

**UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN**

Facultad de Ciencias

Escuela Profesional de Matemática

**EXISTENCIA Y UNICIDAD DE UNA DISTRIBUCIÓN  
ESTACIONARIA PARA CADENAS DE MARKOV  
EN TIEMPO DISCRETO Y SU APLICACIÓN  
EN EL BUSCADOR DE GOOGLE**

**TESIS**

Presentada por:

**Bach. LUZ DELIA QUISPE HERRERA**

Para optar el Título Profesional de:

**LICENCIADO EN MATEMÁTICA**

**TACNA – PERÚ**

2024

# ACTA DE SUSTENTACIÓN DE TESIS N° 426

En la ciudad de Tacna, en el auditorio de la Facultad de Ciencias de la Universidad Nacional Jorge Basadre Grohmann, siendo a las 11:10 a.m. horas del día 27 de setiembre del 2024, estando presente el jurado calificador nominado por Resolución de Facultad N° 11029-2024-FAC/UNJBG, conformado por los siguientes docentes:

- Dr. Roberto Encarnación Supo Hallasi - Presidente
- Dr. Luis Andrés Amaya Cedrón - Secretario
- Dr. Luis Asunción López Puycan - Vocal

Acto seguido se dio lectura a la resolución correspondiente y se leyó el artículo 22 del Reglamento de la Facultad de Ciencias. A continuación, el Presidente del jurado instó al Bachiller Luz Delia Quispe Herrera a exponer la tesis titulada:

“Existencia y Unicidad de una distribución estacionaria para cadenas de Markov en tiempo discreto y su aplicación en el Buscador de Google” para optar por el título profesional de Licenciada en Matemática. Siendo las 11 : 50 a.m., la tesista concluye su exposición. Enseguida, se procedió a la formulación de preguntas por parte de los miembros del jurado calificador. Terminada esta etapa, se invita a los miembros del jurado a emitir su calificación de acuerdo al reglamento. El promedio de calificación dio el siguiente resultado: Aprobado (por unanimidad) con el calificativo de 17 (Diecisiete), de acuerdo al Reglamento de grados y títulos de la Universidad Nacional Jorge Basadre Grohmann. Siendo las 12 : 30 p.m., se dio por concluido el acto de sustentación de la tesis, firmando los señores miembros del jurado calificador en señal de conformidad.



.....  
Dr. Roberto Encarnación Supo Hallasi  
Presidente



.....  
Dr. Luis Andrés Amaya Cedrón  
Secretario



.....  
Dr. Luis Asunción López Puycan  
Vocal

## CERTIFICADO DE SIMILITUD

Yo, Luis Andrés Amaya Cedrón, en mi condición de asesor acreditado por la Resolución de Facultad R.F. N° 10636 – 2024–FACI-UN/JBG de la tesis titulada: **“EXISTENCIA Y UNICIDAD DE UNA DISTRIBUCIÓN ESTACIONARIA PARA CADENAS DE MARKOV EN TIEMPO DISCRETO Y SU APLICACIÓN EN EL BUSCADOR DE GOOGLE”**, presentada por la Bachiller Luz Delia Quispe Herrera para optar el título profesional de Licenciado en Matemática, habiendo cumplido con lo establecido en el Reglamento de originalidad y de similitud de trabajo de investigación y producción intelectual, considerando que según la revisión, evaluación y análisis realizado a través de software de similitud textual TURNITIN cuenta con el nivel de similitud cuyo porcentaje es de 7%, por lo que **CERTIFICO LA SIMILARIDAD** de la tesis, la cual está de acuerdo al nivel **PERMITIDO**, para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional**.

Se emite el presente certificado con fines de continuar con los trámites respectivos para la obtención del título profesional.



.....  
Dr. Luis Andrés Amaya Cedrón  
DNI N° 17906098 - Asesor



.....  
Bach. Luz Delia Quispe Herrera  
Tesisista

## DEDICATORIA

A Dios, mi guía y fortaleza, quien con su infinita sabiduría y amor me ha acompaña cada día, brindándome la perseverancia y la claridad para superar los desafíos que se me presentan.

A mi familia, por su amor incondicional, su apoyo constante y confianza en mí. Gracias por estar siempre a mi lado, brindándome fortaleza en los momentos de incertidumbre.

A mi hermano Jose, cuyo apoyo ha sido fundamental en este camino.

Dedico también a la persona que, con su amor incondicional, paciencia y comprensión, ha sido mi mayor apoyo en mi formación profesional. Gracias por creer en mí. Tu aliento, paciencia, orientación y dedicación han sido fundamentales en este logro.

## AGRADECIMIENTO

Al Dr. Elmer Rusbert Calderón Beltrán, por su apoyo incondicional, revisión y observaciones constantes a lo largo de esta tesis, ya que sin ello este logro no hubiera sido posible.

A mi asesor, el Dr. Luis Andrés Amaya Cedrón, por su apoyo en el desarrollo de esta investigación.

A mis profesores de Escuela Profesional de Matemática por el tiempo, dedicación y compromiso en mi formación profesional.

## RESUMEN

Las cadenas de Markov modelan procesos que evolucionan en el tiempo y tienen la propiedad de que la ocurrencia de un estado futuro depende únicamente del estado actual, un concepto de gran interés es el estudio de las condiciones necesarias para la existencia de distribución estacionaria. En este trabajo presentamos una base teórica de las cadenas de Markov a tiempo discreto con espacio de estados numerables, así como las condiciones necesarias para la existencia y unicidad de distribuciones estacionarias. Además, mostramos como esta teoría se aplica en el algoritmo del buscador de Google.

**Palabras clave:** Algoritmo pagerank, Cadenas de Markov a tiempo discreto, distribución estacionaria, distribución asintótica.

## ABSTRACT

Markov chains model processes that evolve in time and have the property that the occurrence of a future state depends only on the current state. A concept of great interest is the study of the conditions necessary for the existence of a stationary distribution. In this work we present a theoretical basis for discrete-time Markov chains with space of countable states, as well as the necessary conditions for the existence and uniqueness of stationary distributions. Furthermore, we show how this theory is applied in the Google search algorithm.

**Keywords:** Pagerank algorithm, discrete-time Markov chains, stationary distribution, asymptotic distribution.

## ÍNDICE DE CONTENIDO

<b>INTRODUCCIÓN</b>	<b>1</b>
<b>PRELIMINARES</b>	<b>6</b>
1.1. Relación de equivalencia . . . . .	6
1.2. Nociones básicas de teoría de la medida . . . . .	7
1.3. Nociones de teoría de probabilidad . . . . .	10
1.3.1. Esperanza de una variable aleatoria . . . . .	14
1.3.2. Esperanza condicional de una variable aleatoria . . . . .	16
1.4. Procesos estocásticos . . . . .	17
<b>CADENA DE MARKOV EN TIEMPO DISCRETO</b>	<b>20</b>
2.1. Introducción a las cadenas de Markov . . . . .	20
2.2. Ejemplos de modelos de cadenas de Markov . . . . .	24
2.2.1. Cadena Ehrenfest . . . . .	24
2.2.2. Problema del laberinto cerrado . . . . .	25
2.2.3. Caminata aleatoria . . . . .	26
2.3. Matriz de transición . . . . .	30
2.4. Probabilidad de Llegada . . . . .	36
2.5. Primer tiempo de retorno . . . . .	45
2.6. Número de Retorno . . . . .	51
<b>CLASIFICACIÓN DE ESTADOS</b>	<b>56</b>
3.1. Propiedad fuerte de Markov . . . . .	56
3.2. Estados Recurrentes . . . . .	61
3.3. Estados Transitorios . . . . .	65
3.4. Estados Positivamente y Nulamente Recurrentes . . . . .	74

3.5. Periodicidad y Aperiodicidad . . . . .	80
<b>EXISTENCIA Y UNICIDAD DE LAS DISTRIBUCIONES ESTACIONARIAS EN LAS CADENAS DE MARKOV</b>	<b>85</b>
4.1. Distribución Estacionaria . . . . .	86
4.1.1. Existencia de la Distribución Estacionaria . . . . .	92
4.2. Distribución de Probabilidad Asintótica . . . . .	97
4.2.1. Existencia de Distribuciones Asintóticas . . . . .	101
<b>APLICACIONES</b>	<b>106</b>
5.1. Buscador de Google . . . . .	106
5.1.1. Definición intuitiva del PageRank . . . . .	107
5.1.2. Aspectos Matemáticos . . . . .	110
5.1.3. Interpretación del vector de estado estacionario . . . . .	115
<b>CONCLUSIONES</b>	<b>121</b>
<b>REFERENCIAS BIBLIOGRÁFICAS</b>	<b>122</b>

## LISTA DE SÍMBOLOS

$\mathbb{N}$	:	Conjunto de los números naturales $\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	:	$\{0, 1, 2, 3, \dots\}$
$S$	:	Conjunto de espacio de estados
$\pi$	:	Distribución estacionaria
$\Omega$	:	Espacio muestral
$(\Omega, \mathcal{F})$	:	Espacio de medible
$(\Omega, \mathcal{F}, \mathbb{P})$	:	Espacio de probabilidad
$\mathbb{E}(X)$	:	Esperanza de la variable aleatoria $X$
$P$	:	Matriz de transición
$R_j$	:	Número de retornos al estado $j$
$Pr$	:	PageRank
$P_{ij}$	:	Probabilidad de transición del estado $i$ al $j$
$T_j^r$	:	Primer tiempo de retorno al estado $j$
$T_j$	:	Tiempo de llegada al estado $j$
$X_n$	:	Variable aleatoria en el tiempo $n$

## INTRODUCCIÓN

La cadena de Markov es una rama de la teoría de la probabilidad y procesos estocásticos, utilizado ampliamente para resolver problemas en una gran cantidad de dominios, como investigación operativa, informática y sistemas distribuidos, redes de comunicación, biología, física, química, economía, finanzas y ciencias sociales (Sericola, 2013). Este concepto se originó a finales del siglo XIX y principios del XX, destacándose el trabajo pionero de Andréi Markov, quien estableció las bases teóricas cruciales para su desarrollo. Según Çilar, (2011) (Çinlar, 2011), “Andréi Markov (1856-1922) fue uno de los primeros matemáticos en estudiar procesos estocásticos que cambian de estado en forma discreta” (*p*,210). Markov introdujo la idea de que la probabilidad de transitar a un estado futuro depende únicamente del estado presente, una propiedad conocida como la propiedad de Markov. Esta característica simplifica significativamente el análisis y modelado de sistemas complejos al permitir la predicción del comportamiento futuro bajo condiciones de incertidumbre.

La cadena de Markov es una clase fundamental de procesos estocásticos donde “el futuro es independiente del pasado dado el presente” (Ross, 2014). Esta propiedad de “falta de memoria” significa que la distribución de probabilidad de un estado futuro solo depende del estado actual, haciendo que los estados pasados sean irrelevantes para la predicción futura.

La importancia de estudiar las cadenas de Markov, es que un gran número de fenómenos tienen la propiedad de Markov, esto llevó a una gran cantidad de investigaciones en la teoría de los procesos estocásticos ver (Kolmogorov, s.f.), (Ching y Ng, 2006). Un ejemplo típico es el pronóstico del tiempo. Usando la cadena de Markov, solo se pueden usar datos dinámicos recientes o actuales para predecir el futuro, de modo que el propósito de predecir los cambios climáticos se pueda lograr fácilmente. Otro amplio uso de la Cadena de Markov podría encontrarse en el sector financiero para pronosticar el precio de las acciones. El precio de las acciones de una empresa depende únicamente de la situación general actual de la empresa y de las expectativas del mercado sobre el precio de las acciones y no tiene nada que ver con el precio de las acciones de la

empresa hace un mes. Las cadenas de Markov también se utilizan comúnmente en el campo biológico para estimar la actividad de las moléculas, la reproducción de las células y explicar modelos epidemiológicos como se menciona en las investigaciones de (Blanchet, Gallego, y Goyal, 2016) y (Suchard, Weiss, y Sinsheimer, 2001). Además las cadenas de Markov son útiles en ciertas ramas de la Física como la Termodinámica, la Mecánica Cuántica, en Meteorología, así como en la Estadística y Matemática ver (Ching y Ng, 2006). El concepto de la Cadena de Markov fue sin duda una de las mayores contribuciones de Andrew A. Markov, y motivo de muchas publicaciones y libros que estudiaron este fenómeno, el cual ha enriquecido el conocimiento en el área de la Matemática y especialmente en la Teoría de la Probabilidad.

En esta tesis se aborda la cadena de Markov homogénea en tiempo discreto con un espacio de estados numerables infinito que es un concepto importante en la teoría de la probabilidad y la estadística. Se define  $\{X_n : n \geq 0\}$  un proceso estocástico con espacio de estados  $S$ , este proceso estocástico es llamado cadena de Markov si satisface la propiedad de Markov, esta propiedad se refiere a la idea de que la probabilidad de que un sistema o proceso estocástico evolucione en el futuro es decir  $\{X_{n+1} = j\}$  sólo depende de su estado presente  $\{X_n = i\}$  y no de los estados anteriores  $\{X_{n-1} = i_{n-1}, \dots, X_0 = i_0\}$ . En otras palabras, un proceso que cumple con la propiedad de Markov no “recuerda” su historia pasada más allá del estado presente. En el estudio de cadenas de Markov discretas, uno de los elementos fundamentales es la matriz de transición o también llamada matriz estocástica  $[P_{ij}]_{i,j \in S}$ , es una matriz cuadrada de  $S \times S$ , que consisten en las probabilidades de transición entre los diferentes estados del sistema. También se estudia el primer tiempo de llegada en una cadena de Markov al subconjunto  $A \subset S$  como  $T_A = \inf\{n \geq 0 : X_n \in A\}$ , esto refiere a la probabilidad de que la cadena de Markov, empezando en un estado inicial llegue a otro estado en específico en algún momento, luego se estudia el tiempo de primer retorno, referida al número de pasos que toma la cadena de Markov para retornar a un estado específico por primera vez, lo cual es denotado por  $T_j^r := \inf\{n \geq 1 : X_n = j\}$ , este concepto nos ayuda analizar cuánto tiempo en promedio debemos esperar para que el sistema llegue a un estado dado; se estudia también, el número de retorno, referida a la cantidad de veces que la

cadena de Markov regrese a un estado en particular a lo largo de un número  $n$  de pasos, lo cual denotaremos como  $R_j = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}}$ .

Seguidamente en la tesis se estudia la clasificación de estados de cadenas de Markov, incluyendo los estados recurrentes, que son aquellos estados que siempre regresan al mismo estado infinitas veces y es denotado por  $p_{ii} = \mathbb{P}(T_i^r < \infty | X_0 = i) = \mathbb{P}(X_n = i \text{ para algún } n \geq 1 | X_0 = i) = 1$ . Luego lo estados transitorios, que tienen la propiedad de que el estado no regresa al mismo después de un tiempo finito, denotado por  $p_{ii} = \mathbb{P}(T_i^r < \infty | X_0 = i) = \mathbb{P}(X_n = i \text{ para algún } n \geq 1 | X_0 = i) < 1$ ; los estados positivamente y nulamente recurrentes, que se determinan según el tiempo promedio esperado de regreso al estado  $i$ , una vez que se inicia desde  $i$ ; finalmente se estudia la periodicidad y aperiodicidad de estados que son subclases de los estados recurrentes, resaltando que los estados de una cadena de Markov irreducible tienen el mismo periodo, pues sólo si consta de una clase comunicante.

Uno de los temas principales es el estudio de la existencia y unicidad de las distribuciones estacionarias en las cadenas de Markov, esta referida bajo que condiciones las cadenas de Markov converge a una distribución de probabilidad. Se dice que  $(\pi_i)$  es una distribución de probabilidad si  $(\pi_i) \geq 0$  para cada  $i \in S$ , y  $\sum_{i \in S} \pi_i = 1$ . Se estudia la distribución estacionaria, también conocida como distribución de equilibrio o distribución invariante, que juega un papel crucial en determinar el comportamiento a largo plazo del sistema.

Uno de los teoremas principales de este trabajo es demostrar bajo que condiciones existe la distribución estacionaria para cadenas de Markov en tiempo discreto con espacio numerable. Estas condiciones se muestran en el siguiente teorema el cual se encuentra en el Capítulo IV.

**Teorema:** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta irreducible. Entonces, existe una distribución estacionaria para la cadena de Markov si, y solamente si, la cadena de Markov es positivamente recurrente. En este caso la distribución estacionaria  $\pi = (\pi_i)_{i \in S}$  es única y cada  $\pi_i = \frac{1}{\mu(i)}$ .

Además, en este trabajo se estudia las condiciones para la existencia de la distribución asintótica, esta nos dice que independientemente del estado

inicial las distribuciones en el tiempo  $n$  tienden a estabilizarse a largo plazo. Más concretamente tenemos el siguiente teorema que puede ser encontrado en el Capítulo IV.

**Teorema:** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta irreducible y positivamente recurrente con espacio de estados  $S$ , donde cada estado es aperiódico. Entonces para cualquier distribución inicial  $\pi^{(0)}$  tenemos que  $\lim_{n \rightarrow \infty} \pi^{(n)} = \pi$ , donde  $\pi$  es una distribución estacionaria, es decir existe una única distribución asintótica para la cadena de Markov y está es estacionaria.

Se puede ver claramente que para la existencia de las distribución asintótica se requieren que la Cadena de Markov  $\{X_n: n \geq 0\}$  sea irreducible, aperiódico y positivamente recurrente.

El concepto de distribución estacionaria encuentra diversas aplicaciones en el mundo real, incluyendo la optimización de algoritmos y la modelización de fenómenos naturales y sociales. En particular, empresas como Google utilizan cadenas de Markov y distribuciones estacionarias en sus algoritmos de búsqueda y recomendación para optimizar la relevancia de los resultados que ofrecen a los usuarios. Google utiliza un algoritmo llamado PageRank, basado en cadenas de Markov, para asignar una importancia numérica a cada página web, que conforma la parte aplicativa del presente trabajo de investigación.

Los contenidos de la tesis estan organizados de la siguiente manera:

Capítulo I: El primer capítulo, sienta las bases teóricas necesarias para adentrarse en el estudio de las cadenas de Markov en tiempo discreto, proporcionando una introducción a conceptos clave como la relación de equivalencia, la teoría de medida y los procesos estocásticos, que son fundamentales para comprender en profundidad la teoría probabilística que se desarrollará en los capítulos posteriores.

Capítulo II: Se desarrolla una introducción general a la teoría de cadenas de Markov en tiempo discreto. Se aborda la definición de una cadena de Markov en tiempo discreto, matriz de transición de la cadena de Markov, la cual describe las probabilidades de transición entre los diferentes estados. Se presentan defi-

niciones clave como la probabilidad de llegada, el tiempo de primer retorno y el número de retorno, sentando las bases para comprender la evolución temporal de la cadena y la probabilidad de alcanzar ciertos estados en un momento dado.

Capítulo III: Se centra en el estudio de la clasificación de estados en las cadenas de Markov, analizando la distinción entre estados recurrentes, transitorios, positivamente nulamente recurrentes, así también podemos clasificar los estados en una cadena de Markov por el máximo común divisor de todos los tiempos posibles, que puede tomar la cadena de Markov para regresar al estado después de partir y determinar la aperiodicidad y periodicidad de una cadena de Markov.

Capítulo IV: Se presenta las condiciones necesarias y suficientes para la existencia y unicidad de las distribuciones estacionarias en las Cadenas de Markov. Se aborda en detalle cómo estas condiciones son fundamentales para determinar si una cadena de Markov convergerá hacia una distribución estacionaria después de un número suficientemente grande de pasos. Al establecer condiciones necesarias y suficientes, se proporciona un marco teórico sólido para comprender cuándo una cadena de Markov alcanzará un equilibrio a largo plazo y qué propiedades debe cumplir la distribución estacionaria asociada. Estas condiciones son esenciales para el análisis y la predicción de eventos probabilísticos en sistemas dinámicos modelados mediante cadenas de Markov.

Capítulo V: Finalmente, se exploran las aplicaciones prácticas de las cadenas de Markov, con un enfoque especial en su utilización en el algoritmo Page-Rank de Google. Se analiza cómo las cadenas de Markov son fundamentales para comprender la estructura de enlaces de páginas web y alcanzar un estado estacionario en un número finito de iteraciones, lo que tiene un impacto significativo en la eficiencia de la búsqueda y clasificación de información en la web.

*El Autor*

# CAPÍTULO I

## PRELIMINARES

En este primer capítulo abordaremos brevemente algunas definiciones generales para poder desarrollar en los próximos capítulos la teoría de cadenas de Markov a tiempo discreto. Se considera las nociones básicas de relación de equivalencia, la teoría de medida que es fundamental para entender la teoría de probabilidad y finalmente abarcaremos brevemente procesos estocásticos.

### 1.1 Relación de equivalencia

La relación de equivalencia nos permite identificar elementos de un conjunto que comparten ciertas propiedades en común, para ello deben satisfacer ciertas condiciones. Las principales referencias de este capítulo son (Dummit y Foote, 2004) y (Fraleigh, 2003).

**Definición 1** Sean un conjunto  $X \neq \emptyset$  y la relación  $\sim$ , una relación en  $X$ . Decimos que  $\sim$  es una **relación de equivalencia** en  $X$ , si satisface las siguientes condiciones :

- a) Reflexiva:  $x \sim x$  para cada  $x \in X$ ;
- b) Simetría:  $x \sim y$  entonces  $y \sim x$  para cada  $x, y \in X$ ;
- c) Transitividad:  $x \sim y$  y  $y \sim z$  entonces  $x \sim z$ , para cada  $x, y, z \in X$ .

Cuando  $\sim$  es una relación de equivalencia en  $X$ , para cada  $a \in X$  definimos la *clase de equivalencia* de  $a$  como:

$$[a] := \{x \in X : a \sim x\}.$$

**Lema 1** Sea  $\sim$  una relación de equivalencia en  $X$ ,  $x, y \in X$ . Entonces,  $[x] = [y]$  si, y solamente si,  $x \sim y$ .

**Demostración.** Consideremos que  $[x] = [y]$ , luego  $x \in [y]$  y así  $x \sim y$ . Ahora verificaremos la reciproca, si  $x \sim y$ , entonces  $x \in [y]$ , así por transitividad  $[x] \subset [y]$ , análogamente tenemos que  $y \in [x]$ , luego  $[y] \subset [x]$   $\square$

**Definición 2** Sea  $\sim$  una relación de equivalencia en  $X$ . Definimos el **espacio cociente** de  $X$  por  $\sim$  como

$$X/\sim := \{[x] : x \in X\}.$$

## 1.2 Nociones básicas de teoría de la medida

La Teoría de la Medida constituye un pilar fundamental en el área de la probabilidad y otras áreas de las matemáticas y la ciencia. En esta sección desarrollaremos brevemente algunos conceptos que serán abordados en esta tesis, esto nos ayudarán a comprender la estructura y el funcionamiento de las cadenas de Markov. Para un estudio más detallado, de teoría de la medida, se recomienda ver (Bogachev y Ruas, 2007), (Halmos, 2013), (Cohn, 2013) y (Taylor, 2006).

**Definición 3** Sea  $\Omega$  un conjunto no vacío. Una  $\sigma$ -álgebra  $\mathcal{F}$  de  $\Omega$  es una familia de subconjuntos de  $\Omega$ ,  $\mathcal{F} \subset \mathcal{P}(\Omega)$ , que satisface las siguientes propiedades:

- a)  $\emptyset \in \mathcal{F}$ ,
- b) Si  $A \in \mathcal{F}$ , entonces  $A^c \in \mathcal{F}$
- c) Si  $A_n \in \mathcal{F}$ , entonces  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ .

El par  $(\Omega, \mathcal{F})$  se llama **espacio medible**, cada elemento de  $\mathcal{F}$  es llamado **conjunto medible** o **evento**.

Consideremos  $\{\mathcal{A}_\lambda\}_{\lambda \in \Lambda}$  una familia arbitraria de  $\sigma$ -álgebras de  $\Omega$ , es un ejercicio simple mostrar que

$$\bigcap_{\lambda \in \Lambda} \mathcal{A}_\lambda$$

es una  $\sigma$ -álgebra de  $\Omega$ . Sin embargo  $\cup_{\lambda \in \Lambda} A_\lambda$  no es necesariamente una  $\sigma$ -álgebra de  $\Omega$ . Sea  $\mathcal{C} \subset \mathcal{P}(\Omega)$ , la familia

$$\sigma(\mathcal{C}) := \bigcap_{\substack{\mathcal{A}_i \supset \mathcal{C} \\ \mathcal{A}_i \text{ } \sigma\text{-álgebra}} \mathcal{A}_i$$

es llamada  $\sigma$ -álgebra generada por  $\mathcal{C}$ . Notemos que  $\sigma(\mathcal{C})$  es una  $\sigma$ -álgebra pues es la intersección de  $\sigma$ -álgebras y además es la menor  $\sigma$ -álgebra que contiene al conjunto  $\mathcal{C}$ .

La noción de  $\sigma$ -álgebra introducida anteriormente permite definir una de las más importantes  $\sigma$ -álgebras. Consideremos  $(\Omega, \mathcal{T})$  un espacio topológico, la  $\sigma$ -álgebra de Borel, es la  $\sigma$ -álgebra generada por  $\mathcal{T}$  es decir la menor  $\sigma$ -álgebra que contiene a la topología  $\mathcal{T}$ , denotaremos en este caso simplemente por  $\mathcal{B}(\mathcal{T})$  o  $\mathcal{B}$ , cada elemento de la  $\sigma$ -álgebra de Borel es llamado *conjunto de Borel* o *boreliano*.

A continuación vamos a exhibir una  $\sigma$ -álgebra asociada un conjunto discreto finito:

**Ejemplo 1** Sea  $\Omega = \{a, b, c, d\}$ , y  $\mathcal{F} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{c, d\}, \{a, c, d\}, \{b, c, d\}, \Omega\}$ . Afirmamos que  $(\Omega, \mathcal{F})$  es un espacio medible es decir que  $\mathcal{F}$  es una  $\sigma$ -álgebra en  $\Omega$ .

*Solución:* para eso necesitamos verificar los ítems a) – c) de la Definición 3.

i)  $\emptyset \in \mathcal{F}$ ,

ii)  $A_1 = \emptyset \in \mathcal{F}$ ,  $A_1^c = \Omega \in \mathcal{F}$   
 $A_2 = \{a\} \in \mathcal{F}$ ,  $A_2^c = \{a, c, d\} \in \mathcal{F}$   
 $A_3 = \{a, b\} \in \mathcal{F}$ ,  $A_3^c = \{c, d\} \in \mathcal{F}$   
 $A_4 = \{c, d\} \in \mathcal{F}$ ,  $A_4^c = \{a, b\} \in \mathcal{F}$   
 $A_5 = \{a, c, d\} \in \mathcal{F}$ ,  $A_5^c = \{b\} \in \mathcal{F}$   
 $A_6 = \{b, c, d\} \in \mathcal{F}$ ,  $A_6^c = \{a\} \in \mathcal{F}$   
 $A_7 = \Omega \in \mathcal{F}$ ,  $A_7^c = \emptyset \in \mathcal{F}$

iii)  $\{a\}, \{b\} \in \mathcal{F}$  y  $\{a\} \cup \{b\} = \{a, b\} \in \mathcal{F}$   
 $\{a\}, \{b\}, \{a, b\} \in \mathcal{F}$  y  $\{a\} \cup \{b\} \cup \{a, b\} = \{a, b\} \in \mathcal{F}$

$$\{a\}, \{b\}, \{a, b\}, \{c, d\} \in \mathcal{F} \text{ y } \{a\} \cup \{b\} \cup \{a, b\} \cup \{c, d\} = \Omega \in \mathcal{F}$$

$$\{a, c, d\}, \{b, c, d\} \in \mathcal{F} \text{ y } \{a, c, d\} \cup \{b, c, d\} = \Omega \in \mathcal{F}.$$

Análogamente con los casos restantes.

De los ítems *i) – iii)* concluimos que  $\mathcal{F}$  es una  $\sigma$ -álgebra de  $\Omega$  y así  $(\Omega, \mathcal{F})$  es un espacio medible.  $\triangle$

**Definición 4** Sean  $(\Omega, \mathcal{F})$  un espacio medible y  $f : \Omega \rightarrow \mathbb{R}$  una función real. Decimos que  $f$  es una **función medible** con respecto a la  $\sigma$ -álgebra  $\mathcal{F}$  si  $f^{-1}(B) \in \mathcal{F}$ , para cada conjunto de Borel  $B \subset \mathbb{R}$ .

En el área de procesos estocásticos a una función medible también se le llama *variable aleatoria*. Un resultado clásico de teoría de la medida que nos permite obtener una clase amplia de funciones medibles es que cualquier función continua es medible.

**Ejemplo 2** Sean  $(\Omega, \mathcal{F})$  un espacio medible y  $A \in \mathcal{F}$ . La función  $\mathbb{1}_A : \Omega \rightarrow \mathbb{R}$  definida por

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & , \text{ si } \omega \in A \\ 0 & , \text{ si } \omega \notin A \end{cases}$$

es llamada **función característica o función indicadora**.

**Definición 5** Sea  $(\Omega, \mathcal{F})$  un espacio medible. Una función  $\mu : \mathcal{F} \rightarrow [0, \infty]$  es llamada **medida** en  $\Omega$  siempre que:

- a)  $\mu(\emptyset) = 0$ ,
- b) Es  $\sigma$ -aditiva, es decir si dados  $F_1, F_2, \dots, F_n, \dots \in \mathcal{F}$  disjuntos dos a dos, entonces

$$\mu \left( \bigcup_{n=1}^{\infty} F_n \right) = \sum_{n=1}^{\infty} \mu(F_n).$$

Además la terna  $(\Omega, \mathcal{F}, \mu)$  es llamada **espacio de medida**.

Decimos que una sucesión de conjuntos medibles  $F_n \in \mathcal{F}$  disjuntos dos a dos es una *partición medible* de  $\Omega$  si  $\bigcup F_n = \Omega$ . Además, si cada  $\mu(F_n) < \infty$ , decimos que la medida  $\mu$  es  $\sigma$ -finita.

**Ejemplo 3** Sea  $\Omega = \{a, b, c, d\}$ , y  $\mathcal{F} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{c, d\}, \{a, c, d\}, \{b, c, d\}, \Omega\}$ , definimos  $\mu : \mathcal{F} \rightarrow [0, \infty]$  tal que  $\mu(A) = \text{número de elemento de } A$ . Verificar que  $\mu$  es una medida.

*Solución:*

a)  $\mu(\emptyset) = 0$ ,

b) Es  $\sigma$ -aditiva, es decir:

$$\mu(\{a\} \cup \{b\}) = \mu(\{a, b\}) = 2 = \mu(\{a\}) + \mu(\{b\}),$$

$$\mu(\{a\} \cup \{c, d\}) = \mu(\{a, c, d\}) = 3 = \mu(\{a\}) + \mu(\{c, d\}),$$

$$\mu(\{b\} \cup \{c, d\}) = \mu(\{b, c, d\}) = 3 = \mu(\{b\}) + \mu(\{c, d\}),$$

$$\mu(\{a\} \cup \{b, c, d\}) = \mu(\{a, b, c, d\}) = 4 = \mu(\{a\}) + \mu(\{b, c, d\}),$$

análogamente se puede verificar para los casos restantes. Por tanto  $\mu$  es una medida.

△

### 1.3 Nociones de teoría de probabilidad

La intuición básica en la teoría de la probabilidad es la noción de aleatoriedad, en cualquier experimento aleatorio, siempre existe incertidumbre sobre si un evento específico ocurrirá o no. Así, el concepto de probabilidad fue propuesto originalmente para explicar la incertidumbre involucrada en el resultado de un experimento aleatorio. Se asigna probabilidad a los eventos, y esta asignación de probabilidad a los eventos es la clave para cualquier evaluación probabilística. En esta sección abordaremos brevemente algunas definiciones relacionadas a la teoría de la probabilidad y sus propiedades que serán de gran utilidad en el desarrollo de este trabajo, esto desde un punto riguroso de la matemática, para más detalles el lector puede ver (Klenke, 2013), (Durrett y Durrett, 1999) y (Rolla, 2022).

En esta tesis la función medida de interés es la medida de probabilidad, así a seguir reescribiremos la Definición 5 para este contexto.

**Definición 6** Sea  $(\Omega, \mathcal{F})$  un espacio medible. Una **medida de probabilidad** o simplemente **probabilidad** es una aplicación  $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$  que satisface las siguientes propiedades:

- a)  $\mathbb{P}(\emptyset) = 0$ ,
- b)  $\mathbb{P}(\Omega) = 1$ ,
- c)  $A_1, A_2, \dots, A_n, \dots \in \mathcal{F}$  disjuntos, entonces

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Cada elemento  $A \in \mathcal{F}$  es llamado **evento** y la terna  $(\Omega, \mathcal{F}, \mathbb{P})$  es llamada **espacio de probabilidad**.

En la definición anterior si  $\Omega$  es un conjunto numerable, finito o infinito, decimos que  $(\Omega, \mathcal{F}, \mathbb{P})$  es un **espacio de probabilidad discreto**; en caso contrario es un **espacio de probabilidad continuo**.

A continuación presentaremos algunos ejemplos de espacios de probabilidad discreto y continuo:

**Ejemplo 4** Consideremos un experimento de lanzamiento de un dado, corresponde a un espacio de probabilidad discreto, pues

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

es finito numerable.

**Ejemplo 5** Consideremos el experimento aleatorio que consiste en girar una ruleta y observar el tiempo en segundos que tarda en detenerse. Como la ruleta no puede girar indefinidamente, el espacio muestral se puede formular como

$$\Omega = (0, T],$$

donde  $T \in \mathbb{R}^+$ , es el tiempo máximo que la ruleta tarda en detenerse. Notemos que este ejemplo corresponde a un espacio de probabilidad continuo.

**Ejemplo 6** Consideremos un experimento de lanzamiento de una moneda, con la probabilidad de obtener cara  $p$  y la probabilidad de obtener sello  $1 - p$ . Entonces el espacio muestra  $\Omega$  y el  $\sigma$ -álgebra son

$$\Omega = \{C, S\}, \quad \mathcal{F} = 2^\Omega = \{\emptyset, \{C\}, \{S\}, \{\Omega\}\}$$

respectivamente. La probabilidad de cada evento es

$$\mathbb{P}(\{C\}) = p$$

$$\mathbb{P}(\{S\}) = 1 - p$$

Esta probabilidad es comúnmente llamada, medida de Bernoulli en  $(\Omega, 2^\Omega)$

**Definición 7** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $A$  y  $B$  eventos tal que  $\mathbb{P}(B) > 0$ . La **probabilidad condicional** de que ocurra el evento  $A$ , dado que ha ocurrido  $B$ , se denota  $\mathbb{P}(A|B)$  y está dada por

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \quad (1)$$

Notemos que si  $B \in \mathcal{F}$ , es un evento, tal que  $\mathbb{P}(B) > 0$ , entonces es directo de la definición de probabilidad condicional que:

$$\mathbb{P}(\Omega|B) = 1 \quad (2)$$

**Lema 2** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $\{A_n\}_{n \geq 0} \subset \mathcal{F}$ , tal que  $A_i \cap A_j = \emptyset$ , para  $i \neq j$ . Entonces

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n \middle| B\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n|B), \quad (3)$$

para cada  $B \in \mathcal{F}$  con  $\mathbb{P}(B) > 0$ .

**Demostración.** Usando la definición de la probabilidad condicional se tiene:

$$\begin{aligned} \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n \middle| B\right) &= \frac{\mathbb{P}(B \cap \bigcup_{n=1}^{\infty} A_n)}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(\bigcup_{n=1}^{\infty} B \cap A_n)}{\mathbb{P}(B)}, \quad A_i \cap A_j = \emptyset \end{aligned}$$

$$\begin{aligned}
&= \frac{\sum_{n=1}^{\infty} \mathbb{P}(B \cap A_n)}{\mathbb{P}(B)} \\
&= \sum_{n=1}^{\infty} \mathbb{P}(A_n|B).
\end{aligned}$$

□

**Teorema 1 (Probabilidad total)** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $\{A_n\}_{n \geq 0}$  una partición medible de  $\Omega$ . Entonces para cualquier evento  $B$  de  $\Omega$  con  $\mathbb{P}(B) > 0$ , se cumple

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(B|A_n)\mathbb{P}(A_n).$$

**Demostración.** El evento  $B \in \mathcal{F}$ , se puede representar como la unión disjunta de la siguiente forma:

$$B = B \cap \Omega = B \cap \left( \bigcup_{n=1}^{\infty} A_n \right) = (B \cap A_1) \cup (B \cap A_2) \cup \dots = \bigcup_{n=1}^{\infty} (B \cap A_n).$$

Luego por la propiedad c) de la Definición 6 se tiene:

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(B \cap A_n).$$

Finalmente de la definición de la probabilidad condicional se obtiene:

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(B|A_n)\mathbb{P}(A_n).$$

□

**Definición 8** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $\{A_i\}_{i \geq 0}$  una sucesión de eventos en  $\mathcal{F}$ . Decimos que

a)  $\{A_i\}_{i \geq 0}$  es una **sucesión creciente de eventos** para  $A$ , si

$$A_1 \subset A_2 \subset A_3 \subset \dots \subset A$$

y  $\bigcup_{i=1}^{\infty} A_i = A$ . En este caso escribiremos simplemente por  $A_i \uparrow A$ .

b)  $\{A_i\}_{i \geq 0}$  es una **sucesión decreciente de eventos** para  $A$ , si

$$A_1 \supset A_2 \supset A_3 \supset \dots \supset A$$

y  $\bigcap_{i=1}^{\infty} A_i = A$ . En este caso escribiremos simplemente por  $A_i \downarrow A$ .

**Teorema 2** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $\{A_i\}_{i \geq 0}$  una sucesión de eventos en  $\mathcal{F}$ . Entonces

a) Si  $A_i \uparrow \bigcup_{i=1}^{\infty} A_i$ , entonces

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

b) Si  $A_i \downarrow \bigcap_{i=1}^{\infty} A_i$ , entonces

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

**Definición 9** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $A, B$  y  $C$  eventos. Se dice que los eventos  $A$  y  $B$  son **condicionalmente independientes de  $C$**  si y solo si  $\mathbb{P}(C) > 0$  y se denota como

$$\mathbb{P}(A|B, C) = \mathbb{P}(A|B), \quad (4)$$

donde  $\mathbb{P}(A|B, C)$  es la probabilidad conjunta de  $A$  y  $B$  dado  $C$ .

### 1.3.1. Esperanza de una variable aleatoria

Un principio esencial vinculado a las variables aleatorias es la noción de esperanza, valor esperado o promedio, lo cual será abordado brevemente en la presente sección.

Consideremos  $X : \Omega \rightarrow \mathbb{R}$  una variable aleatoria del espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , para cada  $k \in \mathbb{R}$  vamos a introducir las siguientes notaciones

$$\{X = k\} := \{w \in \Omega : X(w) = k\},$$

$$\{X \neq k\} := \{w \in \Omega : X(w) \neq k\},$$

$$\begin{aligned}\{X \leq k\} &:= \{w \in \Omega : X(w) \leq k\}, \\ \{X < k\} &:= \{w \in \Omega : X(w) < k\}, \\ \{X \geq k\} &:= \{w \in \Omega : X(w) \geq k\}, \\ \{X > k\} &:= \{w \in \Omega : X(w) > k\},\end{aligned}$$

notemos que cada conjunto antes definido es un evento, pues  $X$  es una variable aleatoria, es decir son elementos de  $\mathcal{F}$ .

Una variable aleatoria  $X : \Omega \rightarrow \mathbb{R}$  es llamada discreta si  $X(\Omega)$  es un conjunto numerable.

**Definición 10** Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $X : \Omega \rightarrow \mathbb{R}$  una variable aleatoria discreta, tal que  $X(\Omega) = S$ . La **esperanza de la variable aleatoria  $X$**  está dada por

$$\mathbb{E}(X) = \sum_{k \in S} k \mathbb{P}(X = k)$$

siempre que la suma esté bien definida.

La interpretación de la esperanza (esperanza matemática) de una variable aleatoria es el valor medio que toma una cierta variable aleatoria.

**Ejemplo 7** Notemos que  $\mathbb{1}_A$  es una variable aleatoria discreta. Luego su esperanza está dada por:

$$\mathbb{E}[\mathbb{1}_A] = \sum_{k \in \{0,1\}} k \mathbb{P}(\mathbb{1}_A = k) = 1 \cdot \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A).$$

Esto implica que la esperanza de una función indicadora de un evento es igual a su probabilidad. △

A continuación vamos a listar algunas propiedades referentes a la esperanza de una variable aleatoria.

**Lema 3** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $X, Y : \Omega \rightarrow \mathbb{R}$  variables aleatorias discretas con esperanza finita. Entonces

i) Si  $X = c$ , entonces

$$\mathbb{E}[X] = c,$$

donde  $c$  es una constante.

ii) Si  $X \geq 0$ , entonces

$$\mathbb{E}[X] \geq 0.$$

iii) Sean  $a, b \in \mathbb{R}$ , entonces

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y].$$

iv) Si  $X$  y  $Y$  son variables aleatorias independientes, entonces

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

### 1.3.2. Esperanza condicional de una variable aleatoria

**Definición 11** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X: \Omega \rightarrow S$  una variable aleatoria discreta en  $\Omega$  y  $A$  un evento tal que  $\mathbb{P}(A) > 0$ . La **esperanza condicional de una variable aleatoria  $X$  dado el evento  $A$**  es definido como

$$\mathbb{E}[X|A] = \sum_{k \in S} k\mathbb{P}(X = k|A),$$

siempre que la suma sea finita.

**Lema 4** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X$  una variable aleatoria discreta en  $\Omega$  y  $A$  un evento con  $\mathbb{P}(A) > 0$ . Entonces

$$\mathbb{E}[X|A] = \frac{1}{\mathbb{P}(A)}\mathbb{E}[X\mathbb{1}_A]$$

**Demostración.** Sin pérdida de generalidad consideremos  $S = \mathbb{N}_0$ , sin embargo la prueba es válida para variables aleatorias con valores reales. Notemos que la variable aleatoria discreta  $X$  se puede escribir como

$$X = \sum_{k=0}^{\infty} k\mathbb{1}_{\{X=k\}}.$$

Luego

$$\mathbb{E}[X|A] = \sum_{k=0}^{\infty} k\mathbb{P}(X = k|A)$$

$$\begin{aligned}
&= \frac{1}{\mathbb{P}(A)} \sum_{k=0}^{\infty} k \mathbb{P}(\{X = k\} \cap A) \\
&= \frac{1}{\mathbb{P}(A)} \sum_{k=0}^{\infty} k \mathbb{E}[\mathbb{1}_{\{X=k\} \cap A}] \\
&= \frac{1}{\mathbb{P}(A)} \sum_{k=0}^{\infty} k \mathbb{E}[\mathbb{1}_{\{X=k\}} \mathbb{1}_A] \\
&= \frac{1}{\mathbb{P}(A)} \mathbb{E}[\mathbb{1}_A \sum_{k=0}^{\infty} k \mathbb{1}_{\{X=k\}}] \\
&= \frac{1}{\mathbb{P}(A)} \mathbb{E}[X \mathbb{1}_A].
\end{aligned}$$

□

En el caso que  $X$  es una variable aleatoria independiente del evento  $A$  es fácil verificar que:

$$\mathbb{E}[X|A] = \mathbb{E}[X]. \quad (5)$$

En la Definición 11 el evento  $A$  es arbitrario, sin embargo podemos considerar eventos provenientes de alguna variable aleatoria. Para esto considerando  $Y: \Omega \rightarrow S$  otra variable aleatoria discreta y  $k \in S$  luego  $\{Y = k\} \in \mathcal{F}$  es un evento y la esperanza condicional de la variable aleatoria discreta  $X$  dado el evento  $\{Y = k\}$  esta dada por

$$\mathbb{E}[X|Y = k] = \sum_{n \in S} n \mathbb{P}(X = n | Y = k). \quad (6)$$

## 1.4 Procesos estocásticos

Un proceso estocástico es una herramienta matemática utilizada para el modelado de fenómenos aleatorios que evolucionan en el tiempo. Se puede describir estos fenómenos por medio de una colección de variables aleatorias. El término “*estocástico*” significa aleatorio y “*proceso*” se refiere al estado de evolución temporal de un sistema determinado. Los procesos estocásticos tienen aplicaciones en múltiples campos como la termodinámica, mecánica cuántica, meteorología, epidemiología, teoría de juegos, finanzas, matemática, estadística. etc., (Ching y Ng, 2006) y es útil cada vez que se reconoce el papel

de la aleatoriedad y la imprevisibilidad de los eventos que pueden ocurrir en momentos aleatorios. En esta sección se presentaremos algunas definiciones de procesos estocásticos que se aplicarán en los siguientes capítulos. Las fuentes principales de consulta son (Durrett y Durrett, 1999), (Ross, 1995) y (Karlin, 2014) las referencias principales de consulta son

**Definición 12** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad e  $I$  un sistema de índices. Un **proceso estocástico** es una colección de variables aleatorias  $\{X_t : t \in I\}$  en  $\Omega$ .

En la definición anterior, cuando  $I$  es un conjunto numerable  $\{X_t : t \in I\}$  es llamado de *proceso estocástico discreto*, en caso contrario *proceso estocástico continuo*. En este trabajo de tesis estamos interesados en los procesos estocásticos discretos, sin pérdida de generalidad vamos a considerar  $I$  un conjunto finito o  $\mathbb{N}$ , a menos que indiquemos lo contrario.

**Ejemplo 8** Consideremos el proceso estocástico discreto  $\{X_n : n \geq 0\}$  que modela las condiciones atmosféricas del día  $n$ -ésimo. Supongamos que  $X_n$  puede tomar valores en cuatro estados distintos.

$$A = \{\text{soleado}, \text{ventoso}, \text{nublado}, \text{lluvioso}\}$$

Una realización de este proceso estocástico, será cualquier sucesión:

$$X_0 = \text{soleado}, X_1 = \text{ventoso}, X_2 = \text{lluvioso}, \dots$$

**Ejemplo 9** Suponga que, en ciertos instantes, se está midiendo el nivel de agua de una represa, donde para cualquier tiempo  $t > 0$ .  $X(t)$  representa el nivel de agua, el cual puede tomar cualquier valor real en un intervalo fijo  $[0, L]$ , siendo  $L$  el nivel máximo de la represa. Entonces,  $\{X_t\}_{t \in \mathbb{R}^+}$  es un proceso estocástico de tiempo continuo y con espacio de estados  $S = [0, L]$ .

**Ejemplo 10** En una caja de un banco se cuenta el número de clientes en la fila en un tiempo cualquiera.

$X(t)$  : es el número de clientes que esperan para ser atendidos en la caja en algún tiempo  $t > 0$ .  $\{X_t\}_{t \in \mathbb{R}^+}$  es un proceso estocástico con espacio de estados discreto  $S = \{0, 1, 2, \dots\}$  y de tiempo continuo.

**Definición 13** Sea  $S$  un conjunto numerable. La matriz  $P = [P_{ij}]_{i,j \in S}$  es una **matriz estocástica**, si cumple las siguientes condiciones:

a)  $P_{ij} \geq 0, \quad \forall i, j \in S$

b)  $\sum_{j \in S} P_{ij} = 1, \quad \forall i \in S.$

## CAPÍTULO II

### CADENA DE MARKOV EN TIEMPO DISCRETO

En este capítulo se desarrollara de manera general una introducción a la teoría de cadenas de Markov en tiempo discreto, este es un concepto importante en la teoría de la probabilidad, procesos estocásticos y estadística, ya que estudian secuencias de variables aleatorias, donde la probabilidad de ocurrencia de cada evento depende únicamente del evento anterior, sin considerar el historial completo de eventos. Una cadena de Markov en tiempo discreto está definida por variables aleatorias que toman valores en un conjunto finito o numerable llamados estados. En este capítulo abordaremos la matriz de transición, probabilidad de llegada, tiempo de primer retorno y número de retorno, estableciendo las definiciones básicas de cadenas de Markov para en el próximo capítulo abordar el concepto de clasificación de estados. Existe mucha bibliografía con respecto a este tema la cual puede ser encontrada en textos de procesos estocásticos, para un estudio más detallado se recomienda ver (Neill, 2022; Kelbert, 2008; Norris, 1998; Privault, 2013; Seneta, 2016).

#### 2.1 Introducción a las cadenas de Markov

A lo largo de toda la tesis vamos a trabajar en  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Así cuando nos referimos a una variable aleatoria estamos haciendo referencia a una función  $X: \Omega \rightarrow \mathbb{R}$  medible definida en  $\Omega$ .

Sea  $\{X_n: n \geq 0\}$  un proceso estocástico discreto y  $S \subset \mathbb{R}$  un conjunto fijo, donde el proceso estocástico toma valores. El conjunto  $S$  es llamado *espacio de estados* del proceso estocástico y cada elemento  $i \in S$  de *estado*. En esta tesis vamos a considerar  $S$  un espacio de estados numerable, cuando  $S$  es un conjunto numerable infinito sin pérdida de generalidad vamos a considerar  $S = \mathbb{N}$ .

Para cada  $n \geq 0$  y  $w \in \Omega$  tenemos que  $X_n(w) \in S$ . Así, cuando  $X_n(w) = j$  se suele decir que *el proceso estocástico está en el estado  $j$  en la etapa  $n$*  y se denota simplemente por  $X_n = j$ .

Para cada variable aleatoria  $X_n$  con espacio de estados  $S$ , notemos que

$$\sum_{i \in S} \mathbb{P}(X_n = i) = \sum_{i \in S} \mathbb{P}(X_n^{-1}(i)) = \mathbb{P}\left(\bigcup_{i \in S} X_n^{-1}(i)\right) = \mathbb{P}(\Omega) = 1. \quad (1)$$

**Definición 14** Sean  $\{X_n: n \geq 0\}$  un proceso estocástico discreto con espacio de estados  $S$ . Una **probabilidad de transición de un paso** de la cadena de Markov del estado  $i$  para el estado  $j$  en la etapa  $n$ , denotado por  $p_{ij}(n)$ , está dado por

$$p_{ij}(n) := \mathbb{P}(X_{n+1} = j | X_n = i). \quad (2)$$

Si la probabilidad de transición no depende de  $n$ , entonces el proceso es llamado de **tiempo homogéneo** o simplemente **homogéneo** y denotaremos simplemente  $p_{ij}$  en lugar de  $p_{ij}(n)$ .

**Definición 15** Sean  $\{X_n: n \geq 0\}$  un proceso estocástico discreto con espacio de estados  $S$ . El proceso estocástico es llamado **Cadena de Markov** si satisface la propiedad

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i) \quad (3)$$

para cada  $n \in \mathbb{N}$  y para cualesquiera de estados  $i_0, i_1, \dots, i_{n-1}, i_n, i, j \in S$ .

En particular para  $n = 1$  se tiene:

$$\mathbb{P}(X_2 = j | X_1 = i_1, X_0 = i_0) = \mathbb{P}(X_2 = j | X_1 = i_1)$$

La identidad (3) es conocida como *la propiedad de Markov*, esta propiedad se refiere a la idea de que la probabilidad de que un sistema o proceso estocástico evolucione en el futuro es decir  $\{X_{n+1} = j\}$  sólo depende de su estado presente  $\{X_n = i\}$  y no de los estados anteriores  $\{X_{n-1} = i_{n-1}, \dots, X_0 = i_0\}$ . En otras palabras, un proceso que cumple con la propiedad de Markov no “recuerda” su historia pasada más allá del estado presente.

A continuación exhibiremos dos proposiciones donde las probabilidades de transición de primer orden se pueden utilizar para el calculo de ciertas probabilidades.

**Proposición 1** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov homogénea con espacio de estados  $S$ . Entonces

$$\begin{aligned} \mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) &= \\ &= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) \dots \mathbb{P}(X_1 = i_1 | X_0 = i_0) \cdot \mathbb{P}(X_0 = i_0) \end{aligned} \quad (4)$$

para cada  $n \in \mathbb{N}$  y  $i_0, i_1, \dots, i_{n-1}, i_n \in S$ .

**Demostración.** Usando la definición de probabilidad condicional (1) y la propiedad de Markov (3) repetidamente obtenemos:

$$\begin{aligned} \mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) &= \\ &= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \mathbb{P}(X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\ &= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) \mathbb{P}(X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}, \dots, X_0 = i_0) \\ &\quad \times \mathbb{P}(X_{n-2} = i_{n-2}, \dots, X_0 = i_0) \\ &= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) \mathbb{P}(X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}) \\ &\quad \times \mathbb{P}(X_{n-2} = i_{n-2} | X_{n-3} = i_{n-3}, \dots, X_0 = i_0) \\ &\quad \times \mathbb{P}(X_{n-3} = i_{n-3}, \dots, X_0 = i_0). \end{aligned}$$

Usando repetidamente la propiedad de Markov y la definición de la probabilidad condicional obtenemos la identidad (1).  $\square$

**Proposición 2** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov homogénea con espacio de estados  $S$ . Entonces

$$\mathbb{P}(X_{n+2} = j | X_n = i) = \sum_{l \in S} \mathbb{P}(X_{n+2} = j | X_{n+1} = l) \mathbb{P}(X_{n+1} = l | X_n = i),$$

para cada  $i, j \in S$  y  $n \geq 0$ .

**Demostración.** De la propiedad de Markov y la definición de la probabilidad condicional tenemos:

$$\mathbb{P}(X_2 = j | X_0 = i) = \mathbb{P}\left(X_2 = j, \bigcup_{l \in S} X_1 = l | X_0 = i\right)$$

$$\begin{aligned}
&= \mathbb{P} \left( \bigcup_{l \in S} (X_2 = j, X_1 = l) \mid X_0 = i \right) \\
&= \frac{\mathbb{P} \left( \bigcup_{l \in S} (X_2 = j, X_1 = l) \cap (X_0 = i) \right)}{\mathbb{P}(X_0 = i)} \\
&= \frac{\sum_{l \in S} \mathbb{P}(X_2 = j, X_1 = l, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{l \in S} \frac{\mathbb{P}(X_2 = j, X_1 = l, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{l \in S} \frac{P(X_2 = j, X_1 = l, X_0 = i)}{P(X_1 = l, X_0 = i)} \cdot \frac{P(X_1 = l, X_0 = i)}{P(X_0 = i)} \\
&= \sum_{l \in S} \mathbb{P}(X_2 = j \mid X_1 = l, X_0 = i) \mathbb{P}(X_1 = l \mid X_0 = i) \\
&= \sum_{l \in S} \mathbb{P}(X_2 = j \mid X_1 = l) \mathbb{P}(X_1 = l \mid X_0 = i)
\end{aligned}$$

□

Notemos que no todo proceso estocástico homogéneo discreto es una cadena de Markov, a estos procesos se les llaman usualmente *procesos no Markovianos*. A continuación exhibiremos un ejemplo.

### Ejemplo 11 (Proceso de Bernoulli con Memoria)

*En este proceso, cada ensayo tiene dos posibles resultados: éxito (1) o fracaso (0). La probabilidad de éxito en cada ensayo  $n$ , denotamos  $p_n$ , depende de los resultados de los ensayos anteriores.*

*Solución: Definimos  $S_n$  como la cantidad de éxitos en los primeros  $n - 1$  ensayos, entonces tendremos la siguiente probabilidad condicionada*

$$p_n = \mathbb{P}(X_n = 1 \mid S_n),$$

*donde  $X_n = 1$  si hay éxito en el ensayo  $n$ , y  $X_n = 0$  si hay fracaso.*

△

Este proceso no satisface la propiedad de Markov porque la probabilidad de éxito en el ensayo  $n$  depende explícitamente de la historia de resultados anteriores  $S_n$ , no solo del resultado más reciente  $X_{n-1}$ . A pesar de no ser Markoviano, este proceso sigue siendo homogéneo en el sentido de que las reglas para calcular  $p_n$  son las mismas para cada ensayo  $n$ .

## 2.2 Ejemplos de modelos de cadenas de Markov

En esta sección exhibiremos tres modelos matemáticos los cuales pueden ser abordados como cadenas de Markov homogéneas, estos modelos son importantes y han sido ampliamente estudiados en el contexto de procesos estocásticos y en este trabajo serán ejemplos recurrentes en las próximas secciones.

### 2.2.1. Cadena Ehrenfest

Es un modelo matemático discreto desarrollado por la pareja de esposos rusos Paul y Tatiana Ehrenfest en 1906 con la finalidad de estudiar conceptos de la mecánica estadística y la teoría de la probabilidad. Mas específicamente, el modelo discreto estudia el intercambio de moléculas de un gas entre dos contenedores para describir la dinámica del sistema de partículas, esto es una simplificación de sistemas más complejos, como los gases ideales, para más detalles del modelo de cadena de Ehrenfest en tiempo discreto ver (Seneta, 2016; Lalley, 2009) y para un estudio en el tiempo continuo se recomienda revisar (Karlin y McGregor, 1965).

La cadena de Ehrenfest puede formularse como un modelo más simple (*el modelo básico de Ehrenfest*), supongamos que tenemos dos urnas conectadas por un conducto y un total de  $N$  bolas, como se muestra en la Figura 1, donde las bolas corresponden a las moléculas y las urnas a los dos contenedores, en cada unidad de tiempo discreto, independientemente del pasado, se selecciona una bola al azar y se mueve a la otra urna.

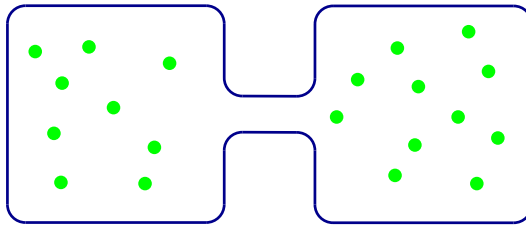


Figura 1: Modelo básico de Ehrenfest

Para este modelo vamos a considerar el espacio de estados  $S = \{0, 1, 2, \dots, N\}$  y  $X_n$  denota el número de bolas que hay en la urna izquierda en la unidad

de tiempo  $n \in \mathbb{N}$ . Notemos que  $\{X_n: n \geq 1\}$  es un proceso estocástico discreto, además es fácil ver que para cada instante de tiempo  $n \in \mathbb{N}$  tenemos las siguientes probabilidades de transición:

$$i) \mathbb{P}(X_{n+1} = s | X_n = t) = 0, \quad \text{cuando } |s - t| \geq 2 \text{ o } s = t;$$

$$ii) \mathbb{P}(X_{n+1} = k + 1 | X_n = k) = \frac{N-k}{N}, \quad \text{para } k = 0, 1, \dots, N - 1;$$

$$iii) \mathbb{P}(X_{n+1} = k - 1 | X_n = k) = \frac{k}{N}, \quad \text{para } k = 0, 1, \dots, N - 1.$$

De los ítems  $i) - iii)$  tenemos que el proceso estocástico es homogéneo pues las probabilidades de transición no depende del tiempo  $n \in \mathbb{N}$ . Además podemos notar que el proceso estocástico  $\{X_n: n \geq 1\}$  es una cadena de Markov, pues el resultado de un estado futuro solo depende del estado presente y no de un estado pasado, satisfaciendo así la identidad (3).

### 2.2.2. Problema del laberinto cerrado

Consideremos un laberinto cerrado, el cual está conformado por 9 compartimentos, las cuales están organizadas en 3 filas y 3 columnas. Algunas de estos compartimentos están conectadas entre sí y dentro del laberinto hay un *ratón*, lo que se pretende es modelar el recorrido del ratón a través del laberinto. Para fijar ideas consideremos el laberinto dado por la Figura 2

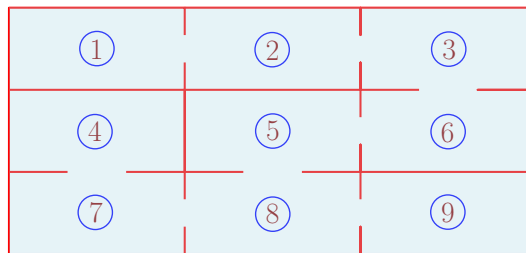


Figura 2: Problema del laberinto cerrado

Notemos que en cada compartimento hay 1, 2 o 3 salidas que conectan con otros compartimentos. El ratón se mueve aleatoriamente en el laberinto y en cada unidad de tiempo cambia de compartimento.

Para contextualizar este modelo dentro de una cadena de Markov coloquemos el conjunto de estados  $S = \{1, 2, 3, \dots, 9\}$  el cual representa los compartimentos del laberinto. Además,  $X_n$  es el compartimento en la cual se encuentra el ratón en la unidad de tiempo  $n \in \mathbb{N}_0$ . Así,  $\{X_n: n \geq 0\}$  es un proceso estocástico discreto y de la Figura 2 tenemos las siguientes probabilidades de transición:

- $\mathbb{P}(X_{n+1} = 2|X_n = 1) = 1$
- $\mathbb{P}(X_{n+1} = 1|X_n = 2) = \mathbb{P}(X_{n+1} = 3|X_n = 2) = 1/2$
- $\mathbb{P}(X_{n+1} = 2|X_n = 3) = \mathbb{P}(X_{n+1} = 6|X_n = 3) = 1/2$
- $\mathbb{P}(X_{n+1} = 7|X_n = 4) = 1$
- $\mathbb{P}(X_{n+1} = 6|X_n = 5) = \mathbb{P}(X_{n+1} = 8|X_n = 5) = 1/2$
- $\mathbb{P}(X_{n+1} = 3|X_n = 6) = \mathbb{P}(X_{n+1} = 5|X_n = 6) = 1/2$
- $\mathbb{P}(X_{n+1} = 4|X_n = 7) = \mathbb{P}(X_{n+1} = 8|X_n = 7) = 1/2$
- $\mathbb{P}(X_{n+1} = 5|X_n = 8) = \mathbb{P}(X_{n+1} = 7|X_n = 8)$   
 $= \mathbb{P}(X_{n+1} = 9|X_n = 8) = 1/2$
- $\mathbb{P}(X_{n+1} = 8|X_n = 9) = 1$

para cada  $n \geq 0$ . Notemos que el proceso estocástico  $\{X_n: n \geq 0\}$  es homogéneo pues las probabilidades de transición no dependen del tiempo  $n \in \mathbb{N}$ . Además, se puede verificar que el proceso estocástico es una cadena de Markov, pues el resultado de un estado futuro solo depende del estado presente y no de un estado pasado, satisfaciendo así la identidad (3).

### 2.2.3. Caminata aleatoria

La *caminata aleatoria*, o “*random walk*”, es un modelo matemático que describe una sucesión de pasos aleatorios en una red infinita, ya sea en  $\mathbb{Z}^n$ , para  $n \geq 1$ , o  $\mathbb{Z}^+ \cup \{0\}$ , la característica principal en este modelo es que solo se puede ir de un estado a otro estado contiguo de la red, es decir, a partir de un estado solo puedes ir al estado del vecino cercano más próximo. Este

concepto es fundamental en la teoría de probabilidad y procesos estocásticos proporcionando una estructura para analizar fenómenos complejos en una red, su aplicación se puede encontrar en diferentes contextos como el movimiento browniano, la dispersión de partículas y las fluctuaciones en los mercados financieros. Desde su introducción por Karl Pearson (Pearson, 1905) a principios del siglo *XX*, las caminatas aleatorias han encontrado aplicaciones en física, biología y economía, entre otras disciplinas. Para una lectura general sobre los modelos de caminata aleatoria se recomienda ver (Lawler y Limic, 2010), (Woess, 2000) y (Gut, 2009). A continuación veremos algunos casos de caminata aleatoria.

### Caminata aleatoria con frontera reflectante

Este primer modelo de caminata aleatoria es en la red unidimensional  $\mathbb{Z}^+ \cup \{0\}$  y se analiza solamente el comportamiento de moverse de un nodo  $i$  a un nodo contiguo, para  $i \geq 0$ , y en  $i = 0$  regresa al mismo nodo, por ello el nombre de caminata aleatoria con frontera reflectante. En este trabajo vamos a considerar la caminata aleatoria uniforme con frontera reflectante, es decir, con probabilidad  $p$  de moverse del nodo  $i$  para  $i + 1$ , probabilidad  $p - 1$  de moverse del nodo  $i + 1$  para  $i$ , cuando  $i \geq 0$ , y en el nodo  $i = 0$  tiene una probabilidad de  $1 - p$  de regresar al mismo nodo. En la Figura 3 se muestra este modelo.

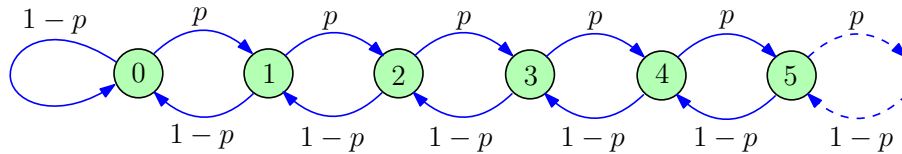


Figura 3: Caminata aleatoria uniforme con frontera reflectante.

Vamos a enmarcar este modelo como una cadena de Markov. Notemos que en este caso el conjunto de estados está dado por los nodos de  $\mathbb{Z}^+ \cup \{0\}$  esto es  $S = \{0, 1, 2, 3, \dots\}$ , el cual es un conjunto numerable e infinito. Además,  $X_n$  denota estar en un determinado estado en el tiempo  $n$ , para cada  $n \in \mathbb{N}_0$ . Las probabilidades de transición para la caminata aleatoria reflectante están

dadas por

$$\mathbb{P}(X_{n+1} = s | X_n = t) = \begin{cases} p & \text{si } s = t + 1, \quad t \geq 0 \\ 1 - p & \text{si } t = s + 1, \quad s \geq 0 \\ 1 - p & \text{si } t = s = 0 \\ 0 & \text{si } |s - t| \geq 2. \end{cases}$$

para cada  $n \geq 0$ , así tenemos que  $\{X_n : n \geq 0\}$  es un proceso estocástico homogéneo. Además, se puede verificar que  $\{X_n : n \geq 0\}$  satisface la condición (3) y por lo tanto tenemos que es una cadena de Markov discreta.

## Caminata aleatoria en $\mathbb{Z}^n$

Definamos

$$\mathbb{Z}^n := \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_i \in \mathbb{Z}, \text{ para } i = 1, 2, 3, \dots, n\},$$

en este espacio definiremos la siguiente métrica

$$d_n(x, y) := \sum_{i=1}^n |x_i - y_i|,$$

donde  $x = (x_1, x_2, \dots, x_n)$  y  $y = (y_1, y_2, \dots, y_n)$ . Es fácil verificar que  $(\mathbb{Z}^n, d_n)$  es un espacio métrico. Ahora vamos a considerar un caminata aleatoria uniforme en la red  $\mathbb{Z}^n$ , donde  $n \in \mathbb{N}$ , notemos que en este caso el conjunto  $\mathbb{Z}^n$  es numerable e infinito, para cada  $n \in \mathbb{N}$ . Para realizar el estudio vamos a separar en casos.

### Caso $n = 1$

Consideremos una caminata aleatoria en la red  $\mathbb{Z}$ , donde se analiza el comportamiento de moverse de un nodo  $i$  a un nodo contiguo, es decir al nodo  $i + 1$  o  $i - 1$ , donde  $i \in \mathbb{Z}$ . En este trabajo vamos a considerar una caminata aleatoria uniforme, así consideraremos que la probabilidad de moverse del nodo  $i$  al  $i + 1$  es  $p$  y de moverse del nodo  $i$  al  $i - 1$  es  $1 - p$ . En la Figura 4 se muestra este modelo.

Para enmarcar este modelo dentro de la teoría de cadenas de Markov, consideremos el conjunto de estados  $S = \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$  el cual

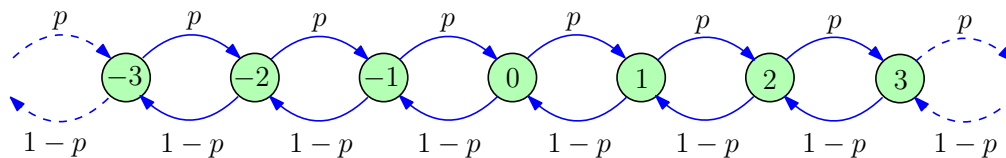


Figura 4: Caminata aleatoria uniforme en  $\mathbb{Z}$

es numerable e infinito y  $X_n$  denota estar en un determinado estado en el tiempo  $n$ , para cada  $n \geq 0$ . Las probabilidades de transición para esta caminata aleatoria es

$$\mathbb{P}(X_{n+1} = s | X_n = t) = \begin{cases} p & \text{si } s = t + 1, \\ 1 - p & \text{si } t = s + 1, \\ 0 & \text{si } |s - t| \geq 2. \end{cases}$$

para cada  $n \geq 0$ . Así tenemos que  $\{X_n : n \geq 0\}$  es un proceso estocástico discreto homogéneo. De manera análoga a la caminata aleatoria reflectante se puede verificar que  $\{X_n : n \geq 0\}$  satisface la condición (3), por consiguiente es una cadena de Markov.

### Caso $n \geq 2$

Ahora consideraremos una caminata aleatoria en la red  $\mathbb{Z}^n$ , para  $n \geq 2$ , de manera análoga analizaremos solamente el comportamiento de moverse de un nodo a un nodo contiguo, sin embargo en este modelo cada nodo tiene  $2n$  nodos contiguos, es decir depende de la dimensión del espacio. Para  $n \geq 2$  consideraremos la caminata aleatoria uniforme, donde la probabilidad de ir de un nodo a un nodo contiguo es de  $\frac{1}{2n}$ . En la Figura 5 mostramos la caminata aleatoria en  $\mathbb{Z}^2$ .

En el contexto de cadenas de Markov, el conjunto de estados está dado por  $S = \mathbb{Z}^n$ , el cual es un conjunto numerable e infinito y  $X_m$  denota estar en un determinado estado en el tiempo  $m \geq 0$ . Las probabilidades de transición para esta caminata aleatoria están dadas por

$$\mathbb{P}(X_{m+1} = s | X_m = t) = \begin{cases} \frac{1}{2n} & \text{si } d_n(s, t) = 1, \\ 0 & \text{si } d_n(s, t) \neq 1. \end{cases}$$

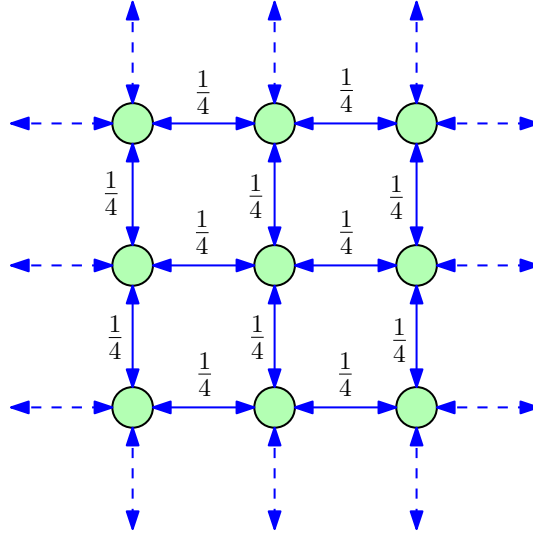


Figura 5: Caminata aleatoria uniforme en  $\mathbb{Z}^2$ .

para cada  $n \geq 0$ . Así el proceso estocástico discreto homogéneo  $\{X_m : m \geq 0\}$  es una cadena de Markov discreta.

### 2.3 Matriz de transición

En el estudio de cadenas de Markov discretas, uno de los elementos fundamentales es la matriz de transición o también llamada matriz estocástica. Una matriz de transición, asociada a una cadena de Markov homogénea, es una matriz cuadrada de  $S \times S$ , donde cada entrada encapsula las probabilidades de transición entre los diferentes estados del sistema. Más específicamente tenemos la siguiente definición:

**Definición 16** *Sea la cadena de Markov homogénea  $\{X_n : n \geq 0\}$  con espacio de estados  $S = \{0, 1, 2, \dots\}$ . La **matriz de transición** asociada a la cadena de Markov es definida por:*

$$P = [P_{ij}]_{i,j \in S} = \begin{pmatrix} P_{00} & P_{01} & P_{02} & \dots & P_{0n} & \dots \\ P_{10} & P_{11} & P_{12} & \dots & P_{1n} & \dots \\ P_{20} & P_{21} & P_{22} & \dots & P_{2n} & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}_{S \times S},$$

donde  $P_{ij} = \mathbb{P}(X_1 = j | X_0 = i)$ .

En caso el espacio de estados  $S$  sea finito,  $S = \{0, 1, 2, \dots, n\}$ , la matriz de transición estará dada por:

$$P = [P_{ij}]_{i,j \in S} = \begin{pmatrix} P_{00} & P_{01} & P_{02} & \dots & P_{0n} \\ P_{10} & P_{11} & P_{12} & \dots & P_{1n} \\ P_{20} & P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{n0} & P_{n1} & P_{n2} & \dots & P_{nn} \end{pmatrix}_{S \times S}.$$

**Lema 5** *Sea la cadena de Markov homogénea  $\{X_n: n \geq 0\}$  con espacio de estados  $S$ . La matriz de transición  $P$  asociada a la cadena de Markov es una matriz estocástica, es decir*

$$\sum_{j \in S} P_{ij} = 1, \quad \forall i \in S.$$

Además,  $P^n$  también es estocástica, para cada  $n \geq 2$ .

**Demostración.** Para cada  $k \in S$ , definamos el evento

$$A_k = \{X_1 = k\} = \{w \in \Omega: X_1(w) = k\} \in \mathcal{F}.$$

Luego tenemos:

i)  $A_i \cap A_j = \emptyset$ , cuando  $i \neq j$ ;

ii)  $\Omega = \bigcup_{k \in S} A_k$ .

De los ítems anteriores, la identidad (2) y el Lema 2 tenemos:

$$\begin{aligned} \sum_{j \in S} P_{ij} &= \sum_{j \in S} \mathbb{P}(X_1 = j | X_0 = i) \\ &= \sum_{j \in S} \mathbb{P}(A_j | X_0 = i) \\ &= \mathbb{P}\left(\bigcup_{j \in S} A_j | X_0 = i\right) \\ &= \mathbb{P}(\Omega | X_0 = i) = 1. \end{aligned}$$

□

A partir de ahora en adelante a la matriz de transición asociada a la cadena de Markov, la llamaremos simplemente por matriz de transición para no sobrecargar el nombre, pues se sobreentiende que esta asociada a una cadena de Markov.

**Definición 17** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov homogénea con espacio de estados  $S$ . Una **probabilidad de transición de  $n$ -pasos** para la cadena de Markov del estado  $i$  al estado  $j$  de  $n$  pasos en la etapa  $k$ , está dado por

$$p_{ij}^{(n)} := \mathbb{P}(X_{k+n} = j | X_k = i). \quad (5)$$

Notemos que (5) tiene la siguiente equivalencia

$$p_{ij}^{(n)} = \mathbb{P}(X_{k+n} = j | X_k = i) = \mathbb{P}(X_n = j | X_0 = i)$$

pues la cadena de Markov  $\{X_n: n \geq 0\}$  es homogénea.

En la siguiente proposición vamos a ver que los elementos de la matriz  $P^n$ , para cada  $n \geq 0$ , están dados por probabilidades de transición de  $n$ -pasos. Cuando  $n = 0$  consideremos  $P^0 = I$ , donde  $I$  es la matriz identidad.

**Proposición 3** Sean la cadena de Markov homogénea  $\{X_n: n \geq 0\}$  con espacio de estados  $S$  y matriz de transición  $P$ . Entonces

$$P_{ij}^n = p_{ij}^{(n)}, \quad (6)$$

para cada  $n \in \mathbb{N}_0$  e  $i, j \in S$ , donde  $P^n$  es la  $n$ -ésima potencia de la matriz  $P$ .

**Demostración.** Notemos que para  $n \in \{0, 1\}$  la identidad (6) se cumple automáticamente. La demostración lo realizaremos por inducción matemática. Para esto usaremos la propiedad de Markov y la definición de probabilidad condicional.

- Primero probaremos para  $n = 2$ . De la Proposición 2 tenemos

$$\begin{aligned} p_{ij}^{(2)} &= \mathbb{P}(X_2 = j | X_0 = i) \\ &= \sum_{l \in S} \mathbb{P}(X_2 = j | X_1 = l) \mathbb{P}(X_1 = l | X_0 = i) \\ &= \sum_{l \in S} P_{il} P_{lj} = P_{ij}^2 \end{aligned}$$

- Por la hipótesis inductiva, asumiremos que se cumple para  $n = k$

$$p_{ij}^{(k)} = P_{ij}^k.$$

- Ahora mostraremos para  $n = k + 1$ ,

$$\begin{aligned}
p_{ij}^{(k+1)} &= \mathbb{P}(X_{k+1} = j | X_0 = i) \\
&= \mathbb{P}\left(X_{k+1} = j, \bigcup_{l \in S} X_k = l | X_0 = i\right) \\
&= \mathbb{P}\left(\bigcup_{l \in S} (X_{k+1} = j, X_k = l) | X_0 = i\right) \\
&= \frac{\mathbb{P}\left(\bigcup_{l \in S} (X_{k+1} = j, X_k = l) \cap (X_0 = i)\right)}{\mathbb{P}(X_0 = i)} \\
&= \frac{\sum_{l \in S} \mathbb{P}(X_{k+1} = j, X_k = l, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{l \in S} \frac{\mathbb{P}(X_{k+1} = j, X_k = l, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{l \in S} \frac{P(X_{k+1} = j, X_k = l, X_0 = i)}{P(X_k = l, X_0 = i)} \frac{P(X_k = l, X_0 = i)}{P(X_0 = i)} \\
&= \sum_{l \in S} \mathbb{P}(X_{k+1} = j | X_k = l, X_0 = i) \mathbb{P}(X_k = l | X_0 = i) \\
&= \sum_{l \in S} \mathbb{P}(X_{k+1} = j | X_k = l) \mathbb{P}(X_k = l | X_0 = i) \\
&= \sum_{l \in S} \mathbb{P}(X_{n+k} = l | X_n = i) P_{lj} \\
&= \sum_{l \in S} P_{il}^k P_{lj} = P_{ij}^{k+1}.
\end{aligned}$$

Por lo tanto para todo  $k \geq 0$  se tiene la identidad

$$[p_{ij}^{(k)}]_{ij \in S} = [\mathbb{P}(X_{n+k} = j | X_n = i)]_{ij \in S} = P^k.$$

□

El siguiente teorema es una generalización de la proposición anterior.

**Teorema 3 (Ecuación de Chapman-Kolmogorov)** *En cadena de Markov homogénea  $\{X_n : n \geq 0\}$  con espacio de estados  $S$  se cumple*

$$\begin{aligned} \mathbb{P}(X_{n+m} = j | X_0 = i) &= \sum_{l \in S} \mathbb{P}(X_m = j | X_0 = l) \mathbb{P}(X_n = l | X_0 = i) \\ &= \sum_{l \in S} \mathbb{P}(X_n = j | X_0 = l) \mathbb{P}(X_m = l | X_0 = i) \end{aligned} \quad (7)$$

para cada  $i, j \in S$ .

**Demostración.** Para demostrar la identidad (7) basta notar:

$$p_{ij}^{(n+m)} = P_{ij}^{n+m} = (P^n P^m)_{ij} = \sum_{l \in S} P_{il}^n P_{lj}^m.$$

□

A continuación vamos a exhibir las matrices de transición de los modelos de las cadenas de Markov introducidas en la Sección 2.2, algunas de éstas tienen espacio de estados finitos y otras infinitas numerable.

**Ejemplo 12 (Cadena Ehrenfest)** *En el modelo de la cadena de Ehrenfest presentada en la Subsección 2.2.1 obtuvimos las siguientes probabilidades de transición*

$$i) P_{t,s} = \mathbb{P}(X_{n+1} = s | X_n = t) = 0, \quad \text{cuando } |s - t| \geq 2 \text{ o } s = t;$$

$$ii) P_{k,k+1} = \mathbb{P}(X_{n+1} = k + 1 | X_n = k) = \frac{N-k}{N}, \quad \text{para } k = 0, 1, \dots, N-1;$$

$$iii) P_{k,k-1} = \mathbb{P}(X_{n+1} = k - 1 | X_n = k) = \frac{k}{N}, \quad \text{para } k = 0, 1, \dots, N-1,$$

donde  $S = \{0, 1, 2, \dots, N\}$ . Reemplazando las probabilidades de un paso en la matriz de transición obtenemos:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \dots & 0 & 0 \\ \frac{1}{N} & 0 & \frac{N-1}{N} & 0 \dots & 0 & 0 \\ 0 & \frac{2}{N} & 0 & \frac{N-2}{N} \dots & 0 & 0 \\ 0 & 0 & \frac{3}{N} & 0 \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \frac{1}{N} \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}_{S \times S}.$$

Considerando  $N = 8$ , la matriz de transición es expresada como:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/8 & 0 & 7/8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2/8 & 0 & 6/8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3/8 & 0 & 5/8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4/8 & 0 & 4/8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5/8 & 0 & 3/8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6/8 & 0 & 2/8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 7/8 & 0 & 1/8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

**Ejemplo 13 (Caminata aleatoria con frontera reflectante)** *En el modelo de la caminata aleatoria con frontera reflectante presentada en la Subsección 2.2.3 obtuvimos las siguientes probabilidades de transición*

$$P = [P_{ij}]_{i,j \in S} = \begin{pmatrix} 1-p & p & 0 & 0 & 0 & 0 & 0 & \dots \\ 1-p & 0 & p & 0 & 0 & 0 & 0 & \dots \\ 0 & 1-p & 0 & p & 0 & 0 & 0 & \dots \\ 0 & 0 & 1-p & 0 & p & 0 & 0 & \dots \\ 0 & 0 & 0 & 1-p & 0 & p & 0 & \dots \\ 0 & 0 & 0 & 0 & 1-p & 0 & p & \dots \\ 0 & 0 & 0 & 0 & 0 & 1-p & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}_{S \times S}.$$

donde  $S = \mathbb{Z}^+ \cup \{0\}$ .

**Ejemplo 14 (Caminata aleatoria en  $\mathbb{Z}$ )** *En el modelo de la caminata aleatoria en  $\mathbb{Z}$  presentada en la Subsección 2.2.3 obtuvimos las siguientes probabilidades de transición*

$$P = [P_{ij}]_{i,j \in S} = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \dots & 0 & p & 0 & 0 & 0 & 0 & \dots \\ \dots & 1-p & 0 & p & 0 & 0 & 0 & \dots \\ \dots & 0 & 1-p & 0 & p & 0 & 0 & \dots \\ \dots & 0 & 0 & 1-p & 0 & p & 0 & \dots \\ \dots & 0 & 0 & 0 & 1-p & 0 & p & \dots \\ \dots & 0 & 0 & 0 & 0 & 1-p & 0 & \dots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}_{\mathbb{Z} \times \mathbb{Z}}.$$

**Observación 1** De aquí en adelante cuando nos referimos a una cadena de Markov estamos asumiendo que está es homogénea.

## 2.4 Probabilidad de Llegada

En esta sección estudiaremos la probabilidad de llegada en una cadena de Markov, esto se refiere a la probabilidad de que la cadena de Markov, empezando en un estado inicial, llegue a otro estado específico en algún momento. Este concepto es fundamental porque nos permite entender si es posible que el sistema alcance ciertos estados y con qué probabilidad ocurre.

**Definición 18** Sea  $\{X_n: n \geq 0\}$  la cadena de Markov con espacio de estados  $S$  y  $A \subset S$ . Definimos el **tiempo de llegada**<sup>1</sup> de la cadena de Markov al subconjunto  $A$  como

$$T_A := \inf\{n \geq 0: X_n \in A\}. \quad (8)$$

Además, consideremos  $T_A = \infty$  cuando  $\{X_n: n \geq 0 : X_n \in A\} = \emptyset$ , es decir si  $X_n \notin A$  para cada  $n \geq 0$ .

En la teoría de cadenas de Markov el caso de mayor estudio es cuando  $A = \{j\}$ , es decir el conjunto está compuesto por un único estado, así

$$T_j = \inf\{n \geq 0: X_n = j\}.$$

<sup>1</sup>En ingles es conocido como *hitting time*.

**Definición 19** Sea  $\{X_n: n \geq 0\}$  la cadena de Markov con espacio de estados  $S$  y  $A \subset S$ . Definimos la **probabilidad de llegada**<sup>2</sup> al conjunto  $A$  iniciando en el estado  $i$  como

$$h_{iA} := \mathbb{P}(X_n \in A \text{ para algún } n \geq 0 \mid X_0 = i). \quad (9)$$

De la definición anterior  $h_{iA}$  también se puede escribir como

$$h_{iA} = \mathbb{P}(T_A < \infty \mid X_0 = i),$$

notemos que  $h_{iA}$  representa la probabilidad de llegar al conjunto  $A$  partiendo del estado  $i$ . De manera análoga, el caso de mayor interés en probabilidad de llegada es cuando  $A = \{j\}$ , así tenemos

$$h_{ij} = \mathbb{P}(X_n = j \text{ para algún } n \geq 0 \mid X_0 = i) = \mathbb{P}(T_j < \infty \mid X_0 = i),$$

además es inmediato que

$$h_{ii} = 1. \quad (10)$$

Estamos interesados en derivar fórmulas para el cálculo de la probabilidad de llegada cuando esta es finita. Para esto haremos uso de la técnica comúnmente llamada *análisis del primer paso*, la idea es descomponer el evento de interés en términos de lo que sucede luego de dar el primer paso en la cadena de Markov. La siguiente proposición nos da una fórmula recursiva para hacer el cálculo de las probabilidades de llegada y la idea es la del análisis del primer paso.

**Proposición 4** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$  y  $A \subset S$ . Entonces  $h_{iA}$  satisface

$$\left. \begin{aligned} h_{iA} &= 1 && \text{para cada } i \in A \\ h_{iA} &= \sum_{j \in S} P_{ij} \cdot h_{jA} && \text{para cada } i \notin A \end{aligned} \right\} \quad (11)$$

Además, si  $\{x_i\}_{i \in S}$  es otra solución no-negativa de (11), entonces  $h_{iA}$  es la mínima solución, en el sentido de  $x_i \geq h_{iA}$ , para cada  $i \in S$ .

---

<sup>2</sup>En ingles es conocido como *hitting probability*

**Demostración.** Para demostrar (11) primero vamos a asumir que  $i \in A$ . Luego  $T_A = 0$ , por tanto

$$h_{iA} = \mathbb{P}(X_n \in A \text{ para algún } n \geq 0 | X_0 = i) = 1.$$

Ahora supongamos que  $i \notin A$ , luego  $T_A \geq 1$ . Sea

$$T_A = 1 + T'_A,$$

donde  $T'_A$  es el tiempo restante, luego de dar un paso, hasta llegar a  $A$ . Es importante notar que  $T'_A$  está en función de  $\{X_1, X_2, \dots\}$  es decir la cadena de Markov inicia en  $X_1$ , mientras que  $T_A$  de  $\{X_0, X_1, X_2, \dots\}$ , por lo cual

$$\mathbb{P}(T'_A < \infty | X_1 = j) = \mathbb{P}(T_A < \infty | X_0 = j). \quad (12)$$

De la definición de probabilidad condicional tenemos

$$\begin{aligned} h_{iA} &= \mathbb{P}(T_A < \infty | X_0 = i) \\ &= \sum_{j \in S} \mathbb{P}(T_A < \infty, \mathbb{1}_{X_1=j} | X_0 = i) \\ &= \sum_{j \in S} \mathbb{P}(T_A < \infty | X_1 = j, X_0 = i) \mathbb{P}(X_1 = j | X_0 = i) \\ &= \sum_{j \in S} P_{ij} \mathbb{P}(T'_A + 1 < \infty | X_1 = j, X_0 = i) \\ &= \sum_{j \in S} P_{ij} \mathbb{P}(T'_A < \infty | X_1 = j, X_0 = i). \end{aligned}$$

Desde que  $T'_A$  es condicionalmente independiente de  $X_0$  y por (12) tenemos

$$\mathbb{P}(T'_A < \infty | X_1 = j, X_0 = i) = \mathbb{P}(T'_A < \infty | X_1 = j) = \mathbb{P}(T_A < \infty | X_0 = j).$$

Por tanto

$$h_{iA} = \sum_{j \in S} P_{ij} \cdot h_{jA}, \quad \text{si } i \notin A.$$

Ahora consideremos  $\{x_i\}_{i \in S}$  una solución no-negativa del sistema (11). El caso más simple es cuando  $i \in A$ , en este caso  $x_i = h_{iA}$ . Supongamos que  $i \notin A$ , entonces tenemos

$$x_i = \sum_{j \in S} P_{ij} \cdot x_j = \sum_{j \in A} P_{ij} + \sum_{j \notin A} P_{ij} \cdot x_j.$$

Sustituyendo  $x_j$  tenemos

$$\begin{aligned}
x_i &= \sum_{j \in A} P_{ij} + \sum_{j \notin A} P_{ij} \cdot \left( \sum_{k \in S} P_{jk} x_k \right) \\
&= \sum_{j \in A} P_{ij} + \sum_{j \notin A} P_{ij} \cdot \left( \sum_{k \in A} P_{jk} + \sum_{k \notin A} P_{jk} x_k \right) \\
&= \sum_{j \in A} P_{ij} + \sum_{j \notin A} P_{ij} \cdot \sum_{k \in A} P_{jk} + \sum_{j \notin A} P_{ij} \cdot \sum_{k \notin A} P_{jk} x_k \\
&= \sum_{j \in A} P_{ij} + \sum_{j \notin A} P_{ij} \cdot \sum_{k \in A} P_{jk} + \sum_{j \notin A} \sum_{k \notin A} P_{ij} \cdot P_{jk} x_k \\
&= \mathbb{P}(X_1 \in A | X_0 = i) + \mathbb{P}(X_1 \notin A, X_2 \in A | X_0 = i) + \sum_{j \notin A} \sum_{k \notin A} P_{ij} \cdot P_{jk} x_k.
\end{aligned}$$

Repitiendo la sustitución de  $x_k$  en el termino final de la última igualdad, luego de  $n$  pasos obtenemos

$$\begin{aligned}
x_i &= \mathbb{P}(X_1 \in A | X_0 = i) + \dots + \mathbb{P}(X_1 \notin A, X_2 \notin A, \dots, X_{n-1} \notin A, X_n \in A | X_0 = i) \\
&\quad + \sum_{j_1 \notin A} \dots \sum_{j_{n-1} \notin A} \sum_{j_n \notin A} P_{ij_1} \dots P_{j_{n-2}j_{n-1}} P_{j_{n-1}j_n} x_{j_n} \\
&\geq \mathbb{P}(X_1 \in A | X_0 = i) + \dots + \mathbb{P}(X_1 \notin A, X_2 \notin A, \dots, X_{n-1} \notin A, X_n \in A | X_0 = i) \\
&= \mathbb{P}(T_A = 1 | X_0 = i) + \mathbb{P}(T_A = 2 | X_0 = i) + \dots + \mathbb{P}(T_A = n | X_0 = i) \\
&= \mathbb{P}(T_A \leq n | X_0 = i).
\end{aligned}$$

Desde que

$$x_i \geq \mathbb{P}(T_A \leq n | X_0 = i),$$

para cada  $n \geq 1$ . Tomando el limite tenemos

$$x_i \geq \lim_{n \rightarrow \infty} \mathbb{P}(T_A \leq n | X_0 = i) = \mathbb{P}(T_A < \infty | X_0 = i).$$

Por tanto

$$x_i \geq h_{iA},$$

para cada  $i \notin A$ . □

**Definición 20** Sea  $\{X_n : n \geq 0\}$  la cadena de Markov con espacio de estados  $S$  y  $A \subset S$ . Definimos el **tiempo de llegada promedio**<sup>3</sup> al conjunto  $A$

<sup>3</sup>En ingles es conocido como *expected hitting time*

iniciando en el estado  $i$  como

$$\eta_{iA} := \mathbb{E}[T_A | X_0 = i]. \quad (13)$$

Notemos que  $\eta_{iA}$  representa el tiempo esperado que lleguemos al conjunto  $A$  partiendo del estado  $i$ . Cuando el estado  $i \in A$ , obviamente tenemos que  $T_A = 0$  y por consiguiente

$$\eta_{iA} = \mathbb{E}[T_A | X_0 = i] = 0. \quad (14)$$

En el caso que  $i \notin A$ , el valor de  $\eta_{iA}$  se puede calcular de manera recursiva realizando el análisis de primer paso, como se muestra en la siguiente proposición.

**Proposición 5** *Sea  $\{X_n : n \geq 0\}$  la cadena de Markov con espacio de estados  $S$  y  $A \subset S$  no vacío. Entonces  $\eta_{iA}$  satisface*

$$\left. \begin{array}{ll} \eta_{iA} = 0 & \text{para cada } i \in A \\ \eta_{iA} = 1 + \sum_{j \notin A} P_{ij} \cdot \eta_{jA} & \text{para cada } i \notin A \end{array} \right\} \quad (15)$$

Además, si  $\{x_i\}_{i \in S}$  es otra solución no-negativa de (15), entonces  $\eta_{iA}$  es la mínima solución, en el sentido de  $x_i \geq \eta_{iA}$ , para cada  $i \in S$ .

**Demostración.** Si  $i \in A$ , por (14) tenemos que  $\eta_{iA} = 0$ . Antes de seguir la vamos a ver dos identidades que serán necesarias en la demostración. Notemos que

$$\mathbb{1}(\cdot) = \sum_{j \in S} \mathbb{1}_{\{X_1(\cdot)=j\}}, \quad (16)$$

donde  $\mathbb{1}(\cdot)$  es la función constante igual a 1. Por otro lado tenemos la siguiente identidad

$$\mathbb{E}[T_A | X_1 = j, X_0 = i] = \begin{cases} 1 & \text{para cada } j \in A, i \notin A \\ 1 + \mathbb{E}[T_A | X_0 = j] & \text{para cada } j \notin A, i \notin A \end{cases} \quad (17)$$

Ahora consideremos  $i \notin A$ . De las identidades (16), (17) y el Lema 4 tenemos

$$\eta_{iA} = \mathbb{E}[T_A | X_0 = i] = \mathbb{E} \left[ T_A \cdot \left( \sum_{j \in S} \mathbb{1}_{\{X_1=j\}} \right) \middle| X_0 = i \right]$$

$$\begin{aligned}
&= \sum_{j \in S} \mathbb{E}[T_A \mathbb{1}_{\{X_1=j\}} | X_0 = i] \\
&= \frac{1}{\mathbb{P}(X_0 = i)} \sum_{j \in S} \mathbb{E}[T_A \mathbb{1}_{\{X_1=j\}} \mathbb{1}_{\{X_0=i\}}] \\
&= \frac{1}{\mathbb{P}(X_0 = i)} \sum_{j \in S} \mathbb{E}[T_A \mathbb{1}_{\{X_1=j\} \cap \{X_0=i\}}] \\
&= \sum_{j \in S} \mathbb{E}[T_A | X_1 = j, X_0 = i] \frac{\mathbb{P}(X_1 = j, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{j \in S} \mathbb{E}[T_A | X_1 = j, X_0 = i] \mathbb{P}(X_1 = j | X_0 = i)
\end{aligned}$$

Usando (17) tenemos

$$\begin{aligned}
\eta_{iA} &= \sum_{j \in A} \mathbb{E}[T_A | X_1 = j, X_0 = i] \mathbb{P}(X_1 = j | X_0 = i) \\
&+ \sum_{j \notin A} \mathbb{E}[T_A | X_1 = j, X_0 = i] \mathbb{P}(X_1 = j | X_0 = i) \\
&= \sum_{j \in A} \mathbb{P}(X_1 = j | X_0 = i) + \sum_{j \notin A} (\mathbb{E}[T_A | X_0 = j] + 1) \mathbb{P}(X_1 = j | X_0 = i) \\
&= \sum_{j \in S} \mathbb{P}(X_1 = j | X_0 = i) + \sum_{j \notin A} \mathbb{E}[T_A | X_0 = j] \mathbb{P}(X_1 = j | X_0 = i) \\
&= 1 + \sum_{j \notin A} P_{ij} \cdot \eta_{jA}.
\end{aligned}$$

Consideremos  $\{x_i\}_{i \in S}$  otra solución del sistema (15). Si  $i \in A$ , tenemos

$$x_i = 0 = \eta_{iA}.$$

Consideremos  $i \notin A$ , de (15) tenemos

$$\begin{aligned}
x_i &= 1 + \sum_{j \notin A} P_{ij} x_j \\
&= 1 + \sum_{j \notin A} P_{ij} \left( 1 + \sum_{k \notin A} P_{jk} x_k \right) \\
&= 1 + \sum_{j \notin A} P_{ij} + \sum_{j \notin A} \sum_{k \notin A} P_{ij} P_{jk} x_k \\
&= \mathbb{P}(T_A \geq 1 | X_0 = i) + \mathbb{P}(T_A \geq 2 | X_0 = i) + \sum_{j \notin A} \sum_{k \notin A} P_{ij} P_{jk} x_k.
\end{aligned}$$

Repitiendo iterativamente el proceso  $n$  veces tenemos

$$\begin{aligned}
 x_i &= \mathbb{P}(T_A \geq 1|X_0 = i) + \mathbb{P}(T_A \geq 2|X_0 = i) + \dots + \mathbb{P}(T_A \geq n|X_0 = i) \\
 &\quad + \sum_{j_1 \notin A} \dots \sum_{j_{n-1} \sum_{j_n} \notin A} P_{ij_1} \dots P_{j_{n-2}j_{n-1}} P_{j_{n-1}j_n} x_{j_n} \\
 &\geq \mathbb{P}(T_A \geq 1|X_0 = i) + \mathbb{P}(T_A \geq 2|X_0 = i) + \dots + \mathbb{P}(T_A \geq n|X_0 = i) \\
 &= \sum_{m=1}^n \mathbb{P}(T_A \geq m|X_0 = i).
 \end{aligned}$$

Luego

$$x_i \geq \sum_{m=1}^n \mathbb{P}(T_A \geq m|X_0 = i),$$

para cada  $n \in \mathbb{N}$ . Tomando el limite tenemos

$$x_i \geq \lim_{n \rightarrow \infty} \sum_{m=1}^n \mathbb{P}(T_A \geq m|X_0 = i) = \mathbb{E}(T_A|X_0 = i) = \eta_{iA},$$

por tanto  $x_i \geq \eta_{iA}$ , para cada  $i \in S$ .

□

**Ejemplo 15** Consideremos la cadena de Markov con estados  $S = \{1, 2, 3, 4\}$  la cual está representada por el siguiente diagrama:

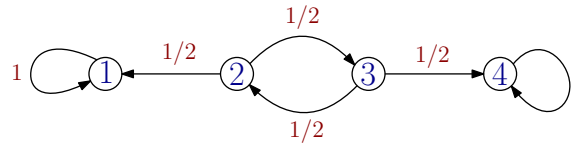


Figura 6: Cadena de Markov con 4 estados.

Iniciando en el estado  $X_0 = 2$  vamos a calcular

$$h_{24} = \mathbb{P}(T_4 < \infty|X_0 = 2) \quad \text{y} \quad \eta_{24} = \mathbb{E}(T_4|X_0 = 2).$$

Supongamos que iniciamos en el estado  $X_0 = 2$ , por el principio de multiplicación el diagrama 6 nos proporciona un sistema de ecuaciones para calcular

la probabilidad de llegar al estado 4 iniciando en el estado  $X_0 = 2$ , el cual está dado por:

$$\left. \begin{aligned} h_{24} &= \frac{1}{2}h_{14} + \frac{1}{2}h_{34} \\ h_{14} &= 0 \\ h_{34} &= \frac{1}{2}h_{24} + \frac{1}{2}h_{44} \\ h_{44} &= 1. \end{aligned} \right\} \quad (18)$$

Del sistema de ecuaciones 18 obtenemos que  $h_{24} = 1/3$ .

Para calcular el tiempo de llegada promedio al estado 4 iniciando en el estado  $X_0 = 2$ , del diagrama 18 obtenemos el siguiente sistema de ecuaciones:

$$\left. \begin{aligned} \eta_{24} &= 1 + \frac{1}{2}\eta_{14} + \frac{1}{2}\eta_{34} \\ \eta_{14} &= 0 \\ \eta_{34} &= 1 + \frac{1}{2}\eta_{24} + \frac{1}{2}\eta_{44} \\ \eta_{44} &= 0. \end{aligned} \right\} \quad (19)$$

donde el valor de 1 en cada ecuación representará el tiempo de transición de un estado a otro, Esto se debe a que el tiempo de salto se define como el número esperado de pasos necesarios para llegar a un estado específico desde otro estado inicial. Además, del diagrama podemos ver que obviamente  $\eta_{44} = 0$  y

$$\begin{aligned} \eta_{14} &= \sum_{n < \infty} n \mathbb{P}(T_4 = n | X_0 = 1) + \infty \cdot \mathbb{P}(T_4 = \infty | X_0 = 2) \\ &= \sum_{n < \infty} n \cdot 0 + \infty \cdot (1 - \mathbb{P}(T_4 < \infty | X_0 = 2)) \\ &= 0 + \infty(1 - 1) = 0. \end{aligned}$$

Así, resolviendo el sistema de ecuaciones (19) obtenemos el tiempo de llegada promedio al estado  $j = 4$  partiendo del estado  $X_0 = 2$ ,  $\eta_{24} = 2$ .  $\triangle$

**Definición 21** Sean  $\{X_n : n \geq 0\}$  una cadena de Markov con espacio de estados  $S$ ,  $i \in S$  un estado y  $P$  su matriz de transición. Decimos que el estado  $i$  es **absorvente** si

$$P_{i,i} = 1.$$

Además, decimos que  $A \subset S$  es un conjunto absorvente si  $P_{i,i} = 1$ , para cada elemento  $i \in A$ .

**Lema 6** Si el estado  $j \in S$ , es absorvente. Entonces  $\forall i \in S$  se tiene

$$h_{i,j} = \mathbb{P}(T_j < \infty | X_0 = i) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i)$$

**Demostración.** Como  $j \in S$  es absorvente entonces se tiene

$$\{T_j < \infty\} = \bigcup_{n \geq 1} \{X_n = j\}$$

además se tiene

$$\{X_n = j\} \subset \{X_{n+1} = j\}, \quad \forall n \in S$$

Dado que  $X_0 = i$ , y por la propiedad de convergencia de probabilidad se tiene

$$\begin{aligned} \mathbb{P}(T_j < \infty | X_0 = i) &= \mathbb{P}\left(\bigcup_{n \geq 1} \{X_n = j\} | X_0 = i\right) \\ &= \mathbb{P}\left(\lim_{n \rightarrow \infty} \{X_n = j\} | X_0 = i\right) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i). \end{aligned}$$

□

Para no sobrecargar la notación definamos

$$\{T_A = \infty\} := \bigcup_{n \geq 1} \{X_n \notin A\}.$$

El siguiente ejemplo muestra que la probabilidad  $\mathbb{P}(T_A = \infty | X_0 = k)$  puede ser no nula, es decir

$$\mathbb{P}(T_A = \infty | X_0 = k) > 0.$$

**Ejemplo 16** Consideremos la cadena de Markov con espacio de estados  $S = \{0, 1, 2\}$  dada por la Figura 7. Sean  $A = 0$  y  $k = 1$

luego:

$$\mathbb{P}(T_A = \infty | X_0 = 1) = 0,2 > 0.$$

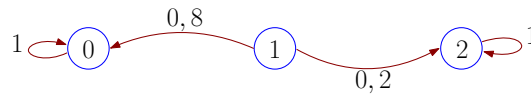


Figura 7: Cadena de Markov con 3 estados.

Solución. En efecto, si  $A = \{0\}$  y  $k = 1$  tenemos

$$\begin{aligned}
 \mathbb{P}(T_A = \infty | X_0 = 1) &= \mathbb{P}\left(\bigcup_{n \geq 1} \{X_n \notin \{0\}\} | X_0 = 1\right) \\
 &= \sum_{n \geq 1} \mathbb{P}(X_n \notin \{0\} | X_0 = 1) \\
 &= \sum_{n \geq 1} \mathbb{P}(X_n \in \mathbb{N} | X_0 = 1) \\
 &= \sum_{n \geq 1} \sum_{l \in \{1,2\}} \mathbb{P}(X_n = l | X_0 = 1) \\
 &= \sum_{l \in \{1,2\}} \sum_{n \geq 1} \mathbb{P}(X_n = l | X_0 = 1) \\
 &= \sum_{l \in \{1,2\}} \mathbb{P}(X_1 = l | X_0 = 1) + \sum_{l \in \{1,2\}} \sum_{n \geq 2} \mathbb{P}(X_n = l | X_0 = 1) \\
 &= \sum_{l \in \{1,2\}} \mathbb{P}(X_1 = l | X_0 = 1) \\
 &= \mathbb{P}(X_1 = 1 | X_0 = 1) + \mathbb{P}(X_1 = 2 | X_0 = 1) \\
 &= P_{11} + P_{21} = 0 + 0,2 = 0,2.
 \end{aligned}$$

△

## 2.5 Primer tiempo de retorno

En esta sección abordaremos un indicador útil en el estudio de Cadenas de Markov, este es el de primer tiempo de retorno. El primer tiempo de retorno es el número de pasos que toma la cadena de Markov para llegar (retornar) a un estado específico por primera vez, partiendo de algún estado. Este concepto nos ayuda a analizar cuánto tiempo, en promedio, debemos esperar para que el sistema llegue a un estado dado. Esta es una medida de recurrencia y es

especialmente útil para determinar si ciertos estados son recurrentes, es decir, si la cadena tiende a ir a esos estados una y otra vez.

**Definición 22** Sea  $\{X_n: n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . El **primer tiempo de retorno** al estado  $j \in S$  es definido por

$$T_j^r := \inf\{n \geq 1: X_n = j\},$$

donde  $T_j^r = \infty$  si  $X_n \neq j$  para todo  $n \geq 1$ .

En la definición anterior notemos que el ínfimo es tomado para los valores de  $n \geq 1$  a diferencia de la Definición 18 que incluye el 0.

**Definición 23** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$ . La **probabilidad de retorno**  $p_{ij}$  al estado  $j \in S$  en un tiempo finito partiendo desde el estado  $i \in S$ , está dada por

$$p_{ij} := \mathbb{P}(T_j^r < \infty | X_0 = i)$$

Notemos que

$$\{T_j^r < \infty\} = \{X_n = j \text{ para algún } n \geq 1\}.$$

Así de la definición anterior tenemos la siguiente equivalencia

$$p_{ij} = \mathbb{P}(T_j^r < \infty | X_0 = i) = \mathbb{P}(X_n = j \text{ para algún } n \geq 1 | X_0 = i). \quad (20)$$

**Proposición 6** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$  e  $i, j \in S$ . Entonces

$$p_{ij} = \sum_{k \in S} P_{ik} \cdot h_{kj}. \quad (21)$$

**Demostración.** Por la propiedad de Markov tenemos

$$\mathbb{P}(T_j^r < \infty | X_1 = k, X_0 = i) = \mathbb{P}(T_j^r < \infty | X_1 = k) = \mathbb{P}(T_j^r < \infty | X_0 = k) = h_{kj}, \quad (22)$$

para cada  $j, k \in S$ .

De (22) y de la definición de probabilidad condicional tenemos

$$p_{ij} = \mathbb{P}(T_j^r < \infty | X_0 = i)$$

$$\begin{aligned}
&= \sum_{k \in S} \mathbb{P}(T_j^r < \infty, \mathbb{1}_{\{X_1=k\}} | X_0 = i) \\
&= \sum_{k \in S} \mathbb{P}(T_j^r < \infty | X_1 = k, X_0 = i) \mathbb{P}(X_1 = k | X_0 = i) \\
&= \sum_{k \in S} \mathbb{P}(X_1 = k | X_0 = i) \mathbb{P}(T_j < \infty | X_0 = k) \\
&= \sum_{k \in S} P_{ik} \cdot h_{kj}.
\end{aligned}$$

□

Vamos a introducir la siguiente notación para  $n \geq 2$

$$\{T_j^r = n\} := \{X_1 \neq j, X_2 \neq j, \dots, X_{n-1} \neq j, X_n = j\},$$

y

$$\{T_j^r = 1\} := \{X_1 = j\}.$$

A continuación obtendremos una expresión para  $p_{ii}$ , donde  $i \in S$

$$\begin{aligned}
p_{ii} &= \mathbb{P}(X_n = i \text{ para algún } n \geq 1 | X_0 = i) \\
&= \mathbb{P}\left(\bigcup_{n \geq 1} \{T_i^r = n\} | X_0 = i\right) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(\{T_i^r = n\} | X_0 = i) \\
&= \sum_{n=1}^{\infty} f_{ii}^{(n)},
\end{aligned}$$

donde

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

Por lo anteriormente expuesto tenemos que

$$p_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)}. \quad (23)$$

**Definición 24** Sea  $\{X_n : n \geq 0\}$  una cadena de Markov con espacio de estados  $S$ . El **tiempo de retorno promedio** al estado  $j \in S$  después de iniciar en el estado  $X_0 = i$  es definido como

$$\mu_{ij} := \mathbb{E}[T_j^r | X_0 = i].$$

**Proposición 7** Sea  $\{X_n: n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . Para cada  $i, j \in S$  tenemos

$$\mu_{ij} = 1 + \sum_{\substack{l \in S \\ l \neq j}} P_{il} \mu_{lj} \quad (24)$$

**Demostración.** Sean  $i, j, k \in S$ , de manera análoga a (17) tenemos

$$\mathbb{E}[T_j^r | X_1 = k, X_0 = i] = \begin{cases} 1 & \text{para cada } k = j, i \neq j \\ 1 + \mathbb{E}[T_j^r | X_0 = k] & \text{para cada } k \neq j, i \neq j \end{cases} \quad (25)$$

De (25) y de manera similar a la demostración de la Proposición 5 tenemos:

$$\mu_{ij} = 1 + \sum_{\substack{l \in S \\ l \neq j}} P_{il} \mu_{lj}.$$

□

**Observación 2** A continuación vamos hacer una comparación sobre el tiempo de llegada promedio y el tiempo de retorno promedio:

- 1) Las Ecuaciones (24) del tiempo de retorno promedio no tienen condiciones de frontera como si las tiene las Ecuaciones (15) del tiempo de llegada promedio.
- 2) Cuando  $i \neq j$ , tenemos la igualdad entre el tiempo de retorno y de llegada, es decir:

$$h_{ij} = \mathbb{E}[T_j | X_0 = i] = \mathbb{E}[T_j^r | X_0 = i] = \mu_{ij}.$$

- 3) Para cada  $i \in S$  tenemos

$$h_{ii} = 0, \quad \text{mientras que} \quad \mu_{ii} \geq 1.$$

- 4) Ambos tiempos están relacionadas por la siguiente identidad:

$$\mu_{ii} = 1 + \sum_{\substack{l \in S \\ l \neq i}} P_{il} h_{li} \quad (26)$$

para cada  $i \in S$ .

5) En la practica es más fácil hallar los valores del tiempo de llegada promedio  $h_{ij}$ , como se exhibió en el Ejemplo 15. Así combinando los items 2) y 4) podemos calcular los valores del tiempo de retorno promedio  $\mu_{ij}$ , para cada  $i, j \in S$ .

**Ejemplo 17 (Problema del laberinto)** Consideremos el problema del laberinto detallado en la Subsección 2.2.2. Recordemos que en el problema, se considera un ratón colocado en un laberinto con 9 compartimentos, enumerados del 1 al 9 los cuales están comunicados con algunos compartimentos contiguos como se muestra en la Figura 8. Calcular el tiempo de retorno promedio

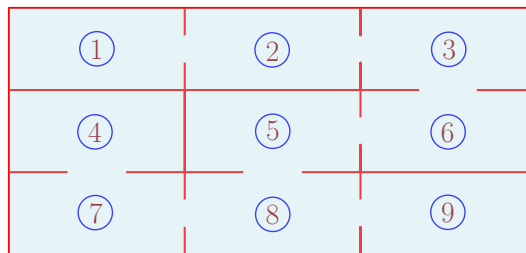


Figura 8: Problema del laberinto en un acuario

al estado 1 después de iniciar en los compartimentos  $j \in \{1, 2, \dots, 9\}$ .

Solución: Del problema del laberinto tenemos,  $S = \{1, 2, 3, \dots, 9\}$  nuestro conjunto de estados y  $X_n$  es el compartimento en la cual se encuentra el ratón en la unidad de tiempo  $n \in \mathbb{N}_0$ , las probabilidades de transición se muestran en la siguiente matriz

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}_{S \times S}$$

Luego debemos calcular  $\mu_{j1}$  donde  $j \in \{1, 2, 3, \dots, 9\}$ , para la cadena de Markov  $\{X_n : n \geq 0\}$ .

- Para  $j = 1$  de la identidad (26) tenemos

$$\mu_{11} = 1 + \sum_{\substack{l \in S \\ l \neq 1}} P_{1l} \mu_{l1} = 1 + P_{12} \mu_{21} = 1 + \frac{1}{2} \mu_{21}.$$

- Para  $j = 2$  de (24) tenemos

$$\mu_{21} = 1 + \sum_{\substack{l \in S \\ l \neq 1}} P_{2l} \mu_{l1} = 1 + P_{23} \mu_{31} = 1 + \frac{1}{2} \mu_{31}.$$

- Para  $j = 3$  de (24) tenemos

$$\mu_{31} = 1 + \sum_{\substack{l \in S \\ l \neq 1}} P_{3l} \mu_{l1} = 1 + P_{32} \mu_{21} + P_{36} \mu_{61} = 1 + \frac{1}{2} \mu_{21} + \frac{1}{2} \mu_{61}.$$

Análogamente se puede realizar para los estados  $j \in \{4, 5, \dots, 9\}$ , formando así el siguiente sistema de ecuaciones

$$\left. \begin{aligned} \mu_{11} &= 1 + \frac{1}{2} \mu_{21} \\ \mu_{21} &= 1 + \frac{1}{2} \mu_{31} \\ \mu_{31} &= 1 + \frac{1}{2} \mu_{21} + \frac{1}{2} \mu_{61} \\ \mu_{41} &= 1 + \mu_{71} \\ \mu_{51} &= 1 + \frac{1}{2} \mu_{81} + \frac{1}{2} \mu_{61} \\ \mu_{61} &= 1 + \frac{1}{2} \mu_{31} + \frac{1}{2} \mu_{51} \\ \mu_{71} &= 1 + \frac{1}{2} \mu_{41} + \frac{1}{2} \mu_{81} \\ \mu_{81} &= 1 + \frac{1}{3} \mu_{71} + \frac{1}{3} \mu_{51} + \frac{1}{3} \mu_{91} \\ \mu_{91} &= 1 + \mu_{81} \end{aligned} \right\}. \quad (27)$$

Resolviendo el sistema de ecuaciones (27) obtenemos las siguientes soluciones

$$\begin{array}{lll} \mu_{11} = 16 & \mu_{21} = 15 & \mu_{31} = 28 \\ \mu_{41} = 59 & \mu_{51} = 48 & \mu_{61} = 39 \\ \mu_{71} = 58 & \mu_{81} = 55 & \mu_{91} = 56 \end{array}$$

En consecuencia se necesitan en promedio 16 pasos para retornar al compartimento 1 partiendo de el mismo, se necesitan en promedio 15 pasos para retornar al compartimento 1 partiendo del compartimento 2, y así sucesivamente.  $\triangle$

## 2.6 Número de Retorno

En esta sección abordaremos el concepto de número de retorno, esto se refiere a la cantidad de veces que la cadena de Markov regresa a un estado particular a lo largo de un número grande de pasos. Este número nos da una idea de la frecuencia con la que el sistema visita un estado específico en el largo plazo. Es importante en la teoría ergódica y en la evaluación del comportamiento a largo plazo de la cadena de Markov, ya que nos permite entender cómo se distribuye el tiempo de la cadena entre los diferentes estados.

**Definición 25** Sea  $\{X_n: n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . Denotaremos

$$R_j = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}} \quad (28)$$

como el **número de retornos** al estado  $j \in S$  a lo largo de la cadena de Markov.

**Observación 3** Notemos que si  $R_j = \infty$  para algún estado  $j$ , entonces  $\mathbb{1}_{\{X_n=j\}}$  no necesariamente es 1 para todo  $n \in \mathbb{N}$ , si no que puede ser nulo para una cantidad finita de  $n \in \mathbb{N}$ , en este sentido diremos que  $R_j$  es infinito casi siempre.

**Lema 7** Sea  $\{X_n: n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . Entonces la distribución de probabilidad del número de retornos  $R_j$  al estado

$j$ , teniendo como condición inicial  $\{X_0 = i\}$ , está dado por:

$$\mathbb{P}(R_j = m | X_0 = i) = \begin{cases} 1 - p_{ij} & , \text{ si } m = 0, \\ p_{ij}(p_{jj})^{m-1}(1 - p_{jj}) & , \text{ si } m \geq 1. \end{cases}$$

**Demostración.** Vamos primero a considerar el caso  $m = 0$ , es decir cuando  $R_j = 0$ , en otras palabras la cadena de Markov  $\{X_n : n \geq 0\}$  nunca visita al estado  $j$ , empezando desde el estado  $X_0 = i$ . Así tenemos:

$$\begin{aligned} \mathbb{P}(R_j = 0 | X_0 = i) &= \mathbb{P}(T_i^r = \infty | X_0 = i) \\ &= 1 - \mathbb{P}(T_i^r < \infty | X_0 = i) \\ &= 1 - p_{ij}. \end{aligned}$$

Ahora vamos a considerar el caso  $m \neq 0$ . Notemos que en este caso realiza una primera visita al estado  $j$  con probabilidad  $p_{ij}$ , luego hace únicamente  $m - 1$  retornos (visitas) al estado  $j$ , después de las  $m$  visitas al estado  $j$ , nunca más regresa al estado  $j$ , y este evento ocurre con probabilidad  $1 - p_{jj}$ . Así obtenemos:

$$\mathbb{P}(R_j = m | X_0 = i) = p_{ij} \cdot (p_{jj})^{m-1} \cdot (1 - p_{jj}).$$

□

**Corolario 1** Sea  $\{X_n : n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . Entonces la distribución de probabilidad del número de retornos  $R_i$  al estado  $i$ , teniendo como condición inicial  $\{X_0 = i\}$ , está dado por

$$\mathbb{P}(R_i = m | X_0 = i) = (p_{ii})^m (1 - p_{ii}),$$

para cada  $m \geq 0$ .

**Lema 8** Sea  $\{X_n : n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . Entonces la probabilidad de  $R_j < \infty$  teniendo como condición inicial  $\{X_0 = i\}$ , está determinada por

$$\mathbb{P}(R_j < \infty | X_0 = i) = \begin{cases} (1 - p_{jj}) & , \text{ si } p_{jj} = 1 \\ 1 & , \text{ si } p_{jj} < 1 \end{cases}$$

**Demostración.** Consideremos que  $p_{jj} < 1$ , luego

$$\begin{aligned}
 \mathbb{P}(R_j < \infty | X_0 = i) &= \mathbb{P}(R_j = 0 | X_0 = i) + \sum_{m \geq 1} \mathbb{P}(R_j = m | X_0 = i) \\
 &= 1 - p_{ij} + (p_{ij}) \cdot (1 - p_{jj}) \sum_{m \geq 1} (p_{jj})^{m-1} \\
 &= 1 - p_{ij} + (p_{ij}) \cdot (1 - p_{jj}) \frac{1}{1 - p_{jj}} \\
 &= 1.
 \end{aligned}$$

En el caso que  $p_{jj} = 1$

$$\begin{aligned}
 \mathbb{P}(R_j < \infty | X_0 = i) &= \mathbb{P}(R_j = 0 | X_0 = i) + \sum_{m \geq 1} \mathbb{P}(R_j = m | X_0 = i) \\
 &= 1 - p_{ij} + \sum_{m \geq 1} (p_{ij}) \cdot (p_{jj})^{m-1} (1 - p_{jj}) \\
 &= 1 - p_{ij}.
 \end{aligned}$$

Por lo tanto

$$\mathbb{P}(R_j < \infty | X_0 = i) = \begin{cases} 1 - p_{ij} & , \text{ si } p_{jj} = 1 \\ 1 & , \text{ si } p_{jj} < 1 \end{cases}$$

□

Notemos que en el lema anterior  $i$  puede ser igual a  $j$  y la caracterización no se altera.

**Lema 9** Sea  $\{X_n : n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . Entonces la probabilidad de  $R_j = \infty$  teniendo como condición inicial  $\{X_0 = i\}$ , está determinada por

$$\mathbb{P}(R_j = \infty | X_0 = i) = \begin{cases} 1 & , \text{ si } p_{jj} = 1 \\ 0 & , \text{ si } p_{jj} < 1 \end{cases}$$

**Demostración.** En particular, si  $p_{jj} = 1$ , tenemos que  $R_j = \infty$ , así

$$\mathbb{P}(R_j = m | X_0 = i) = 0, \quad \text{para } m \geq 1.$$

Usando la identidad anterior y el Lema 7 tenemos

$$\begin{aligned}
\mathbb{P}(R_j = \infty | X_0 = i) &= 1 - \mathbb{P}(R_j < \infty | X_0 = i) \\
&= 1 - \left( \mathbb{P}(R_j = 0 | X_0 = i) + \sum_{m \geq 1} \mathbb{P}(R_j = m | X_0 = i) \right) \\
&= 1 - (1 - p_{jj}) = p_{jj}.
\end{aligned}$$

Asumamos que  $p_{jj} < 1$ , por el Lema 8 tenemos

$$\mathbb{P}(R_j = \infty | X_0 = i) = 1 - \mathbb{P}(R_j < \infty | X_0 = i) = 0.$$

□

En el lema anterior notemos que el resultado es independiente de  $i$ .

**Lema 10** *Sea  $\{X_n : n \geq 0\}$  una cadena de Markov con espacio de estados  $S$  e  $i, j \in S$ . Cuando  $p_{jj} < 1$ , se tiene que*

$$\mathbb{E}[R_j | X_0 = i] = \frac{p_{ij}}{1 - p_{jj}} < \infty.$$

*Además, si  $p_{jj} = 1$ ,  $\mathbb{E}[R_j | X_0 = i] = \infty$ , a menos que  $p_{ij} = 0$ , en tal caso  $\mathbb{E}[R_j | X_0 = i] = 0$ .*

**Demostración.** Por la definición de la esperanza condicional y el Lema 7, tenemos

$$\begin{aligned}
\mathbb{E}[R_j | X_0 = i] &= \sum_{m=0}^{\infty} m \mathbb{P}(R_j = m | X_0 = i) \\
&= \sum_{m=0}^{\infty} m \mathbb{P}(R_j = m | X_0 = i) \\
&= (1 - p_{jj}) p_{ij} \sum_{m=0}^{\infty} m (p_{jj})^{m-1} \\
&= \frac{p_{ij}}{1 - p_{jj}}.
\end{aligned}$$

□

El lema anterior nos deja el siguiente corolario.

**Corolario 2** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$  y  $i \in S$ . Entonces

$$\mathbb{E}[R_i|X_0 = i] = \frac{p_{ii}}{1 - p_{ii}} < \infty$$

si, y solamente si,  $p_{ii} < 1$ .

**Lema 11** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$  e  $i, j \in S$ . El número de retorno promedio al estado  $j$  viene dado por

$$\mathbb{E}[R_j|X_0 = i] = \sum_{n=1}^{\infty} P_{ij}^n, \quad (29)$$

donde  $P$  es la matriz de transición de la cadena de Markov.

**Demostración.** Por la Definición 28, se tiene

$$\begin{aligned} \mathbb{E}[R_j|X_0 = i] &= \mathbb{E}\left(\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}}|X_0 = i\right) \\ &= \sum_{n=1}^{\infty} \mathbb{E}(\mathbb{1}_{\{X_n=j\}}|X_0 = i) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(X_n = j|X_0 = i) = \sum_{n=1}^{\infty} P_{ij}^n \end{aligned}$$

□

## CAPÍTULO III

### CLASIFICACIÓN DE ESTADOS

En esta sección del trabajo, clasificaremos los estados de las Cadenas de Markov como estados recurrentes, estados transitorios, estados positivamente y nulamente recurrentes, así como la periodicidad y aperiodicidad de estados. Estas clasificaciones analizan características fundamentales del comportamiento de las Cadenas de Markov  $\{X_n: n \geq 0\}$ , que son fundamentales para establecer la presencia de distribuciones estacionarias y asintóticas que se desarrollan en el Capítulo 4 de la tesis. Previo a la introducción de estos conceptos, ampliaremos nuestra comprensión de la Propiedad fuerte de Markov, las referencias consultadas para esta sección son (Norris, 1998), (Privault, 2013) y (Grimmett y Welsh, 2014).

### 3.1 Propiedad fuerte de Markov

**Definición 26** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$ , definida en el espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ . Una variable aleatoria  $T: \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$  es un tiempo de parada si para todo  $n \in \mathbb{N}_0$  el evento  $T = n$  es independiente de  $X_{n+1}, X_{n+2}, \dots$  (depende de  $X_0, X_1, \dots, X_n$  solamente).

La definición conduce al siguiente teorema:

**Teorema 4 (Propiedad fuerte de Markov)** Sea  $\{X_n: n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ , con matriz de probabilidad de transición  $P$ , sea  $T$  un tiempo de parada y sea  $\{Y_n: n \geq 0\}$  una secuencia de variables aleatorias definidas por  $Y_n = X_{T+n}$  para todo  $n \in \mathbb{N}_0$ . Si  $T < \infty$  y  $X_T = i_0$ , entonces la cadena de Markov  $\{Y_n: n \geq 0\}$  con matriz de probabilidad de transición  $P$  y estado inicial  $Y_0 = i_0$ , es independiente de  $X_0, X_1, \dots, X_T$ .

**Demostración.** Sea  $i_1, i_2, \dots \in S$ , para todo  $n \in \mathbb{N}$

$$\begin{aligned} & \mathbb{P}(Y_n = i_n | Y_{n-1} = i_{n-1}, Y_{n-2} = i_{n-2}, \dots, Y_0 = i_o) \\ &= \mathbb{P}(X_{T+n} = i_n | X_{T+n-1} = i_{n-1}, X_{T+n-2} = i_{n-2}, \dots, X_T = i_o) \\ &= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \dots, X_0 = i_o), \end{aligned}$$

asumiendo que la cadena de Markov es homogénea. Por la identidad (3) tenemos

$$\begin{aligned} \mathbb{P}(Y_n = i_n | Y_{n-1} = i_{n-1}, Y_{n-2} = i_{n-2}, \dots, Y_0 = i_o) &= \mathbb{P}(Y_n = i_n | Y_{n-1} = i_{n-1}) \\ &= \mathbb{P}(X_{T+n} = i_n | X_{T+n-1} = i_{n-1}), \end{aligned}$$

donde el paso final es asumir que la cadena de Markov es homogénea. Es decir

$$\mathbb{P}(Y_n = i_n | Y_{n-1} = i_{n-1}, Y_{n-2} = i_{n-2}, \dots, Y_0 = i_o) = \mathbb{P}(Y_n = i_n | Y_{n-1} = i_{n-1}).$$

□

Por lo tanto  $\{Y_n : n \geq 0\}$  es una cadena de Markov con matriz de transición  $P$  y estado inicial  $Y_0 = i_o$ .

Sea  $H$  un evento denotado en términos de  $X_0, X_1, \dots, X_{T-1}$ . Se requiere demostrar que

$$\begin{aligned} & \mathbb{P}(Y_1 = i_1, Y_2 = i_2, \dots, H | T < \infty, X_T = i_o) \\ &= \sum_{m=0}^{\infty} P(Y_1 = i_1, Y_2 = i_2, \dots, H, T = m | T < \infty, X_T = i_o) \\ &= \sum_{m=0}^{\infty} P(X_{T+1} = i_1, X_{T+2} = i_2, \dots, H, T = m | T < \infty, X_T = i_o) \\ &= \sum_{m=0}^{\infty} P(X_{m+1} = i_1, X_{m+2} = i_2, \dots, H \cap \{T = m\} | T < \infty, X_T = i_o). \end{aligned}$$

Como  $T$  es un tiempo de parada, el evento  $H \cap \{T = m\}$  depende solamente de  $X_0, X_1, \dots, X_{m-1}$ . Por lo tanto

$$\mathbb{P}(Y_1 = i_1, Y_2 = i_2, \dots, H | T < \infty, X_T = i_o)$$

$$\begin{aligned}
&= \sum_{m=0}^{\infty} P(X_{m+1} = i_1, X_{m+2} = i_2, \dots) \mathbb{P}(H \cap \{T = m\} | T < \infty, X_T = i_0) \\
&= \sum_{m=0}^{\infty} P(X_{T+1} = i_1, X_{T+2} = i_2, \dots, T = m) \mathbb{P}(H, T = m | T < \infty, X_T = i_0) \\
&= \sum_{m=0}^{\infty} \mathbb{P}(Y_1 = i_1, Y_2 = i_2, \dots, T = m) \mathbb{P}(H, T = m | T < \infty, X_T = i_0) \\
&= P(Y_1 = i_1, Y_2 = i_2, \dots) \mathbb{P}(H | T < \infty, X_T = i_0),
\end{aligned}$$

Esto implica que  $\{Y_n : n \geq 0\}$  es independiente de  $X_0, X_1, \dots, X_T$ , según sea necesario. Esto significa que la propiedad de Markov se mantiene en un tiempo de parada aleatorio  $T$ , en lugar de simplemente en un estado conocido  $i$  esto se conoce como propiedad de Markov fuerte.

**Definición 27** Sea  $\{X_n : n \geq 0\}$  la cadena de Markov con espacio de estados  $S$ . Un estado  $j \in S$  se llama **accesible** a partir de un estado  $i \in S$ , si:

$$\mathbb{P}(X_n = j | X_0 = i) > 0,$$

para algún  $n \in \mathbb{N}$  y se denota como  $i \rightarrow j$ .

La definición anterior quiere decir que luego de  $n$  pasos el estado  $i$  conecta al estado  $j$ , con probabilidad distinta de cero.

Cuando  $i \neq j$  se tiene

$$\mathbb{P}(T_i^r < \infty | X_0 = i) \geq \mathbb{P}(T_i^r \leq n | X_0 = i) \geq \mathbb{P}(X_n = j | X_0 = i) > 0$$

**Observación 4** Desde que  $P^0 = I_d$  y

$$P_{ij}^0 = \mathbb{P}(X_0 = j | X_0 = i) = \mathbb{1}_{\{i=j\}} = \begin{cases} 1 & , \text{ si } i = j \\ 0 & , \text{ si } i \neq j \end{cases}$$

así de la Definición 27 tenemos que cualquier estado  $i$  siempre es accesible consigo mismo (en este caso en 0 pasos), esto incluso si  $P_{ii} = 0$ .

En el caso que  $i \rightarrow j$  y  $j \rightarrow i$  decimos que los estados  $i, j \in S$  están comunicados y esto es denotado por  $i \leftrightarrow j$ .

**Proposición 8** La relación binaria “ $\leftrightarrow$ ” en  $S$  es una relación de equivalencia.

**Demostración.** Para demostrar que “ $\leftrightarrow$ ” es una relación de equivalencia, debemos verificar que:

a) *Reflexiva:* Para todo  $i \in S$  tenemos  $i \leftrightarrow i$ .

La relación es reflexiva por la observación (4).

b) *Simétrica:* Para todo  $i, j \in S$ , si  $i \leftrightarrow j$  entonces  $j \leftrightarrow i$ .

Si  $i \leftrightarrow j$ , esto quiere decir que  $i \rightarrow j$  y  $j \rightarrow i$ , equivalentemente  $j \rightarrow i$  y  $i \rightarrow j$ , así tenemos que  $j \leftrightarrow i$ .

c) *Transitiva:* Para todo  $i, j, k \in S$ , si  $i \leftrightarrow j$  y  $j \leftrightarrow k$ , entonces  $i \leftrightarrow k$ .

Como  $i \leftrightarrow j$  y  $j \leftrightarrow k$ , existen  $a \geq 1$  y  $b \geq 1$  tal que

$$P_{ij}^a > 0 \quad \text{y} \quad P_{jk}^b > 0.$$

Para  $n \geq a + b$  y usando la Proposición 1 tenemos

$$\begin{aligned} \mathbb{P}(X_n = k | X_0 = i) &= \sum_{l, m \in S} \mathbb{P}(X_n = k, X_{n-b} = l, X_a = m | X_0 = i) \\ &= \sum_{l, m \in S} \mathbb{P}(X_n = k | X_{n-b} = l) \mathbb{P}(X_{n-b} = l | X_a = m) \mathbb{P}(X_a = m | X_0 = i). \end{aligned}$$

En la suma anterior tomemos el sumando  $l = m = j$ , luego

$$\begin{aligned} \mathbb{P}(X_n = k | X_0 = i) &\geq \mathbb{P}(X_n = k | X_{n-b} = j) \mathbb{P}(X_{n-b} = j | X_a = j) \\ &\quad \times \mathbb{P}(X_a = j | X_0 = i) \\ &= P_{ij}^a P_{jj}^{n-(a+b)} P_{jk}^b. \end{aligned} \tag{1}$$

En la anterior desigualdad, tomando  $n = a + b$

$$\mathbb{P}(X_n = k | X_0 = i) \geq P_{ij}^a P_{jk}^b > 0, \tag{2}$$

es decir que  $i \rightarrow k$ . Análogamente, se obtiene  $k \rightarrow i$ . De esto concluimos que

$$i \leftrightarrow k.$$

□

Del Lema 1 tenemos que la relación “ $\leftrightarrow$ ” induce una partición de clases disjuntas, es decir existen  $A_1, A_2, \dots, A_n \subset S$  tal que

- $S = A_1 \cup A_2 \cup \dots \cup A_m$
- $A_i \cap A_j = \emptyset$ , para cada  $i \neq j$ .

Los conjuntos  $A_1, A_2, \dots, A_n$  se denominan *clases comunicantes* de la cadena de Markov. Desde que “ $\leftrightarrow$ ” es una relación de equivalencia, entonces para cada  $i, j \in A_q$  tenemos que

$$i \leftrightarrow j.$$

En el caso que  $i \in A_p$  y  $j \in A_q$  con  $p \neq q$  entonces los estados  $i$  y  $j$  no están comunicados y lo denotaremos por

$$i \nleftrightarrow j.$$

**Ejemplo 18** Consideremos la cadena de Markov discreta con 3 estados  $S = \{0, 1, 2\}$ , dada por la Figura 1

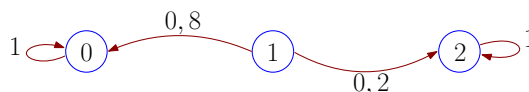


Figura 1: Cadena de Markov con 3 estados.

Por lo mencionado anteriormente en la observación (4), tenemos que esta cadena de Markov está conformada por 3 clases comunicantes

$$A_1 = \{0\}, \quad A_2 = \{1\}, \quad \text{y} \quad A_3 = \{2\}.$$

**Definición 28** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov. Decimos que la cadena de Markov es **irreducible** si esta tiene una única clase comunicante, de lo contrario decimos que la cadena es reducible.

Notemos que la cadena de Markov del Ejemplo 18 es una cadena de Markov reducible, pues esta tiene 3 clases comunicantes.

**Ejemplo 19 (Caminata aleatoria infinita en  $\mathbb{Z}$ )** Consideremos la cadena de Markov discreta dada por el modelo de la caminata aleatoria infinita en  $\mathbb{Z}$  descrita en la Sección 2.2.3 (esta se muestra en la Figura 6). Esto podemos interpretarlo como las decisiones aleatorias de un jugador, donde en cada paso gana un dólar (con probabilidad  $p > 0$ ) o lo pierde (con probabilidad  $q = 1 - p$ ). Determinar si la cadena de Markov es irreducible o reducible.

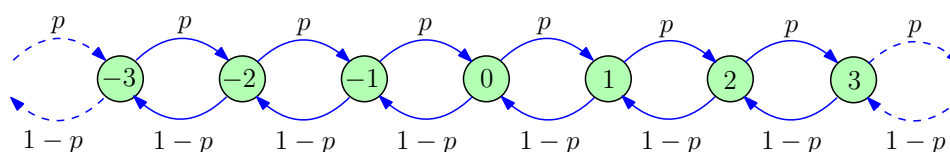


Figura 2: Caminata aleatoria infinita en  $\mathbb{Z}$

*Solución:* De la Figura 2 es fácil ver que siempre existe  $n \in \mathbb{N}$  tal

$$\mathbb{P}[X_n = j | X_0 = i] > 0$$

para cada  $i, j \in S$ , es decir  $i \rightarrow j$  para todo  $i, j \in S$ , luego la cadena de Markov tiene una única clase comunicante, está es  $\mathbb{Z}$ . Por tanto es una cadena de Markov irreducible.  $\triangle$

### 3.2 Estados Recurrentes

En esta sección abordaremos a detalle los estados recurrentes de una cadena de Markov, estos se pueden entender como estados que siempre regresan al mismo estado infinitas veces, de ahí el nombre.

**Definición 29** Sean  $\{X_n : n \geq 0\}$  una Cadena de Markov con espacio de estados  $S$ . Un estado  $i \in S$  es llamado **estado recurrente**, si:

$$p_{ii} = \mathbb{P}(T_i^r < \infty | X_0 = i) = \mathbb{P}(X_n = i \text{ para algún } n \geq 1 | X_0 = i) = 1.$$

La siguiente proposición nos da una caracterización para los estados recurrentes.

**Proposición 9** Sea  $\{X_n : n \geq 0\}$  una Cadena de Markov con espacio de estados  $S$ . Para cualquier estado  $i \in S$  en una cadena de Markov  $\{X_n : n \geq 0\}$  las siguientes afirmaciones son equivalentes:

i) El estado  $i \in S$  es recurrente, es decir  $p_{ii} = 1$ ;

ii) El número de retornos  $i \in S$  es casi siempre infinito, es decir

$$\mathbb{P}(R_i = \infty | X_0 = i) = 1;$$

iii) El número de retorno promedio del estado  $i \in S$  es infinito, es decir

$$\mathbb{E}[R_i | X_0 = i] = \infty;$$

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

**Demostración.** La demostración será realizada de la siguiente manera  $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$ . Consideremos  $i \in S$  un estado recurrente, el ítem ii) es una consecuencia inmediata del Lema 9 tomando  $j = i$ . tenemos que se cumple. Asumiendo que se cumple el ítem ii) por el Lema 9 tenemos que  $p_{ii} = 1$ , así el ítem iii) es una consecuencia del Lema 10. Desde que  $\mathbb{E}[R_i | X_0 = i] = \infty$  por el Corolario 2 tenemos que  $p_{ii} = 1$ , así por identidad (23) tenemos que el ítem iv) es valido, lo cual concluye la demostración de las equivalencias.  $\square$

El ítem ii) nos dice que un estado  $i$  es recurrente, si

$$\mathbb{P}(R_i = \infty | X_0 = i) = \mathbb{P}(X_n = i \text{ para una cantidad infinita de } n \geq 1 | X_0 = i) = 1$$

es decir que si la cadena regresa al estado  $i$  infinitas veces, es decir, si lo visita a ese estado de forma recurrente.

El siguiente corolario es una consecuencia inmediata de la proposición anterior y caracteriza los estados recurrentes. Además, en muchos textos esta caracterización aparece como definición.

**Corolario 3** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov y con espacio de estados  $S$ . Un estado  $i \in S$  es recurrente si, y sólo si,

$$\sum_{n=1}^{\infty} P_{ii}^n = \infty. \quad (3)$$

**Demostración.** Para todo  $i, j \in S$  y por (11) tenemos

$$\begin{aligned}\mathbb{E}[R_j|X_0 = i] &= \mathbb{E}\left(\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}}|X_0 = i\right) \\ &= \sum_{n=1}^{\infty} \mathbb{E}(\mathbb{1}_{\{X_n=j\}}|X_0 = i) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(X_n = j|X_0 = i) \\ &= \sum_{n=1}^{\infty} P_{ij}^n.\end{aligned}$$

En la suma anterior tomando  $j = i$  y por el ítem *iii*) de la Proposición 9 concluimos que

$$\sum_{n=1}^{\infty} P_{ii}^n = \infty.$$

□

**Corolario 4** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov con espacio de estados  $S$  y  $A$  una clase comunicante. Si  $j \in A$  es un estado recurrente, entonces cualquier estado  $i \in A$  también es recurrente.

**Demostración.** Consideremos  $|A| > 1$ . Sea  $i \in A$  diferente de  $j$ , mostraremos que el estado  $i$  es recurrente. Por definición de clase comunicante tenemos que  $i \longleftrightarrow j$ , es decir, existe  $a \geq 1$  y  $b \geq 1$  tal que

$$P_{ij}^a > 0 \quad \text{y} \quad P_{ji}^b > 0$$

Por la desigualdad (1), cuando  $k = i$  se tiene

$$\begin{aligned}\sum_{n=a+b}^{\infty} P_{ii}^n &= \sum_{n=a+b}^{\infty} \mathbb{P}(X_n = i|X_0 = i) \\ &\geq P_{ij}^a P_{ji}^b \sum_{n=a+b}^{\infty} P_{jj}^{n-(a+b)} \\ &= P_{ij}^a P_{ji}^b \sum_{n=0}^{\infty} P_{jj}^n = \infty,\end{aligned}$$

pues el estado  $j$  es recurrente, ver Corolario 3. Del mismo corolario tenemos que el estado  $i \in A$  también es recurrente.  $\square$

Esto quiere decir que la recurrencia es una propiedad que comparten todos los estados dentro de una misma clase comunicante, ya que si uno de ellos es recurrente, entonces todos los demás también lo son. Por tanto, diremos que una clase comunicante  $A \in S$  es recurrente, si alguno de sus estados es recurrente.

**Corolario 5** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov irreducible, con espacio de estados  $S$  y  $j \in S$  un estado recurrente. Entonces cualquier estado  $i \in S$  también es recurrente.

**Demostración.** La demostración es inmediata del corolario anterior.  $\square$

**Lema 12** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov irreducible donde cada estado es recurrente, Entonces

$$\mathbb{P}(T_j < \infty) = 1,$$

para todo  $j \in S$ .

**Demostración.** Vamos a presentar la prueba dada en (Norris, 1998). Dado que  $\{X_n: n \geq 0\}$  una Cadena de Markov irreducible, para cualquier  $i, j \in S$  tenemos algún  $m \in \mathbb{N}$  tal que  $P_{ij}^m > 0$ . Además, dado que todo estado es recurrente, por la Definición 29 tenemos que

$$\mathbb{P}(X_n = j \text{ para algún } n \geq 1 | X_0 = j).$$

Esto significa

$$\begin{aligned} 1 &= \mathbb{P}(X_n = j \text{ para algún } n \geq m + 1 | X_0 = j) \\ &= \sum_{i \in S} \mathbb{P}(X_n = j \text{ para algún } n \geq m + 1 | X_m = i, X_0 = j) \mathbb{P}(X_m = i | X_0 = j) \\ &= \sum_{i \in S} \mathbb{P}(T_j < \infty | X_0 = i) P_{ji}^{(m)}. \end{aligned}$$

Como sabemos  $\sum_{i \in S} P_{ji}^{(m)} = 1$ , entonces tenemos  $\mathbb{P}(T_j < \infty | X_0 = i) = 1$ . Dado que esto es válido para todo  $i \in S$ , significa que  $\mathbb{P}(T_j < \infty) = 1$ .  $\square$

### 3.3 Estados Transitorios

En esta sección abordaremos los estados transitorios, estos a diferencia de los estados recurrentes tienen la propiedad de que el estado no regresa al mismo después de un tiempo finito.

**Definición 30** Sea  $\{X_n: n \geq 0\}$  una Cadena de Markov y con espacio de estados  $S$ . Un estado  $i \in S$  es llamado **estado transitorio**, si no es recurrente, esto es:

$$p_{ii} = \mathbb{P}(T_i^r < \infty | X_0 = i) = \mathbb{P}(X_n = i \text{ para algún } n \geq 1 | X_0 = i) < 1.$$

De igual manera como en el caso de estados recurrentes tenemos la siguiente proposición que es análoga a la Proposición 9.

**Proposición 10** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov y con espacio de estados  $S$ . Para cualquier estado  $i \in S$ , las siguientes sentencias son equivalentes:

i) El estado  $i \in S$  es transitorio, es decir

$$p_{ii} < 1;$$

ii) El número de retornos de  $i \in S$  es casi siempre es finito, es decir

$$\mathbb{P}(R_i = \infty | X_0 = i) = 0 \text{ o equivalentemente } \mathbb{P}(R_i < \infty | X_0 = i) = 1; \quad (4)$$

iii) El número promedio de retornos de  $i \in S$  es finito, es decir

$$\mathbb{E}[R_i | X_0 = i] < \infty; \quad (5)$$

iv) También tenemos

$$\sum_{n=1}^{\infty} f_{ii}^{(n)} < 1. \quad (6)$$

**Demostración.** La demostración será organizada de la siguiente manera  $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$ . Consideremos  $i \in S$  un estado recurrente, el ítem ii) es una

consecuencia inmediata del Lema 9 tomando  $j = i$ . tenemos que se cumple. Asumiendo que se cumple el ítem *ii*) por el Corolario 2 tenemos que  $p_{ii} < 1$ , así el ítem *iii*) es una consecuencia del Corolario 2. Desde que  $\mathbb{E}[R_i|X_0 = i] < \infty$  por el Corolario 2 tenemos que  $p_{ii} < 1$ , así por identidad (23) tenemos que el ítem *iv*) es valido, lo cual concluye la demostración de las equivalencias.  $\square$

El ítem *ii*) nos dice que un estado  $i$  es transitorio, si

$$\mathbb{P}(R_i = \infty|X_0 = i) = \mathbb{P}(X_n = i \text{ para una cantidad infinita de } n \geq 1|X_0 = i) = 0$$

es decir que si la cadena de Markov no regresa al estado  $i$  infinitas veces.

El siguiente corolario es una consecuencia inmediata de la proposición anterior y caracteriza a los estados transitorios. Además, en muchos textos esta caracterización aparece como definición.

**Corolario 6** *Sea  $\{X_n: n \geq 0\}$  una Cadena de Markov con espacio de estados  $S$ . Un estado  $i \in S$  es transitorio si, sólo si,*

$$\sum_{n=1}^{\infty} P_{ii}^n < \infty,$$

*es decir la serie converge.*

**Demostración.** Si  $i \in S$  es un estado transitorio, por la Definición 30  $p_{ii} < 1$  y por el Corolario 2 se tiene

$$\mathbb{E}[R_i|X_0 = i] < \infty$$

y por el Lema 11, cuando  $j = i$  se tiene

$$\mathbb{E}[R_i|X_0 = i] = \sum_{n=1}^{\infty} P_{ii}^n < \infty$$

$\square$

**Corolario 7** *Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov con espacio de estados  $S$  y  $A$  una clase comunicante. Si  $j \in A$  es un estado transitorio, entonces cualquier estado  $i \in A$  también es transitorio.*

**Demostración.** Consideremos  $|A| > 2$ . Notemos que los estados o solamente son recurrentes o transitorios. Por contradicción, supongamos que el estado  $i$  es recurrente, entonces  $j$  también debe ser recurrente, pues pertenece a la misma clase comunicante, esto contradice la hipótesis.  $\square$

Por el corolario anterior notemos que si un estado es transitorio entonces todos los estados dentro de esa misma clase comunicante son transitorios. Por tanto, diremos que una clase comunicante  $A \in S$  es transitoria, si alguno de sus estados es transitorio.

**Corolario 8** Sean  $\{X_n : n \geq 0\}$  una Cadena de Markov irreducible, con espacio de estados  $S$  y  $j \in S$  un estado transitorio. Entonces cualquier estado  $i \in S$  también es transitorio.

**Demostración.** La demostración es inmediata del corolario anterior.  $\square$

El siguiente resultado nos da un mecanismo para poder determinar si un estado es recurrente o transitorio cuando la cadena de Markov es irreducible, la demostración puede ser encontrada en (Lawler, 1995).

**Lema 13** Sea  $\{X_n : n \geq 0\}$  una cadena de Markov irreducible con espacio de estados  $S$ . Fijemos  $j \in S$ , consideremos

$$0 \leq h_{ij} \leq 1; \tag{7}$$

$$h_{jj} = 1, \quad \inf\{h_{ij} : i \in S\} = 0; \tag{8}$$

Si la cadena de Markov es transitoria, entonces existe una única solución para (7) - (8). Además, si la cadena de Markov es recurrente no existe solución para (7) - (8).

El siguiente resultado dice que una cadenas de Markov con espacio de estados finito siempre tiene un estado recurrente.

**Teorema 5** Sea  $\{X_n : n \geq 0\}$  una Cadena de Markov con espacio de estados  $S$  tal que  $|S| < \infty$ . Entonces la cadena de Markov tiene al menos un estado recurrente.

**Demostración.** Sea  $j \in S$  es un estado transitorio, por el Lema 11 y el Corolario 2 tenemos

$$\mathbb{E}[R_j|X_0 = i] = \sum_{n=1}^{\infty} P_{ij}^n < \infty,$$

así la serie converge, tenemos

$$\lim_{n \rightarrow \infty} P_{ij}^n = 0. \quad (9)$$

Ahora por contradicción, supongamos que todos los estados son transitorios (es decir que no existen estados recurrentes). Como  $S$  es finito y por (9) tenemos

$$\begin{aligned} 0 &= \sum_{j \in S} \lim_{n \rightarrow \infty} P_{ij}^n = \lim_{n \rightarrow \infty} \sum_{j \in S} P_{ij}^n, \\ &= \lim_{n \rightarrow \infty} \sum_{j \in S} \mathbb{P}[X_n = j|X_0 = i], \\ &= \lim_{n \rightarrow \infty} \mathbb{P}[\cup_{j \in S} \{X_n = j\}|X_0 = i], \\ &= \lim_{n \rightarrow \infty} \mathbb{P}[\Omega|X_0 = i] = 1. \end{aligned}$$

Notemos que el cambio de la sumatoria por el limite es posible ya que el conjunto  $S$  tiene cardinalidad finita, así hemos obtenido una contradicción. Por lo tanto todos los estados no pueden ser transitorios, esto garantiza la existencia de al menos un estado recurrente para la cadena de Markov  $\{X_n: n \geq 0\}$ .  $\square$

**Lema 14** *Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados finito  $S$  tal que  $S = C_1 \cup C_2 \cup \dots \cup C_m \cup T$ , donde  $C_1, C_2, \dots, C_m$  son clases comunicantes recurrentes y  $T$  es una clase transitoria. Entonces haciendo una permutación en los estados, la matriz de transición asociada a la cadena de Markov se puede representar*

$$P = \begin{pmatrix} P_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & 0 \\ 0 & 0 & 0 & P_{m-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & P_m & 0 \\ & & R & & & Q \end{pmatrix}$$

donde cada  $P_k$  y  $Q$  son bloques cuadrados de orden  $|C_k|$ ,  $|T|$  respectivamente para  $k = 1, 2, \dots, m$  y  $R$  es una matriz rectangular no nula.

## Ejemplos de cadenas de Markov con estados recurrentes y transitorios

### Ejemplo 20 (Cadena de Markov con estados recurrentes y transitorios.)

Consideremos la cadena de Markov discreta con 3 estados  $S = \{0, 1, 2\}$ , representada por la Figura 3.

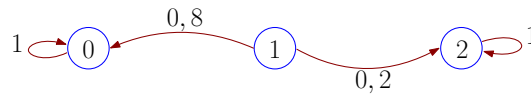


Figura 3: Cadena de Markov con 3 estados.

La matriz de transición asociada a la cadena de Markov está dada por

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0,8 & 0 & 0,2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Para cada  $n \geq 1$ , tenemos

$$P^n = \begin{pmatrix} 1 & 0 & 0 \\ 0,8 & 0 & 0,2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Luego, por los Corolarios 3 y 6 tenemos que los estados 0 y 2 son recurrentes y el estado 1 es transitorio, como afirmaba el Teorema 5.  $\triangle$

**Ejemplo 21 (Cadenas de Markov con dos estados)** Consideremos la Cadena de Markov discreta, con espacio de estados  $S = \{0, 1\}$ , cuya matriz de transición está dada por

$$P = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

donde  $\alpha, \beta \in (0, 1)$ . La representación de la cadena de Markov está dada en la Figura 4. Notemos que para cada  $n \in \mathbb{N}$ , se tiene

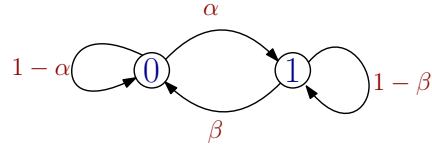


Figura 4: Cadena de Markov con 2 estados

$$P^n = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta + \alpha(1 - \alpha - \beta)^n & \alpha(1 - (1 - \alpha - \beta)^n) \\ \beta(1 - (1 - \alpha - \beta)^n) & \alpha + \beta(1 - \alpha - \beta)^n \end{pmatrix}$$

denotaremos  $\lambda = 1 - \alpha - \beta$ , así

$$P^n = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta + \alpha\lambda^n & \alpha(1 - \lambda^n) \\ \beta(1 - \lambda^n) & \alpha + \beta\lambda^n \end{pmatrix} \quad (10)$$

Desde que  $|\lambda| < 1$ , tenemos que

$$\sum_{n=1}^{\infty} P_{00}^n = \sum_{n=1}^{\infty} P_{11}^n = \infty.$$

Por lo tanto los estados 0 y 1 son recurrentes.  $\triangle$

**Observación 5** En el Ejemplo 21 se ha omitido los casos cuando  $\alpha, \beta \in \{0, 1\}$ , esto es para evitar los casos triviales.

**Ejemplo 22 (Caminata aleatoria infinita con frontera reflectante en  $\mathbb{Z}_0^+$ .)**

Consideremos el modelo de la caminata aleatoria infinita con frontera reflectante en  $\mathbb{Z}_0^+$  descrita en la Sección 2.2.3, esta se muestra en la Figura 5. Determinar que estados de la cadena de Markov son recurrente o transitorios.

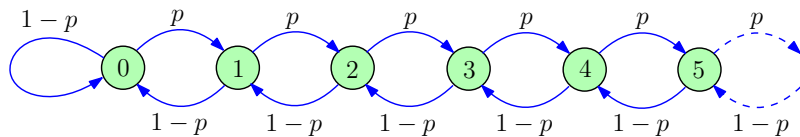


Figura 5: Caminata aleatoria infinita con frontera reflectante.

*Solución:* Desde que la cadena de Markov es irreducible tenemos que todos los estados son recurrentes o transitorios. Escojamos un estado inicial  $k \in \mathbb{N}$  arbitrario, de la Figura 5 por el análisis de un paso tenemos

$$h_{k0} = (1 - p) \cdot h_{(k-1)0} + p \cdot h_{(k+1)0}$$

para cada  $k > 0$ . Para no sobrecargar la notación denotemos  $f(k) := h_{k0}$  para cada  $k > 0$ , luego resulta

$$f(k) = (1 - p)f(k - 1) + pf(k + 1). \quad (11)$$

La solución de la ecuación en diferencias lineales (11) es de la siguiente forma

$$f(k) = \begin{cases} c_1 + c_2 \left(\frac{1-p}{p}\right)^k & \text{si } p \neq 1/2 \\ c_1 + c_2 k & \text{si } p = 1/2 \end{cases}$$

donde  $c_1, c_2 \in \mathbb{R}$ . Como  $f(0) = h_{00} = 1$  tenemos que la solución se simplifica a

$$f(k) = \begin{cases} 1 - c_2 + c_2 \left(\frac{1-p}{p}\right)^k & \text{si } p \neq 1/2 \\ 1 + c_2 k & \text{si } p = 1/2 \end{cases}$$

donde  $c_2 \in \mathbb{R}^+$ .

Notemos que si  $c_2 = 0$ , entonces  $f(k) = h_{k0} = 1$  para todo  $k > 0$ , lo cual contradice a (8) del Lema 13.

- Caso  $p < 1/2$

Supongamos que  $c_2 > 0$ , luego (7) del Lema 13 no se satisface. Luego para  $p < 1/2$  no existe solución para (7) - (8), así por el Lema 13 tenemos que la cadena de Markov es recurrente.

- Caso  $p = 1/2$

Para este caso notemos que si  $c_2 > 0$ , luego (7) del Lema 13 no se satisface. Luego para  $p = 1/2$  no existe solución para (7) - (8), de igual manera por el Lema 13 la cadena de Markov es recurrente.

- Caso  $p > 1/2$

En este caso existe solución para

$$f(k) = 1 - c_2 + c_2 \left( \frac{1-p}{p} \right)^k,$$

luego por el Lema 13 tenemos que la cadena de Markov es transitoria. Además, en este caso tenemos que  $f(k) = 0$  cuando  $k \rightarrow \infty$ , luego

$$f(k) = \left( \frac{1-p}{p} \right)^k.$$

△

Los resultados obtenidos en el ejemplo anterior están resumidos en la siguiente proposición.

**Proposición 11** *En el modelo de la caminata aleatoria infinita con frontera reflectante en  $\mathbb{Z}_0^+$ , dada en la Figura 5, cada estado es*

- i) *Recurrente, si  $p \leq 1/2$ ;*
- iii) *Transitorio, si  $p > 1/2$ .*

**Ejemplo 23 (Caminata aleatoria infinita en  $\mathbb{Z}$ )** *Consideremos el modelo de la caminata aleatoria infinita en  $\mathbb{Z}$  descrita en la Sección 2.2.3, esta se muestra en la Figura 6. Determinar que estados de la cadena de Markov son recurrente o transitorios.*

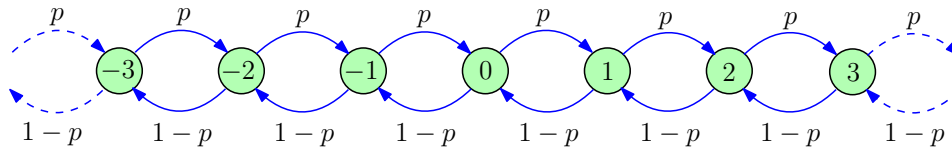


Figura 6: Caminata aleatoria infinita en  $\mathbb{Z}$

Solución: En este caso el conjunto de estados está dado por

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

y  $X_n$  se define como la posición de la caminata aleatoria en el tiempo  $n$ ,  $n \in \mathbb{N}_0$ . Por el Ejemplo 19 tenemos que la cadena de Markov es irreducible, así por los

Corolarios 5 y 7 tenemos que todos los estados son recurrentes o transitorios, sin perdida de generalidad elegiremos el estado  $i = 0$ . Sea

$$L = \sum_{n=1}^{\infty} P_{00}^n,$$

luego

$$\begin{aligned} L &= \sum_{n \text{ impar}} P_{00}^n + \sum_{n \text{ par}} P_{00}^n \\ &= \sum_{n=1}^{\infty} P_{00}^{2n} \\ &= \sum_{n=1}^{\infty} \binom{2n}{n} p^n q^n, \end{aligned}$$

donde  $q = 1 - p$ . En la primera igualdad el primer sumando es igual a cero, pues no hay manera como ir del estado 0 a el mismo en una cantidad impar de pasos.

Antes de proseguir con la resolución, notemos que

$$\frac{4^n}{2n+1} < \binom{2n}{n} < 4^n, \quad (12)$$

para cada  $n \geq 1$ . En efecto, la afirmación es verdadera desde que  $\binom{2n}{n}$  es el termino central del desarrollo del binomio de Newton  $(1+1)^{2n} = 4^n$  que está conformado por  $2n+1$  términos.

De (12) tenemos cotas para el valor de  $L$

$$\sum_{n=1}^{\infty} \frac{(4pq)^n}{2n+1} < L < \sum_{n=1}^{\infty} (4pq)^n. \quad (13)$$

i) Si  $p = q = 1/2$ , de (13) tenemos

$$L > \sum_{n=1}^{\infty} \frac{1}{2n+1} = \infty.$$

ii) Si  $p \neq q$ , se puede verificar que  $pq < 1/4$ , luego de (12) tenemos

$$L < \sum_{n=1}^{\infty} (4pq)^n < \infty,$$

pues la serie es geométrica de razón menor a 1.

Por tanto, si  $p = q = 1/2$  todos los estados son recurrentes y en otro caso todos los estados son transitorios.  $\triangle$

Los resultados obtenidos en el ejemplo anterior se van a resumir en la siguiente proposición.

**Proposición 12** *En el modelo de la caminata aleatoria infinita en  $\mathbb{Z}$ , dada por la Figura 6, la cadena es recurrente únicamente cuando  $p = 1/2$  en otro caso es transitoria.*

Un resultado interesante sobre caminos aleatorios infinitos en  $\mathbb{Z}^d$ , es que para altas dimensiones no existen estados recurrentes, para los detalles de la prueba se sugiere ver (Lawler, 1995; Neill, 2022).

**Teorema 6** *Consideremos el modelo de la caminata aleatoria infinita en  $\mathbb{Z}^d$  descrita en la Sección 2.2.3. Entonces:*

- a) *Si  $d = 2$ , cada estado es recurrente;*
- b) *Si  $d \geq 3$ , cada estado es transitorio.*

### 3.4 Estados Positivamente y Nulamente Recurrentes

En esta subsección vamos a dividir a los estados recurrentes en dos subclases de estados lo cual más adelante llamaremos de positivamente recurrente y nulamente recurrente. Estas categorías o subclases se determinan según el tiempo promedio esperado de regreso al estado  $i$ , una vez que se iniciado en el estado  $i$ .

Recordemos que el tiempo de retorno promedio al estado  $j \in S$  después de iniciar en el estado  $i \in S$  fue denotado por

$$\mu_{ij} = \mathbb{E}[T_j^r | X_0 = i],$$

en el caso del tiempo de retorno promedio al estado  $i \in S$  después de iniciar en el estado  $i \in S$  denotaremos simplemente por  $\mu(i)$  para no sobrecargar la notación, además notemos que tenemos las siguientes equivalencias

$$\mu(i) = \mathbb{E}[T_i^r | X_0 = i] = \sum_{n=1}^{\infty} n \mathbb{P}(T_i^r = n | X_0 = i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)}.$$

Notemos que cuando un estado  $i \in S$  es recurrente, tenemos que

$$\mathbb{P}(T_i^r < \infty | X_0 = i) = 1.$$

Sin embargo, esto no proporciona información sobre la finitud de su esperanza

$$\mu(i) = \mathbb{E}[T_i^r | X_0 = i].$$

La discusión anterior justifica la siguiente definición:

**Definición 31** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov con espacio de estados  $S$  e  $i \in S$  un estado recurrente de la cadena de Markov. Decimos que el estado  $i$  es

- a) **positivamente recurrente**, si el tiempo promedio de retorno al estado  $i$  después de iniciar en  $i$  es finito, es decir

$$\mu(i) = \sum_{n=1}^{\infty} n\mathbb{P}(T_i^r = n | X_0 = i) < \infty$$

- b) **nulamente recurrente**, si el tiempo promedio de retorno al estado  $i$  después de iniciar en  $i$  es infinito, es decir

$$\mu(i) = \sum_{n=1}^{\infty} n\mathbb{P}(T_i^r = n | X_0 = i) = \infty.$$

Podemos concluir que un estado  $i$  se considera positivamente recurrente si el tiempo promedio de regreso al estado  $i$ , después de haber iniciado en él, es finito. Es decir, el tiempo promedio hasta que regresemos a  $i$  es finita. Este concepto implica que, aunque puede llevar tiempo, eventualmente se regresará al estado  $i$  con certeza. Mientras que un estado  $i$  es nulamente recurrente si el tiempo promedio de regreso a  $i$  desde  $i$  es infinito. Esto significa que la probabilidad de regresar a  $i$  nula, reflejando una dinámica donde se espera que la cadena se mueva significativamente lejos de  $i$ , sin la posibilidad de un regreso.

Ahora veremos algunos resultados clásicos para el caso de cadena de Markov con espacios de estados finitos. El siguiente teorema, muestra que en toda cadena de Markov  $\{X_n: n \geq 0\}$  con espacio de estados de cardinalidad finita, no contiene clase nulamente recurrente, es decir la cadena de Markov es transitoria o positivamente recurrente.

**Teorema 7** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov y con espacio de estados  $S$ . Si  $|S| < \infty$ , entonces todos los estados recurrentes en  $S$  son positivamente recurrentes.

**Demostración.** Supongamos que algún un estado fuera nulamente recurrente, entonces todos los estados en la misma clase  $C$  seran nulamente recurrentes, ver Corolario 10. Por el Teorema 8 se tiene

$$\lim_{n \rightarrow \infty} P_{ij}^n = 0, \quad \forall i, j \in C.$$

Además, como  $C$  es una clase recurrente, para cada  $i \in C$  tenemos que  $\sum_{j \in C} P_{ij}^n = 1$  para todo  $n = 0, 1, 2, \dots$ , luego entonces

$$1 = \lim_{n \rightarrow \infty} \sum_{j \in C} P_{ij}^n = \sum_{j \in C} \lim_{n \rightarrow \infty} P_{ij}^n = 0.$$

Por lo tanto no existen estados nulamente recurrentes.  $\square$

**Corolario 9** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov irreducible con espacio de estados finito  $S$ . Entonces todos los estados de  $\{X_n: n \geq 0\}$  son positivamente recurrentes.

**Demostración.** Desde que la cadena de Markov es irreducible tenemos que está conformada por una única clase comunicante. Además, esa clase comunicante contiene un estado recurrente, ver el Teorema 5, pero como solo existe una clase comunicante todos los estados son recurrentes, así por el Teorema 7 tenemos que todos los estados recurrentes son positivamente recurrentes.  $\square$

El siguiente teorema muestra una característica que tienen los estados que no son positivamente recurrentes.

**Teorema 8** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov y con espacio de estados  $S$ . Si el estado  $j \in S$  es transitorio o nulamente recurrente, entonces

$$\lim_{n \rightarrow \infty} P_{ij}^n = 0, \tag{14}$$

para cada  $i \in S$ .

**Demostración.** Por la Proposición 10 y la Definición 3.4, siendo el estado  $j$  transitorio o nulamente recurrente, implica que la probabilidad de estar en el estado  $j$  después de  $n$  pasos,  $P_{jj}^n$  tiende a cero cuando  $n \rightarrow \infty$ . Ahora la probabilidad de estar en el estado  $j$  después de  $n$  pasos es

$$P_{ij}^n = \sum_{k \in S} \mathbb{P}(X_n = j | X_0 = i, X_{n-1} = k) \mathbb{P}(X_{n-1} = k | X_0 = i),$$

por la identidad (3) de las cadenas de Markov se tiene

$$P_{ij}^n = \sum_{k \in S} \mathbb{P}(X_n = j | X_0 = i, X_{n-1} = k) = \mathbb{P}(X_n = j | X_{n-1} = k) = P_{kj}$$

, donde  $P_{kj}$  es la probabilidad de transición de  $k$  a  $j$ . Entonces,

$$\begin{aligned} P_{ij}^n &= \sum_{K \in S} P_{kj} P_{ik}^{n-1} \\ \lim_{n \rightarrow \infty} P_{ij}^n &= \lim_{n \rightarrow \infty} \sum_{K \in S} P_{kj} P_{ik}^{n-1} \end{aligned}$$

Como  $\lim_{n \rightarrow \infty} P_{ik}^{n-1} = 0$  debido a la propiedad de que  $j$  es transitorio o nulamente recurrente y  $P_{kj}$  es una probabilidad de transición que está acotada entre 0 y 1, se concluye que

$$\lim_{n \rightarrow \infty} P_{ij}^n = 0$$

□

La siguiente proposición muestra que si en una clase tenemos un estado positivamente recurrente, entonces cada estado de la clase también lo es, es por eso que la clase será llamada clase positivamente recurrente.

**Proposición 13** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov, con espacio de estados  $S$  y  $A$  una clase recurrente. Entonces la clase  $A$  está compuesta solamente por estados positivamente recurrente o nulamente recurrentes.

**Demostración.** Sean  $i, j \in A$  tal que  $i$  es un estado nulamente recurrente. Por la desigualdad (2) tenemos

$$P_{ii}^n \geq P_{ij}^a P_{jj}^{n-(a+b)} P_{ji}^b, \quad (15)$$

donde  $P_{ij}^a > 0$  y  $P_{ji}^b > 0$ , así, aplicando limite en (15) obtenemos que

$$\lim_{n \rightarrow \infty} P_{jj}^n = 0,$$

ver Teorema 8 tenemos que  $i$  es transitorio o nulamente recurrente. □

**Corolario 10** Sean  $\{X_n: n \geq 0\}$  una Cadena de Markov irreducible y recurrente. Entonces todos los estados son positivamente recurrente o nulamente recurrentes.

**Demostración.** El corolario nos dice

- a) Si  $i$  es un estado positivamente recurrente e  $i \leftrightarrow j$ , entonces  $j$  es positivamente recurrente.
- b) Si  $i$  es un estado nulamente recurrente e  $i \leftrightarrow j$ , entonces  $j$  es nulamente recurrente.

Por tanto es suficiente mostrar cualesquiera de las afirmaciones. Demostraremos el item a). Supongamos que  $i$  es un estado positivamente recurrente, según la Definición 31 se tiene que es decir  $\mu(i) < \infty$ . Como  $i \leftrightarrow j$ , se tiene que  $j$  también es un estado recurrente. Además existen entero  $n$  y  $m$  tales que  $P_{ij}^n > 0$  y  $P_{ji}^m > 0$ . Entonces para cualquier  $k \in \mathbb{N}$ ,

$$P_{jj}^{n+m+k} \geq P_{ji}^m P_{ii}^k P_{ij}^n.$$

Sumando para  $k = 1, 2, \dots, N$  y dividiendo entre  $N$ , tenemos

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

Luego haciendo  $N \rightarrow \infty$

$$\frac{1}{\mu(i)} \geq P_{ji}^m \frac{1}{\mu(i)} P_{ij}^n > 0.$$

Por lo tanto  $\frac{1}{\mu(i)}$  es estrictamente positivo, lo que significa que  $\mu(i)$  es finito, es decir,  $j$  es positivamente recurrente.  $\square$

## Ejemplos de cadenas de Markov con estados positivamente recurrentes

**Ejemplo 24 (Caminata aleatoria infinita en  $\mathbb{Z}$ )** *El modelo de la caminata aleatoria infinita en  $\mathbb{Z}$  mostrada en la Figura 7 corresponde a una cadena de Markov, donde cada estado es recurrente para  $p = 1/2$ , ver Teorema 12. Vamos a determinar que esta cadena de Markov es nulamente recurrente..*

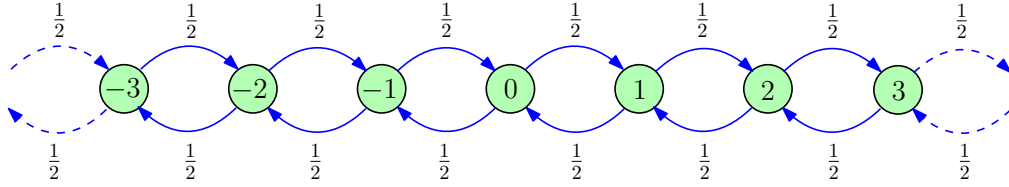


Figura 7: Caminata aleatoria infinita en  $\mathbb{Z}$

*Solución:* Como la cadena de Markov está compuesta por una única clase comunicante, por la Proposición 13 solo necesitamos analizar en un estado, sea este  $i = 0$ . De (24) tenemos

$$\mu_{00} = 1 + \frac{1}{2}\mu_{10} + \frac{1}{2}\mu_{-10}. \quad (16)$$

Otra vez usando (24), obtenemos

$$\left. \begin{aligned} \mu_{10} &= 1 + \frac{1}{2}\mu_{00} + \frac{1}{2}\mu_{20} \\ \mu_{-10} &= 1 + \frac{1}{2}\mu_{00} + \frac{1}{2}\mu_{-20} \end{aligned} \right\}$$

y remplazando en (16)

$$\mu_{00} = 4 + \frac{1}{2}\mu_{20} + \frac{1}{2}\mu_{-20}. \quad (17)$$

Análogamente podemos obtener

$$\mu_{00} = 9 + \frac{1}{2}\mu_{30} + \frac{1}{2}\mu_{-30}. \quad (18)$$

Realizando inducción matemática, para cada  $n \in \mathbb{N}$  obtenemos

$$\mu_{00} = n^2 + \frac{1}{2}\mu_{n0} + \frac{1}{2}\mu_{-n0}, \quad (19)$$

es decir

$$\mu_{00} \geq n^2,$$

para cada  $n \in \mathbb{N}$ , tomando limite cuando  $n \rightarrow \infty$ , obtenemos

$$\mu_{00} = \infty.$$

Por tanto cada estado de la cadena de Markov es nulamente recurrente, y no tiene estados positivamente recurrentes.  $\triangle$

Los resultados obtenidos en el ejemplo anterior se van a resumir en el siguiente teorema lo cual es una extensión de la proposición 12.

**Teorema 9** *Consideremos el modelo de la caminata aleatoria infinita en  $\mathbb{Z}$  descrita en la Subsección 2.2.3. Entonces:*

- a) *Si  $p = 1/2$ , cada estado es nulamente recurrente;*
- b) *Si  $p \neq 1/2$ , cada estado es transitorio.*

El siguiente teorema es una extensión de la Proposición 11.

**Teorema 10** *Consideremos el modelo de la caminata aleatoria infinita con frontera reflectante en  $\mathbb{Z}_0^+$  descrita en la Subsección 2.2.3. Entonces:*

- a) *Positivamente recurrente, si  $0 < p < 1/2$ ;*
- b) *Nulamente recurrente, si  $p \in \{0, 1/2\}$ ;*
- c) *Transitorio, si  $p > 1/2$ .*

**Demostración.** Por la Proposición 11 solo resta demostrar que la caminata aleatoria infinita con frontera reflectante en  $\mathbb{Z}_0^+$  es positivamente recurrente para  $p < 1/2$ , esto se mostrará más adelante en el Ejemplo 30 y el Teorema 11.  $\square$

### 3.5 Periodicidad y Aperiodicidad

Sea un estado  $i \in S$ , consideremos la secuencia de números enteros

$$\{n \geq 1: P_{ii}^n > 0\}$$

notemos que cada  $n$  del conjunto representa los pasos para el cual el estado  $i$  regresa al mismo.

**Definición 32** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$  e  $i \in S$ . Llamaremos **periodo del estado**  $i \in S$  al máximo común divisor de la secuencia  $\{n \geq 1 : P_{ii}^n > 0\}$  y lo denotaremos por

$$d_i := \text{mcd}\{n \geq 1 : P_{ii}^n > 0\}.$$

Cuando  $d_i = 1$  decimos que el estado  $i$  es **aperiódico** y si  $d_i > 1$  decimos que el estado es **periódico**. Además, si

$$\{n \geq 1 : P_{ii}^n > 0\} = \emptyset.$$

entonces  $d_i = 0$ .

Notemos que de la definición de estado absorbente y recurrente tenemos que estos son aperiódicos. Además, si para un estado  $i \in S$  tenemos  $d_i = 0$ , entonces por el Corolario 6 tenemos que  $i$  es un estado transitorio.

**Proposición 14** Sea una cadena de Markov  $\{X_n: n \geq 0\}$  con espacio de estados  $S$  y  $A$  una clase comunicante. Entonces cada elemento de  $A$  tiene el mismo periodo.

**Demostración.** Supongamos que  $|A| > 1$ . Sean  $i, j \in A$  con periodos  $d_i$  y  $d_j$ , respectivamente. Luego existen  $a$  y  $b$  enteros positivos tal que

$$P_{ij}^a > 0 \quad \text{y} \quad P_{ji}^b > 0,$$

entonces por la desigualdad (2) se tiene

$$P_{ii}^{a+b} \geq P_{ij}^a P_{ji}^b > 0,$$

es decir  $P_{ii}^{a+b} > 0$ , luego  $(a+b) \in \{n \geq 1 : P_{ii}^n > 0\}$  eso quiere decir que  $(a+b)$  es múltiplo de  $d_i$ .

Es fácil ver que  $\{n \geq 1 : P_{jj}^n > 0\} \neq \emptyset$ . Escojamos  $m \in \{n \geq 1 : P_{jj}^n > 0\}$ , notemos que  $m$  es múltiplo de  $d_j$ . Por la desigualdad (1) tenemos

$$P_{ii}^{m+a+b} \geq P_{ij}^a P_{jj}^m P_{ji}^b > 0,$$

esto quiere decir que  $(m+a+b) \in \{n \geq 1 : P_{ii}^n > 0\}$ , luego es múltiplo de  $d_i$ . Así

$$m = (m+a+b) - (a+b) = \overset{\circ}{d}_i$$

y como  $d_i = \text{mcd}$ , por lo que  $d_i$  es un divisor de  $a + b$ .

Si  $P_{jj}^n > 0$ , entonces tenemos

$$P_{ii}^{n+k+l} \geq P_{jj}^n P_{ij}^k P_{ji}^l > 0,$$

así  $d_i$  es divisor de  $n + k + l$ , y como  $d_i$  es divisor de  $k + l$ , entonces también será divisor de  $n$ . Esto implica que  $d_i \leq d_j$ . Obtenemos de manera similar si  $i \leftrightarrow j$  que  $d_j \leq d_i$ . Se concluye que una cadena de Markov es aperiódica cuando todos sus estados son aperiódicos.  $\square$

Por la proposición anterior los estados de una misma clase tienen el mismo periodo, consideremos esta  $d$ ; en este caso decimos que esa clase comunicante de la cadena de Markov tiene periodo  $d$ .

**Lema 15** *Sea  $i \in S$  un estado aperiódico. Entonces existe  $N \in \mathbb{N}$ , tal que  $P_{ii}^n > 0$  para todo  $n > N$ .*

**Demostración.** Sabemos que  $\text{mcd}\{n \geq 1: P_{ii}^n > 0\} = 1$ , luego podemos considerar  $n_1, n_2, \dots$ , donde  $P_{ii}^{n_j} > 0$  para todo  $j \geq 1$ . Entonces existe  $p \in \mathbb{N}$ , tal que  $\text{mcd}\{n_1, \dots, n_p\} = 1$ , bajo un re-ordenamiento de los índices. Entonces, desde la teoría de números existen  $l_1, l_2, \dots, l_p \in \mathbb{N}_0$  y  $N \in \mathbb{N}$ , tal que para cualquier  $n \geq N$  se tiene

$$n = l_1 n_1 + l_2 n_2 + \dots + l_p n_p.$$

Esto significa que tenemos

$$P_{ii}^n \geq (P_{ii}^{n_1})^{l_1} (P_{ii}^{n_2})^{l_2} \dots (P_{ii}^{n_p})^{l_p} > 0$$

para todo  $n \geq N$ .  $\square$

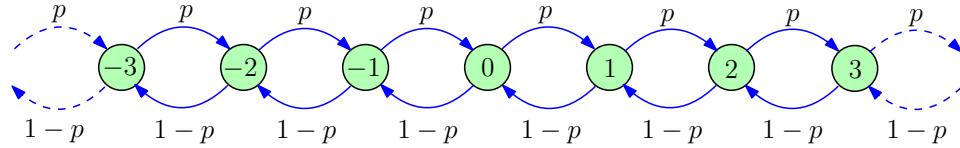
**Lema 16** *Sea una cadena de Markov  $\{X_n: n \geq 0\}$  irreducible con espacio de estados  $S$ , donde cada estado es aperiódico. Entonces para cualquier  $j, k \in S$ , existe un  $M \in \mathbb{N}$  tal que  $P_{jk}^m > 0$  para todo  $m \geq M$ .*

**Demostración.** Del Lema 15, para cada  $i \in S$ , existe  $N \in \mathbb{N}$  tal que  $P_{ii}^n > 0$  para todo  $n \geq N$ . Sean  $j, k \in S$ , desde que  $\{X_n : n \geq 0\}$  es irreducible, existen  $r, s \in \mathbb{N}$  tal que  $P_{ji}^r, P_{ik}^s > 0$ . Entonces

$$P_{jk}^{r+n+s} \geq P_{ji}^r P_{ii}^n P_{ik}^s > 0$$

para todo  $n \geq N$ . Entonces  $P_{jk}^m > 0$  para todo  $m \geq M = r + N + s$ .  $\square$

**Ejemplo 25 (Caminata aleatoria infinita en  $\mathbb{Z}$ )** Consideremos la cadena de Markov discreta dada por el modelo de la caminata aleatoria infinita en  $\mathbb{Z}$ , descrita en la Sección 2.2.3



Notemos que

$$d_0 = \text{mcd}\{n \geq 1 : P_{00}^n > 0\} = \text{mcd}\{2, 4, 6, 8, \dots\} = 2$$

$$d_1 = \text{mcd}\{n \geq 1 : P_{11}^n > 0\} = \text{mcd}\{2, 4, 6, 8, \dots\} = 2$$

De la figura, podemos ver que la cadena de Markov está compuesta por una única clase comunicante, así por la Proposición 14 todos los estados tienen el mismo periodo igual a 2.

$\triangle$

**Ejemplo 26** Consideremos la cadena de Markov discreta, dada en la Figura 8, con espacio de estados  $S = 0, 1, 2, 3$ .

De la Figura 8, podemos ver que

$$d_0 = \text{mcd}\{1, 2, 3, 4, \dots\} = 1$$

$$d_1 = \text{mcd}\{2, 4, 6, 8, \dots\} = 2$$

$$d_2 = \text{mcd}\{2, 4, 6, 8, \dots\} = 2$$

$$d_3 = \text{mcd}\{2, 4, 6, 8, \dots\} = 2$$

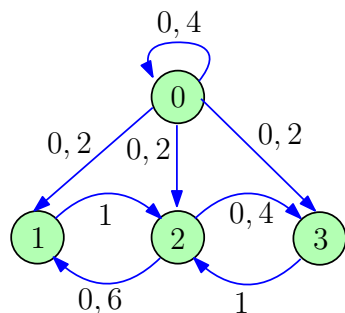


Figura 8: Cadena de Markov con dos clases comunicantes.

En la figura anterior podemos ver que el estado 0 es aperiódico y los estados  $\{1, 2, 3\}$  son periódicos, con periodo 2.

△

**Ejemplo 27 (Caminata aleatoria infinita con frontera reflectante en  $\mathbb{Z}_0^+$ .)**

Consideremos la cadena de Markov discreta dada por el modelo de la caminata aleatoria infinita con frontera reflectante en  $\mathbb{Z}_0^+$  descrita en la Sección 2.2.3 con  $p > 0$ .

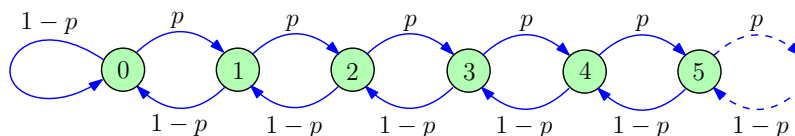


Figura 9: Caminata aleatoria infinita con frontera reflectante.

Notemos de la figura anterior que

$$d_0 = \text{mcd}\{1, 2, 3, 4, 5, \dots\} = 1$$

$$d_1 = \text{mcd}\{2, 3, 4, 5, 6, 8, \dots\} = 1$$

$$d_2 = \text{mcd}\{2, 4, 5, 6, 7, \dots\} = 1$$

$$d_3 = \text{mcd}\{2, 4, 6, 7, 8, \dots\} = 1.$$

Así los estados 0, 1, 2, 3 tienen periodo  $d = 1$ , un razonamiento análogo muestra que todos los estados tienen periodo 1 por causa de la frontera reflectante. Notemos que la cadena de Markov tiene una sola clase comunicante y por la Proposición 14 todos los estados deben tener el mismo periodo.

## CAPÍTULO IV

### EXISTENCIA Y UNICIDAD DE LAS DISTRIBUCIONES ESTACIONARIAS EN LAS CADENAS DE MARKOV

En esta sección estudiaremos como a partir de cualquier distribución de probabilidad inicial  $\pi^{(0)}$ , las distribuciones  $\pi^{(0)}P^n$  convergen a  $\pi$ , donde esta última distribución de probabilidad tiene la propiedad  $\pi = \pi P$ , es decir la distribución  $\pi$  se mantiene invariante bajo la matriz de transición. Además, como mencionamos anteriormente  $\pi$  puede ser obtenida mediante la convergencia de distribuciones donde no importa quien es la distribución inicial. Las referencias principales de este capítulo son (Norris, 1998; Privault, 2013; Neill, 2022).

**Definición 33** Sea  $S$  un conjunto numerable. Decimos que  $\pi = (\pi_i)_{i \in S}$  es una **distribución de probabilidad** si  $\pi_i \geq 0$ , para cada  $i \in S$ , y

$$\sum_{i \in S} \pi_i = 1. \tag{1}$$

Cuando  $\pi = (\pi_i)_{i \in S}$  no satisface (1) decimos simplemente que es un *vector*. Para cada  $n \in \mathbb{N}_0$  notemos que  $(\mathbb{P}(X_n = i))_{i \in S}$  es una distribución de probabilidad, ver (1), en este caso diremos que  $X_n$  *tiene distribución*

$$\pi^{(n)} = (\mathbb{P}(X_n = i))_{i \in S},$$

donde

$$\pi_i^{(n)} := \mathbb{P}(X_n = i),$$

para cada  $i \in S$ . Notemos que  $\pi^{(n)}$  describe la distribución de probabilidad de la variable aleatoria  $X_n$  en el paso  $n$ .

El siguiente lema muestra que las distribuciones de probabilidad  $\{\pi^{(n)}\}_{n \geq 0}$  están relacionadas, es decir, a partir de la distribución  $\pi^{(0)}$  y la matriz de transición  $P$  podemos determinar cualquier distribución  $\pi^{(n)}$  mediante

$$\pi^{(n)} = \pi^{(0)} P^n,$$

para  $n \geq 1$ . A  $\pi^{(0)}$  la llamaremos *distribución inicial* de la cadena de Markov, esta distribución es independiente de la matriz de transición.

**Lema 17** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta con espacio de estados  $S$  y  $\pi^{(n)}$  la distribución de  $X_n$ , para cada  $n \geq 0$ . Para cualquier  $n, m \in \mathbb{N}_0$  tenemos

$$\pi^{(n+m)} = \pi^{(m)} P^n,$$

donde  $P$  es la matriz de transición asociada a la cadena de Markov.

**Demostración.** Desde que la cadena de Markov es homogénea, para cada  $i \in S$  tenemos

$$\begin{aligned} (\pi^{(m)} P^n)_i &= \sum_{j \in S} \pi_j^{(m)} P_{ji}^n \\ &= \sum_{j \in S} \mathbb{P}(X_m = j) \mathbb{P}(X_n = i | X_0 = j) \\ &= \sum_{j \in S} \mathbb{P}(X_{m+n} = i | X_m = j) \mathbb{P}(X_m = j) \\ &= \mathbb{P}(X_{n+m} = i) \\ &= \pi_i^{(n+m)}. \end{aligned}$$

Por tanto  $\pi^{(n+m)} = \pi^{(m)} P^n$ . □

## 4.1 Distribución Estacionaria

Además de la matriz de probabilidades de transición, otro recurso fundamental para el análisis de las cadenas de Markov son las distribuciones. El Teorema 11 da las condiciones necesarias y suficientes para la existencia de distribución estacionaria asociada a una cadena de Markov.

**Definición 34** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta con espacio de estados  $S$ ,  $P$  su matriz de transición y  $\pi^{(0)}$  una distribución inicial. La distribución  $\pi^{(0)}$  es llamada **distribución estacionaria** si es también distribución de  $X_1$ , en otras palabras  $\pi^{(0)} = \pi^{(1)}$ .

De modo más general, si  $\pi^{(0)}$  es distribución estacionaria de la cadena de Markov, entonces

$$\pi^{(0)} = \pi^{(1)} = \dots = \pi^{(n)} = \pi^{(n+1)} = \dots \quad (2)$$

En efecto, como  $X_n$  tiene una distribución  $\pi^{(n)} = (\pi_i^{(n)})_{i \in S}$ , es decir

$$\mathbb{P}(X_n = i) = \pi_i^{(n)}, \quad \text{para cada } i \in S,$$

luego tenemos

$$\begin{aligned} \pi^{(n+1)} &= \mathbb{P}(X_{n+1} = j) \\ &= \sum_{i \in S} \mathbb{P}(X_{n+1} = j | X_n = i) \mathbb{P}(X_n = i) \\ &= \sum_{i \in S} \mathbb{P}(X_n = i) \mathbb{P}(X_1 = j | X_0 = i) \\ &= \sum_{i \in S} \pi_i^{(n)} P_{ij} \\ &= [\pi^{(n)} P]_j \\ &= \pi_j^{(n)} \end{aligned}$$

es decir  $\pi^{(n)} = \pi^{(n+1)}$ , para cada  $n \geq 1$ . △

Cuando la distribución inicial  $\pi^{(0)}$  es una distribución estacionaria de la cadena de Markov entonces tenemos que

$$\pi^{(0)} = \pi^{(1)} = \dots = \pi^{(n)} = \pi^{(n+1)} = \dots,$$

así para no sobrecargar la notación lo denotaremos simplemente por  $\pi$ .

**Proposición 15** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta con espacio de estados  $S$  y  $P$  su matriz de transición. El vector de probabilidad  $\pi$  es una distribución estacionaria si, y solamente si,

$$\pi = \pi P. \quad (3)$$

**Demostración.** Supongamos que  $\pi = (\pi_i)_{i \in S}$  es una distribución estacionaria para la cadena de Markov, luego para cada  $i \in S$  tenemos

$$\pi_i = \mathbb{P}(X_1 = i) = \sum_{j \in S} \mathbb{P}(X_1 = i | X_0 = j) \mathbb{P}(X_0 = j) = \sum_{j \in S} \pi_j P_{ji} = (\pi P)_i,$$

es decir  $\pi = \pi P$ . Ahora supongamos que  $\pi = \pi P$ , automáticamente tenemos que  $\pi$  es una distribución estacionaria, pues  $\pi = \pi P = \pi^{(1)}$ , notemos que por definición de distribución estacionaria aquí se está asumiendo que  $\pi^{(0)} = \pi$ .  $\square$

**Proposición 16** Sean  $\{X_n : n \geq 0\}$  una cadena de Markov con espacio de estados finito  $S$  tal que  $S = C_1 \cup C_2 \cup \dots \cup C_m \cup T$ , donde  $C_1, C_2, \dots, C_m$  representan a las clases recurrentes y  $T$  el conjunto de estados transitorios. Además, sea

$$P = \begin{pmatrix} P_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & 0 \\ 0 & 0 & 0 & P_{m-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & P_m & 0 \\ & & R & & & Q \end{pmatrix}_{S \times S}$$

la representación de la matriz de transición asociada a la cadena de Markov donde cada  $P_k$  y  $Q$  son bloques cuadrados de orden  $|C_k|$ ,  $|T|$  respectivamente para  $k = 1, 2, \dots, m$ . Entonces las siguientes afirmaciones son equivalentes:

- i)  $\pi$  es una distribución estacionaria,
- ii)  $\pi_{C_k} = \pi_{C_k} P_k$ , para cada  $k = 1, 2, \dots, m$ , y  $\pi_T = (0, 0, \dots, 0)$ , donde  $\pi_{C_k} = (\pi_i)_{i \in C_k}$  para  $k = 1, 2, \dots, m$  y  $\pi_T = (\pi_i)_{i \in T}$ .

A continuación vamos a ver ejemplos de cadena de Markov con espacio de estados finitos y notaremos que en general la distribución estacionaria puede no existir y esta no necesariamente es única.

**Ejemplo 28 (La distribución estacionaria puede ser única.)** Consideremos la cadena de Markov discreta con espacio de estados  $S = \{0, 1\}$ , cuya matriz de transición está dada por

$$P = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

donde  $\alpha, \beta \in (0, 1)$ . Notemos que cada estado de  $S$  es recurrente, ver Ejemplo 21.

Si  $\pi = (\pi_0, \pi_1)$  es una distribución estacionaria entonces  $\pi = \pi P$  y  $\pi_0 + \pi_1 = 1$ , es decir

$$\begin{aligned} \pi_0 &= (1 - \alpha)\pi_0 + \beta\pi_1 \\ \pi_1 &= \alpha\pi_0 + (1 - \beta)\pi_1 \\ 1 &= \pi_0 + \pi_1, \end{aligned}$$

resolviendo el sistema de ecuaciones anterior obtenemos

$$\pi = \left( \frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right).$$

△

El siguiente ejemplo muestra que la distribución estacionaria puede no ser única en una cadena de Markov discreta.

**Ejemplo 29 (La distribución estacionaria no es única.)** Consideremos la cadena de Markov discreta  $\{X_n : n \geq 0\}$  con espacio de estados  $S = \{0, 1, 2, 3\}$ , cuya matriz de transición está dada por

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

donde  $\alpha, \beta \in (0, 1)$ . Notemos que la cadena de Markov no es irreducible, pues tiene dos clases comunicantes

$$A = \{0, 1\}, \quad y \quad B = \{2, 3\},$$

es decir, la cadena de Markov está compuesta por dos sub-cadenas de Markov con espacio de estados  $A$  y  $B$ . Del Ejemplo 28 tenemos que existe una

distribución estacionaria asociada a cada clase comunicante, dada por

$$\pi_A = \left( \frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right), \quad \text{y} \quad \pi_B = \left( \frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right),$$

estas son las distribuciones estacionarias asociadas a las sub-cadenas de Markov con espacio de estados  $A$  y  $B$  respectivamente.

Además, para cada  $\gamma \in [0, 1]$ , es fácil verificar que

$$\pi := (\gamma\pi_A, (1 - \gamma)\pi_B) = \left( \frac{\gamma\beta}{\alpha + \beta}, \frac{\gamma\alpha}{\alpha + \beta}, \frac{(1 - \gamma)\beta}{\alpha + \beta}, \frac{(1 - \gamma)\alpha}{\alpha + \beta} \right)$$

es una distribución estacionaria de la cadena de Markov con espacio de estados  $S = \{0, 1, 2, 3\}$ , pues  $\pi = \pi P$ . Desde que  $\gamma$  es arbitrario tenemos que la distribución estacionaria no es única.

Ahora exhibiremos un ejemplo de una cadena de Markov con estados numerables infinitos, donde existen y no existen estados estacionarios.

**Ejemplo 30 (La distribución estacionaria existe.)** Consideremos el modelo de la caminata aleatoria con frontera reflectante presentada en la Subsección 2.2.3, con espacio de estados  $S = \mathbb{Z}^+ \cup \{0\}$  como se muestra en la Figura 1. Entonces para valores de  $0 < p < 1/2$  existe estado estacionario.

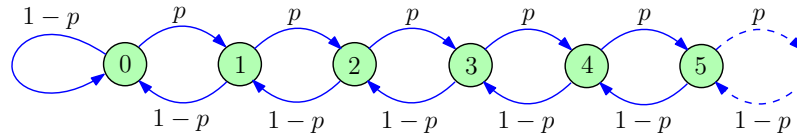


Figura 1: Caminata aleatoria infinita con frontera reflectante.

Solución: Si existe estado estacionario  $\pi$  entonces

$$(\pi_0, \pi_1, \dots) = (\pi_0, \pi_1, \dots) \begin{pmatrix} 1-p & p & 0 & 0 & 0 & 0 & 0 & \dots \\ 1-p & 0 & p & 0 & 0 & 0 & 0 & \dots \\ 0 & 1-p & 0 & p & 0 & 0 & 0 & \dots \\ 0 & 0 & 1-p & 0 & p & 0 & 0 & \dots \\ 0 & 0 & 0 & 1-p & 0 & p & 0 & \dots \\ 0 & 0 & 0 & 0 & 1-p & 0 & p & \dots \\ 0 & 0 & 0 & 0 & 0 & 1-p & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}_{S \times S}.$$

Es decir

$$(\pi_0, \pi_1, \pi_2, \dots) = \left( (1-p)\pi_0 + (1-p)\pi_1, p\pi_0 + (1-p)\pi_2, p\pi_1 + (1-p)\pi_3, \dots \right),$$

de donde se tiene el siguiente sistema de ecuaciones

$$\begin{cases} \pi_0 &= (1-p)\pi_0 + (1-p)\pi_1 \\ \pi_1 &= p\pi_0 + (1-p)\pi_2 \\ \pi_2 &= p\pi_1 + (1-p)\pi_3 \\ \pi_3 &= p\pi_2 + (1-p)\pi_4 \\ &\vdots \end{cases}$$

Resolviendo obtenemos que cada entrada  $\pi_i$  queda en términos de  $\pi_0$

$$\pi_i = \left( \frac{p}{1-p} \right)^i \pi_0 \quad \forall i \geq 1.$$

Como

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \dots = 1,$$

tenemos

$$\pi_0 \sum_{i=0}^{\infty} \left( \frac{p}{1-p} \right)^i = 1,$$

es decir  $\pi$  es una distribución estacionaria si, y solamente si,

$$\sum_{i=0}^{\infty} \left( \frac{p}{1-p} \right)^i < \infty,$$

como es una serie geométrica, solo es finita para valores  $0 < p < 1/2$ .

Para valores de  $0 < p < 1/2$ ,

$$\sum_{i=0}^{\infty} \left( \frac{p}{1-p} \right)^i = \frac{1-p}{1-2p}.$$

Luego la distribución estacionaria  $\pi = (\pi_i)_{i \in S}$  está dada por

$$\pi_0 = \frac{1-2p}{1-p}, \quad \pi_i = \left( \frac{p}{1-p} \right)^i \left( \frac{1-2p}{1-p} \right) \quad \forall i \geq 1.$$

Como caso particular, para  $p = 1/3$  tenemos la siguiente distribución estacionaria

$$\pi = \left( \frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \frac{1}{2^4}, \dots \right).$$

△

Notemos que en el Ejemplo 30 la distribución estacionaria solo existen para los valores de  $p$  para los cuales la cadena de Markov es positivamente recurrente. En la próxima subsección veremos que la propiedad de positivamente recurrente es esencial para la existencia de distribución estacionaria.

### 4.1.1. Existencia de la Distribución Estacionaria

En esta sección exhibiremos las condiciones necesarias y suficientes para la existencia y unicidad de la distribución estacionaria asociada a una cadena de Markov discreta con espacio de estados numerables e infinitos. El Teorema 11 es uno de los teoremas principal de este trabajo.

Para cada  $k \in \mathbb{N}$ , el vector  $\gamma^k = (\gamma_i^k)_{i \in S}$ , es definido como:

$$\gamma_i^k = \mathbb{E} \left( \sum_{n=0}^{T_k^r - 1} \mathbb{1}_{\{X_n = i\}} | X_0 = k \right), \quad \forall i \in S. \quad (4)$$

La sumatoria en la parte de arriba está contabilizando el número de veces que la cadena de Markov visita al estado  $i$ , desde que la cadena inicia en  $k$  y regresa por primera vez a  $k$ . Así, la cantidad  $\gamma_i^k$  definida en (4) es el número promedio que la cadena de Markov visita al estado  $i$ , desde que inicia en el estado  $k$  y regresa al estado  $k$ .

**Observación 6** De la definición anterior tenemos que  $\gamma_k^k = 1$ , para cada  $k$ .

**Observación 7** Para cada  $k \in S$ , tenemos que:

$$\sum_{i \in S} \gamma_i^k = \mathbb{E}(T_k^r | X_0 = k) = \mu(k).$$

ver

**Lema 18** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta irreducible con espacio de estados  $S$  y matriz de transición asociada a la cadena de Markov  $P$ , donde cada estado es recurrente y  $k \in S$ . Entonces,  $\gamma^k = (\gamma_i^k)_{i \in S}$  satisface

$$\gamma^k = \gamma^k P.$$

**Demostración.** Sea  $n \in \mathbb{N}$ . Entonces el evento  $\{n \leq T_k^r\}$  depende solo de  $X_0, X_1, X_2, \dots, X_{n-1}$ , entonces por la propiedad de Markov (3) se tiene:

$$\mathbb{P}(X_n = j | X_{n-1} = i, n \leq T_k^r, X_0 = k) = \mathbb{P}(X_n = j | X_{n-1} = i), \quad (5)$$

para algunos  $i \in S$ . Luego

$$\mathbb{P}(X_n = j, n \leq T_k^r | X_0 = k) = \sum_{i \in S} \mathbb{P}(X_n = j, X_{n-1} = i, n \leq T_k^r | X_0 = k),$$

así por la Proposición 2 tenemos:

$$\begin{aligned} \mathbb{P}(X_n = j, n \leq T_k^r | X_0 = k) &= \sum_{i \in S} \mathbb{P}(X_n = j | X_{n-1} = i, X_0 = k, n \leq T_k^r) \\ &\quad \mathbb{P}(X_{n-1} = i, n \leq T_k^r | X_0 = k) \end{aligned}$$

Por la ecuación (5) se tiene:

$$\begin{aligned} \mathbb{P}(X_n = j, n \leq T_k^r | X_0 = k) &= \sum_{i \in S} \mathbb{P}(X_n = j | X_{n-1} = i) \mathbb{P}(X_{n-1} = i, n \leq T_k^r | X_0 = k) \\ &= \sum_{i \in S} P_{ij} \mathbb{P}(X_{n-1} = i, n \leq T_k^r | X_0 = k) \end{aligned} \quad (6)$$

Para  $j \in S$ .

Por lo tanto, utilizando el hecho de que  $X_0 = X_{T_k^r}$ , para cambiar la suma, para cualquier  $j \in S$ ,

$$\begin{aligned} \gamma_j^k &= \mathbb{E} \left( \sum_{n=0}^{T_{k-1}^r} \mathbb{1}_{\{X_n=j\}} | X_k = 0 \right) \\ &= \mathbb{E} \left( \sum_{n=1}^{T_k^r} \mathbb{1}_{\{X_n=j\}} | X_k = 0 \right) \\ &= \mathbb{E} \left( \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j, n \leq T_k^r\}} | X_k = 0 \right) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(X_n = j, n \leq T_k^r | X_k = 0) \end{aligned}$$

luego por la ecuación (6) se tiene:

$$\gamma_j^k = \sum_{i \in S} \sum_{n=0}^{\infty} P_{ij} \mathbb{P}(X_{n-1} = i, n \leq T_k^r | X_0 = k)$$

$$= \sum_{i \in S} P_{ij} \sum_{n=0}^{\infty} \mathbb{P}(X_{n-1} = i, n \leq T_k^r | X_0 = k)$$

Al cambiar el límite inferior de la sumatoria tenemos:

$$\begin{aligned} \gamma_j^k &= \sum_{i \in S} \sum_{m=0}^{\infty} P_{ij} \mathbb{P}(X_{m-1} = i, m \leq T_{k-1}^r | X_0 = k) \\ &= \sum_{i \in S} P_{ij} \mathbb{E} \left( \sum_{m=0}^{\infty} \mathbb{1}_{\{X_m=i\}}, m \leq T_{k-1}^r | X_k = 0 \right) \\ &= \sum_{i \in S} P_{ij} \mathbb{E} \left( \sum_{m=0}^{T_{k-1}^r} \mathbb{1}_{\{X_m=i\}} | X_k = 0 \right) \\ &= \sum_{i \in S} P_{ij} \gamma_j^k \\ &= (\gamma^k P)_j, \end{aligned}$$

es decir  $\gamma^k = \gamma^k P$ . □

Del Lema anterior podemos ver que

$$0 < \gamma_i^k < \infty, \quad \forall i \in S. \quad (7)$$

**Lema 19** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta irreducible con espacio de estados  $S$ , matriz de transición asociada a la cadena de Markov  $P$  y  $\lambda = (\lambda_i)_{i \in S}$  un vector con  $\lambda_k = 1$ , para algún  $k \in S$ . Si  $\lambda = \lambda P$ , entonces  $\lambda \geq \gamma^k$ .

**Demostración.**

Como  $\lambda = \lambda P$ , para algunos  $j \in S \setminus \{k\}$  y  $n \in \mathbb{N}$  se tiene:

$$\lambda_j = \sum_{i_1 \in S} \lambda_{i_1} P_{i_1 j} = P_{kj} + \sum_{i_1 \in S \setminus \{k\}} \lambda_{i_1} P_{i_1 j}$$

ya que  $\lambda_k = 1$ , entonces:

$$\begin{aligned} \lambda_j &= P_{kj} + \sum_{i_1 \in S \setminus \{k\}} \left( \sum_{i_2 \in S} \lambda_{i_2} P_{i_2 i_1} \right) P_{i_1 j} \\ &= P_{kj} + \sum_{i_1 \in S \setminus \{k\}} \sum_{i_2 \in S} \lambda_{i_2} P_{i_2 i_1} P_{i_1 j} \end{aligned}$$

$$\begin{aligned}
&= P_{kj} + \sum_{i_1 \in S \setminus \{k\}} P_{ki_1} P_{i_1j} + \sum_{i_1 \in S \setminus \{k\}, i_2 \in S \setminus \{k\}} \lambda_{i_2} P_{i_2i_1} P_{i_1j} \\
&= \mathbb{P}(X_1 = j, T_k^r \geq 1 | X_0 = k) + \mathbb{P}(X_2 = j, T_k^r \geq 2 | X_0 = k) + \\
&\quad \dots + \sum_{i_1 \in S \setminus \{k\}, i_2 \in S \setminus \{k\}} \lambda_{i_2} P_{i_2i_1} P_{i_1j}.
\end{aligned}$$

Después de  $n$  pasos tenemos:

$$\begin{aligned}
\lambda_j &= \mathbb{P}(X_1 = j, T_k^r \geq 1 | X_0 = k) + \mathbb{P}(X_2 = j, T_k^r \geq 2 | X_0 = k) + \dots \\
&\quad + \mathbb{P}(X_n = j, T_k^r \geq n | X_0 = k) + \sum_{i_1, i_2, \dots, i_n \in S \setminus \{k\}, i_2 \in S \setminus \{k\}} \lambda_{i_2} P_{i_2i_1} P_{i_1j} \\
&\geq \mathbb{P}(X_1 = j, T_k^r \geq 1 | X_0 = k) + \mathbb{P}(X_2 = j, T_k^r \geq 2 | X_0 = k) + \dots \\
&\quad + \mathbb{P}(X_n = j, T_k^r \geq n | X_0 = k) \\
&= \mathbb{E}(\mathbb{1}_{\{X_1=j, T_k^r \geq 1\}} | X_0 = k) + \mathbb{E}(\mathbb{1}_{\{X_2=j, T_k^r \geq 2\}} | X_0 = k) + \dots \\
&\quad + \mathbb{E}(\mathbb{1}_{\{X_n=j, T_k^r \geq n\}} | X_0 = k).
\end{aligned}$$

Como  $n \rightarrow \infty$ ,

$$\begin{aligned}
&\mathbb{E}(\mathbb{1}_{\{X_1=j, T_k^r \geq 1\}} | X_0 = k) + \mathbb{E}(\mathbb{1}_{\{X_2=j, T_k^r \geq 2\}} | X_0 = k) + \dots \\
&\quad \dots + \mathbb{E}(\mathbb{1}_{\{X_n=j, T_k^r \geq n\}} | X_0 = k) \rightarrow \gamma_j^k
\end{aligned}$$

entonces  $\lambda_j \geq \gamma_j^k$ , para cualquier  $j \neq k$ , claramente  $\lambda_k = 1 = \gamma_k^k$ .  $\square$

**Lema 20** Sean  $\{X_n : n \geq 0\}$  una cadena de Markov discreta irreducible con espacio de estados  $S$  donde cada estado es recurrente, matriz de transición asociada a la cadena de Markov  $P$  y  $\lambda = (\lambda_i)_{i \in S}$  un vector con  $\lambda_k = 1$  para algún  $k \in S$ . Si  $\lambda = \lambda P$ , entonces  $\lambda = \gamma^k$ .

**Demostración.** Por el Lema 19 se tiene,  $\lambda \geq \gamma^k \Rightarrow \mu(k) = \lambda - \gamma^k \geq 0$ . Como  $\lambda_k = 1 = \gamma_k^k$ , se tiene  $\mu(k) = \lambda - \gamma^k = 1 - 1 = 0$ .

Por el Lema 18 un estado es recurrente si  $\lambda_k = \lambda_k P$ . Esto significa que  $\mu = \mu P$ . Como  $X_n$  es una cadena de Markov Irreducible, por lo que para cualquier  $j \in S$ , y  $j \rightarrow k$ . Esto significa que existen  $n \in \mathbb{N}$  tal que  $P_{jk}^n > 0$ . Por lo tanto

$$0 = \mu(k) = \sum_{i \in S} \mu(i) P_{jk}^n \geq \mu(j) P_{jk}^n$$

Como  $P_{jk}^n > 0$  y  $\mu(j) \geq 0$ , tenemos  $\mu = 0$ , lo que significa  $\lambda = \gamma_k$ .  $\square$

Con los Lemas anteriores podemos demostrar el teorema principal de esta sección.

**Teorema 11** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta irreducible. Entonces, existe una distribución estacionaria para la cadena de Markov si, y solamente si, la cadena de Markov es positivamente recurrente. En este caso la distribución estacionaria  $\pi = (\pi_i)_{i \in S}$  es única y cada  $\pi_i = \frac{1}{\mu(i)}$ .

**Demostración.** Sea  $k \in S$  cualquier estado y sea  $\pi$  una distribución estacionaria. Por lo tanto se cumple  $\sum_{i \in S} \pi_i = 1$ , entonces  $\pi_j > 0$ , para algún  $j \in S$ . Como  $\{X_n: n \geq 0\}$  es irreducible,  $j \rightarrow k$  ( $\forall k \in S$ ), fijando un  $k$ , existe  $n \in \mathbb{N}$  tal que  $P_{jk}^n > 0$ . Por lo tanto se tiene:

$$\pi_k = \sum_{i \in S} \pi_i P_{ik}^n \geq \pi_j P_{jk}^n > 0$$

Definamos el vector  $\lambda$  como  $\lambda = \frac{\pi}{\pi_k}$ , luego  $\lambda_k = \frac{\pi_k}{\pi_k} = 1$  y

$$\lambda P = \frac{\pi P}{\pi_k} = \frac{\pi}{\pi_k} = \lambda,$$

entonces por el Lema 19, se tiene  $\lambda \geq \gamma^k$ . Además, por la Observación 7, tenemos:

$$\mu(k) = \sum_{i \in S} \gamma_i^k \leq \sum_{i \in S} \lambda_i = \sum_{i \in S} \frac{\pi_i}{\pi_k} = \frac{\sum_{i \in S} \pi_i}{\pi_k} = \frac{1}{\pi_k} < \infty. \quad (8)$$

Así, por la Definición 31 tenemos que el estado  $k$  es positivamente recurrente. Además, como la cadena de Markov es irreducible luego esta es positivamente recurrente.

De (8) tenemos que  $\mu(k) \leq \frac{1}{\pi_k}$ , verificaremos que en realidad se tiene que  $\mu(k) = \frac{1}{\pi_k}$ . En efecto, desde que la cadena de Markov es positivamente recurrente el Lema 18 garantiza que  $\lambda = \gamma^k$  así, siguiendo los pasos de (8) tenemos que

$$\mu(k) = \frac{1}{\pi_k}. \quad (9)$$

De la última identidad tenemos que cuando existe distribución estacionaria esta es única, pues cada entrada está únicamente determinado por los valores de  $\mu(k)$ , es decir

$$\pi = \left( \frac{1}{\mu(k)} \right)_{k \in S}.$$

Ahora probaremos la reciproca. Dado que  $\{X_n: n \geq 0\}$  una cadena de Markov irreducible, entonces cada estado es positivamente recurrente, ver Corolario 10. Por el Lema 18 tenemos que  $\gamma^k = \gamma^k P$ , para todo  $k \in S$ . Además

$$\sum_{i \in S} \gamma_i^k = \mu(k) < \infty, \quad (10)$$

pues cada estado es positivamente recurrente. Así, definamos el vector

$$\pi^{(k)} = (\pi_i^{(k)})_{i \in S},$$

donde  $\pi_i^{(k)} = \frac{\gamma_i^k}{\mu(k)}$  para cada  $i \in S$ . Luego

$$\sum_{i \in S} \pi_i^{(k)} = \frac{1}{\mu(k)} \sum_{i \in S} \gamma_i^k = \frac{\mu(k)}{\mu(k)} = 1,$$

es decir  $\pi^{(k)}$  es una distribución. Dado que  $\gamma^k = \gamma^k P$ , dividimos por  $\mu(k)$  obtenemos  $\pi^{(k)} = \pi^{(k)} P$ . Esto significa que  $\pi^{(k)}$  es una distribución estacionaria, mostrando la existencia. Por la primera parte sabemos que cuando la distribución estacionaria existe esa es única, e.d.  $\pi^k = \left( \frac{1}{\mu(i)} \right)_{i \in S}$  para cada  $k \in S$ .  $\square$

## 4.2 Distribución de Probabilidad Asintótica

Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$ , notemos que nada garantiza la existencia del siguiente limite

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

para  $i, j \in S$  y en caso que exista puede depender tanto de  $i$  como de  $j$ .

**Definición 35** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$ . Decimos que el vector no-negativo  $\pi^\infty = (\pi_j^\infty)_{j \in S}$  es una **distribución de probabilidad asintótica**, si

$$\pi_j^\infty = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i) \quad (11)$$

para todo  $i, j \in S$ , y

$$\sum_{j \in S} \pi_j^\infty = 1. \quad (12)$$

Notemos que en la condición (11) la coordenada  $\pi_j^\infty$  no depende de  $i$ . Luego,  $\pi^\infty$  es una distribución de probabilidad asintótica si, es un vector de probabilidad y además:

$$\begin{aligned} \lim_{n \rightarrow \infty} P^n &= \lim_{n \rightarrow \infty} \begin{pmatrix} \mathbb{P}(X_n = 1 | X_0 = 1) & \mathbb{P}(X_n = 2 | X_0 = 1) & \dots & \mathbb{P}(X_n = m | X_0 = 1) & \dots \\ \mathbb{P}(X_n = 1 | X_0 = 2) & \mathbb{P}(X_n = 2 | X_0 = 2) & \dots & \mathbb{P}(X_n = m | X_0 = 2) & \dots \\ \mathbb{P}(X_n = 1 | X_0 = 3) & \mathbb{P}(X_n = 2 | X_0 = 3) & \dots & \mathbb{P}(X_n = m | X_0 = 3) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}_{S \times S} \\ &= \begin{pmatrix} \pi_1^\infty & \pi_2^\infty & \dots & \pi_m^\infty & \dots \\ \pi_1^\infty & \pi_2^\infty & \dots & \pi_m^\infty & \dots \\ \pi_1^\infty & \pi_2^\infty & \dots & \pi_m^\infty & \dots \\ \vdots & \vdots & \dots & \vdots & \ddots \end{pmatrix}_{S \times S} = \begin{pmatrix} \pi^\infty \\ \pi^\infty \\ \pi^\infty \\ \vdots \end{pmatrix}_{S \times S}. \end{aligned}$$

A lo largo de esta sección a la distribución de probabilidad asintótica la llamaremos simplemente asintótica.

**Lema 21** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados finitos  $S$ . Si la condición (11) se cumple, entonces (12) es válida.

**Demostración.**

Fijemos  $i \in S$ , luego

$$\sum_{j \in S} \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i) = \lim_{n \rightarrow \infty} \sum_{j \in S} \mathbb{P}(X_n = j | X_0 = i) = \lim_{n \rightarrow \infty} \mathbb{P}(\Omega | X_0 = i) = 1.$$

□

**Ejemplo 31** Consideremos la Cadena de Markov del Ejemplo 21, cuya matriz de transición está dada por

$$P = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

donde  $\alpha, \beta \in (0, 1)$ . Verificaremos que esta cadena de Markov tiene distribución asintótica.

Solución: Del Ejemplo 21 tenemos

$$P^n = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta + \alpha\lambda^n & \alpha(1 - \lambda^n) \\ \beta(1 - \lambda^n) & \alpha + \beta\lambda^n \end{pmatrix},$$

para cada  $n \in \mathbb{N}$ , donde  $\lambda = 1 - \alpha - \beta$ . Como  $0 < -\lambda < 1$  es fácil ver que

$$\lim_{n \rightarrow \infty} \lambda^n = 0. \quad (13)$$

Verificaremos la condición (11). Para  $i = 0$  y  $j = 1$  tenemos

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 1 | X_0 = 0) &= \lim_{n \rightarrow \infty} \frac{\alpha(1 - \lambda^n)}{\alpha + \beta} \\ &= \frac{\alpha}{\alpha + \beta} - \frac{\alpha}{\alpha + \beta} \lim_{n \rightarrow \infty} \lambda^n \\ &= \frac{\alpha}{\alpha + \beta}, \end{aligned}$$

análogamente obtenemos

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 1 | X_0 = 1) = \frac{\alpha}{\alpha + \beta}$$

y

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0 | X_0 = 0) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0 | X_0 = 1) = \frac{\beta}{\alpha + \beta}.$$

Además

$$\lim_{n \rightarrow \infty} P^n = \lim_{n \rightarrow \infty} \frac{1}{\alpha + \beta} \begin{pmatrix} \beta + \alpha\lambda^n & \alpha(1 - \lambda^n) \\ \beta(1 - \lambda^n) & \alpha + \beta\lambda^n \end{pmatrix} = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}.$$

Por lo tanto la Cadena de Markov admite la Distribución asintótica

$$\pi^\infty = \left( \frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right)$$

△

**Observación 8** *Notemos que la distribución asintótica es igual a la distribución estacionaria, ver Ejemplo 28.*

Decimos que una matriz  $P$  es *regular*, si existe  $n \geq 1$  tal que todas las entradas de la matriz  $P^n$  son distintas de cero.

**Proposición 17** *Consideremos  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados finito  $S$ , cuya matriz de transición  $P$  es regular. Entonces  $\{X_n: n \geq 0\}$  tiene una distribución asintótica.*

**Demostración.** Una cadena de Markov  $\{X_n: n \geq 0\}$  con espacio de estados finito es regular si es aperiódica e irreducible, cf. por el Lema 16.  $\square$

La siguiente proposición dice que en una cadena de Markov con estados finitos toda distribución asintótica en caso que exista es estacionaria, más adelante (Teorema 12) veremos que esta afirmación es válida para espacio con estados numerables, además se mostrará bajo que condiciones existe la distribución asintótica.

**Proposición 18** *Consideremos  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados finito  $S$ . Entonces si existe distribución asintótica esta es estacionaria.*

**Demostración.** Consideremos  $\pi^\infty = (\pi_j^\infty)_{j \in S}$  la distribución asintótica, luego

$$\begin{aligned}
 \pi_j^\infty &= \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i) \\
 &= \lim_{n \rightarrow \infty} \mathbb{P}(X_{n+1} = j | X_0 = i) \\
 &= \lim_{n \rightarrow \infty} \sum_{l \in S} \mathbb{P}(X_{n+1} = j | X_n = l) \mathbb{P}(X_n = l | X_0 = i) \\
 &= \lim_{n \rightarrow \infty} \sum_{l \in S} P_{lj} \mathbb{P}(X_n = l | X_0 = i) \\
 &= \sum_{l \in S} P_{lj} \lim_{n \rightarrow \infty} \mathbb{P}(X_n = l | X_0 = i) \\
 &= \sum_{l \in S} P_{lj} \pi_l^\infty = \pi^\infty P_j,
 \end{aligned}$$

donde  $P_j$  es la columna  $j$  de  $P$ . Por tanto

$$\pi^\infty = \pi^\infty P.$$

□

La siguiente proposición nos dice que en una cadena de Markov irreducible con estados transitorios no admite distribución asintótica.

**Proposición 19** *Sea una Cadena de Markov  $\{X_n: n \geq 0\}$  con espacio de estados  $S$ , tal que  $j \in S$  es un estado transitorio. Entonces se tiene*

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i) = 0,$$

para todo  $i \in S$ .

**Demostración.** Dado que  $j \in S$  es un estado transitorio tenemos que  $p_{jj} < 1$ . Así, por los Lemas 10 y 11 tenemos

$$\mathbb{E}[R_j | X_0 = i] = \sum_{n=1}^{\infty} \mathbb{P}(X_n = j | X_0 = i) < \infty$$

de donde tenemos que

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i) = 0,$$

para todo  $i \in S$ .

□

### 4.2.1. Existencia de Distribuciones Asintóticas

Nuestro interés en esta sección, es exhibir bajo que condiciones existen distribuciones asintóticas para una cadena de Markov, estas condiciones se muestran en el Teorema 12, el cual es el teorema principal de esta sección.

Ahora vamos a presentar una caracterización de la distribución asintótica la cual es a veces tomada como definición en los textos correspondientes al área de cadena de Markov, esta caracterización será importante al momento de demostrar el teorema principal de esta sección.

**Lema 22** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov con espacio de estados  $S$ . Entonces  $\pi$  es una distribución asintótica si, y solamente si

$$\pi = \lim_{n \rightarrow \infty} \pi^{(n)},$$

donde cada  $\pi^{(n)}$  está determinada a partir de cualquier distribución inicial  $\pi^{(0)}$ , e.d.  $\pi^{(n)} = \pi^{(0)} P^n$ .

**Demostración.** Consideremos  $j \in S$ . Luego

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

existe y es independiente de  $i \in S$ . Además, sea  $\pi^{(0)}$  una distribución inicial arbitraria.

Dado que  $\pi^{(n)} = \pi^{(0)} P^n$ , tenemos

$$\pi_j^{(n)} = \sum_{i \in S} \pi_i^{(0)} P_{ij}^n,$$

tomando  $n \rightarrow \infty$  obtenemos

$$\begin{aligned} \lim_{n \rightarrow \infty} \pi_j^{(n)} &= \lim_{n \rightarrow \infty} \sum_{i \in S} \pi_i^{(0)} P_{ij}^n \\ &= \sum_{i \in S} \pi_i^{(0)} \lim_{n \rightarrow \infty} P_{ij}^n \\ &= \sum_{i \in S} \pi_i^{(0)} \pi_j \\ &= \left( \sum_{i \in S} \pi_i^{(0)} \right) \pi_j = \pi_j. \end{aligned}$$

Por lo tanto  $\pi = \lim_{n \rightarrow \infty} \pi^{(n)}$ .

Ahora demostraremos la recíproca. Asumamos que  $\pi = \lim_{n \rightarrow \infty} \pi^{(n)}$  existe, donde  $\pi^{(n)} = \pi^{(0)} P^n$ , luego

$$\begin{aligned} \pi_j &= \lim_{n \rightarrow \infty} \pi_j^{(n)} \\ &= \lim_{n \rightarrow \infty} \sum_{t \in S} \pi_t^{(0)} P_{tj}^n \\ &= \sum_{t \in S} \pi_t^{(0)} \left( \lim_{n \rightarrow \infty} P_{tj}^n \right), \end{aligned}$$

como  $\pi^{(0)}$  es arbitrario tomemos  $\pi^{(0)} = (0, 0, \dots, \underbrace{1}_i, 0, \dots)$ , donde  $i \in S$ , luego

$$\pi_j = \lim_{n \rightarrow \infty} P_i^n j = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i).$$

Por lo tanto  $\pi$  es una distribución asintótica.  $\square$

**Teorema 12** Sean  $\{X_n: n \geq 0\}$  una cadena de Markov discreta irreducible y positivamente recurrente con espacio de estados  $S$ , donde cada estado es aperiódico. Entonces para cualquier distribución inicial  $\pi^{(0)}$  tenemos que

$$\lim_{n \rightarrow \infty} \pi^{(n)} = \pi, \quad (14)$$

donde  $\pi$  es una distribución estacionaria, es decir existe una única distribución asintótica para la cadena de Markov y está es estacionaria.

**Demostración.** Sea  $\{X_n: n \geq 0\}$  una cadena de Markov con una distribución inicial  $\pi^{(0)}$  y  $P$  su matriz de transición. Por el Teorema 11  $\{X_n: n \geq 0\}$  tiene una distribución estacionaria  $\pi$ . Sea  $\{Y_n: n \geq 0\}$  una cadena de Markov con una distribución inicial  $\pi$  y  $P$  su matriz de transición de probabilidad, independiente de  $\{X_n: n \geq 0\}$ .

Formando la secuencia de pares ordenados de variables aleatorias  $\{Z_n: n \geq 0\}$ , donde  $Z_n = (X_n, Y_n)$  para todo  $n \in \mathbb{N}_0$ .

Claramente  $Z_n = (X_n, Y_n)$  es un espacio de estados  $S \times S$  para cualquier

$$h_0 = (i_0, j_0), h_1 = (i_1, j_1), \dots, h_n = (i_n, j_n) \in S \times S,$$

$n \in \mathbb{N}$  luego

$$\begin{aligned} & \mathbb{P}(Z_n = h_n | Z_{n-1} = h_{n-1}, \dots, Z_0 = h_0) \\ &= \mathbb{P}(X_n = i_n, Y_n = j_n | X_{n-1} = i_{n-1}, Y_{n-1} = j_{n-1}, \dots, X_0 = i_0, Y_0 = j_0) \\ &= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \mathbb{P}(Y_n = j_n | Y_{n-1} = j_{n-1}, \dots, Y_0 = j_0). \end{aligned}$$

y por la propiedad de Markov 15 tenemos

$$\begin{aligned} & \mathbb{P}(Z_n = h_n | Z_{n-1} = h_{n-1}, \dots, Z_0 = h_0) \\ &= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) \mathbb{P}(Y_n = j_n | Y_{n-1} = j_{n-1}) \end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}(X_n = i_n, Y_n = j_n | X_{n-1} = i_{n-1}, Y_{n-1} = j_{n-1}) \\
&= \mathbb{P}(Z_n = h_n | Z_{n-1} = h_{n-1}),
\end{aligned}$$

entonces  $\{Z_n : n \geq 0\}$  es una Cadena de Markov.

De las ecuaciones anteriores vemos que  $\{Z_n : n \geq 0\}$  tiene una matriz de transición de probabilidad  $Q$ , donde

$$Q_{(ik)(jl)} = P_{ij}P_{kl}$$

para todo  $i, j, k \in S$ . Además  $\{Z_n : n \geq 0\}$  tiene una distribución inicial  $\mu^{(0)}$ , donde

$$\mu_{(ik)}^{(0)} = \pi_i^{(0)} \pi_k$$

para todo  $i, k \in S$ .

Sean  $i, j, k, l \in S$ . Dado que  $\{X_n : n \geq 0\}$  es una Cadena de Markov irreducible y todo estado es aperiódico, por el Lema 16 existe  $N, M \in \mathbb{N}$  tal que  $P_{ij}^n, P_{kl}^m > 0$  para todo  $n \geq N, m \geq M$ . Esto significa  $P_{ij}^r, P_{kl}^r > 0$  para todo  $r \neq R = \max\{N, M\}$ .

Por lo tanto

$$Q_{(i,k)(j,l)} = P_{ij}^r P_{kl}^r > 0,$$

para todo  $r \geq R$ . Esto significa  $(i, k) \rightarrow (j, l)$ , por lo que  $\{X_n : n \geq 0\}$  es irreducible.

Claramente  $\{Z_n : n \geq 0\}$  tiene una distribución invariante  $v$ , donde  $v_{i,k} = \pi_i \pi_k$  para todo  $i, k \in S$ . Por lo tanto por el Teorema 11 todo estado recurrente de  $\{Z_n : n \geq 0\}$  positivamente recurrente.

Sea  $b \in S$  y considere  $T_{b,b}$  con respecto  $\{Z_n : n \geq 0\}$ , el primer paso a  $(b, b)$ . Esto es

$$T_{b,b} = \inf\{n \in \mathbb{N} | Z_n = (b, b)\} = \inf\{n \in \mathbb{N} | X_n = Y_n = b\}.$$

Ahora consideremos las secuencias de las variables aleatorias  $(A_n), (B_n), (C_n)$  donde para todo  $n \in \mathbb{N}_0$

$$A_n = \begin{cases} X_n & \text{si } n < T_{(b,b)}, \\ Y_n & \text{si } n \geq T_{(b,b)}, \end{cases}$$

$$B_n = \begin{cases} Y_n & \text{si } n < T_{(b,b)}, \\ X_n & \text{si } n \geq T_{(b,b)}, \end{cases}$$

$$C_n = (A_n, B_n)$$

Aplicando el Teorema de la propiedad fuerte de Markov 4 a  $Z_n$  en el tiempo  $T_{(b,b)}$ , vemos que  $(X_{T_{(b,b)}+n}, Y_{T_{(b,b)}+n})$  es una cadena de Markov con matriz de probabilidad de transición  $Q$ , con una distribución inicial  $\mu^{T_{(b,b)}}$ , y es independiente de  $(X_0, Y_0), (X_1, Y_1), \dots, (X_{T-1}, Y_{T-1})$ . Por lo tanto  $(C_n)$  es una cadena de Markov con matriz de transición  $Q$  y una distribución inicial  $\mu^{T_{(b,b)}}$ , por tanto  $(A_n)$  es una cadena de Markov con matriz de probabilidad de transición  $P$ , con una distribución inicial  $\pi^{(0)}$ .

Entonces tenemos que para algunos  $n \in \mathbb{N}, i \in S$ ,

$$\begin{aligned} |\pi^{(n)} - \pi_i| &= |\mathbb{P}(A_n = i) - \mathbb{P}(Y_n = i)| \\ &= |\mathbb{P}(A_n = i, n < T_{(b,b)}) + \mathbb{P}(A_n = i, n \geq T_{(b,b)}) - \mathbb{P}(Y_n = i)| \\ &= |\mathbb{P}(X_n = i, n < T_{(b,b)}) + \mathbb{P}(Y_n = i, n \geq T_{(b,b)}) - \mathbb{P}(Y_n = i)| \\ &= |\mathbb{P}(X_n = i, n < T_{(b,b)}) - \mathbb{P}(Y_n = i, n \geq T_{(b,b)})|, \end{aligned}$$

luego por la Proposición 2 tenemos

$$\begin{aligned} |\pi^{(n)} - \pi_i| &= |\mathbb{P}(X_n = i | n < T_{(b,b)})\mathbb{P}(n < T_{(b,b)}) - \mathbb{P}(Y_n = i | n < T_{(b,b)})\mathbb{P}(n < T_{(b,b)})| \\ &= \mathbb{P}(n < T_{(b,b)})|\mathbb{P}(X_n = i | n < T_{(b,b)}) - \mathbb{P}(Y_n = i | n < T_{(b,b)})| \\ &\leq \mathbb{P}(n < T_{(b,b)}) \\ &= 1 - \mathbb{P}(T_{(b,b)} \leq n). \end{aligned}$$

Tomando límite se tiene:

$$\begin{aligned} \lim_{n \rightarrow \infty} |\pi^{(n)} - \pi_i| &\leq \lim_{n \rightarrow \infty} (1 - \mathbb{P}(T_{(b,b)} \leq n)) \\ &= 1 - \mathbb{P}(T_{(b,b)} < \infty). \end{aligned}$$

Dado que todo estado de  $Z_n$  es recurrente, según el Lema 12,  $\mathbb{P}(T_{(b,b)} < \infty) = 1$  luego se tiene

$$\lim_{n \rightarrow \infty} |\pi^{(n)} - \pi_i| \leq 0.$$

Por lo tanto, independientemente de la distribución inicial  $\pi^{(0)}, \pi^{(n)} \rightarrow \pi$  como  $n \rightarrow \infty$ , por lo que  $\pi$  es la asintótica de  $X_n$ .  $\square$

## CAPÍTULO V

### APLICACIONES

En el presente capítulo mostraremos una explicación básica sobre la aplicación de las Cadenas de Márkov en el buscador de Google centrándonos en el análisis de la estructura de enlaces de las páginas web, especialmente aquellos que alcanzan un estado estacionario después de un número finito de iteraciones (o clicks) desde un estado inicial, dado que estas son esenciales para el algoritmo PageRank de Google, utilizado para la búsqueda y clasificación de las páginas.

#### **5.1 Buscador de Google**

La historia de Google como motor de búsqueda se remonta a 1996, cuando Larry Page y Sergey Brin eran estudiantes de doctorado en la Universidad de Stanford, mientras trabajaban en un proyecto para mejorar la clasificación de páginas web, idearon un algoritmo llamado PageRank. Este algoritmo analiza los enlaces entre páginas web y otorga una mayor clasificación a aquellas que están vinculadas con numerosas otras páginas, en 1997, Page y Brin registraron el nombre de dominio “google.com” y lanzaron la primera versión de su motor de búsqueda. Inicialmente, el motor se llamaba BackRub, pero en 1998 fue rebautizado como “Google”. Desde entonces, Google se ha consolidado como el motor de búsqueda más utilizado en el mundo. El éxito de Google se atribuye a varios factores, entre ellos, el eficiente Algoritmo PageRank, que desempeña un papel crucial en la clasificación de páginas web, para mas detalle ver (Langville y Meyer, 2006).

### 5.1.1. Definición intuitiva del PageRank

PageRank es un algoritmo clave utilizado por Google para medir la importancia de las páginas web (Ishii y Tempo, 2014). Lo hace considerando la cantidad y calidad de enlaces que apuntan a una página (Machado, Rocha, Magalhães, y Silva, 2005). Este método proporciona una forma objetiva y mecánica de calificar las páginas web, midiendo efectivamente el interés y la atención humana (Page, Brin, Motwani, y Winograd, 1999). A pesar de la naturaleza subjetiva de la importancia de la página web, PageRank proporciona una medida de la autoridad de la página basada únicamente en la estructura topológica de la web (Bianchini, Gori, y Scarselli, 2005). Cuando se escribe una consulta en el buscador de Google hay millones de páginas que nos conducen a esa consulta, lo que hace el PageRank es clasificar esas páginas, de tal modo del que aparece en primer lugar sea probablemente el que nos interese. *Es decir PageRank es un valor numérico que permite clasificar a las páginas.*

Y según (Pretto, 2002), la clasificación de una página depende de las clasificaciones de todas las páginas que apuntan a ella, donde cada clasificación se divide por el número de enlaces externos o salientes que tienen esas páginas. En el presente trabajo el PageRank de una página ( $i$ ) denotaremos por  $Pr(i)$ .

Ahora imaginemos a un usuario de internet navegando aleatoriamente por la web, yendo de una página a otra seleccionando un enlace destacado de una página para llegar a la siguiente. Esto a veces puede llevar a callejones sin salida, es decir, páginas sin enlaces salientes, o a ciclos alrededor de un grupo de páginas interconectadas. Por lo tanto, una cierta fracción del tiempo, el navegante elige una página aleatoria de la web. Esta caminata aleatoria teórica se conoce como cadena de Markov o proceso de Markov y el PageRank es una distribución de probabilidad que representa la posibilidad de el usuario de internet, al hacer (clic) aleatoriamente en enlaces, llegue a una determinada página.

Para una mejor comprensión de la clasificación de una página en Google veremos el siguiente ejemplo.

**Ejemplo 32** Consideremos un universo de 4 páginas:  $A, B, C, D$ . Si todas las páginas enlazan a la página  $A$ , entonces el  $Pr$  (PageRank) de la página  $A$  sería

la suma del  $Pr$  de las páginas  $B, C$  y  $D$ .

$$Pr(A) = Pr(B) + Pr(C) + Pr(D)$$

graficamente tenemos:

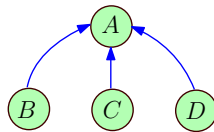


Figura 1: Página Web con 3 enlaces.

Ahora supongamos que la página web  $B$  también está vinculado a  $C$  y que  $D$  tiene enlaces a las otras 3 páginas. Cada página no puede emitir dos votos, por lo que se interpreta que  $B$  otorga medio voto a  $A$  y medio voto a  $C$ . Del mismo modo, solo se cuenta un tercio del voto de  $D$  para la página  $A$ :

$$Pr(A) = \frac{Pr(B)}{2} + Pr(C) + \frac{Pr(D)}{3}$$

graficamente tenemos:

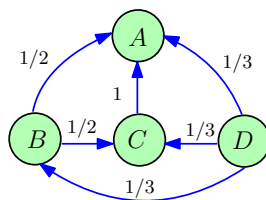


Figura 2: Página Web con vínculos entrantes.

En resumen, se distribuye el PageRank de una página web dividiéndolo entre la cantidad total de enlaces que se originan desde esa página.

$$Pr(A) = \frac{Pr(B)}{O(B)} + \frac{Pr(C)}{O(C)} + \frac{Pr(D)}{O(D)}$$

$$Pr(A) = \frac{Pr(B)}{2} + Pr(C) + \frac{Pr(D)}{3}$$

Según Langville et al.(2004), el PageRank básico comienza con la ecuación (1), donde el PageRank  $Pr(j)$  de una página  $j$ , es a la suma de los PageRank de todas las páginas que enlazan o apuntan a la página  $P_j$ .

$$Pr(j) = \sum_{i \rightarrow j} \frac{Pr(i)}{|O(i)|}, \quad j = 1, 2, 3, \dots, n. \quad (1)$$

dónde:

$Pr(j)$ : es el PageRank de la página web  $j$ .

$Pr(i)$ : es el PageRank de la página web  $i$ .

$O(i)$  : es el número de enlaces salientes o externos de la página  $i$ .

$n$ : es el número de páginas en la web.

Mensiona (Pretto, 2002), "Si una página  $i$  tiene muchos enlaces externos a la misma página  $j$ , todos estos enlaces externos cuentan como uno. Según esta definición, la clasificación de una página depende no sólo del número de páginas que apuntan a ella sino también de su importancia".

Asi mismo (Möller y Markarian, 2004), señalan que una página web obtiene una alta clasificación si la suma de las clasificaciones de los enlaces entrantes es alta, lo cual abarca tanto el caso de tener muchos enlaces entrantes como pocos con una alta clasificación. No obstante, la definición plantea ciertos inconvenientes significativos como menciona (Pretto, 2002) en relación con la posibilidad de que se produzca lo que se conoce como un "sumidero de rango"; un sumidero de clasificación es un conjunto de páginas que se enlazan entre sí, con algunos enlaces internos que apuntan hacia el grupo pero ninguno que salga de él como se muestra en la siguiente figura:

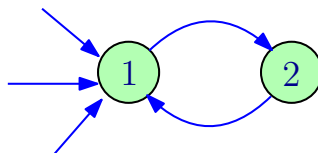


Figura 3: Ejemplo de un hundimiento de rango

En una situación de la Figura 3 , una vez que calculamos el PageRank utilizando la ecuación 1, las páginas dentro del grupo incrementan anormalmente

su PageRank, como demostraremos más adelante. Algunas consideraciones adicionales que sugieren la necesidad de modificar la ecuación 1 se presentará en la siguiente sección.

### 5.1.2. Aspectos Matemáticos

Según (Pretto, 2002), para obtener una comprensión más profunda de las dificultades asociadas con la definición intuitiva de la ecuación 1 del PageRank y la razón por la cual necesita ser modificada, nos enfocaremos en una posible manera de interpretar dicha ecuación. Si miramos el gráfico que representa la Web como la representación gráfica de una cadena de Markov de tiempo discreto ver (Karlin y McGregor, 1965), la ecuación 1 puede considerarse como el sistema de ecuaciones de equilibrio de una cadena de Markov. Esta cadena de Markov tendrá probabilidades de transición denotada por:

$$P_{ij} = \begin{cases} \frac{1}{O(i)} & \text{si existe un enlace de la página } i \text{ a } j \\ 0 & \text{si no existe} \end{cases}$$

siendo

$O(i)$  : es el número de enlaces salientes o externos que salen de la página  $i$ .

Dado que solo nos interesan los valores relativos de PageRank, PageRank  $Pr$ ,  $i = 1, 2, 3, \dots, n$ , puede ser visto como el límite y la probabilidad estacionaria de estar en el estado  $i$ , es decir, cada PageRank solo puede tomar un valor entre 0 y 1, y  $\sum_{i=1}^n Pr(i) = 1$ , donde,  $n$  es el número de todas las páginas web. Si este límite y probabilidad estacionaria existe, puede calcularse con un procedimiento iterativo, partieron del supuesto de que, al principio, todas las páginas tienen un PageRank igual  $Pr(i, 0) = \frac{1}{n}$ , donde  $n$  es el número de páginas en la Web, para mas detalle ver (Pretto, 2002). El siguiente algoritmo de PageRank itera es:

$$Pr(i; t+1) = \sum_{h \rightarrow i} \frac{Pr(h, t)}{|O(h)|}, \quad k = 1, 2, 3, \dots, n. \quad t = 0, 1, 2, 3, \dots \quad (2)$$

donde:

$Pr(i; t)$  : es el PageRank de la página  $i$  en el paso  $t$ .

El vector  $Pr(0) = [Pr(1;0), Pr(2;0), \dots, Pr(n,0)]$  da los valores del PageRank de todas las páginas en el paso 0.

**Ejemplo 33** *Supongamos que un usuario aleatorio quiere navegar en este universo de cuatro páginas, asumiremos que esta web es admisible, es decir, que cada página apunta al menos a otra (ver figura a continuación). Calcularemos el PageRank de cada página.*

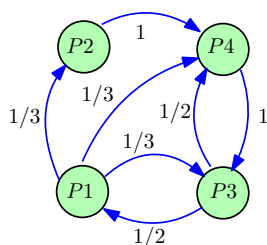


Figura 4: Red de 4 páginas

El vector  $Pr(0) = [Pr(1;0), Pr(2;0), Pr(3;0), Pr(4;0)]$  son los valores del PageRank de todas las páginas en el paso 0, es decir:

$$Pr(1;0) = Pr(2;0) = Pr(3;0) = Pr(4;0) = \frac{1}{4}$$

El valor de  $Pr(i;t+1)$  cuando  $t=0$ :

$$Pr(1;1) = \frac{Pr(3;0)}{2} = \frac{1}{8} = 0,125$$

$$Pr(2;1) = \frac{Pr(1;0)}{3} = \frac{1}{12} = 0,08333$$

$$Pr(3;1) = \frac{Pr(1;0)}{3} + \frac{Pr(4;0)}{1} = \frac{1}{12} + \frac{1}{4} = \frac{1}{3} = 0,333$$

$$Pr(4;1) = \frac{Pr(1;0)}{3} + \frac{Pr(2;0)}{1} + \frac{Pr(3;0)}{2} = \frac{1}{12} + \frac{1}{4} + \frac{1}{8} = \frac{11}{24} = 0,45833$$

El valor de  $Pr(i;t+1)$  cuando  $t=1$ :

$$Pr(1;2) = \frac{Pr(3;1)}{2} = \frac{1}{6} = 0,16667$$

$$Pr(2; 2) = \frac{Pr(1; 1)}{3} = \frac{1}{24} = 0,04167$$

$$Pr(3; 2) = \frac{Pr(1; 1)}{3} + \frac{Pr(4; 1)}{1} = \frac{1}{24} + \frac{11}{24} = \frac{1}{2} = 0,5$$

$$Pr(4; 2) = \frac{Pr(1; 1)}{3} + \frac{Pr(2; 1)}{1} + \frac{Pr(3; 1)}{2} = \frac{1}{24} + \frac{1}{12} + \frac{1}{6} = \frac{7}{24} = 0,29167$$

Tabla .1 Primeras iteraciones usando la Ecuación 2 de la Figura 4.

Iteración 0	Iteración 1	Iteración 2
$Pr(1; 0) = 0,25$	$Pr(1; 1) = 0,125$	$Pr(1; 2) = 0,1667$
$Pr(2; 0) = 0,25$	$Pr(2; 1) = 0,0833$	$Pr(2; 2) = 0,04167$
$Pr(3; 0) = 0,25$	$Pr(3; 1) = 0,3333$	$Pr(3; 2) = 0,5$
$Pr(4; 0) = 0,25$	$Pr(4; 1) = 0,4583$	$Pr(4; 2) = 0,29167$

Cuadro .1: Importancia de las páginas en las 2 primeras iteraciones.

Se puede concluir que la página P3 es más improtante que las demás páginas. Por lo que al realizar una búsqueda en internert la probabilidad de estar en la página 3 es 50 %.

**Ejemplo 34** En la siguiente Figura 5 tenemos una cadena de Markov irreducible, donde existe solo dos estados, 1 y 2 .

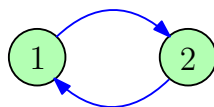


Figura 5: Cadena de Markov irreducible

En el cálculo del PageRank pueden surgir problemas si consideramos un Page-Rank inicial de:

$$Pr(1; 0) = 1, Pr(2; 0) = 0,$$

El valor de  $Pr(h; t + 1)$  cuando  $t = 0$ :

$$Pr(1; 1) = \frac{Pr(2; 0)}{1} = \frac{0}{1} = 0$$

$$Pr(2; 1) = \frac{Pr(1; 0)}{1} = \frac{1}{1} = 1$$

Cuando  $t = 1$ :

$$\begin{aligned}Pr(1;2) &= \frac{Pr(2;1)}{1} = 1 \\Pr(2;2) &= \frac{Pr(1;1)}{1} = 0\end{aligned}$$

Si seguimos iterando llegaremos en una especie de ping pong de rango sin convergencia. Y cuando consideramos un PageRank inicial

$$Pr(1;0) = Pr(2;0) = \frac{1}{2}.$$

El valor de  $Pr(i; t + 1)$  cuando  $t = 0$ :

$$\begin{aligned}Pr(1;1) &= \frac{Pr(2;0)}{1} = \frac{1}{2} \\Pr(2;1) &= \frac{Pr(2;0)}{1} = \frac{1}{2}\end{aligned}$$

Cuando  $t = 1$ :

$$\begin{aligned}Pr(1;2) &= \frac{Pr(2;1)}{1} = \frac{1}{2} \\Pr(2;2) &= \frac{Pr(1;1)}{1} = \frac{1}{2}\end{aligned}$$

Si seguimos iterando notaremos una convergencia:

$$\lim_{t \rightarrow +\infty} Pr(1;t) = \lim_{t \rightarrow +\infty} Pr(2;t) = \frac{1}{2}$$

Se concluye que no se puede definir la importancia de las páginas.

**Ejemplo 35** Consideremos la siguiente cadena de Markov:

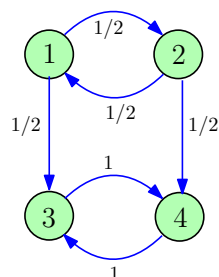


Figura 6: Cadena de Markov con fregadero de rango (de 4 páginas web).

En la Figura 6 se puede observar que existe un sumidero de rango, formado por las páginas 3 y 4, a partir de los siguientes valores de PageRank:

$$Pr(1; 0) = Pr(2; 0) = Pr(3; 0) = Pr(4; 0) = \frac{1}{4}$$

es fácil conseguir que en un determinado  $t$ :

$$Pr(1; t) = Pr(2; t) = \frac{1}{4} \frac{1}{2^t}, \quad Pr(3; t) = Pr(4; t) = \frac{1}{4} \left(1 + \frac{1}{2} + \dots + \frac{1}{2^t}\right)$$

entonces:

$$\lim_{t \rightarrow +\infty} Pr(1; t) = \lim_{t \rightarrow +\infty} Pr(2; t) = 0, \quad \lim_{t \rightarrow +\infty} Pr(3; t) = \lim_{t \rightarrow +\infty} Pr(4; t) = \frac{1}{2}$$

De estas probabilidades podemos observar que el PageRank ha sido totalmente tragado, esto ocurre cada vez que se tiene un estado finito en una cadena de Markov que no es rreducible, como se muestra en la figura 6, donde todos los estados que se encuentren fuera del círculo cerrado de la clase comunicante tendrá un PageRank igual a cero y por lo tanto será indistinguible de entre sí. Otro problema de las cadenas de Markov que no son irreducibles es que; si tenemos dos o más clases comunicantes cerradas, los valores del PageRank de las distintas páginas calculadas a través de la Ecuación 2 son diferentes, dependiendo de la elección del  $Pr(0)$ . En este caso, el orden de los valores del PageRank de las distintas páginas pueden cambiar, dependiendo de la elección del  $Pr(0)$ .

Basándonos en los ejemplos anteriores, podemos sostener que para tener una definición sólidamente establecida de PageRank, es necesario que la cadena

de Markov tenga ciertas propiedades que garanticen la existencia de distribuciones asintóticas es decir que tengan límite único estrictamente positivo y una distribución estacionaria. Esto implica que la cadena de Markov debe ser irreducible, aperiódica y positivamente recurrente, ver Teoremas 11 y 12. La presencia de estados absorbentes implica que los vectores de estado no tienden a un límite único cuando  $t \rightarrow \infty$ .

### 5.1.3. Interpretación del vector de estado estacionario

Según Langville et al. (2006), las ecuaciones 1 y 2 calculan el PageRank de una página a la vez y usando matrices, calcularemos un vector PageRank, que utiliza un único vector de  $1 \times n$  para contener los valores del PageRank para todas las páginas del índice. Para esto introduciremos una matriz  $P$  de  $n \times n$  y un vector fila  $\pi$  de  $1 \times n$ . La matriz  $P$  es una matriz de hiperenlace normalizada por filas, donde:

$$P_{ij} = \begin{cases} \frac{1}{O(i)} & \text{si existe un enlace de la página } i \text{ a } j \\ 0 & \text{si no existe} \end{cases} \quad (3)$$

siendo

$O(i)$ : es el número de enlaces salientes o externos que salen de la página  $i$ .

Cuando realizamos una búsqueda en internet, Google solo tiene en cuenta la búsqueda anterior y no las anteriores búsquedas, lo que implica que comenzamos con un valor inicial de PageRank  $\pi^{(0)}$ , Si continuamos navegando y realizamos una nueva búsqueda, el siguiente valor de PageRank sería  $\pi^{(1)} = \pi^{(0)}P$ , ver Proposición 15. Este proceso puede expresarse como:

$$\pi^{(k+1)} = \pi^{(k)}P,$$

donde:

$\pi^{(k)}$ : es el vector que nos indica la importancia de cada página. Esto está de acuerdo con lo descrito por Langville et al. (2006). Llamaremos  $\pi_i = Pr(i)$  al PageRank de la página  $i$ .

**Ejemplo 36** Consideramos el Ejemplo 33 y la ecuación (3) cuya matriz de transición  $P$  para el grafo de la figura 4 es

$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}_{4 \times 4} .$$

Calcularemos el PageRank de las páginas a notación matricial  $\pi = \pi P$ , sea  $\pi = (\pi_0, \pi_1, \pi_2, \pi_3)$  es decir

$$\left( \pi_0 \quad , \quad \pi_1 \quad , \quad \pi_2 \quad , \quad \pi_3 \right) = \left( \pi_0 \quad , \quad \pi_1 \quad , \quad \pi_2 \quad , \quad \pi_3 \right) \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\left( \pi_0 \quad , \quad \pi_1 \quad , \quad \pi_2 \quad , \quad \pi_3 \right) = \left( \frac{1}{2}\pi_2 \quad , \quad \frac{1}{3}\pi_0 \quad , \quad \frac{1}{3}\pi_0 + \pi_3 \quad , \quad \frac{1}{3}\pi_0 + \pi_1 + \frac{1}{2}\pi_2 \right),$$

luego se tiene el siguiente sistema de ecuaciones:

$$\left\{ \begin{array}{l} \pi_0 = \frac{1}{2}\pi_2 \\ \pi_1 = \frac{1}{3}\pi_0 \\ \pi_2 = \frac{1}{3}\pi_0 + \pi_3 \\ \pi_3 = \frac{1}{3}\pi_0 + \pi_1 + \frac{1}{2}\pi_2 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 = 1 \end{array} \right.$$

Luego el vector PageRank es  $(\pi_1, \pi_2, \pi_3, \pi_4) = (0,2 \quad 0,066 \quad 0,4 \quad 0,33)$ , son las puntuaciones respectivas del PageRank de las páginas.

Se concluye que la página web P3 tiene mayor relevancia que las otras páginas, con una probabilidad de 40%, seguida de la página P4 con una probabilidad de 33%, la página P1 con 20% y la página P2 con 7%, Es decir que un usuario de internet tendrá más posibilidad de acceder a la página P3.

**Ejemplo 37** Considere el siguiente modelo de web de 6 páginas como se muestra en la siguiente figura:

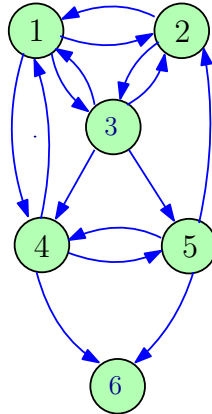


Figura 7: Grafico dirigido de 6 páginas web.

Y la matriz de transición  $P$  para el grafo de la figura 7 es:

$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}_{6 \times 6},$$

Notemos que en una fila de la matriz suma es cero, eso nos indica que tenemos una página sin enlaces externos. Este tipo de páginas se denominan nodo colgante. Los nodos colgantes representan un problema al intentar configurar una cadena de Markov. Para evitar este problema se reemplaza con  $\frac{e}{n}$ , donde  $e$  es el vector fila de todos unos y  $n$  es el orden de la matriz  $P$  a toda

la fila de ceros, creando una nueva matriz  $\bar{P}$ , para mas detalle ver (Atherton y Hogben, 2005).

$$\bar{P} = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}_{6 \times 6},$$

el resultado es una matriz estocástica.

Sin embargo, la estocasticidad no garantiza por sí sola que nuestro modelo de Markov tenga una distribución asintótica y que exista una distribución estacionaria. El otro problema que enfrenta nuestra matriz de transición  $\bar{P}$ , es que la matriz no sea regular y tengamos que hacer ajustes.

Brian y Page fuerzan que la matriz de transición sea regular asegurándose que cada entrada satisfaga  $0 < p_{ij} < 1$ . Esto garantiza la convergencia  $\pi^n$  hacia un vector único y positivo de estado estacionario. Para ello se añade una matriz de perturbación  $E = \frac{ee^T}{n}$  para formar lo que generalmente se llama **Matriz Google**, para mas detalle ver (Langville y Meyer, 2005), (Langville y Meyer, 2006) y (Page y cols., 1999).

$$\bar{\bar{P}} = \alpha \bar{P} + (1 - \alpha)E$$

Esta nueva matriz  $\bar{\bar{P}}$ , tiende a modelar mejor a un usuario que navega en internet, es decir el usuario tiene una probabilidad de  $1 - \alpha$  de saltar a un lugar aleatorio en la web, es decir escribiendo un URL en linea de comando y una probabilidad de  $\alpha$  de decidir de hacer click en un enlace externo en una página actual. Ver (Hyvönen, 2007).

Se considera el valor de  $\alpha = 0,85$ , para el cálculo de la matriz Google.

La matriz Google  $\bar{\bar{P}}$  para el Ejemplo 37 es:

$$\bar{\bar{P}} = 0,85 \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} + 0,15 \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} .$$

$$\bar{\bar{P}} = \begin{pmatrix} 0,025 & 0,308 & 0,308 & 0,308 & 0,025 & 0,025 \\ 0,45 & 0,025 & 0,45 & 0,025 & 0,025 & 0,025 \\ 0,238 & 0,238 & 0,025 & 0,238 & 0,238 & 0,025 \\ 0,308 & 0,025 & 0,025 & 0,025 & 0,308 & 0,308 \\ 0,025 & 0,308 & 0,025 & 0,025 & 0,308 & 0,308 \\ 0,167 & 0,167 & 0,167 & 0,167 & 0,167 & 0,167 \end{pmatrix}$$

Como la matriz de transición Google  $\bar{\bar{P}}$  es una matriz estocástica y por la Proposición 5, la matriz  $\bar{\bar{P}}^k$  también es estocástica, por tanto se puede determinar el vector de probabilidad estacionaria  $q$  de tamaño  $1 \times n$ , lo cual nos brindará información de la importancia de las páginas.

$$\bar{\bar{P}}^{20} = \begin{pmatrix} 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \end{pmatrix}$$

$$\bar{P}^{150} = \begin{pmatrix} 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \\ 0,199 & 0,186 & 0,179 & 0,138 & 0,168 & 0,130 \end{pmatrix}.$$

De la matriz de transición Google  $\bar{P}^k$ , se puede ver que los valores de la columna convergen para una única matriz, por tanto la matriz  $\bar{P}$ , tiene un único vector estacionario  $q$

$$q = (0,199 \quad 0,186 \quad 0,179 \quad 0,138 \quad 0,168 \quad 0,130).$$

Por lo tanto, un usuario de internet tiene más posibilidad de ingresar a la página  $P1$  en un 20 %, seguida de la página  $P2$  con un 19 %.

## CONCLUSIONES

Al finalizar este trabajo, se obtuvieron las siguientes conclusiones:

- a) Las cadenas de Markov tienen aplicaciones en campos como la estadística, la economía, la biología y la informática, permitiendo predecir comportamientos, optimizar sistemas y tomar decisiones bajo incertidumbre. Además, son la base para métodos avanzados como la simulación Monte Carlo y los modelos de aprendizaje automático. Su versatilidad las convierte en una herramienta clave para comprender fenómenos complejos en contextos reales.
- b) Se ha encontrado que en el contexto de las cadenas de Markov discretas e irreducibles, la distribución estacionaria únicamente existe cuando la cadena es positivamente recurrente, además las entradas de la distribución estacionaria están determinadas por el tiempo promedio de retorno a cada estado. Esta propiedad garantiza no sólo la existencia, sino también la unicidad de la distribución estacionaria. La investigación sobre las condiciones de existencia y unicidad de las distribuciones estacionarias en cadenas de Markov es fundamental para comprender la evolución de los procesos estocásticos.
- c) Para que exista comportamiento asintótico en las distribuciones de la cadena de Markov en tiempo discreto una condición fundamental es que la cadena sea aperiódica.
- d) En el buscador de Google, a través del algoritmo PageRank, se destaca la relevancia práctica de la aplicación de los conceptos teóricos de la teoría de cadenas de Markov. La capacidad de las cadenas de Markov para modelar la estructura de enlaces de páginas web y alcanzar un estado estacionario después de un número finito de iteraciones demuestra su utilidad en escenarios del mundo real y su impacto en la tecnología actual.

## REFERENCIAS BIBLIOGRÁFICAS

- Atherton, R., y Hogben, L. (2005). A look at markov chains and their use in google. *Iowa State University, MSM Creative Component*.
- Bianchini, M., Gori, M., y Scarselli, F. (2005). Inside pagerank. *ACM Transactions on Internet Technology (TOIT)*, 5(1), 92–128.
- Blanchet, J., Gallego, G., y Goyal, V. (2016). A markov chain approximation to choice modeling. *Operations Research*, 64(4), 886–905.
- Bogachev, V. I., y Ruas, M. A. S. (2007). *Measure theory* (Vol. 1). Springer.
- Ching, W.-K., y Ng, M. K. (2006). Markov chains. *Models, algorithms and applications*.
- Çinlar, E. (2011). *Probability and stochastics*. Springer.
- Cohn, D. L. (2013). *Measure theory* (Vol. 2). Springer.
- Dummit, D. S., y Foote, R. M. (2004). *Abstract algebra* (Vol. 3). Wiley Hoboken.
- Durrett, R., y Durrett, R. (1999). *Essentials of stochastic processes* (Vol. 1). Springer.
- Fraleigh, J. B. (2003). *A first course in abstract algebra*. Pearson Education India.
- Grimmett, G., y Welsh, D. J. (2014). *Probability: an introduction*. Oxford University Press.
- Gut, A. (2009). *Stopped random walks*. Springer.
- Halmos, P. R. (2013). *Measure theory* (Vol. 18). Springer.
- Hyvönen, S. (2007). *Linear algebra methods for data mining*. Springer, University of Helsinki.
- Ishii, H., y Tempo, R. (2014). The pagerank problem, multiagent consensus, and web aggregation: A systems and control viewpoint. *IEEE Control Systems Magazine*, 34(3), 34–53.

- Karlin, S. (2014). *A first course in stochastic processes*. Academic press.
- Karlin, S., y McGregor, J. (1965). Ehrenfest urn models. *Journal of Applied Probability*, 2(2), 352–376.
- Kelbert, M. (2008). *Probability and statistics by example: Markov chains: a primer in random processes and their applications*. Cambridge University Press.
- Klenke, A. (2013). *Probability theory: a comprehensive course*. Springer Science & Business Media.
- Kolmogorov, A. (s.f.). *Foundations of the theory of probability: Second english edition*.
- Lalley, S. P. (2009). Convergence rates of markov chains. *Course notes for Statistics*, 313.
- Langville, A. N., y Meyer, C. D. (2005). A survey of eigenvector methods for web information retrieval. *SIAM review*, 47(1), 135–161.
- Langville, A. N., y Meyer, C. D. (2006). *Google's pagerank and beyond: The science of search engine rankings*. Princeton university press.
- Lawler, G. F. (1995). *Introduction to stochastic processes*. New York: Chapman & Hall/CRC.
- Lawler, G. F., y Limic, V. (2010). *Random walk: a modern introduction* (Vol. 123). Cambridge University Press.
- Machado, C. A., Rocha, C. M. M., Magalhães, E. J., y Silva, M. D. P. (2005). *Utilização do algoritmo pagerank para obtenção de qualidade no ranqueamento de páginas web*. Anápolis - Brasil.
- Möller, N., y Markarian, R. (2004). La importancia de cada nodo en una estructura de enlaces: Google-pagerank. *Boletín de la Asociación Matemática Venezolana*, 11(2), 233–252.
- Neill, J. (2022). *Discrete-time markov chains*. Lancaster University, Lancaster, England, United Kingdom. (Tesis de Maestría)
- Norris, J. R. (1998). *Markov chains* (Vol. 2). Cambridge: Cambridge University Press.
- Page, L., Brin, S., Motwani, R., y Winograd, T. (1999). *The pagerank citation ranking: Bringing order to the web*. (Inf. Téc.). Stanford infolab.
- Pearson, K. (1905). The problem of the random walk. *Nature*, 72(1867), 342–342.

- Pretto, L. (2002). A theoretical analysis of google's pagerank. En *String processing and information retrieval: 9th international symposium, spire 2002 lisbon, portugal, september 11–13, 2002 proceedings 9* (pp. 131–144).
- Privault, N. (2013). Understanding markov chains. *Examples and Applications, Publisher Springer-Verlag Singapore, 357, 358.*
- Rolla, L. T. (2022). *Measure theory for probability.*
- Ross, S. M. (1995). *Stochastic processes.* New York: John Wiley & Sons.
- Ross, S. M. (2014). *Introduction to probability models.* Academic press.
- Seneta, E. (2016). Markov chains as models in statistical mechanics.
- Sericola, B. (2013). Front matter. En B. Sericola (Ed.), *Markov chains.*
- Suchard, M. A., Weiss, R. E., y Sinsheimer, J. S. (2001). Bayesian selection of continuous-time markov chain evolutionary models. *Molecular biology and evolution, 18(6), 1001–1013.*
- Taylor, M. E. (2006). *Measure theory and integration.* American Mathematical Soc.
- Woess, W. (2000). *Random walks on infinite graphs and groups* (n.º 138). Cambridge university press.