1 \end{pmatrix},$$

vector de distribución inicial  $\pi$

$$(0,2 \quad 0,4 \quad 0,2 \quad 0,2),$$

vector de espacio de estados  $E$

$$(1 \quad 2 \quad 3 \quad 4),$$

y horizonte de tiempo  $m=30$ . En la Figura ?? podemos observar la gráfica de la trayectoria simulada.

**Ejemplo 2.8.2. El problema de la lluvia.** Suponga que una persona se traslada todos los días de su casa a la oficina por las mañanas y regresa por las tardes. Cada vez que está a punto de trasladarse comprueba antes si está lloviendo (lo cual ocurre independientemente cada mañana o tarde con la misma probabilidad  $p$ ) para buscar uno de los dos paraguas que posee para salir. Cuando no está lloviendo no se lleva ningún paraguas.

Una motivación natural para modelar este problema es conocer la probabilidad de que la persona se moje debido a que no tiene paraguas disponibles para salir.

Notemos que la cantidad de paraguas que tendrá disponible al realizar un traslado, no dependerá de cuántos traslados haya realizado con anterioridad. De modo que podemos el siguiente proceso resultará ser una cadena de Markov:

$X_n$  = La cantidad de paraguas disponibles en el destino del  $n$ -ésimo traslado.

Notemos que el espacio de estados está dado por  $E = \{0, 1, 2\}$  y la matriz de transición correspondiente a la dinámica es:

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

Una observación es que esta modelación puede al caso en que la persona posea  $r$  paraguas, y se lleva un paraguas cada vez que está lloviendo. En tal caso el espacio de estados es  $E = \{0, 1, 2, \dots, r\}$  y la matriz de transición es de tamaño  $r + 1$ .

## 2.9. Inferencia para Cadenas de Markov

Al tratar de hacer inferencia sobre una trayectoria observada de un Cadena de Markov hay que tener cierto conocimiento previo sobre la misma. Las condiciones mínimas que se deben de cumplir la Cadena es ser irreducible, finita y aperiódica.

Supongamos una Cadena de Markov con espacio de estados  $E$ , entonces la matriz de transiciones se convierte en los parámetros que buscamos encontrar. De esta forma la verosimilitud definida para trayectoria observada a través de los estados  $\{x_i\}_{i=0,1,\dots,n}$  puede definirse como

$$L(P) = \pi(x_0) \prod_{k=1}^n P[X_k = x_k | X_{k-1} = x_{k-1}] = \pi(x_0) \prod_{i \in E} \prod_{j \in E} p_{ij}^{n_{ij}}$$

donde  $n_{ij}$  son los saltos al estado  $j$  desde el estado  $i$ , entonces la log-verosimilitud es

$$\ell_n(P) = \log(\pi(x_0)) + \sum_{i \in E} \sum_{j \in E} n_{ij} \log(p_{ij})$$

con lo cual al maximizar con respecto de las probabilidades encontraremos los estimadores  $\hat{p}_{ij}$ .

Utilizaremos el método de maximización mediante multiplicadores de Lagrange debido a que tenemos una condición determinante sobre estos valores estimados de la matriz estocástica que es cumplir que  $\sum_{j \in E} p_{ij} = 1, \forall i, j \in E$ .

Así pues vamos a plantear de nuevo el problema como maximizar

$$\ell_n(P) = \log(\pi(x_0)) + \sum_{i \in E} \sum_{j \in E} n_{ij} \log(p_{ij}) + \sum_{i \in E} \lambda_i (1 - \sum_{j \in E} p_{ij}),$$

sea  $n_i = \sum_{j \in E} n_{ij}$ , entonces al derivar con respecto de  $p_{ij}$  y de  $\lambda_i$  se tiene que

$$\frac{\partial \ell_n}{\partial p_{ij}} = \frac{n_{ij}}{p_{ij}} - \lambda_i = 0$$

o bien

$$p_{ij} = \frac{n_{ij}}{\lambda_i}$$

por tanto

$$1 = \sum_{j \in E} p_{ij} = \sum_{j \in E} \frac{n_{ij}}{\lambda_i} = \frac{n_i}{\lambda_i}$$

entonces  $\lambda_i = n_i$ , y por tanto  $\hat{p}_{ij} = \frac{n_{ij}}{n_i}$ .

## 2.10. Ejemplos

### Ejemplo 2.10.1. Calificaciones crediticias

En los contratos de crédito, las posibles pérdidas económicas involucran la estimación de las probabilidades de incumplimiento por parte de los acreditados.

El acreditado es clasificado según la capacidad de cumplir con sus obligaciones en un contrato de crédito. En el campo de incumplimiento por parte de las empresas, en un contrato de crédito, las matrices de



|     | Aaa   | Aa    | A     | Baa   | Ba    | B     | C     | D    |
|-----|-------|-------|-------|-------|-------|-------|-------|------|
| Aaa | 0.95  | 0.02  | 0.015 | 0.009 | 0.006 | 0     | 0     | 0    |
| Aa  | 0.06  | 0.91  | 0.01  | 0.007 | 0.005 | 0.004 | 0.004 | 0    |
| A   | 0.01  | 0.04  | 0.88  | 0.03  | 0.019 | 0.011 | 0.01  | 0.0  |
| Baa | 0.002 | 0.003 | 0.005 | 0.88  | 0.05  | 0.03  | 0.02  | 0.01 |
| Ba  | 0.001 | 0.005 | 0.005 | 0.009 | 0.85  | 0.05  | 0.05  | 0.04 |
| B   | 0     | 0     | 0.003 | 0.007 | 0.02  | 0.88  | 0.05  | 0.04 |
| C   | 0     | 0     | 0.002 | 0.003 | 0.005 | 0.09  | 0.79  | 0.11 |
| D   | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 1    |

Cuadro 2.1: Matriz de Transición de Calificación Crediticia Moody's 2000.

transición representan las probabilidades de pasar de una calificación a otra en un intervalo de tiempo.

Las clasificaciones son emitidas por las compañías calificadoras. El papel de estas compañías en los mercados globales de capital es tener una medición del riesgo de crédito de manera independiente, creíble y objetiva; una cobertura comprensible y consistencia global; una transparencia crediticia y un aumento de la eficiencia y la liquidez.

El mercado de calificadoras de crédito está dominado por dos compañías: **Standard and Poors (S&P)** y **Moody's Investor Services (Moody's)**. Una calificación de crédito representa una tasa global del cumplimiento de o e las obligaciones por parte del acreditado.

La tabla muestra una matriz de transición anual con las calificaciones de crédito otorgadas por Moody's.

## 2.11. Cadenas de Markov a tiempo discreto en Seguros

Es esta sección presentamos un modelo para la estimación de las primas de seguro de automóviles.

### 2.11.1. El sistema Bonus-Malus

En Hong Kong y en otros lugares del mundo, se usa un sistema para fijar las primas de seguro de automóvil conocido como Bonus-Malus que consiste de 6 clases de tarificación, de 1 fort bonus a 6 fort malus, que se rige por las siguientes reglas:

- si un asegurado no tuvo siniestros durante el periodo, entonces pasa de la categoría  $i$  a la categoría  $\max\{1, i - 1\}$ ,
- si el asegurado tiene al menos un siniestro entonces pasa de la categoría  $i$  a la 6.

Si  $X_n$  denota la categoría en cual se encuentra un individuo al  $n$ -ésimo periodo entonces  $\{X_n\}_{n \geq 0}$  es una cadena de Markov con espacio de estados  $\{1, 2, \dots, 6\}$ . Si  $p \in (0, 1)$  es la probabilidad de no presentar

sinistros en cada periodo entonces la correspondiente matriz de transición es:

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

Supongamos que estamos interesados en resolver los dos siguientes problemas

1. **¿Cual es la proporción de tiempo que un cliente pasa en el estado  $j$  después de  $n$  pasos, dado que  $X_0 = j$ ?**

Dado que la cadena es irreducible y con espacio de estados finito la única distribución invariante está dada por  $\nu = (p^5, p^4(1-p), p^3(1-p), p^2(1-p), p(1-p), (1-p))$ , que corresponde a la proporción buscada.

2. **¿Cual es la prima promedio que paga un asegurado?**

Denotemos por  $g$  una función que determina la prima a pagar en función de la categoría.  $g$  es una función definida en  $\{1, 2, \dots, 6\}$  no-decreciente que toma sólo valores positivos. La prima promedio que pagará un individuo en  $n$  periodos será entonces:

$$\frac{1}{n} \sum_{i=1}^n g(X_i) = \sum_{i=1}^6 g(i) \frac{V_i^{(n)}}{n} \sim \sum_{i=1}^6 \nu_i g(i),$$

con probabilidad 1.

# Capítulo 3

## Proceso de Poisson

En esta capítulo estudiaremos las principales características del proceso de conteo más usado en el contexto actuarial, el proceso de Poisson.

### 3.1. Definiciones

Podemos definir, en general, a un proceso de conteo  $\mathbf{N} = \{N_t\}_{t \geq 0}$  como un procesos estocástico con trayectorias constantes por pedazos con valores en los naturales y que va incrementando de uno en uno en tiempos aleatorios. Entonces, interpretamos a  $N_t$  como el número de eventos ocurridos hasta el tiempo  $t$ .

Si denotamos por  $0 = S_0 < S_1 < \dots$  a los tiempos en los que el proceso incrementa su valor, entonces tenemos que

$$N_t = \text{máx}\{n \in \mathbb{N} : S_n \leq t\}.$$

**Nota 3.1.** *Las siguientes equivalencias de eventos son importantes:*

- $\{N_t = n\} = \{S_n \leq t < S_{n+1}\}.$
- $\{N_t \geq n\} = \{S_n \leq t\}.$

**Nota 3.2.** *A los tiempos  $T_i = S_i - S_{i-1}$  les llamaremos tiempos de interarribo que son los tiempo aleatorios entre la ocurrencia de dos eventos seguidos*

Ahora daremos una definición formal de un proceso de conteo.

**Definición 3.1.** *Un proceso estocástico  $\mathbf{N} = \{N_t\}_{t \geq 0}$  se dice que es un proceso de conteo si  $N_t$  representa el número total de eventos que han ocurrido hasta el tiempo  $t$ .*

**Nota 3.3.** *De la definición se tiene que:*

1.  $N_t \geq 0.$
2.  $N_t$  es un entero no negativo.

3. Si  $s < t$  entonces  $N_s \leq N_t$ .
4. Para  $s < t$ ,  $N_t - N_s$  es el número de eventos que ocurren en el intervalo de tiempo  $(s, t)$ .

Ahora daremos un par de definiciones del proceso de conteo que es nuestro foco de interés.

**Definición 3.2. (Definición 1 P.P.)** Un proceso de conteo  $N = \{N_t\}_{t \geq 0}$  es de Poisson con parámetro  $\lambda > 0$  (intensidad del proceso) si satisface las siguientes condiciones:

1.  $N_0 = 0$ .
2. Tiene incrementos independientes.
3. El número de eventos en cualquier intervalo de longitud  $t$  se distribuye Poisson de parámetro  $\lambda t$ .

**Definición 3.3. (Definición 2 P.P.)** Un proceso de conteo  $N = \{N_t\}_{t \geq 0}$  es de Poisson con parámetro  $\lambda > 0$  (intensidad del proceso) si satisface las siguientes condiciones:

1.  $N_0 = 0$ .
2. Tiene incrementos independientes.
3. Tiene incrementos estacionarios.
4.  $P(N_{t+h} - N_t \geq 1) = \lambda h + o(h)$ .
5.  $P(N_{t+h} - N_t \geq 2) = o(h)$  para  $t \geq 0$  cuando  $h \rightarrow 0$ .

**Proposición 3.1.** Las definiciones 3.2 y 3.3 son equivalentes.

Veamos algunos resultados importantes del proceso de Poisson.

**Proposición 3.2.** La sucesión de variables aleatorias  $\{T_n\}$  (tiempos de interarribo entre eventos) de un proceso de Poisson de parámetro  $\lambda$  son independientes y con la misma distribución exponencial de parámetro  $\lambda$ .

Consecuencia inmediata de esta proposición se tiene el siguiente corolario.

**Corolario 3.1.** Para sucesión de variables aleatorias  $\{S_n\}$  (tiempos de ocurrencia de eventos) definidas anteriormente se tiene que:

$$S_n \sim \text{Gamma}(n, \lambda).$$

Hay una definición alternativa de proceso de Poisson que facilita hacer cálculos.

**Definición 3.4.** Un proceso de conteo  $N = \{N_t\}_{t \geq 0}$  es de Poisson si y sólo si existe  $\lambda > 0$  tal que los tiempos de interarribo son variables aleatorias exponenciales independientes de parámetro  $\lambda$ .

Consecuencias inmediatas de esta definición son que se pueden calcular explícitamente las distribuciones de tiempos de interarribo, los tiempos de ocurrencia y que se tiene conocida la distribución exacta del proceso de conteo.

Veamos un ejemplo.

**Ejemplo 3.1.1.** *Suponga que las reclamaciones es una compa?ia de seguros ocurren de acuerdo a un proceso de Poisson con intensidad diaria  $\lambda = 2$*

1. ¿Cuál es la probabilidad que en una semana se presenten 5 reclamaciones?

$$P(N_7 = 5) = \frac{e^{-14}14^5}{5!}.$$

2. ¿Cuál es el número de reclamaciones esperadas en un mes?

$$E[N_{30}] = 60.$$

3. ¿Cuál es el tiempo esperado hasta la octava reclamación?

$$E[S_8] = 4.$$

### 3.1.1. Propiedades principales

Sea  $\{N_t\}_{t \geq 0}$  un proceso de Poisson con parámetro  $\lambda$ , supongamos que cada tiempo de ocurrencia de un evento está clasificado como un evento tipo I o tipo II. Además supongamos que es tipo I con probabilidad  $p$  y tipo II con probabilidad  $1 - p$ . Si  $N_t^1$  y  $N_t^2$  denotan el número de eventos tipo I y II respectivamente que ocurren en  $[0, t]$ . El siguiente resultado modela el comportamiento en el tiempo de un escenario como el que acabamos de describir.

**Proposición 3.3.**  $\{N_t^1\}_{t \geq 0}$  y  $\{N_t^2\}_{t \geq 0}$  son procesos de Poisson independientes de parámetros  $p\lambda$  y  $(1-p)\lambda$  respectivamente.

**Nota 3.4.** Es fácil generalizar este resultado cuando se tienen  $k$  tipos de evento, cada uno con probabilidad  $p_i \in (0, 1)$  de ocurrencia y  $\sum_{i=1}^k p_i = 1$ . Entonces se tiene que  $\{N_t^1\}_{t \geq 0}, \{N_t^2\}_{t \geq 0}, \dots, \{N_t^k\}_{t \geq 0}$  procesos de Poisson independientes de parámetros  $p_1\lambda, p_2\lambda, \dots, p_k\lambda$  respectivamente.

**Ejemplo 3.1.2.** *Suponga que las reclamaciones se han categorizado según el monto de reclamación. Si se tienen 5 diferentes categorías con vector de probabilidades  $p = (1/9, 1/18, 3/18, 1/18, 10/18)$  para cada categoría. Si además suponemos que las reclamaciones se presentan de acuerdo a un proceso de Poisson con intensidad diaria  $\lambda = 3$ .*

1. ¿Cuál es la probabilidad que en una semana se presenten 5 reclamaciones de la categoría 5?
2. ¿Cuál es el número de reclamaciones esperadas en un mes de cada categoría?
3. ¿Cuál es el tiempo esperado hasta la octava reclamación de la categoría 1?
4. Si en una semana se presentan 3 reclamaciones, ¿cuál es la probabilidad que todas sean del mismo tipo?

Ahora, supongamos que sabemos que exactamente un evento ha ocurrido hasta el tiempo  $t$  y estamos interesados en la distribución del tiempo al cual éste ha ocurrido. Es razonable pensar que el tiempo está uniformemente distribuido en  $[0, t]$ . Supongamos que  $s \leq t$ , entonces

$$P(T_1 \leq s | N_t = 1) = \frac{P(N_s = 1, N_{t-s} = 0)}{P(N_t = 1)} = \frac{s}{t}.$$

Generalizando esta idea tenemos el siguiente resultado.

**Teorema 3.1.** *Dado que  $N_t = n$ , los  $n$  tiempos de ocurrencia tienen la misma distribución que las estadísticas de orden correspondientes a  $n$  variables aleatorias independientes e idénticamente distribuidas uniformes en  $[0, t]$ , es decir,*

$$f_{S_1, \dots, S_n | N_t}(s_1, \dots, s_n | n) = \frac{n!}{t^n}.$$

### 3.1.2. El proceso Poisson compuesto

**Definición 3.5.** *Sea  $\{N_t\}_{t \geq 0}$  un proceso Poisson con tasa  $\lambda$  y sean  $\{Y_i\}_{i \geq 0}$  una familia de variables aleatorias independientes entre sí e idénticamente distribuidas. Suponga que el proceso Poisson  $\{N_t\}_{t \geq 0}$  y la sucesión  $\{Y_i\}_{i \geq 0}$  son independientes. Definimos al proceso aleatorio  $\{X_t\}_{t \geq 0}$  como un **proceso Poisson compuesto** si para  $t \geq 0$ ,*

$$X_t = \sum_{i=1}^{N_t} Y_i.$$

Entonces,  $X_t$  es una suma Poisson de variables aleatorias.

**Proposición 3.4.**  $E[X_t] = \lambda t E[Y]$ .

Veamos algunos ejemplos.

#### Ejemplo 3.1.3. Retiro de efectivo del banco

Suponga que los clientes de cierto banco ingresan a una sucursal para realizar retiros de efectivo en ventanilla de acuerdo a un proceso Poisson de tasa  $\lambda = 1/3$  (tres clientes por minuto). Suponga que cada cliente retira efectivo de dicho banco en sólo una ocasión al día y que el monto retirado es una variable aleatoria independiente e idénticamente distribuida log-normal con media  $\mu = 7$  y desviación típica  $\sigma = 1,5$ . El monto total de los retiros al tiempo  $t$  es un proceso Poisson compuesto

$$X(t) = \sum_{i=0}^{N_t} Y_i, \quad \text{con} \quad Y_0 = 0.$$

Ahora, supongamos que en este caso se desea calcular el monto total de los retiros en el transcurso de una hora,  $T = 60$ , es decir,  $E[X_{60}]$ . Sabemos que la distribución log-normal tiene función de densidad

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{(\log(x) - \mu)^2}{2\sigma^2}\right\},$$

donde  $\mu$  y  $\sigma$  son la media y la desviación típica del logaritmo. Para cualquier variable aleatoria log-normal:

- la media es:  $E(X) = \exp\{\mu + \sigma^2/2\}$ , y
- la varianza es:  $Var(X) = (e^{\sigma^2} - 1) e^{2\mu + \sigma^2}$

Entonces

$$\begin{aligned}
E(X(t)) &= E \left[ \sum_{i=1}^{N(t)} Y_i \right] \\
&= E \left[ E \left[ \sum_{i=1}^{N(t)} Y_i \mid N(t) \right] \right] \\
&= E [N(t) E [Y_i]] = E [Y_i] E [N(t)].
\end{aligned}$$

Por otra parte, sabemos que:  $N(t) \sim Poi(\lambda t)$ , entonces  $E[N(t)] = \lambda t$  y  $Var[N(t)] = \lambda t$  y también que  $E[Y_i] = e^{\mu+\sigma^2/2}$  y  $Var[Y_i] = (e^{\sigma^2} - 1) e^{2\mu+\sigma^2}$  entonces,

$$\begin{aligned}
E[X(t)] &= E[Y_i] E[N(t)] \\
&= e^{7+(1,5)^2/2} \lambda t \\
&= e^{8,125} (1/3) 60 \\
&\approx 67557,36.
\end{aligned}$$

**Ejemplo 3.1.4. El modelo clásico de Cramér-Lundberg.**

Este modelo tiene sus orígenes en la tesis doctoral de Filip Lundberg defendida en el año de 1903. En este trabajo Lundberg analiza el reaseguro de riesgos colectivos y presenta el proceso de Poisson compuesto. En 1930 Harald Cramér retoma las ideas originales de Lundberg y las pone en el contexto de los procesos estocástico. El modelo ha sido estudiado en extenso y se han propuesto varias formas de generalizarlo.

El modelo que estudiaremos es el proceso a tiempo continuo  $\{C_t\}_{t \geq 0}$  dado por

$$C_t = u + ct - \sum_{j=1}^{N_t} Y_j$$

donde:

- $u$  es el capital inicial de la compañía aseguradora
- $ct$  es la entrada por primas hasta el tiempo  $t$  con  $c$  una constante positiva
- $Y_j$  es el moneto de la  $j$ -ésima reclamación, y
- $N_t$  es un proceso de Poisson de parámetro  $\lambda$ .

$C_t$  representa el balance más sencillo de ingresos menos egresos de una compañía aseguradora. Al proceso  $C_t$  se le llama proceso de riesgo o proceso de superávit.

La variable aleatoria  $C_t$  se puede interpretar como el capital de la compañía aseguradora al tiempo  $t$  y por razones naturales y legales es importante que  $C_t$  permanezca por arriba de cierto nivel mínimo. Cuando

$C_t < 0$  se dice que hay ruina. La ruina casi nunca sucede en la práctica, es solamente un término técnico que produce alguna toma de decisión.

Por ejemplo, si el capital de una compañía aseguradora asignado a una cartera decrece en forma significativa, ésta automáticamente puede tomar ciertas medidas para subsanar esta situación y no se trata de un evento insalvable.

**Ejemplo 3.1.5.** *Supongamos que en el modelo de Cramer Lundberg el monto de reclamaciones siguen una distribución uniforme en  $(0, 100)$  y éstas se presentan con una intensidad  $\lambda = 3$ . Modelar el capital esperado de la compañía aseguradora en el tiempo.*

### 3.1.3. Proceso de Poisson no homogéneo

Un proceso de conteo extremadamente importante para propósitos de modelación es el proceso Poisson no homogéneo, el cual omite la suposición de incrementos estacionarios del proceso Poisson. Esto abre la posibilidad a que la tasa de arribo no necesariamente sea constante, sino que pueda variar en el tiempo.

**Definición 3.6.** *Un **Proceso de Poisson no homogéneo** es un proceso a tiempo continuo  $\{N_t\}_{t \geq 0}$  y espacio de estados  $E = \{0, 1, 2, \dots\}$  con parámetro una función positiva y localmente integrable  $\lambda(t)$ , que satisface:*

1.  $N_0 = 0$
2. Tiene incremento independientes
3. Para cualquier  $t \geq 0$  y  $h > 0$  decreciente a cero se tiene que
  - a)  $P(N_{t+h} - N_t \geq 1) = \lambda(t)h + o(h)$
  - b)  $P(N_{t+h} - N_t \geq 2) = o(h)$

El resultado análogo en cuanto a la distribución del número de eventos en un proceso de Poisson no homogéneo es el siguiente:

**Proposición 3.5.** *La variable aleatoria  $N_t$  en un proceso de Poisson no homogéneo tiene distribución Poisson( $\Lambda(t)$ ), donde  $\Lambda(t) = \int_0^t \lambda(s)ds$  donde  $\lambda(t)$  es la intensidad al tiempo  $t$ .*

De este resultado se sigue inmediatamente el siguiente resultado.

**Proposición 3.6.**  $N_{t+s} - N_t \sim \text{Poisson}(\Lambda(t+s) - \Lambda(t))$ .

**Ejemplo 3.1.6.** *Para los ejemplos 3.1.1 y 3.1.2 consideremos el supuesto que  $\lambda(t) = t/2$ . Resolver numericamente los ejemplos 3.1.1 y 3.1.2.*

## 3.2. Simulación de un proceso de Poisson

### 3.2.1. Simulación de un proceso de Poisson homogéneo

Suponga que se quieren generar los primeros  $n$  eventos de un proceso Poisson de intensidad  $\lambda$ . Para ello nos apoyaremos del resultado de que el tiempo transcurrido entre dos eventos sucesivos cualesquiera son



una sucesión de variables aleatorias independientes exponenciales de parámetro  $\lambda$ . De manera que para generar un proceso Poisson basta con generar esos tiempos de interarribo. Podemos usar el Algoritmo 2 para dicho propósito.

---

**Algorithm 2** Simulación de tiempos de ocurrencia de  $n$  eventos en un proceso de Poisson de parámetro  $\lambda > 0$

---

- 1: Hacemos  $k = 0$  y  $S_0 = 1$ .
  - 2: Generamos  $T_{k+1} \sim \text{Exp}(\lambda)$ .
  - 3: Hacemos  $S_{k+1} = S_k + T_{k+1}$
  - 4: Hacemos  $k = k + 1$ .
  - 5: Si  $k = n$  terminamos, en otro caso volvemos al paso 2.
- 

Supongamos que queremos generar una trayectoria de un proceso de Poisson en el intervalo finito  $[0, T]$ , podemos seguir el Algoritmo 3

---

**Algorithm 3** Genera trayectorias de un proceso de Poisson en  $[0, T]$

---

- 1: Hacer  $t = 0, n = 0$
  - 2: Generar  $s \sim \text{Exp}(\lambda)$ .
  - 3: Hacer  $t = t + s$ . Si  $t > T$ , detenerse.
  - 4: En caso contrario, hacer  $n = n + 1$  y  $N_n = t$
  - 5: Ir al paso 2.
- 

### 3.2.2. Simulación de Proceso de Poisson no homogéneo

El proceso de conteo extremadamente importante para propósitos de modelación es el proceso Poisson no homogéneo, el cual relaja la suposición de incrementos estacionarios del proceso Poisson. Esto abre la posibilidad a que la tasa de arribo no necesariamente sea constante, sino que pueda variar en el tiempo, en esta sección se presenta un algoritmo para simular trayectorias de este proceso.

**Proposición 3.7.** *Consideremos un proceso de Poisson no homogéneo  $N = \{N_t\}_{t \geq 0}$  como en la Definición 3.6. Supongamos que  $\sup_{t > 0} \lambda(t) \leq \lambda$ . Sea  $N^* = \{N_t^*\}_{t \geq 0}$  un proceso de Poisson con parámetro  $\lambda$  con tiempos de ocurrencia  $\{S_n^*\}_{n \geq 1}$  e independiente de una sucesión de variables aleatorias  $\{U_n\}_{n \geq 1}$  independientes e idénticamente distribuidas uniformes en el intervalo  $(0, 1)$ . Entonces, se puede simular a  $N$  haciendo*

$$N_t = \sum_{n \geq 1} \mathbb{I}_{\{S_n^* \leq t, U_n \leq \frac{\lambda(T_n^*)}{\lambda}\}}$$

La demostración se deja como ejercicio para el lector.

Suponga que deseamos simular una trayectoria de un proceso de Poisson no homogéneo en un intervalo de tiempo  $[0, T]$  con función de intensidad  $\lambda(t)$ . El método que revisaremos a continuación se conoce con el nombre de **muestreo aleatorio**.

Este comienza eligiendo un valor  $\lambda$  el cual es tal que

$$\lambda(t) \leq \lambda \quad \text{para toda } t \leq T.$$

Tal proceso Poisson no homogéneo puede generarse seleccionando aleatoriamente el evento tiempos de un proceso Poisson con tasa  $\lambda$ . Es decir, si un evento de un proceso Poisson con tasa  $\lambda$  que ocurre al tiempo  $t$

es contabilizado con probabilidad  $\lambda(t)/\lambda$ , entonces el proceso de los eventos contabilizados es un proceso Poisson no homogéneo con función de intensidad  $\lambda(t)$  para  $0 \leq t \leq T$ .

Por lo tanto, simulando un proceso Poisson y contabilizando aleatoriamente sus eventos, podemos generar un proceso Poisson no homogéneo. De esta forma podemos describir el algoritmo deseado como sigue

---

**Algorithm 4** Genera Proceso de Poisson no homogéneo

---

- 1: Hacer  $t = 0$ ,  $n = 0$
  - 2: Generar  $s \sim \text{Exp}(\lambda)$ .
  - 3: Hacer  $t = t + s$ . Si  $t > T$ , detenerse.
  - 4: En caso contrario, En caso contrario, generar otro número aleatorio  $U \sim U(0, 1)$
  - 5: Si  $U \leq \lambda(t)/\lambda$ , hacer  $n = n + 1$  y  $S_n = t$
  - 6: Regresar al paso 2.
- 

**Nota 3.5.** La justificación de que el Algoritmo 4 genera trayectorias de un proceso de Poisson no homogéneo se sigue de la Proposición 3.7.

En la Figura 3.1 podremos apreciar la forma en que actúa la función de intensidad con la ocurrencia de saltos cuando tenemos  $\lambda(t) = t/2$  y  $\lambda(t) = 1 + \text{sen}(t)$ .

### 3.2.3. Simulación de Proceso de Poisson compuesto

En esta sección formalizamos de manera algorítmica la simulación de un proceso de Poisson compuesto  $\mathbf{X} = \{X_t\}_{t \geq 0}$  donde

$$X_t = \sum_{i=1}^{N_t} Y_i,$$

con  $N = \{N_t\}_{t \geq 0}$  un proceso de Poisson y  $\{Y_i\}_{i \geq 1}$  una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución  $F$ . El siguiente algoritmo describe la manera de simular a  $X$  en el intervalo de tiempo  $[0, T]$ .

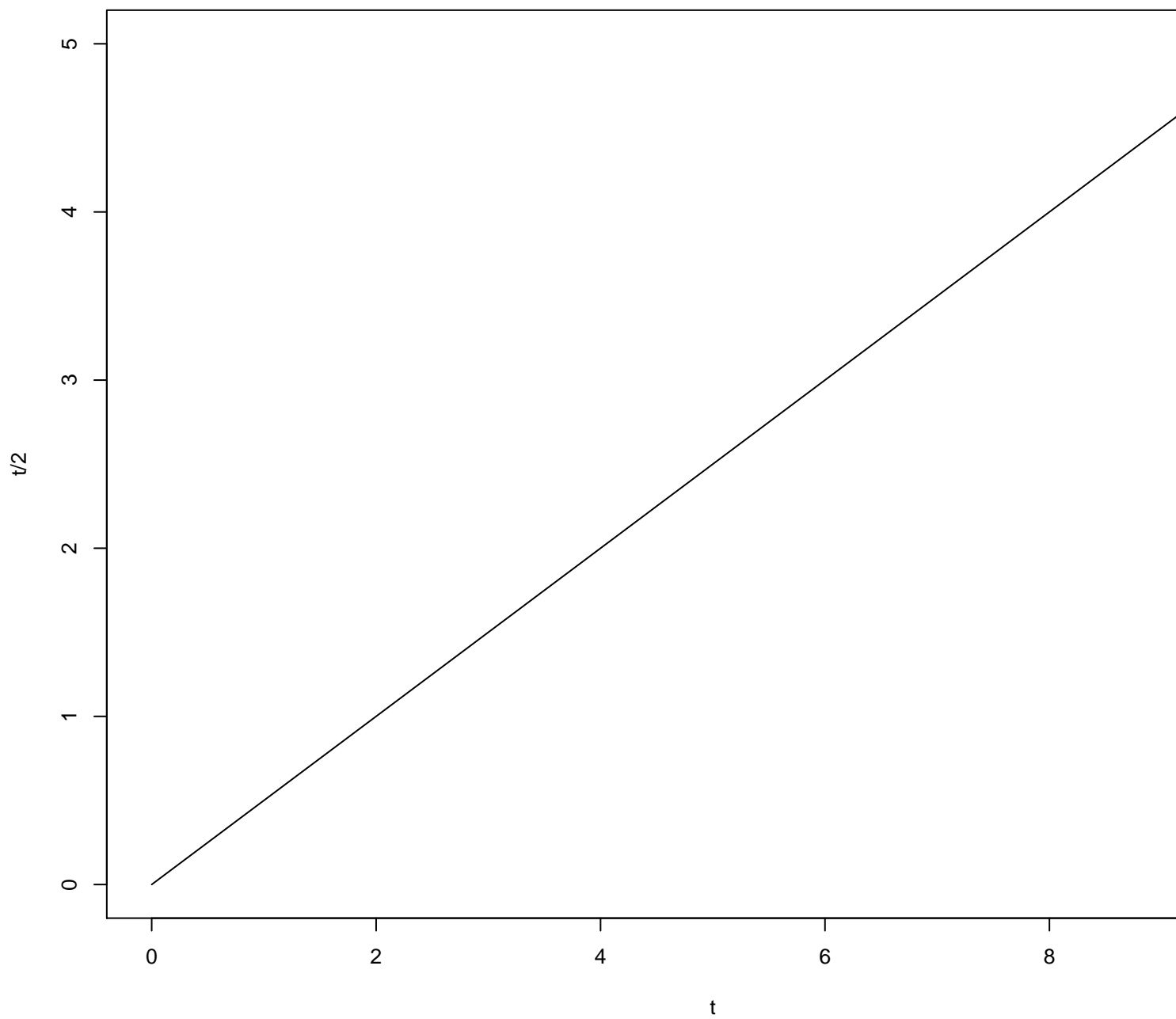
---

**Algorithm 5** Genera Proceso de Poisson Compuesto

---

- 1: Hacer  $t = 0$ ,  $n = 0$ .
  - 2: Generar  $s \sim \text{Exp}(\lambda)$ .
  - 3: Hacer  $t = t + s$ . Si  $t > T$ , detenerse.
  - 4: En caso contrario, generar hacer  $n = n + 1$  ,y  $S_n = t$
  - 5: Generar  $Y_n \sim F$ .
  - 6: Hacer  $X_t = \sum_{i=1}^n Y_i$
  - 7: al paso 2.
- 

Con el Algoritmo 5 podemos también implementar la simulación del modelo de Cramér-Lundberg, con el que es posible obtener trayectorias del capital de una compañía de seguros a través del tiempo para estimar resultados de interés, como la probabilidad de ruina o tiempo esperado de ruina, como se muestra en la Figura 3.2.3 para el caso de arribos  $\text{exp}(.4)$  y saltos  $\text{exp}(.8)$  con un capital inicial de 10 unidades.



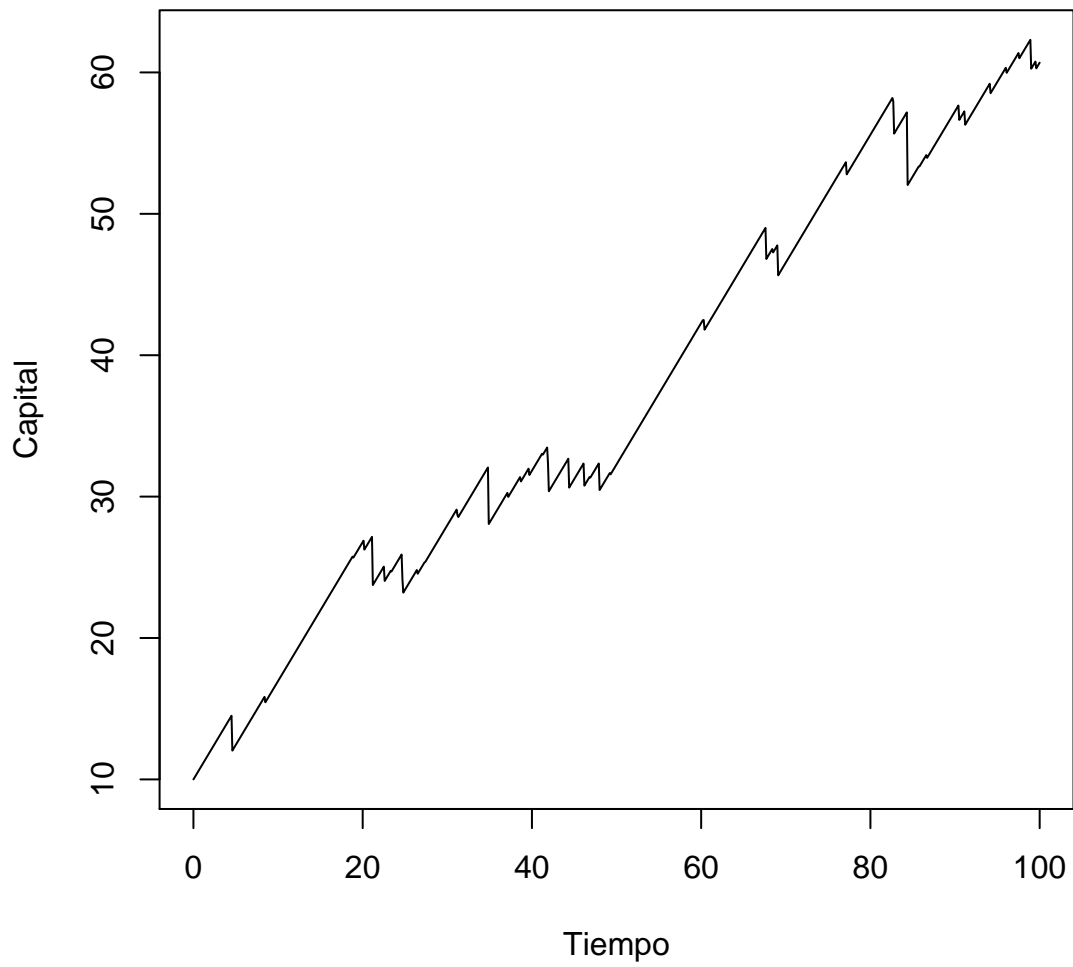


Figura 3.2: Modelo de Cramer-Lundberg

# Capítulo 4

## Martingalas

**Definición 4.1.** Una sucesión creciente de  $\sigma$ -álgebras  $\{\mathcal{F}_n\}_{n \geq 1}$  es llamada filtración.

**Definición 4.2.** Dada una sucesión de variables aleatorias  $\{X_n\}_{n \geq 1}$ , su filtración natural es la sucesión de  $\sigma$ -álgebras dada por  $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$ .

**Definición 4.3.** Un proceso estocástico  $\{X_n\}_{n \geq 1}$  es adaptado a una filtración  $\{\mathcal{F}_n\}_{n \geq 1}$  si la variable aleatoria  $X_n$  es  $\mathcal{F}_n$ -medible para cada  $n \geq 1$ .

**Definición 4.4.** Un proceso estocástico  $\{X_n\}_{n \geq 1}$  es predecible respecto a una filtración  $\{\mathcal{F}_n\}_{n \geq 1}$  si la variable aleatoria  $X_n$  es  $\mathcal{F}_{n-1}$ -medible para cada  $n \geq 1$ .

**Definición 4.5.** Un **tiempo aleatorio** es una variable aleatoria

$$T : \Omega \rightarrow \mathbb{N} \cup \{\infty\}.$$

**Definición 4.6.** Un tiempo aleatorio  $T$  es un **tiempo de paro** si para  $n \in \mathbb{N}$  existe  $A_n \subset E^{n+1}$  tal que

$$\{T = n\} = \{(X_0, \dots, X_n) \in A_n\}.$$

A la variable aleatoria tiempo de paro la podemos interpretar como el tiempo que se obtiene de observar una trayectoria hasta que se cumpla cierta condición. Veamos algunos ejemplos de tiempo de paro.

**Ejemplo 4.0.1.** Sea  $T$  el tiempo en el que una cadena de Markov regresa a su estado inicial, es decir

$$T = \begin{cases} \infty & X_n \neq X_0 \text{ para toda } n \\ \min\{n \geq 1 | X_n = X_0\} & \text{e.o.c} \end{cases}$$

**Ejemplo 4.0.2.** Sea  $T_n$  el tiempo en el ocurre la  $n$ -ésima visita al estado inicial es un tiempo de paro.

**Definición 4.7.** Un proceso estocástico  $\{X_n\}_{n \geq 1}$  es una martingala respecto a una filtración  $\{\mathcal{F}_n\}_{n \geq 1}$  si:

1. Es integrable.
2. Es adaptado a la filtración  $\{\mathcal{F}_n\}_{n \geq 1}$ .
3. Si  $n \leq m$  entonces  $E[X_m | \mathcal{F}_n] = X_n$ .

**Nota 4.1.** Si  $E[X_m | \mathcal{F}_n] \geq X_n$  entonces el proceso es una submartingala y si  $E[X_m | \mathcal{F}_n] \leq X_n$  es una supermartingala.

**Ejemplo 4.0.3.** *Juego de apuestas*

**Ejemplo 4.0.4.** *Sea  $\{N_t\}_{t \geq 0}$  un proceso de Poisson de parámetro  $\lambda$ . El proceso  $\{X_t - \lambda t\}_{t \geq 0}$  es una martingala.*

**Teorema 4.1. Teorema de paro opcional.** *Sea  $\{X_n\}_{n \geq 1}$  una martingala y  $\tau$  un tiempo de paro finito con respecto a la misma filtración  $\{\mathcal{F}_n\}_{n \geq 1}$  tales que:*

1.  $X_\tau$  es integrable.
2.  $\lim_{n \rightarrow \infty} E[X_n 1_{\tau > n}] = 0$ .

*Entonces  $E(X_\tau) = E(X_n)$  para cualquier  $n \geq 1$ .*

# Capítulo 5

## Procesos con valores continuos

### 5.1. El Movimiento Browniano y sus transformaciones

**Definición 5.1.** *Movimiento Browniano Estándar*

Un Movimiento Browniano Estándar (MB) es un proceso  $\mathbf{B} = \{B_t\}_{t \geq 0}$  markoviano a tiempo continuo con incrementos estacionarios e independientes tal que

1.  $\mathbb{P}[B_0 = 0] = 1$ ,
2.  $\mathbf{B}$  tiene trayectorias continuas casi seguramente, y
3.  $B_t - B_s \sim \text{Normal}(0, t - s)$ ,  $\forall 0 \leq s \leq t$ .

Este proceso es un proceso fundamental en la teoría de procesos estocásticos. La primera razón para ello es que es un proceso continuo análogo a la caminata aleatoria simple; su dinámica es simétrica en la posición y equiprobable. La segunda razón es su **universalidad**. En matemáticas, la universalidad de un modelo significa que aparece constantemente: así como la distribución normal aparece como un límite de sumas (centradas y escaladas) de variables aleatorias independientes y arbitrarias (bajo algunos supuestos generales), el Movimiento Browniano aparece como el límite (centrado y escalado) de procesos arbitrarios (bajo algunos supuestos generales).

Un Movimiento Browniano (no estándar) de parámetros  $\mu$  (media o deriva), y  $\sigma^2$  (coeficiente de difusión o volatilidad) es un proceso que satisface las mismas propiedades que el Movimiento Browniano estándar pero tal que

$$B_t - B_s \sim \text{Normal}(\mu(t - s), \sigma^2(t - s)), \quad \forall 0 \leq s \leq t.$$

A partir del Movimiento Browniano podemos definir más procesos como transformaciones del MB.

**Definición 5.2.** *Un movimiento Browniano geométrico  $\mathbf{G} = \{G_t\}_{t \geq 0}$  está definido como la siguiente transformación del Movimiento Browniano estándar  $\mathbf{B}$ :*

$$G_t := G_0 \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right\}.$$

Cabe mencionar, que en el modelo de Black-Scholes-Merton, se supone que el activo financiero tiene un comportamiento aleatorio modelado por un Movimiento Browniano Geométrico.

**Definición 5.3.** *Un proceso de Orstein-Uhlenbeck o proceso de Vasicek es un proceso  $\mathbf{X} = \{X_t\}_{t \geq 0}$  definido como la siguiente transformación del Movimiento Browniano estándar  $\mathbf{B}$ :*

$$X_t := X_0 e^{-at} + b(1 - e^{-at}) + \frac{\sigma e^{-at}}{\sqrt{2a}} B_t(e^{2at} - 1).$$

Este modelo se utiliza ampliamente para modelar tasas de interés o demográficas. Tiene la desventaja de que tiene posibilidad de tomar valores negativos. El parámetro  $b$  representa la media a largo plazo, al parámetro  $a$  se le conoce como velocidad de reversión hacia la media, y  $\sigma$  es la volatilidad del proceso.

## 5.2. Ecuaciones diferenciales estocásticas

En esta sección presentaremos de forma no rigurosa el concepto de una Ecuación Diferencial Estocástica (SDE, por sus siglas en inglés). Presentemos primero a los procesos Markovianos con valores continuos en un cierto nivel de generalidad.

**Definición 5.4.** *Proceso de difusión*

Un proceso de difusión  $\{X_t\}_{t \geq 0}$  unidimensional es un proceso estocástico a tiempo continuo con propiedad fuerte de Markov y trayectorias continuas casi en todas partes.

Si  $I = [r, s]$ , es el espacio de estados de  $\{X_t\}_{t \geq 0}$  entonces

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}[|X_{t+h} - x| > \epsilon | X_t = x] = 0$$

$\forall x \in I$  y  $\epsilon > 0$ .

Casi todos los procesos de difusión son caracterizados por dos condiciones básicas que describen su media y varianza infinitesimales. Sea  $\Delta h X_t = X_{t+h} - X_t$  y  $x \in I$ , dadas por

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[\Delta h X_t | X_t = x] = \mu(x, t)$$

y

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}[(\Delta h X_t)^2 | X_t = x] = \sigma(x, t).$$

Asociados a los procesos estocásticos anteriores, tenemos ecuaciones diferenciales a través de los cuáles podemos definirlos (tal como algunas funciones reales pueden definirse como soluciones de ecuaciones diferenciales ordinarias).

**Definición 5.5.** *Ecuación Diferencial Estocástica*



Sean  $\mu(x, t)$  y  $\sigma(x, t)$  dos funciones  $\mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ . Una SDE es una ecuación de la forma

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t,$$

definida para los valores de  $t$  en  $[0, T]$  y con condición inicial  $X_0$  independiente del Movimiento Browniano, cuya solución representa un proceso de difusión con tales condiciones infinitesimales.

Notemos que en la ecuación diferencial aparece un diferencial relativo al Movimiento Browniano. Este diferencial no tiene sentido dentro de la teoría de integración de cálculo, ni de la teoría de integración de Stieljes o de Lebesgue, debido a la gran oscilación que tienen las trayectorias de un Movimiento Browniano. El **cálculo estocástico**, desarrollado principalmente por Kiyoshi Itô, desarrolla una teoría en donde tal diferencial adquiere sentido. Su profundidad es demasiada para exponerse aquí; tan sólo mencionaremos que es una teoría adecuada para funciones aleatorias (procesos estocásticos) que generaliza al cálculo diferencial e integral estándar al incluir la posibilidad de integrar y derivar funciones que tienen una gran oscilación (como el movimiento Browniano).

**Definición 5.6.** *Tiempo de Llegada*

Sea  $\{X_t\}_{t \geq 0}$  un Proceso de Difusión, definimos a la variable aleatoria  $T_z$ , para  $z \in I$ , como

$$T_z = \begin{cases} \inf\{t \geq 0 \mid X_t = z\} & \{t \geq 0 \mid X_t = z\} \neq \emptyset \\ \infty & e.o.c \end{cases}$$

**Definición 5.7.** *Proceso regular*

Un proceso de difusión  $\{X_t\}_{t \geq 0}$  es regular si

$$\mathbb{P}[T_z < \infty \mid X_0 = x] > 0, \forall x, z \in I.$$

**Teorema 5.1.** *Transformaciones de Procesos de Difusión*

Sea  $\{X_t\}_{t \geq 0}$  un Proceso de Difusión regular en el espacio de estados  $I$  y parámetros  $\mu(x, t)$  y  $\sigma^2(x, t)$ . Sea  $g$  estricta monótona en  $I$  con segunda derivada continua, entonces  $Y_t = g(X_t)$  define un Proceso de Difusión Regular en  $I' = [g(r), g(s)]$  y tiene parámetros

$$\mu_Y(x, t) = \frac{1}{2}\sigma^2(x, t)g''(x) + \mu(x, t)g'(x)$$

y

$$\sigma_Y^2(x, t) = \sigma^2(x, t)[g'(x)]^2$$

Es importante mencionar algunos conceptos que son importantes para abordar de manera más amplia el cálculo estocástico,

**Lema 5.1.** *Lema de Itô.*

Sea  $dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$  entonces  $Y_t = f(X_t, t)$  satisface la EDE:

$$dY_t = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}dX_t dX_t = \left(\frac{\partial f}{\partial t} + b\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}\right)dt + \sigma\frac{\partial f}{\partial x}dW_t$$

**Definición 5.8.** *Puente Browniano*

Se define como Puente Browniano aquel proceso  $\{X_t : t \in [0, 1]\}$  solución de la ecuación diferencial estocástica

$$dX_t = -\frac{X_t}{1-t}dt + dB_t$$

con condición inicial  $X_0 = 0$ .

este proceso tiene distintas caracterizaciones y cumple que

$$\lim_{t \rightarrow 1^-} X_t = 0 \text{ c.s.}$$

# Bibliografía

- [1] Caballero, Rivero, Uribe y Velarde. Cadenas de Markov. Un enfoque elemental. Number 29 in Textos, Nivel Medio. SMM, 2004.
- [2] Conover, W. J. Practical Nonparametric Statistics. Wiley and Sons, second edition, 1980.
- [3] Durrett, R. (1999) Essential of Stochastic Processes. Springer.
- [4] Halton, J. (1970) A retrospective and prospective survey of Monte Carlo method. SIAM Rev. 12, page 1-63.
- [5] Harvey, C. R. . Time-varying conditional covariances in tests of asset pricing models. Journal of Financial Economics, 24, page 289-317., 1989.
- [6] Fishman, G. (2006).A First Course in Monte Carlo. Belmont, CA : Thomson Brooks.
- [7] Gilks, W.R., Richardson, S. y Spiegelhalter, D.J. (1996). Markov Chain Monte Carlo in Practice. Chapman and Hall.
- [8] Iacus, Stefano M. Simulation and Inference for Stochastic Differential Equations. Springer, 2008.
- [9] Karlin, Samuel y Taylor A First Course in Stochastics Processes. Academia Press, New York-London, second edition, 1975.
- [10] Laguna, M. and Marklund, J. (2005). Business process modeling, simulation and design. USA: Prentice Hall.
- [11] L'Ecuyer, Pierre Efficient and Portable Combined Number Generators. Research Contributions, Montreal, Canada, 2010. <http://www.iro.umontreal.ca/~lecuyer/myftp/papers/cacm88.pdf>
- [12] Ripley, B. D. (1987) Stochastic Simulation. Wiley.
- [13] Ross, S. M. y Palmas Velasco, O. A. (1999). Simulation. México: Prentice- Hall.
- [14] Ross, S.M. (1996) Stochastic Processes. John Wiley.
- [15] Rubinstein y Kroese Simulation and the Monte Carlo Method. Wiley Series in Probability and Statistics, second edition, 2007.