



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

Una técnica de separación para espacios de
estados de una cadena de Markov

REPORTE DE SEMINARIO DE TITULACIÓN

QUE PARA OBTENER EL TÍTULO DE:

Actuario

PRESENTA:

JORGE ALBERTO SÁNCHEZ HERNÁNDEZ

DIRECTOR DE TESIS:

DRA. ANA MEDA GUARDIOLA



Septiembre 2009



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Hoja de datos del jurado

1. Datos del alumno

Sánchez
Hernández
Jorge Alberto
58 45 28 10
Universidad Nacional Autónoma de México
Facultad de Ciencias
Actuaría
300197088

2. Datos del tutor

Dra.
Ana
Meda
Guardiola

3. Datos del sinodal 1

Act.
Jaime
Vázquez
Alamilla

4. Datos del sinodal 2

M. en C.
Juan Martín
Barrios
Vargas

5. Datos del sinodal 3

M. en C.
Sergio Iván
López
Ortega

6. Datos del sinodal 4

Dr.
Nelson Omar
Muriel
Torrero

7. Datos del trabajo escrito

Una técnica de separación para espacios de estados de una cadena de Markov
120 p
2009

Santiago y Fidelia.

Gracias por todo su apoyo, amor y comprensión, los cuales han sido indispensables para lograr esta meta.

Sus consejos, ejemplos y regaños han forjado al hombre que ahora soy.

A ustedes les dedico este trabajo como muestra del amor y respeto que les tengo a los dos.

¡Gracias!

Ara.

Gracias amor por haberme acompañado a lo largo de esta aventura, gracias por tu comprensión, amor, consejos y por haberme apoyado en todo momento.

Y ahora que compartimos nuestras vidas, sabemos que todos nuestros esfuerzos valieron la pena. A ti dedico este trabajo con todo mi amor.

¡Gracias chaparrita!

Lety, Mary y Lupita.

Finamente a ustedes dedico este trabajo ya que siempre me han apoyado y han estado apoyandome en toda circunstancia. Gracias por todos los bonitos momentos que hemos compartido a lo largo de nuestras vidas.

¡Las quiero!

*Ana, Sergio, Juan, Jaime y Nelson.
Gracias por su apoyo, ayuda y conse-
jos. Y por haberme apoyado para lle-
gar hasta aquí.*

¡Gracias!

Índice general

Introducción	VII
1. Cadenas de Markov con espacio de estados discreto	1
1.1. Definiciones básicas.	1
1.2. Recurrencia y transitoriedad.	3
1.2.1. Recurrencia nula y positiva.	11
1.3. Distribuciones estacionarias.	15
1.3.1. Propiedades de las distribuciones estacionarias.	16
1.3.2. Existencia y unicidad de las distribuciones estacionarias.	19
1.3.3. Cadenas Reducibles	27
2. Cadenas de Markov en espacios generales	29
2.1. Preliminares	29
2.1.1. Kernels	30
2.1.2. Cadenas de Markov	31
2.2. Kernels irreducibles	33
2.2.1. φ -irreducibilidad	38
2.3. Funciones pequeñas	43
2.4. Recurrencia y transitoriedad	46
2.4.1. Elementos básicos de teoría potencial	46
2.4.2. R-transitoriedad y R-recurrencia	50

2.4.3. Tiempos de paro para una cadena de Markov	53
2.4.4. Las partes disipativa y conservativa	56
2.5. Recurrencia	58
3. Teoría de Renovación	61
3.1. Elementos generales de Teoría de renovación	61
3.1.1. Proceso de conteo de renovación	62
3.1.2. Ecuación de renovación	68
3.2. Sucesiones y procesos de renovación.	71
3.3. Un átomo para kernels y cadenas de Markov	75
3.4. Esquema general de regeneración.	78
4. Técnica de separación.	83
Conclusiones	97
Apéndice	99

Introducción

Cuando se estudian cadenas de Markov recurrentes en espacios de estados discreto, se fija un estado y se investigan las propiedades de la cadena con base en las visitas a este estado fijo. Pero cuando el espacio de estados es un espacio general medible, este tipo de técnicas fallan, ya que en general no existe ningún punto que sea visitado con probabilidad positiva desde otro punto en el espacio de estados.

Sea $\{X_n : n = 0, 1, \dots\}$ una cadena de Markov en un espacio de estados general o infinito no numerable E . Se supondrá que $\{X_n\}_{n \geq 0}$ es una cadena de Markov Harris recurrente (que se definirá más adelante).

Este trabajo tiene como objetivo presentar una técnica que permita manipular las cadenas de Markov en un espacio de estados general de una manera similar a como se usan en espacios de estados finito o infinito numerable, esta técnica consiste en "*separar*" un espacio de estados general de una cadena de Markov Harris recurrente e introducir en el espacio separado un *átomo* (se entenderá como un átomo: un conjunto de puntos para los cuales todas sus transiciones son idénticas).

Cuando se tenga el nuevo espacio de estados, es decir, el espacio de estados original con el átomo incluido, se tiene un nuevo objetivo, el cual consiste en usar la teoría de renovación para contar los regresos de la cadena de Markov al átomo construido, de manera similar a la

técnica en espacios de estados discretos.

Con lo dicho anteriormente se tendrá un proceso bivariado, formado por la cadena de Markov $\{X_n\}_{n \geq 0}$ y una variable aleatoria asociada a la cadena que llevará los valores 0 y 1, los cuales nos indicará la dinámica de la cadena.

Este trabajo comienza con la teoría en espacios de estados finitos o infinitos numerables para ilustrar los resultados principales y posteriormente llegar a espacios de estados generales.

En el Capítulo 1 se analizarán las cadenas de Markov en espacios finitos, haciendo el mayor énfasis en propiedades de comunicación, recurrencia, transitoriedad y comportamiento límite de distribuciones estacionarias, tratando de utilizar lo indispensable para cubrir los objetivos. La bibliografía principal para este capítulo fue: [3, Capítulo 1 y 2].

En el Capítulo 2, se analizarán los fundamentos de cadenas de Markov en espacios generales, comenzando por definir qué es un kernel y cuando éste es subestocástico o estocástico (en este caso se dice que es una probabilidad de transición para una cadena de Markov, en donde ahora las transiciones son de un estado a un conjunto de estados). Se define una cadena irreducible, una cadena R -recurrente, R -transitoria, tiempos de paro, la parte conservativa y disipativa de un espacio de estados y finalmente Recurrencia y Harris Recurrencia para una cadena de Markov. Este trabajo está enfocado a estas dos últimas propiedades, ya que la técnica de separación es para cadenas Recurrentes y en particular para Harris recurrentes. La bibliografía principal para este Capítulo fue: [6, Capítulos 1-3] y [9].

En el Capítulo 3 se verá la Teoría de Renovación en donde se analizarán aspectos de la función de conteo de renovaciones, la función de renovación y la ecuación de renovación; posteriormente se analizará esta teoría ahora aplicada a los espacios generales de una cadena de Markov, que es muy similar a la primera parte de este Capítulo. Es aquí donde se definirá el concepto de átomo, que es el que se usará en la separación del espacio de estados. Y con

esto estaremos listos para poder analizar la Técnica de separación para espacios de estados de una cadena de Markov. La bibliografía principal para este Capítulo fue: [8, Capítulo 3] y [6, Capítulo 4].

Para finalizar, en el Capítulo 4 se aplica toda la teoría analizada. Tomando una cadena irreducible se construye un kernel subestocástico en el cual existe una función que tiene un comportamiento igual a la distribución Bernoulli, es esta función la que decide hacia dónde se mueve la siguiente transición de la cadena, y así se tendrá una nueva cadena $\{X_n, Y_n\}$. En este Capítulo encontraremos el Teorema central, el cual nos dice que la distribución marginal de la cadena separada $\{X_n, Y_n\}$ es igual a la distribución de la cadena original $\{X_n\}$; éste es un hecho de gran importancia, ya que $\{X_n, Y_n\}$ contiene un átomo propio y así podremos aplicar la teoría de renovación para contar el número de regresos a dicho átomo. Mediante este Teorema las propiedades de $\{X_n, Y_n\}$ serán heredadas a $\{X_n\}$. Por otro lado, al suponer $\{X_n\}$ una cadena Harris-recurrente, $\{X_n, Y_n\}$ también lo será. La bibliografía principal para este Capítulo fue: [6, Capítulo 4], [5] y [10, Capítulo 1].

Cadenas de Markov con espacio de estados discreto

En este capítulo se analizarán cadenas de Markov en un espacio de estados discreto, se describen Propiedades y Teoremas que se utilizarán de referencia cuando pasemos a una teoría más general, es decir, a un espacio de estados general o infinito no numerable de una cadena de Markov. Ya que la teoría en los espacios de estados discretos o infinitos numerables es muy amplia, nos enfocamos en las propiedades más importantes como son recurrencia, transitoriedad y distribuciones de probabilidad. La bibliografía principal para este Capítulo fue: [3, Capítulo 1 y 2]

En particular para este capítulo consideremos $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, E un conjunto finito o infinito numerable, y \mathcal{F} la σ -álgebra natural, esto es, la potencia de E . Todo proceso será observado en tiempos discretos $n = 0, 1, 2, \dots$, y X_n denotará el estado del sistema al tiempo n .

1.1. Definiciones básicas.

Comenzaremos con definiciones básicas para cadenas de Markov estacionarias en un espacio de estados discreto.

Definición 1.1. Una cadena de Markov es una sucesión de variables aleatorias a tiempo

discreto $\{X_n : n = 0, 1, \dots\}$, que satisface la propiedad de Markov, esto es, para cualquier entero $n \geq 0$, y para cualesquiera estados $x_0, x_1, \dots, x_{n+1} \in E$, se cumple

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

Ahora se define $\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n)$.

Definición 1.2. Sea $\{X_n\}_{n \geq 0}$ una cadena de Markov con espacio de estados E . Se define la función o probabilidad de transición como

$$P(x, y) = \mathbb{P}(X_1 = y | X_0 = x), \quad x, y \in E,$$

que cumple con

$$P(x, y) \geq 0 \quad y \quad \sum_y P(x, y) = 1 \quad x, y \in E.$$

También se define la distribución inicial de la cadena como

$$\pi_0(x) = \mathbb{P}(X_0 = x) \quad x \in E,$$

que cumple con

$$\pi_0(x) \geq 0 \quad y \quad \sum_x \pi_0(x) = 1 \quad x \in E.$$

La cadena es homogénea en el tiempo, es decir

$$\mathbb{P}(X_{n+1} = y | X_n = x) = P(x, y) \quad x, y \in E, \quad n \geq 1.$$

y denotemos a

$$\mathbb{P}(X_n = y | X_0 = x) = P^n(x, y) \quad x, y \in E, \quad n \geq 1.$$

como la probabilidad de transición de x a y en n pasos.

Proposición 1.3. (Ecuación de Chapman-Kolmogorov). Para cualesquiera números enteros r y n tales que $0 \leq r \leq n$, y cualesquiera estados $x, y \in E$ se cumple

$$P^n(x, y) = \sum_k P^r(x, z) P^{n-r}(z, y), \quad z \in E.$$

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

$$\begin{aligned}
 P^n(x, y) &= \mathbb{P}(X_n = y \mid X_0 = x) = \frac{\mathbb{P}(X_n = y, X_0 = x)}{\mathbb{P}(X_0 = x)} \\
 &= \sum_{z \in E} \frac{\mathbb{P}(X_n = y, X_r = z, X_0 = x)}{\mathbb{P}(X_0 = x)} \\
 &= \sum_{z \in E} \frac{\mathbb{P}(X_n = y, X_r = z, X_0 = x)}{\mathbb{P}(X_r = z, X_0 = x)} \frac{\mathbb{P}(X_r = z, X_0 = x)}{\mathbb{P}(X_0 = x)} \\
 &= \sum_{z \in E} \mathbb{P}(X_n = y \mid X_0 = x, X_r = z) \mathbb{P}(X_r = z \mid X_0 = x) \\
 &= \sum_{z \in E} \mathbb{P}(X_n = y \mid X_r = z) \mathbb{P}(X_r = z \mid X_0 = x) \\
 &= \sum_{z \in E} P^{n-r}(z, y) P^r(x, z).
 \end{aligned}$$

□

Ahora veamos como se relacionan los estados de una cadena de Markov.

Definición 1.4. (*Accesibilidad y comunicación*). Sea una cadena de Markov $\{X_n\}_{n \geq 0}$, con espacio de estados E y función de transición P .

1. Para cualesquiera estados $x, y \in E$, se dice que el estado y es accesible desde el estado x si $P^n(x, y) > 0$ para algún $n \geq 0$. Y lo denotamos por $x \rightarrow y$.
2. Se dice que los estados x, y , son comunicantes o se comunican entre sí, $x \rightarrow y$ y $y \rightarrow x$. Y lo denotamos por $x \longleftrightarrow y$.

Definición 1.5. Se dice que un subconjunto $C \subseteq E$ es *cerrado* si

$$P^n(x, y) = 0 \quad \forall x \in C, y \notin C, n \geq 0.$$

Definición 1.6. Una cadena de Markov es *irreducible* si todos los estados se comunican entre sí.

1.2. Recurrencia y transitoriedad.

En esta sección se verá que podemos clasificar los estados de una cadena de Markov, y esta clasificación se basa en si la cadena es capaz de regresar al estado de partida.

Definición 1.7. (Recurrencia, transitoriedad y absorbencia.)

1. Se dice que un estado x es **recurrente**, si la probabilidad de que la cadena partiendo de un estado x eventualmente regrese a x es uno, es decir

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

2. Si el estado no es recurrente entonces se llama **transitorio**, en este caso

$$\mathbb{P}(X_n = x \text{ para algún } n \geq 1 \mid X_0 = x) < 1.$$

3. Un estado x de la cadena de Markov se llama **absorbente** si $P(x, x) = 1$.

Definición 1.8. Se define el primer tiempo en que una cadena de Markov visita el estado x , como

$$T_x = \min\{n \geq 0 : X_n = x\}.$$

Definición 1.9. Se define la probabilidad de que una cadena de Markov que comienza en un estado x llegue por primera vez al estado y en un tiempo finito, como

$$\rho_{xy} = \mathbb{P}_x(T_y < \infty)$$

Por las definiciones anteriores se puede decir que el estado y es recurrente si $\rho_{yy} = 1$ y es transitorio si $\rho_{yy} < 1$.

Sea $1_y(z)$, $z \in E$, la función indicadora de y , definida por

$$1_y(z) = \begin{cases} 1 & \text{si } z = y, \\ 0 & \text{si } z \neq y. \end{cases}$$

Se define a $N(y)$ como el número de visitas de la cadena al estado y , dado que $1_y(X_n) = 1$, si la cadena está en y al tiempo n y $1_y(X_n) = 0$ en otro caso; entonces se tiene

$$N(y) = \sum_{n=1}^{\infty} 1_y(X_n).$$

El evento $\{N(y) \geq 1\}$ es el mismo que $\{T_y < \infty\}$, entonces

$$\mathbb{P}_x(N(y) \geq 1) = \mathbb{P}_x(T_y < \infty) = \rho_{xy}.$$

Más adelante ocuparemos las siguientes igualdades.

1.2 Recurrencia y transitoriedad.

- $\mathbb{P}_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1}$
- $\mathbb{P}_x(N(y) = m) = \rho_{xy}\rho_{yy}^{m-1}(1 - \rho_{yy})$.

Por ejemplo en la segunda ecuación, se observa que si la cadena comienza en x y visita exactamente m veces a y , es igual a decir que x visita a y por primera vez, después regresa $m - 1$ veces adicionales a y y nunca más regresa.

Se usará $E_x()$ y $\mathbb{P}_x()$, para denotar la esperanza y la probabilidad respectivamente, definidas en términos de una cadena de Markov que comienza en un estado x . Entonces

$$\begin{aligned} E_x(1_y(X_n)) &= (1)\mathbb{P}_x(X_n = y) + (0)\mathbb{P}_x(X_n \neq y) \\ &= P^n(x, y). \end{aligned}$$

El número de visitas de una cadena de Markov a un estado y durante los tiempos $m = 1, \dots, n$; y el número esperado de tales visitas para una cadena que comienza en un estado x esta dado como

$$N_n(y) = \sum_{m=1}^n 1_y(X_m)$$

y

$$\begin{aligned} E_x(N_n(y)) &= E_x \left[\sum_{m=1}^n 1_y(X_m) \right] \\ &= \sum_{m=1}^n E_x(1_y(X_m)) \\ &= \sum_{m=1}^n P^m(x, y) := G_n(x, y). \end{aligned}$$

En las siguientes dos Proposiciones, se verá que el número de visitas a un estado transitorio, al igual que su valor esperado son finitos; por el contrario, para un estado recurrente el número de visitas y su valor esperado son infinitos.

Proposición 1.10. *Sea y un estado transitorio. Entonces*

1. $\lim_{n \rightarrow \infty} N_n(y) = N_y < \infty$, con probabilidad 1 y

$$2. \quad \lim_{n \rightarrow \infty} G_n(x, y) = G(x, y) < \infty, \quad x \in E.$$

Demostración.

(1). Supongamos y transitorio, con $\rho_{yy} < 1$

$$\begin{aligned} \mathbb{P}_x(N_y = \infty) &= \lim_{m \rightarrow \infty} \mathbb{P}_x(N_y \geq m) \\ &= \lim_{m \rightarrow \infty} \rho_{xy} \rho_{yy}^{m-1} \\ &= \rho_{xy} \lim_{m \rightarrow \infty} \rho_{yy}^{m-1} \\ &= 0, \end{aligned}$$

$\lim_{m \rightarrow \infty} \rho_{yy}^{m-1}$ es la probabilidad de una cadena que comienza en un estado y , llegue por primera vez al estado y en tiempo finito, una infinidad de veces, como y es transitorio entonces esta probabilidad es cero, por lo tanto $N_y < \infty$ con probabilidad 1.

(2). Veamos que

$$\begin{aligned} \lim_{n \rightarrow \infty} G_n(x, y) &= \lim_{n \rightarrow \infty} E_x[N_n(y)] \quad (\text{por Teorema de convergencia dominada}) \\ &= E_x \left[\lim_{n \rightarrow \infty} N_n(y) \right] \\ &= E_x[N(y)]. \end{aligned}$$

Ahora calculando

$$\begin{aligned} E_x[N(y)] &= \sum_{m=1}^{\infty} m \mathbb{P}_x(N(y) = m) \\ &= \sum_{m=1}^{\infty} m \rho_{xy} \rho_{yy}^{m-1} (1 - \rho_{yy}) \\ &= \rho_{xy} (1 - \rho_{yy}) \sum_{m=1}^{\infty} m \rho_{yy}^{m-1}. \end{aligned}$$

1.2 Recurrencia y transitoriedad.

Por otro lado, sea $t = \rho_{yy}$, entonces

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \sum_{m=1}^n m t^{m-1} &= \lim_{n \rightarrow \infty} \sum_{m=1}^n \frac{\partial}{\partial t} t^m = \frac{\partial}{\partial t} \lim_{n \rightarrow \infty} \sum_{m=1}^n t^m \\
 &= \frac{\partial}{\partial t} \left(\frac{t}{1-t} \right) \\
 &= \frac{1}{(1-t)^2} \\
 &= \frac{1}{(1-\rho_{yy})^2},
 \end{aligned}$$

por lo tanto

$$\begin{aligned}
 \rho_{xy}(1-\rho_{yy}) \lim_{n \rightarrow \infty} \sum_{m=1}^n m \rho_{yy}^{m-1} &= \rho_{xy}(1-\rho_{yy}) \frac{1}{(1-\rho_{yy})^2} \\
 &= \frac{\rho_{xy}}{(1-\rho_{yy})} < \infty.
 \end{aligned}$$

□

Proposición 1.11. *Sea y un estado recurrente. Entonces*

1. $\lim_{n \rightarrow \infty} N_n(y) = N(y) = \infty$, con probabilidad 1 y
2. $\lim_{n \rightarrow \infty} G_n(y, y) = G(y, y) = \infty$, $y \in E$.

Demostración.

(1). Sea y un estado recurrente, entonces $\rho_{yy} = 1$ y usando que $(\mathbb{P}_x(N(y) \geq m) = \rho_{xy} \rho_{yy}^{m-1})$,

se obtiene

$$\begin{aligned}
 \mathbb{P}_x(N(y) = \infty) &= \lim_{m \rightarrow \infty} \mathbb{P}_x(N(y) \geq m) \\
 &= \lim_{m \rightarrow \infty} \rho_{xy} \rho_{yy}^{m-1} \\
 &= \rho_{xy} \lim_{m \rightarrow \infty} \rho_{yy}^{m-1} \\
 &= \rho_{xy},
 \end{aligned}$$

en particular $\mathbb{P}_y(N(y) = \infty) = 1$.

(2). Si una variable aleatoria no negativa tiene probabilidad positiva de ser infinita, entonces, su esperanza es infinita, es decir,

$$\begin{aligned}\lim_{n \rightarrow \infty} G_n(y, y) &= \lim_{n \rightarrow \infty} E_y[N_n(y)] \\ &= E_y[N(y)] = \infty.\end{aligned}$$

□

De la proposición 1.10, pág 5, cuando un estado y es transitorio, podemos decir que:

- $\lim_{n \rightarrow \infty} \frac{N_n(y)}{n} = 0$, con probabilidad 1,
- $\lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = 0$, $x \in E$.

$N_n(y)/n$ es la proporción de las primeras n unidades de tiempo en que la cadena está en el estado y , y $G_n(x, y)/n$ es el valor esperado de esta proporción para la cadena que comienza en un estado x .

Proposición 1.12. *Si x es un estado recurrente y x es comunicante con y , entonces y es un estado recurrente*

Demostración. Como x es comunicante con y , entonces existen enteros $n \geq 1$ y $m \geq 1$, tales que

$$P^n(x, y) > 0 \quad \text{y} \quad P^m(y, x) > 0,$$

entonces para cualquier entero $r \geq 1$

$$P^{n+m+r}(y, y) \geq P^m(y, x)P^r(x, x)P^n(x, y)$$

ahora sumando sobre $r = 1, 2, \dots$

$$\sum_{r=1}^{\infty} P^{n+m+r}(y, y) \geq P^m(y, x) \left[\sum_{r=1}^{\infty} P^r(x, x) \right] P^n(x, y),$$

$$\sum_{r=1}^{\infty} P^{n+m+r}(y, y) \geq P^m(y, x)G(x, x)P^n(x, y),$$

1.2 Recurrencia y transitoriedad.

por la Proposición 1.11, pág 7, como x es un estado recurrente, entonces la suma del lado derecho es infinito, por lo tanto la suma del lado izquierdo también es infinita, es decir, y es un estado recurrente.

□

Se define:

$$1_{\{T_y < \infty\}} = \begin{cases} 1 & \text{si } T_y < \infty, \\ 0 & \text{si } T_y = \infty. \end{cases}$$

Definición 1.13. El *tiempo medio de recurrencia* de un estado recurrente y , de una cadena de Markov que comienza en un estado y , denotado por m_y se define como

$$m_y = E_y[T_y].$$

Teorema 1.14. *Sea y un estado recurrente, entonces*

- $\lim_{n \rightarrow \infty} \frac{N_y(y)}{n} = \frac{1_{\{T_y < \infty\}}}{m_y}$ con probabilidad 1 y ,
- $\lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = \frac{\rho_{xy}}{m_y}$, con $x \in E$.

Demostración. Vamos a introducir algunas variables aleatorias.

Consideremos una cadena de Markov que comienza en un estado recurrente y , entonces con probabilidad 1 regresa a y una infinidad de veces.

Para $r \geq 1$, sea T_y^r el tiempo de la r -ésima visita a y , es decir,

$$T_y^r = \min(n \geq 1 : N_n(y) = r).$$

Sea $W_y^1 = T_y^1 = T_y$ y para $r \geq 2$ sea $W_y^r = T_y^r - T_y^{r-1}$, que denota el tiempo de espera entre la $(r - 1)$ -ésima a y y la r -ésima visita a y .

Entonces

$$T_y^r = W_y^1 + \cdots + W_y^r.$$

Las variables aleatorias W_y^1, W_y^2, \dots , son independientes e idénticamente distribuidas y tienen la misma esperanza $E_y(W_y^1) = E_y(T_y) = m_y$.

La ley fuerte de los grandes números (ver apéndice) implica que:

$$\lim_{k \rightarrow \infty} \frac{W_y^1 + \dots + W_y^k}{k} = m_y \quad \text{con probabilidad 1,}$$

i.e. que

$$\lim_{k \rightarrow \infty} \frac{T_y^k}{k} = m_y \quad \text{con probabilidad 1.}$$

Sea $r = N_n(y)$, con esto decimos que al tiempo n la cadena ha hecho exactamente r visitas al estado y . Entonces la r -ésima visita a y ocurre a lo más en el tiempo n , y la $(r + 1)$ -ésima visita a y ocurre después del tiempo n , esto es

$$T_y^{N_n(y)} \leq n < T_y^{N_n(y)+1},$$

si multiplicamos la ecuación anterior por $\frac{1}{N_n(y)}$, tenemos

$$\frac{T_y^{N_n(y)}}{N_n(y)} \leq \frac{n}{N_n(y)} \leq \frac{T_y^{N_n(y)+1}}{N_n(y)}$$

este resultado es válido para cuando n es grande y además $N_n(y) \geq 1$.

Sabemos que $N_n(y) \rightarrow \infty$ con probabilidad 1, cuando $n \rightarrow \infty$, por la Proposición 1.11, pág 7.

Tomando límites de la desigualdad anterior tenemos:

$$\lim_{n \rightarrow \infty} \frac{T_y^{N_n(y)}}{N_n(y)} \leq \lim_{n \rightarrow \infty} \frac{n}{N_n(y)} \leq \lim_{n \rightarrow \infty} \frac{T_y^{N_n(y)+1}}{N_n(y)},$$

y obtenemos:

$$m_y \leq \lim_{n \rightarrow \infty} \frac{n}{N_n(y)} \leq m_y$$

entonces

$$\lim_{n \rightarrow \infty} \frac{n}{N_n(y)} = m_y \quad \text{con probabilidad 1.}$$

Por lo tanto

$$\lim_{n \rightarrow \infty} \frac{N_n(y)}{n} = \frac{1}{m_y}.$$

Podemos interpretar a $\frac{1}{m_y}$ como el tiempo promedio que la cadena permanece en el estado y a largo plazo.

Sea y un estado recurrente, y supongamos que X_0 tiene una distribución arbitraria. Entonces

1.2 Recurrencia y transitoriedad.

la cadena puede nunca alcanzar a y . Sin embargo, si alcanza a y , el argumento anterior sigue siendo válido, entonces:

$$\lim_{n \rightarrow \infty} \frac{N_n(y)}{n} = \frac{1_{\{T_y < \infty\}}}{m_y} \quad \text{con probabilidad 1.}$$

Por definición sabemos que

$$\begin{aligned} 0 &\leq N_n(y) \leq n & y \\ 0 &\leq \frac{N_n(y)}{n} \leq 1, \end{aligned}$$

por el Teorema de Convergencia dominada (ver apéndice) concluimos que:

$$\begin{aligned} \lim_{n \rightarrow \infty} E_x \left[\frac{N_n(y)}{n} \right] &= E_x \left[\lim_{n \rightarrow \infty} \frac{N_n(y)}{n} \right] \\ &= E_x \left[\frac{1_{\{T_y < \infty\}}}{m_y} \right] \\ &= \frac{\mathbb{P}_x(T_y < \infty)}{m_y} \\ &= \frac{\rho_{xy}}{m_y} \end{aligned}$$

Con esto queda demostrado el teorema 1.14.

□

1.2.1. Recurrencia nula y positiva.

Definición 1.15. Un estado recurrente y es llamado *recurrente nulo* si

$$m_y = \infty.$$

Un estado recurrente y es llamado *recurrente positivo* si

$$m_y < \infty.$$

Proposición 1.16.

1. Si y es recurrente nulo, entonces

$$\lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = \lim_{n \rightarrow \infty} \frac{\sum_{m=1}^n P^m(x, y)}{n} = 0, \quad x \in E.$$

2. Si y es recurrente positivo, entonces

$$\lim_{n \rightarrow \infty} \frac{G_n(y, y)}{n} = \frac{1}{m_y} > 0.$$

Demostración.

(1). Del teorema 1.14, como y es recurrente, entonces

$$\lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = \frac{\rho_{xy}}{m_y}.$$

como $m_y = \infty$,

$$\lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = 0.$$

(2). Análogamente, por el teorema 1.14, como y es recurrente, entonces

$$\lim_{n \rightarrow \infty} \frac{G_n(y, y)}{n} = \frac{1}{m_y}.$$

como y es recurrente positivo, $m_y < \infty$, por lo tanto

$$\lim_{n \rightarrow \infty} \frac{G_n(y, y)}{n} = \frac{1}{m_y} > 0.$$

□

Consideremos una cadena de Markov que comienza fuera de un estado recurrente y . Se sigue del Teorema 1.14, pág 9, que si y es recurrente nulo, entonces con probabilidad 1, la proporción de tiempo en que la cadena este en el estado y durante la primeras n unidades de tiempo se aproxima a cero, cuando $n \rightarrow \infty$. Por otro lado si y es un estado recurrente positivo, entonces, con probabilidad 1, la proporción de tiempo en que la cadena este en el estado y durante las primeras n unidades de tiempo se aproxima al limite positivo $1/m_y$ cuando $n \rightarrow \infty$.

Teorema 1.17. *Sea x un estado recurrente positivo y supongamos que x es comunicante con y , entonces y es recurrente positivo.*

Demostración. Como x es comunicante con y , entonces existen enteros positivos n_1 y n_2 , tales que

$$P^{n_1}(y, x) > 0 \quad \text{y} \quad P^{n_2}(x, y) > 0$$

1.2 Recurrencia y transitoriedad.

Ahora

$$P^{n_1+m+n_2}(y, y) \geq P^{n_1}(y, x)P^m(x, x)P^{n_2}(x, y)$$

esto se da por Chapman-Kolmogorov.

Ahora sumando sobre $m = 1, 2, \dots, n$ y dividiendo por n obtenemos:

$$\begin{aligned} \sum_{m=1}^n \left[\frac{P^{n_1+m+n_2}(y, y)}{n} \right] &\geq \sum_{m=1}^n \left[\frac{P^{n_1}(y, x)P^m(x, x)P^{n_2}(x, y)}{n} \right] \\ \sum_{m=1}^n \left[\frac{P^{n_1+m+n_2}(y, y)}{n} \right] &\geq P^{n_1}(y, x)P^{n_2}(x, y) \frac{\sum_{m=1}^n P^m(x, x)}{n} \end{aligned}$$

completando y restando del lado izquierdo, obtenemos

$$\begin{aligned} \frac{1}{n} \left[\sum_{m=1}^{n_1+n+n_2} P^m(y, y) - \sum_{m=1}^{n_1+n_2} P^m(y, y) \right] &\geq P^{n_1}(y, x)P^{n_2}(x, y) \frac{G_n(x, x)}{n} \\ \frac{G_{n_1+n+n_2}(y, y)}{n} - \frac{G_{n_1+n_2}(y, y)}{n} &\geq P^{n_1}(y, x)P^{n_2}(x, y) \frac{G_n(x, x)}{n} \end{aligned}$$

cuando $n \rightarrow \infty$

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[\frac{G_{n_1+n+n_2}(y, y)}{n} - \frac{G_{n_1+n_2}(y, y)}{n} \right] &\geq \lim_{n \rightarrow \infty} P^{n_1}(y, x)P^{n_2}(x, y) \frac{G_n(x, x)}{n} \\ \lim_{n \rightarrow \infty} \frac{G_{n_1+n+n_2}(y, y)}{n} - \lim_{n \rightarrow \infty} \frac{G_{n_1+n_2}(y, y)}{n} &\geq P^{n_1}(y, x)P^{n_2}(x, y) \lim_{n \rightarrow \infty} \frac{G_n(y, y)}{n}, \end{aligned}$$

por la Proposición 1.12, pág 8, el estado y es recurrente, entonces usando el Teorema 1.14, pág 9 se tiene que

$$\begin{aligned} \frac{1}{m_y} - 0 &\geq P^{n_1}(y, x)P^{n_2}(x, y) \frac{1}{m_x} \\ \frac{1}{m_y} &\geq \frac{P^{n_1}(y, x)P^{n_2}(x, y)}{m_x} > 0 \\ \frac{m_x}{P^{n_1}(y, x)P^{n_2}(x, y)} &\geq m_y > 0, \end{aligned}$$

por lo tanto:

$$m_y < \infty.$$

Esto muestra que y es recurrente positivo.

□

Una cadena de Markov es llamada *cadena recurrente nula* si todos los estados son nulos recurrentes, una cadena es *recurrente positiva* si todos los estados son recurrentes positivos. Entonces una cadena de Markov irreducible puede ser una cadena recurrente nula o una cadena recurrente positiva.

Proposición 1.18. *Sea C un conjunto cerrado finito de estados, entonces C tiene al menos un estado recurrente.*

Demostración. Sabemos que

$$\sum_{y \in C} P^m(x, y) = 1 \quad x \in C,$$

y sumando sobre $m = 1, \dots, n$ y dividiendo por n , tenemos que

$$\sum_{y \in C} \frac{G_n(x, y)}{n} = \frac{\sum_{m=1}^n 1}{n} = 1.$$

Si C es finito y cada estado de C es transitorio o recurrente nulo, entonces por la proposición 1.10(2), pág. 5, $\lim_{n \rightarrow \infty} G_n(x, y)/n = 0$, con $x \in E$,

$$\begin{aligned} 1 &= \lim_{n \rightarrow \infty} 1 = \lim_{n \rightarrow \infty} \sum_{y \in C} \frac{G_n(x, y)}{n} \\ &= \sum_{y \in C} \lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = 0, \end{aligned}$$

es una contradicción, esto por suponer que todos los estados son transitorios o recurrentes nulos.

Entonces, por lo menos existe un estado recurrente.

□

Teorema 1.19. *Sea C un conjunto cerrado, finito e irreducible de estados, entonces todo estado en C es recurrente positivo.*

Demostración. Cuando C es un conjunto finito cerrado, entonces hay al menos un estado recurrente positivo en C .

Como C es irreducible, todos los estados en C se comunican; entonces todos los estados en C son recurrentes positivos por el Teorema 1.17, pág 12.

□

Corolario 1.20. *Una cadena de Markov con un número finito de estados no tiene estados nulos recurrentes.*

Demostración. Si y es recurrente, entonces y está contenido en un conjunto irreducible C de estados recurrentes, donde C es finito, se sigue del Teorema 1.19, pág 14, que todos los estados en C , incluido y , son recurrentes positivos. Entonces todo estado recurrente, es recurrente positivo y no hay estados recurrentes nulos.

□

1.3. Distribuciones estacionarias.

Considere una cadena de Markov con espacio de estados discreto $\{0, 1, 2, \dots, N\}$, y una distribución de probabilidad inicial $\pi_0(y)$, $y \in E$. Esta distribución inicial se transforma en una nueva distribución de probabilidad $\pi_1(y)$, $y \in E$ a través de

$$\pi_1(y) = \sum_{x \in E} \pi_0(x)P(x, y).$$

La distribución π_1 indica la probabilidad de encontrar la cadena en los distintos estados después del primer paso. En general pasa que $\pi_n(y) = \sum_x \pi_{n-1}(x)P(x, y) = \sum_x \pi_0(x)P^n(x, y)$.

Lema 1.21. *Sea π_0 una distribución de probabilidad inicial. Entonces*

$$\mathbb{P}(X_n = y) = \sum_x \pi_0(x)P^n(x, y), \quad y \in E.$$

Demostración. Por inducción, para $n = 1$

$$\begin{aligned} \mathbb{P}(X_1 = y) &= \sum_x \mathbb{P}(X_0 = x)\mathbb{P}(X_1 = y | X_0 = x) \\ &= \sum_x \mathbb{P}(X_0 = x)P(x, y) \\ &= \sum_x \pi_0(x)P(x, y). \end{aligned}$$

Suponemos válido para n , entonces

$$\mathbb{P}(X_n = y) = \sum_x \pi_0(x)P^n(x, y).$$

Por demostrar para $n + 1$, usando la ecuación de Chapman-Kolmogorov

$$\begin{aligned}
 \mathbb{P}(X_{n+1} = y) &= \sum_k \mathbb{P}(X_n = k) \mathbb{P}(X_{n+1} = y | X_n = k) \\
 &= \sum_k \left[\sum_x \mathbb{P}(X_0 = x) \mathbb{P}(X_n = k | X_0 = x) \right] \mathbb{P}(X_{n+1} = y | X_n = k) \\
 &= \sum_x \left[\sum_k \mathbb{P}(X_n = k | X_0 = x) \mathbb{P}(X_{n+1} = y | X_n = k) \right] \mathbb{P}(X_0 = x) \\
 &= \sum_x P^{n+1}(x, y) P(X_0 = x) \\
 &= \sum_x \pi_0(x) P^{n+1}(x, y).
 \end{aligned}$$

□

Hay casos donde la distribución inicial no cambia al ser multiplicada por P (donde P es la matriz de probabilidades de transición). A tales distribuciones se le llama estacionarias o invariantes en el tiempo.

Definición 1.22. Sea $\pi(x)$, $x \in E$ una distribución de probabilidad, con números no negativos y $\sum_x \pi(x) = 1$. $\pi(x)$ es *estacionaria* o *invariante en el tiempo* para una cadena de Markov con probabilidades de transición $P(x, y)$, si

$$\sum_x \pi(x) P(x, y) = \pi(y), \quad y \in E,$$

también se puede ver en términos matriciales, es decir, π es estacionaria si $\pi = \pi P$.

1.3.1. Propiedades de las distribuciones estacionarias.

Supongamos que existe una distribución estacionaria π y

$$\lim_{n \rightarrow \infty} P^n(x, y) = \pi(y), \quad y \in E.$$

Entonces, sin considerar la distribución inicial de la cadena π_0 , la distribución de $\{X_n\}$ se aproxima a π cuando $n \rightarrow \infty$. Este hecho se verá más adelante. Ahora se analizarán algunas propiedades básicas de las distribuciones estacionarias.

Propiedad 1.23. Sea π una distribución estacionaria, entonces

$$\sum_x \pi(x)P^n(x, y) = \pi(y)$$

Demostración. Por inducción.

Para $n=2$

$$\begin{aligned} \sum_x \pi(x)P^2(x, y) &= \sum_x \pi(x) \sum_z P(x, z)P(z, y) \\ &= \sum_z \left[\sum_x \pi(x)P(x, z) \right] P(z, y) \\ &= \sum_z \pi(z)P(z, y) \\ &= \pi(y) \end{aligned}$$

suponemos que es valido para n . *i.e.*

$$\sum_x \pi(x)P^n(x, y) = \pi(y)$$

y ahora demostramos para $n+1$

$$\begin{aligned} \sum_x \pi(x)P^{n+1}(x, y) &= \sum_x \pi(x) \sum_z P^n(x, z)P(z, y) \\ &= \sum_z \left[\sum_x \pi(x)P^n(x, z) \right] P(z, y) \\ &= \sum_z \pi(z)P(z, y) \\ &= \pi(y). \end{aligned}$$

□

Propiedad 1.24. La distribución de X_n es independiente de n , si y sólo si, la distribución inicial es una distribución estacionaria.

Demostración. Primero, supongamos que la distribución de X_n es independiente de n . También se tiene que

$$\pi_0(y) = P(X_0 = y).$$

Por otro lado,

$$P(X_1 = y) = \sum_x \pi_0(x)P(x, y) = \pi(y),$$

usando nuestra hipótesis de independencia del lado derecho de la igualdad, obtenemos

$$P(X_1 = y) = \pi_0(y),$$

entonces

$$\pi_0(y) = P(X_0 = y) = P(X_1 = y) = \sum_x \pi_0(x)P(x, y).$$

por lo tanto π_0 es una distribución estacionaria.

Ahora, supongamos que X_0 tiene una distribución estacionaria π como su distribución inicial; entonces, al utilizar la propiedad 1.23, pág 17 implica que para toda n

$$P(X_n = y) = \pi(y) \quad y \in E,$$

ya que

$$\begin{aligned} P(X_n = y) &= \sum_x \pi(x)P^n(x, y) \\ &= \pi(y). \end{aligned}$$

De aquí se observa que la distribución de X_n es independiente de n , es decir, la probabilidad de que la cadena caiga en un estado $y \in E$ en cualquier tiempo n , no depende del número de pasos para llegar a y .

□

Propiedad 1.25. *Suponga que existe una distribución estacionaria π y*

$\lim_{n \rightarrow \infty} P^n(x, y) = \pi(y)$. Sea π_0 una distribución inicial, entonces

$$\lim_{n \rightarrow \infty} P(X_n = y) = \pi(y), \quad y \in E.$$

Demostración. Sabemos que

$$P(X_n = y) = \sum_x \pi_0(x)P^n(x, y),$$

entonces cuando $n \rightarrow \infty$ se obtiene

$$\lim_{n \rightarrow \infty} P(X_n = y) = \lim_{n \rightarrow \infty} \sum_x \pi_0(x) P^n(x, y),$$

como el $\lim_{n \rightarrow \infty} P^n(x, y) = \pi(y)$, $|P^n(x, y)| \leq 1$, $|\pi(y)| \leq 1$; entonces, por el Teorema de convergencia acotada

$$\lim_{n \rightarrow \infty} \sum_x \pi_0(x) P^n(x, y) = \sum_x \pi_0(x) \pi(y),$$

como $\sum_x \pi_0(x) = 1$ concluimos que:

$$\lim_{n \rightarrow \infty} P(X_n = y) = \pi(y), \quad y \in E.$$

□

Lo anterior nos dice que, sin considerar la distribución inicial para grandes valores de n , la distribución de X_n es aproximadamente igual a la distribución estacionaria π .

1.3.2. Existencia y unicidad de las distribuciones estacionarias.

En esta sección se determinará que cadenas de Markov tienen distribuciones estacionarias y cuando tienen una única distribución.

Sea π una distribución estacionaria y sea m un entero positivo, entonces por la propiedad 1.23, pág 17

$$\sum_z \pi(z) P^m(z, x) = \pi(x).$$

Sumando esta ecuación sobre $m = 1, \dots, n$ y dividiendo por n , se concluye

$$\sum_z \pi(z) \frac{G_n(z, x)}{n} = \pi(x), \quad x \in E. \quad (1.3.1)$$

Teorema 1.26. *Sea π una distribución estacionaria. Si x es un estado transitorio o un estado recurrente nulo, entonces $\pi(x) = 0$.*

Demostración. Si x es un estado transitorio o recurrente nulo

$$\lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} = 0 \quad x \in E,$$

como se vio en la sección 1.2, se sabe que

$$\sum_z \pi(z) = 1, \quad \left| \frac{G_n(z, x)}{n} \right| \leq 1, \quad x, z \in E, n \geq 1.$$

Tomando el límite en (1.3.1)

$$\lim_{n \rightarrow \infty} \pi(x) = \lim_{n \rightarrow \infty} \sum_z \pi(z) \frac{G_n(z, x)}{n}$$

entonces, usando el teorema de convergencia acotada (ver apéndice)

$$\pi(x) = \sum_z \pi(z) \lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} = \sum_z \pi(z)(0) = 0.$$

que es lo que deseábamos.

□

Teorema 1.27. *Una cadena de Markov irreducible, recurrente positiva tiene una única distribución estacionaria π , dada por*

$$\pi(x) = \frac{1}{m_x} \quad x \in E.$$

Demostración. Se sigue de la Proposición 1.16, pág 11

$$\lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} = \frac{1}{m_x} \quad z, x \in E. \quad (1.3.2)$$

Suponga que π es una distribución estacionaria. Tomando el límite a (1.3.2)

$$\lim_{n \rightarrow \infty} \pi(x) = \lim_{n \rightarrow \infty} \sum_z \pi(z) \frac{G_n(z, x)}{n},$$

entonces, debido al Teorema de convergencia acotada

$$\pi(x) = \sum_z \pi(z) \lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n}$$

y sustituyendo en (1.3.2)

$$\pi(x) = \sum_z \pi(z) \frac{1}{m_x} = \frac{1}{m_x}.$$

De este modo si existe una distribución estacionaria, esta es $\pi(x) = 1/m_x$.

Para completar la prueba del teorema en el caso del espacio de estados finito, se necesita mostrar que la función $\pi(x)$, $x \in E$, definida al inicio del teorema, es desde luego una distribución estacionaria. π es no negativa, entonces basta mostrar que:

1.3 Distribuciones estacionarias.

1. $\sum_x \frac{1}{m_x} = 1$ y
2. $\sum_x \frac{1}{m_x} P(x, y) = \frac{1}{m_y}, \quad y \in E.$

Para lograr esto, primero se observa que

$$\sum_x P^m(z, x) = 1,$$

Sumando sobre $m = 1, \dots, n$ y dividiendo por n , se tiene

$$\sum_x \frac{G_n(z, x)}{n} = \frac{\sum_{m=1}^n 1}{n} = 1 \quad z \in E.$$

También observamos por Chapman-Kolmogorov

$$\sum_x P^m(z, x) P(x, y) = P^{m+1}(z, y),$$

una vez más, sumando sobre $m = 1, \dots, n$ y dividiendo por n

$$\frac{1}{n} \sum_{m=1}^n \sum_x P^m(z, x) P(x, y) = \frac{1}{n} \sum_{m=1}^n P^{m+1}(z, y),$$

ahora sumando un cero del lado derecho, como la suma comienza en 2, se obtiene

$$\frac{1}{n} \sum_x \sum_{m=1}^n P^m(z, x) P(x, y) = \frac{1}{n} \sum_{m=1}^n [P^{m+1}(z, y) + P(z, y) - P(z, y)]$$

entonces

$$\sum_x \frac{G_n(z, x)}{n} P(x, y) = \frac{G_{n+1}(z, y)}{n} - \frac{P(z, y)}{n}.$$

Si E es finito, sabemos que

$$\lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = \frac{1}{m_y} \quad \text{y} \quad \sum_x \frac{G_n(z, x)}{n} = 1,$$

si hacemos que $n \rightarrow \infty$ en la segunda ecuación, se obtiene

$$\lim_{n \rightarrow \infty} 1 = \lim_{n \rightarrow \infty} \sum_x \frac{G_n(z, x)}{n},$$

por el Teorema de convergencia acotada

$$1 = \sum_x \lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} = \sum_x \frac{1}{m_x},$$

y con esto (1.) queda demostrado.

Ahora para (2.), usemos

$$\sum_x \frac{G_n(z, x)}{n} P(x, y) = \frac{G_{n+1}(z, y)}{n} - \frac{P(z, y)}{n},$$

tomando el límite se obtiene

$$\lim_{n \rightarrow \infty} \sum_x \frac{G_n(z, x)}{n} P(x, y) = \lim_{n \rightarrow \infty} \left[\frac{G_{n+1}(z, y)}{n} - \frac{P(z, y)}{n} \right],$$

como ambos límites del lado derecho existen, entonces

$$\lim_{n \rightarrow \infty} \sum_x \frac{G_n(z, x)}{n} P(x, y) = \lim_{n \rightarrow \infty} \frac{G_{n+1}(z, y)}{n} - \lim_{n \rightarrow \infty} \frac{P(z, y)}{n}$$

usando el Teorema de convergencia acotada, se obtiene

$$\begin{aligned} \sum_x P(x, y) \lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} &= \lim_{n \rightarrow \infty} \frac{G_{n+1}(z, y)}{n} \\ \sum_x P(x, y) \frac{1}{m_x} &= \frac{1}{m_y}, \end{aligned}$$

con esto queda demostrado (2.), Con esto se completa la prueba en el caso cuando E es finito.

El argumento para completar la prueba para E infinito, es más complicado ya que no se puede intercambiar directamente los límites y las sumas, como se hace en el caso finito (el teorema de convergencia acotada no es aplicable).

Sea E_1 un subconjunto finito de E . Se puede observar de $\sum_x \frac{G_n(z, x)}{n} = 1$ que

$$\sum_{x \in E_1} \frac{G_n(z, x)}{n} \leq 1 \quad z \in E,$$

como E_1 es finito, se puede tomar el límite en esta desigualdad y debido a que

$$\lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} = \frac{1}{m_x}$$

$$\lim_{n \rightarrow \infty} \sum_{x \in E_1} \frac{G_n(z, x)}{n} \leq \lim_{n \rightarrow \infty} 1$$

$$\sum_{x \in E_1} \lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} \leq 1$$

1.3 Distribuciones estacionarias.

$$\sum_{x \in E_1} \frac{1}{m_x} \leq 1,$$

esta última desigualdad fue para todo $E_1 \subseteq E$. Además veremos que

$$\sum_x \frac{1}{m_x} \leq 1, \quad x \in E,$$

ya que si la suma de $1/m_x$ sobre $x \in E$ excede a 1, entonces la suma sobre algún subconjunto finito de E podría exceder a 1, demostremos este hecho.

Supongamos que

$$\sum_{x \in E} \frac{1}{m_x} > 1,$$

Se define

$$\sum_{x \in E} \frac{1}{m_x} = \sum_{n=1}^{\infty} \frac{1}{m_{x_n}} = \lim_{k \rightarrow \infty} \sum_{n=1}^k \frac{1}{m_{x_n}}.$$

Sea $\sum_{n=1}^k \frac{1}{m_{x_n}} = S_k$.

Como el $\lim_{k \rightarrow \infty} S_k > 1$, podemos hacer

$$\lim_{k \rightarrow \infty} S_k = 1 + \varepsilon,$$

entonces, para toda $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que, para toda $k \geq k_0$

$$|S_k - (1 + \varepsilon)| < 1 + \varepsilon - 1$$

sea $\alpha = 1 + \varepsilon$

$$|S_k - \alpha| < \alpha - 1$$

$$-\alpha + 1 < S_k - \alpha < \alpha - 1$$

de la desigualdad del lado izquierdo, se observa

$$-\alpha + 1 < S_k - \alpha$$

$$1 < S_k,$$

pero esto es una contradicción, ya que $S_k = \sum_{n=1}^k \frac{1}{m_{x_n}} \leq 1$, por que se tiene un número finito de términos. Por lo tanto

$$\sum_x \frac{1}{m_x} \leq 1.$$

Una vez más, utilizando

$$\sum_x \frac{G_n(z, x)}{n} P(x, y) = \frac{G_{n+1}(z, y)}{n} - \frac{P(z, y)}{n},$$

obtenemos la siguiente desigualdad:

Si E_1 es un subconjunto finito de E , entonces

$$\sum_{x \in E_1} \frac{G_n(z, x)}{n} P(x, y) \leq \frac{G_n(n+1)(z, y)}{n} - \frac{P(z, y)}{n},$$

tomando el límite en esta desigualdad y también por $\lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} = \frac{1}{m_x}$

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{x \in E_1} \frac{G_n(z, x)}{n} P(x, y) &\leq \lim_{n \rightarrow \infty} \left[\frac{G_n(z, y)}{n} - \frac{P(z, y)}{n} \right] \\ \sum_{x \in E_1} P(x, y) \lim_{n \rightarrow \infty} \frac{G_n(z, x)}{n} &\leq \lim_{n \rightarrow \infty} \frac{G_{n+1}(z, y)}{n} - \lim_{n \rightarrow \infty} \frac{P(z, y)}{n} \\ &= \sum_{x \in E_1} \frac{1}{m_x} P(x, y) \leq \frac{1}{m_y}. \end{aligned}$$

Como en el caso anterior hay que demostrar que

$$\sum_{x \in E} \frac{1}{m_x} P(x, y) \leq \frac{1}{m_y}.$$

Supongamos que $\sum_{x \in E} \frac{1}{m_x} > \frac{1}{m_y}$; se define

$$\sum_{x \in E} \frac{1}{m_x} P(x, y) = \sum_{n=1}^{\infty} \frac{1}{m_{x_n}} P(x_n, y) = \lim_{k \rightarrow \infty} \sum_{n=1}^k \frac{1}{m_{x_n}} P(x_n, y).$$

Sea $\sum_{n=1}^k \frac{1}{m_{x_n}} P(x_n, y) = S_k$.

Como el $\lim_{k \rightarrow \infty} S_k > \frac{1}{m_y}$, podemos hacer

$$\lim_{k \rightarrow \infty} S_k = \frac{1}{m_y} + \varepsilon,$$

entonces, para toda $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que, para toda $k \geq k_0$.

$$|S_k - (1/m_y + \varepsilon)| < 1/m_y + \varepsilon - 1/m_y$$

1.3 Distribuciones estacionarias.

sea $\alpha = 1/m_y + \varepsilon$

$$|S_k - \alpha| < \alpha - 1/m_y$$

$$-\alpha + 1/m_y < S_k - \alpha < \alpha - 1/m_y$$

de la desigualdad del lado izquierdo, se observa

$$-\alpha + 1/m_y < S_k - \alpha$$

$$1/m_y < S_k,$$

pero esto es una contradicción, ya que $S_k = \sum_{n=1}^n \frac{1}{m_{x_n}} P(x_n, y) \leq 1/m_y$, por que se tiene un número finito de términos. Por lo tanto

$$\sum_x \frac{1}{m_x} P(x, y) \leq \frac{1}{m_y}.$$

Continuando con la demostración. Se sigue de $\sum_x 1/m_x \leq 1$, que la suma en y de lado derecho de $\sum_x 1/m_x P(x, y) \leq 1/m_y$ es finita. Si tomamos la desigualdad estricta para alguna y y sumemos sobre este de la siguiente manera

$$\begin{aligned} \sum_y \frac{1}{m_y} &> \sum_y \left[\sum_x \frac{1}{m_x} P(x, y) \right] \\ &= \sum_x \frac{1}{m_x} \sum_y P(x, y) \\ &= \sum_x \frac{1}{m_x} \end{aligned}$$

lo cual es una contradicción, entonces esto demuestra (2), es decir,

$$\sum_x \frac{1}{m_x} P(x, y) = \frac{1}{m_y}.$$

Por otro lado, se sabe que $\sum_x \frac{1}{m_x} \leq 1$.

Sea

$$c = \frac{1}{\sum_x \frac{1}{m_x}},$$

entonces por (2.), se tiene

$$\pi(x) = \frac{c}{m_x} \quad x \in E,$$

que se define como una distribución estacionaria, entonces por la primera parte de este Teorema

$$\pi(x) = \frac{1}{m_x} = \frac{c}{m_x}$$

y de aquí que $c = 1$, esto prueba (1.) y con esto queda demostrado este Teorema.

□

Corolario 1.28. *Una cadena irreducible es recurrente positiva, si y sólo si, esta tiene una distribución estacionaria.*

Demostración.

⇒] Por el Teorema 1.27, pág 20, en la que dice que una cadena irreducible, recurrente positiva tiene una distribución estacionaria y además esta es única.

⇐] Se sigue del Teorema 1.26, pág 19, que nos dice que si π es una distribución estacionaria, en el caso cuando el estado x es transitorio o nulo recurrente, entonces $\pi(x) = 0$, pero por hipótesis la cadena es irreducible y existe una distribución estacionaria, entonces todos los estados son recurrentes positivos, por lo tanto la cadena es recurrente positiva.

□

Corolario 1.29. *Si una cadena de Markov tiene un número finito de estados y es irreducible, entonces, esta tiene una única distribución estacionaria.*

Demostración. Recordemos que $N_n(x)$ que denota el número de visitas a x durante $m = 1, \dots, n$.

Como la cadena es irreducible, entonces, todo estado en E es recurrente positivo. Por el Teorema 1.14, pág 9

$$\lim_{n \rightarrow \infty} \frac{N_n(y)}{n} = \frac{1}{m_x},$$

si decimos que $1/m_x$ es una distribución estacionaria por el corolario 1.28, pág 26, se puede decir que existe una única distribución estacionaria y por el Teorema 1.27, pág 20, se puede decir que $1/m_x$ es una distribución estacionaria y además es única.

□

1.3.3. Cadenas Reducibles

Sea π una distribución en E , es decir, sea $\pi(x)$, con $x \in E$, números no negativos que suman 1 y sea C un subconjunto de E , se dice que π es concentrado en C si

$$\pi(x) = 0 \quad x \notin C.$$

Teorema 1.30. *Sea C un conjunto cerrado irreducible de estados recurrentes positivos, entonces la cadena de Markov tiene una distribución estacionaria π concentrada en C , dada por*

$$\pi(x) = \begin{cases} \frac{1}{m_x} & \text{si } x \in C, \\ 0 & \text{si } x \notin C. \end{cases}$$

Demostración. Por definición si $x \notin C$, entonces $\pi(x) = 0$.

Como C es cerrado e irreducible, por el Teorema 1.27, pág 20, este conjunto tiene una única distribución estacionaria dada por $\pi(x) = \frac{1}{m_x}$ con $x \in C$; si $x \notin C$, entonces $\pi(x) = 0$.

□

Capítulo 2

Cadenas de Markov en espacios generales

En este capítulo, se analiza la teoría de las cadenas de Markov en espacios de estados generales, a través de kernels que son funciones con ciertas propiedades; en particular cuando un kernel es subestocástico se dice que es una probabilidad de transición para una cadena de Markov. Se definirá una cadena de Markov, sus propiedades de comunicación en el espacio de estados, para llegar a recurrencia, transitoriedad y Harris-recurrencia. La bibliografía principal para este Capítulo fue: [6, Capítulos 1-3] y [9].

2.1. Preliminares

Para este capítulo, se utilizará la siguiente notación:

Sea E conjunto y \mathcal{E} una σ -álgebra de subconjuntos de E . Se supondrá que la σ -álgebra es generada por una colección numerable de subconjuntos de E .

El espacio medible (E, \mathcal{E}) es llamado *espacio de estados* y los puntos de E , son llamados estados.

El símbolo \mathcal{G} denotará la colección de funciones medibles valuadas en los reales extendidos sobre (E, \mathcal{E})

Los símbolos x, y, z, \dots denotarán estados; A, B, C, \dots denotarán elementos de la σ -álgebra \mathcal{E} y f, g, h, \dots denotarán funciones en \mathcal{G} .

$\mathcal{M}(\mathcal{E})$ denotará la colección de medidas sobre (E, \mathcal{E}) . Los símbolos λ, μ, \dots , son elementos de $\mathcal{M}(\mathcal{E})$. Cuando \mathcal{A} es una colección de funciones o medidas, se usará \mathcal{A}_+ para la clase de elementos no negativos de \mathcal{A} . Los elementos de \mathcal{M}_+ serán solo medidas. \mathcal{M}^+ denota la clase de medidas positivas, es decir, $\mathcal{M}^+ = \{\lambda \in \mathcal{M}_+ : \lambda(E) > 0\}$.

2.1.1. Kernels

En esta sección definimos qué es un kernel, y cómo opera con funciones y medidas.

Definición 2.1. (*Kernel*).

Un *kernel* (no negativo) sobre (E, \mathcal{E}) es una función $K : E \times \mathcal{E} \rightarrow \overline{\mathbb{R}}_+$ que satisface las siguientes dos condiciones:

1. Para todo conjunto fijo $A \in \mathcal{E}$, la función $K(\cdot, A)$ es medible,
2. Para todo estado fijo $x \in E$, la función conjuntista $K(x, \cdot)$ es una medida en (E, \mathcal{E}) .

Un Kernel K es llamado *σ -finito*, si existe una función $\mathcal{E} \otimes \mathcal{E}$ -medible, $f \in \mathcal{E} \otimes \mathcal{E}$ y $f > 0$, tal que:

$$\int_E K(x, dy) f(x, y) < \infty \quad \forall x \in E.$$

- El kernel K es llamado *finito* si: $K(x, E) < \infty \quad \forall x \in E$;
- *acotado* si: $\sup_{x \in E} K(x, E) < \infty$;
- *subestocástico* si: $K(x, E) < 1 \quad \forall x \in E$;
- *estocástico* si: $K(x, E) = 1 \quad \forall x \in E$.

Definición 2.2. Todo kernel K funciona como un operador lineal no negativo en \mathcal{G}_+ , es decir:

$$Kf(x) = \int_E K(x, dy)f(y), \quad f \in \mathcal{G}_+.$$

De manera parecida K actúa como un operador en \mathcal{M}_+ , como:

$$\lambda K(A) = \int_E \lambda(dx)K(x, A), \quad \lambda \in \mathcal{M}_+.$$

Se denota por I el kernel *identidad* $I(x, A)$, $x \in E$, $A \in \mathcal{E}$:

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

Si K_1 y K_2 son dos kernels, su kernel producto K_1K_2 es

$$K_1K_2(x, A) = \int_E K_1(x, dy)K_2(y, A).$$

La iteración K^n , $n \geq 0$, de un kernel K se define como

$$K^0 = I, \quad K^n = KK^{n-1}, \quad \text{para } n \geq 1.$$

En particular si K es subestocástico, se usa el símbolo P , en donde P denota una *probabilidad de transición*, y en este caso también se cumple que la iteración P^n es subestocástica.

2.1.2. Cadenas de Markov

Se definirá una cadena de Markov en espacios generales a través de kernels subestocásticos.

Sea (Ω, \mathcal{F}) un espacio medible, es llamado un *espacio de estados* y \mathbb{P} una medida de probabilidad sobre (Ω, \mathcal{F}) . Una función medible ξ definida sobre (Ω, \mathcal{F}) y que lleva sus valores en un espacio medible arbitrario (Ω', \mathcal{F}') es llamado un *elemento aleatorio* Ω' -valuado. Una sucesión $\{\xi_n ; n \geq 0\}$ de elementos aleatorios Ω' -valuados es llamado un *proceso estocástico* Ω' -valuado.

Si ξ es un elemento aleatorio Ω' -valuado, se denotará por $\mathcal{L}(\xi)$ la *distribución de probabilidad* de ξ , es decir,

$$\mathcal{L}(\xi)(A') = \mathbb{P}\{\xi \in A'\} = \mathbb{P}\{\omega : \xi(\omega) \in A'\}, \quad A' \in \mathcal{F}'.$$

Una sucesión creciente de sub- σ -álgebras de \mathcal{F} , $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n \subseteq \dots$, es llamada *historia*. Se dice que un proceso estocástico $\{\xi_n ; n \geq 0\}$ es *adaptado a la historia* $\{\mathcal{F}_n ; n \geq 0\}$, si ξ_n es \mathcal{F}_n -medible para cada $n \geq 0$. Un proceso estocástico $\{\xi_n\}_{n \geq 0}$ siempre es adaptado a alguna historia ya que si definimos

$$\mathcal{F}_n^\xi = \sigma(\xi_0, \dots, \xi_n), \quad n \geq 0;$$

entonces $\{\xi_n\}_{n \geq 0}$ es adaptado a $\{\mathcal{F}_n^\xi\}_{n \geq 0}$, donde \mathcal{F}_n^ξ es llamada la *historia interna* de $\{\xi_n\}_{n \geq 0}$.

Definición 2.3. (*Propiedad de Markov*). Una *cadena de Markov con probabilidad de transición P* , es un proceso estocástico E -valuado $\{X_n : n \geq 0\}$, que cumple

$$\mathcal{L}(X_{n+1} | \mathcal{F}_n^X) = \mathcal{L}(X_{n+1} | X_n) = P(X_n, \cdot), \quad \forall n \geq 0.$$

Una cadena de Markov $\{X_n : n \geq 0\}$ es llamada *cadena de Markov con respecto a su historia*, denotada por (X_n, \mathcal{F}_n) , si es adaptada a $\{\mathcal{F}_n ; n \geq 0\}$. Entonces, de la definición anterior \mathcal{F}_n^X será reemplazado por \mathcal{F}_n , es decir,

$$\mathcal{L}(X_{n+1} | \mathcal{F}_n) = \mathcal{L}(X_{n+1} | X_n) = P(X_n, \cdot), \quad \forall n \geq 0.$$

La distribución $\mathcal{L}(X_0)$ de X_0 es llamada la *distribución inicial* de la cadena (X_n) .

La siguiente definición es una consecuencia de la propiedad de Markov.

Definición 2.4. Si $\{X_n : n \geq 0\}$ es una cadena de Markov con distribución inicial $\mathcal{L}(X_0) = \lambda$ y probabilidad de transición P ; la *distribución de probabilidad de la cadena* $\{X_n : n \geq 0\}$ es determinada por:

$$\mathbb{P}\{X_0 \in A_0, \dots, X_n \in A_n\} = \int_{A_0} \lambda(dx_0) \int_{A_1} P(x_0, dx_1) \dots \int_{A_n} P(x_{n-1}, dx_n)$$

para toda $n \geq 0$, $A_0, \dots, A_n \in E$. En particular tenemos

$$\mathcal{L}(X_n) = \lambda P^n.$$

Si la probabilidad de transición P no es *estocástica*, podemos hacer lo siguiente: se da un punto Δ que no pertenece al conjunto E , este punto Δ lo agregamos a E para obtener el *espacio de estados extendido* $(E_\Delta, \mathcal{E}_\Delta) = (E \cup \{\Delta\}, \sigma(\mathcal{E} \cup \{\Delta\}))$. Se extiende la medida λ a $(E_\Delta, \mathcal{E}_\Delta)$ siendo $\lambda(\{\Delta\}) = 0$. La probabilidad de transición P es extendida a $(E_\Delta, \mathcal{E}_\Delta)$ como sigue:

- $P(x, \{\Delta\}) = 1 - P(x, E)$
- $P(\Delta, \{\Delta\}) = 1$

y de esta manera P es un kernel estocástico sobre $(E_\Delta, \mathcal{E}_\Delta)$. El punto Δ es llamado el *cementerio* de la cadena de Markov (X_n) .

Una función medible θ en el espacio de estados (Ω, \mathcal{F}) es llamado *operador shift* (para la cadena de Markov $\{X_n\}$), dado que

$$X_n(\theta(\omega)) = X_{n+1}(\omega) \quad \forall \omega \in \Omega, n \geq 0.$$

Las iteraciones θ_m , $m \geq 0$, de θ son definidas por $\theta_0 = I_\Omega$ que es el operador identidad en Ω , e iterativamente

$$\theta_m = \theta \circ \theta_{m-1}, \quad \text{para } m \geq 1.$$

En lo subsecuente se denotará \mathbb{P}_λ ó \mathbb{P}_x para indicar una distribución inicial específica λ ó un estado inicial específico x de la cadena, respectivamente.

2.2. Kernels irreducibles

En esta sección se describirá la estructura de comunicación inducida por el kernel K en el espacio de estados (E, \mathcal{E}) . Esta estructura es muy parecida a la comunicación de estados en un espacio discreto como fué descrita en el Capítulo 1.

Definición 2.5. (*Accesibilidad.*) Se dice que el conjunto $A \in \mathcal{E}$ es accesible desde el estado $x \in E$, denotado por $(x \rightarrow A)$, si

$$K^n(x, A) > 0 \quad \text{para algún } n \geq 1.$$

Si A no es accesible desde x , denotado por $(x \not\rightarrow A)$, entonces $K^n(x, A) = 0$ para toda $n \geq 1$.

Cuando $K = P$, donde P es la probabilidad de transición de una cadena de Markov $\{X_n\}_{n \geq 0}$, se tiene la siguiente interpretación probabilística para accesibilidad:

$$x \rightarrow A \quad \text{si y sólo si } \mathbb{P}_x\{X_n \in A\} > 0, \quad \text{para algún } n \geq 1.$$

Definición 2.6. (*Conjunto cerrado, no separable y absorbente*).

- Un conjunto no vacío $F \in \mathcal{E}$ es llamado **cerrado**, si

$$K^n(x, F^c) = 0 \quad \forall n \geq 1, x \in F.$$

- Un conjunto $B \in \mathcal{E}$ es llamado **no separable**, si no existen dos conjuntos cerrados disjuntos contenidos en B .
- Un conjunto no vacío $F \in \mathcal{E}$ es llamado **absorbente** si

$$K^n(x, F) = K^n(x, E) = 1 \quad \forall n \geq 1, x \in F.$$

$K^n(x, E) = 1$ sólo se cumple si el kernel es estocástico.

De las definiciones anteriores, $F \in \mathcal{E}$ es cerrado si y sólo si $x \not\rightarrow F^c$ para toda $x \in F$. Un conjunto absorbente siempre es cerrado; un conjunto F puede ser cerrado pero no absorbente, es decir, es posible que $K(x, F) = K(x, E) \neq 1$ para algún $x \in F$.

En particular, cuando $K = P$, donde P es una probabilidad de transición de una cadena de Markov $\{X_n\}_{n \geq 0}$, un conjunto $F \in \mathcal{E}$ es llamado *cerrado* para $\{X_n\}_{n \geq 0}$ si este es cerrado para la probabilidad de transición P de $\{X_n\}_{n \geq 0}$; de la misma manera se define un conjunto no separable y absorbente para una cadena de Markov.

Notemos que F es cerrado para $\{X_n\}_{n \geq 0}$ si y sólo si

$$\mathbb{P}_x\{X_n \in F^c \text{ para algún } n \geq 1\} = 0 \quad \forall x \in F;$$

2.2 Kernels irreducibles

dicho de otra manera F es cerrado para $\{X_n\}_{n \geq 0}$ si y sólo si

$$\mathbb{P}_x\{X_n \in F, \quad \forall n \geq 1\} = 1 \quad \forall x \in F.$$

y esta definición coincide con la de un conjunto absorbente, es decir, cuando $K = P$ la definición de un conjunto cerrado y un conjunto absorbente coinciden.

Definición 2.7. El *kernel potencial* se define como

$$G = \sum_{n=0}^{\infty} K^n.$$

Más adelante se utilizará la siguiente proposición.

Proposición 2.8. Para toda $n \geq 1$:

$$G = \sum_{m=0}^{n-1} K^m + K^n G.$$

Demostración.

$$\begin{aligned} G = \sum_{m=0}^{\infty} K^m &= \sum_{m=0}^{n-1} K^m + \sum_{m=n}^{\infty} K^m \\ &= \sum_{m=0}^{n-1} K^m + K^n \sum_{m=0}^{\infty} K^m \quad (\text{por Teo. de conv. monótona}) \\ &= \sum_{m=0}^{n-1} K^m + K^n G. \end{aligned}$$

□

Se denotará por A^0 el conjunto de estados en A^c desde los cuales A no es accesible:

$$A^0 = \{x \in A^c : x \not\rightarrow A\} = \{x \in A^c : G1_A = 0\},$$

donde $G1_A(x) = \int_E G(x, dy)1_A(y) = \int_A G(x, dy)$.

Proposición 2.9. Para todo conjunto $A \in \mathcal{E}$, A^0 es vacío o cerrado.

Demostración. Para toda $x \in A^0$ se tiene

$$\sum_{n=1}^{\infty} K^n(x, A) \leq \sum_{n=0}^{\infty} K^n(x, A) = 0,$$

$$K \sum_{n=0}^{\infty} K^n(x, A) \leq \sum_{n=0}^{\infty} K^n(x, A) = 0,$$

$$KG(x, A) \leq G(x, A) = 0,$$

de lo anterior se tiene

$$\begin{aligned} 0 = G(x, A) &\geq \int_E K(x, dy)G(y, A) \\ &\geq \int_{A^c} K(x, dy)G(y, A) \\ &= \int_{A^0} K(x, dy)G(y, A) + \int_{(A^0)^c} K(x, dy)G(y, A) \\ &= 0 + \int_{(A^0)^c} K(x, dy)G(y, A) \end{aligned}$$

Para que el segundo término sea 0, entonces $K(x, (A^0)^c) = 0$. Por lo tanto A^0 es vacío o cerrado.

□

Definición 2.10. Sean $\lambda, \mu \in \mathcal{M}_+(\mathcal{E})$ dos medidas. La medida μ es *absolutamente continua* con respecto a λ , ($\mu \ll \lambda$), si y sólo si:

$$\lambda(A) = 0 \Rightarrow \mu(A) = 0 \quad \forall A \in \mathcal{E}.$$

Con la definición anterior, se considera la noción de la equivalencia de kernels.

Definición 2.11. (*Equivalencia de medidas*)

- Se dice que dos medidas $\lambda, \mu \in \mathcal{M}_+(\mathcal{E})$ son *equivalentes*, denotado por $\lambda \sim \mu$ si son absolutamente continuas una con respecto a la otra, es decir, $\mu \ll \lambda$ y $\lambda \ll \mu$.
- Se dice que un kernel \tilde{K} es *equivalente* con el kernel K , si para toda $x \in E$, las medidas $K(x, \cdot)$ y $\tilde{K}(x, \cdot)$ son equivalentes.

Sea $\psi \in \mathcal{M}_+(\mathcal{E})$ una medida de tal manera que ψK sea absolutamente continua con respecto a ψ . Entonces para toda medida $\varphi \in \mathcal{M}_+(\mathcal{E})$, sea

$$\psi = \sum_{n=0}^{\infty} 2^{-(n+1)} \varphi \tilde{K}^n, \tag{2.2.1}$$

donde $\tilde{\varphi}$ es una medida de probabilidad equivalente a φ y \tilde{K} es un kernel subestocástico equivalente a K . (Ver Proposición 4.13, pág 106 del apéndice).

Lema 2.12. 1. Si $\psi K \ll \psi$ entonces $\psi G \sim \psi$.

2. Si $\psi K \ll \psi$ y E es no separable, entonces $\psi(F) > 0$ para todos los conjuntos cerrados $F \in \mathcal{E}$.

Demostración.

(1). Primero, por inducción se comprueba que $\psi K^n \ll \psi$, $\forall n \geq 0$.

Para $n = 1$,

como $\psi K \ll \psi$ se cumple que

$$\text{si } \psi(A) = 0 \quad \Rightarrow \quad \psi K(A) = \int_E \psi(dx) K(x, A) = 0, \quad \forall A \in \mathcal{E}.$$

Para $n = 2$,

si $\psi(A) = 0$, entonces

$$\begin{aligned} \psi K^2(A) &= \int_E \psi(dx) K^2(x, A) = \int_E \int_E \psi(dx) K(x, dy) K(y, A) \\ &= \int_E [\psi K](dy) K(y, A) = 0. \end{aligned}$$

Ahora se supone válido para n , entonces, $\psi K^n(A) \ll \psi(A)$.

Por demostrar que es válido para $n + 1$, es decir, $\psi K^{n+1}(A) \ll \psi(A)$.

Si $\psi(A) = 0$, entonces

$$\begin{aligned} \psi K^{n+1}(A) &= \int_E \psi(dx) K^{n+1}(x, A) = \int_E \int_E \psi(dx) K(x, dy) K^n(y, A) \\ &= \int_E [\psi K](dy) K^n(y, A) = 0. \end{aligned}$$

Por lo tanto $\psi K^n(A) \ll \psi(A)$, para toda $A \in \mathcal{E}$, $n \in \mathbb{N}$.

Ahora sumando sobre $n \in \mathbb{N}$, se obtiene

$$\begin{aligned}
 \sum_{n=0}^{\infty} \psi K^n(A) &= \sum_{n=0}^{\infty} \int_E \psi(dx) K^n(x, A) \\
 &= \int_E \sum_{n=0}^{\infty} \psi(dx) K^n(x, A) \quad (\text{Teo. conv. monótona}) \\
 &= \int_E \psi(dx) G(x, A) \\
 &= \psi G(A) \ll \psi(A),
 \end{aligned}$$

ya que, cuando $\psi(A) = 0$, cada $\psi K^n(A) = 0$ para toda $n \in \mathbb{N}$.

Por otro lado, si $\psi G(A) = 0$

$$0 = \psi G(A) = \psi \left[\sum_{n=0}^{\infty} K^n \right] (A),$$

cada sumando de G es cero, en particular

$$0 = \psi K^0(A) = \int_E \psi(dx) 1_A(x) = \int_A \psi(dx) = \psi(A),$$

entonces $\psi(A) \ll \psi G(A)$.

Por lo tanto podemos concluir que $\psi G \sim \psi$.

(2). Por la proposición 2.9, pág. 35 y por la hipótesis de que F es cerrado y E no separable, entonces F^0 es vacío, como consecuencia $K(x, F) > 0$ para toda $x \in E$; también por hipótesis se sabe que $\psi K \ll \psi$, entonces por (1), $\psi G \sim \psi$, con todo esto se concluye que $\psi G(F) > 0$. Como $\psi G(F) \ll \psi(F)$, entonces $\psi(F) > 0$.

□

2.2.1. φ -irreducibilidad

En esta parte definiremos cuándo un kernel es irreducible en términos de una medida $\varphi \in \mathcal{M}_+(\mathcal{E})$.

Definición 2.13. (*Conjunto φ -positivo, φ -comunicante.*) Sea una medida $\varphi \in \mathcal{M}_+(\mathcal{E})$ sobre (E, \mathcal{E}) y sea $B \in \mathcal{E}$.

- El conjunto B es φ -**positivo** si $\varphi(B) > 0$
- El conjunto B es llamado φ -**comunicante** si todo conjunto $A \subseteq B$, $\varphi(A) > 0$, entonces

$$K^n(x, A) > 0, \quad \text{para algún } n \geq 1, \quad \forall x \in B.$$

Con base en lo anterior ahora se define un kernel φ -irreducible e irreducible.

Definición 2.14. (*Kernel φ -irreducible, φ -comunicante y medida de irreducibilidad*). Sea una medida $\varphi \in \mathcal{M}_+(\mathcal{E})$ en (E, \mathcal{E}) .

- Si E es φ -comunicante, entonces el kernel K es llamado φ -**irreducible**.
- K es llamado **irreducible**, si alguna medida φ cumple con que K sea φ -irreducible; en este caso φ es llamada **medida de irreducibilidad** para el kernel K .

La siguiente proposición nos dice que todo conjunto $B \in \mathcal{E}$ φ -comunicante es no separable.

Proposición 2.15. *Si un conjunto $B \in \mathcal{E}$ es φ -comunicante, entonces, B es no separable.*

Demostración. Supongamos que existen dos conjuntos F_1 y F_2 cerrados y disjuntos en B . Entonces F_1 ó F_2 podrían ser φ -positivos.

Si suponemos que F_1 es φ -positivo, entonces, como $F_1 \subseteq B$ y B es φ -comunicante, entonces

$$K^n(x, F_1) > 0, \quad \text{para algún } n \geq 1, \quad \forall x \in B. \tag{2.2.2}$$

Si $x \in F_2$, como F_2 es cerrado, entonces

$$K^n(x, F_2^c) = 0, \quad \forall n \geq 1,$$

además $F_1 \subseteq F_2^c$, por lo tanto

$$K^n(x, F_1) = 0, \quad \forall n \geq 1,$$

que es una contradicción con (2.2.2).

Por lo tanto, B no es separable.

□

Ocurre que toda medida $\psi \in \mathcal{M}_+(\mathcal{E})$ absolutamente continua con respecto a una medida de irreducibilidad φ , también es una medida de irreducibilidad [6, pág 13]. Con una medida de irreducibilidad φ se puede construir una medida de máxima irreducibilidad.

Definición 2.16. Se dice que una medida de irreducibilidad $\psi \in \mathcal{M}_+(\mathcal{E})$ es una **medida de máxima irreducibilidad** si cualquier otra medida irreducible es absolutamente continua con respecto a ψ .

Proposición 2.17. *Suponga que K es φ -irreducible, entonces:*

1. *Existe una medida de máxima irreducibilidad.*
2. *Una medida de irreducibilidad es máxima si y sólo si, $\psi K \ll \psi$.*
3. *Sea ψ una medida de máxima irreducibilidad. Si $\psi(B) = 0$ entonces, también $\psi(B^+) = 0$ donde:*

$$B^+ =: B \cup \{x \in E : x \rightarrow B\} = \{G1_B > 0\}.$$

Demostración.

(2). \Leftarrow]

Supongamos que ψ es una medida de irreducibilidad que satisface que $\psi K \ll \psi$.

por el Lema 2.12, pág. 37, sabemos que $\psi G \ll \psi$, de este hecho y de que K es φ -irreducible, se sigue que si $\varphi(A) > 0$, entonces $G(x, A) > 0$, para toda $x \in E$, entonces

$$\int_E \psi(dx)G(x, A) > 0,$$

es decir $\psi G(A) > 0$, entonces $\psi(A) > 0$. De aquí concluimos que ψ es máxima.

\Rightarrow]

Si ψ es una medida de irreducibilidad, entonces $\psi K(A) = \int_E \psi(dx)K(x, A) > 0$ para toda $A \in \mathcal{E}$ con A ψ -positivo, entonces la medida ψK también es de irreducibilidad, ahora si se supone que ψ es máxima entonces ψK cumple con que $\psi K \ll \psi$.

2.2 Kernels irreducibles

(1). Se define una medida finita ψ , como en la ecuación (2.2.1), pág. 36, esta es $\psi = \sum_{n=0}^{\infty} 2^{-(n+1)} \tilde{\varphi} \tilde{K}^n$, donde $\tilde{\varphi}$ es una medida de probabilidad equivalente a φ y \tilde{K} es subestocástico equivalente a K . Como φ es de irreducibilidad, entonces para toda $A \in \mathcal{E}$ tal que $\varphi(A) > 0$, se tiene que

$$\begin{aligned} \psi(A) &= \sum_{n=0}^{\infty} 2^{-(n+1)} \tilde{\varphi} \tilde{K}^n(A) \\ &= \sum_{n=0}^{\infty} 2^{-(n+1)} \int_E \tilde{\varphi}(dx) \tilde{K}^n(x, A) > 0, \end{aligned}$$

entonces ψ es una medida de irreducibilidad, y por la Proposición 4.13, pág. 106 $\psi K \ll \psi$, por lo tanto es máxima por (2).

(3). Por definición $B^+ = B \cup \{x \in E : x \rightarrow B\} = \{G1_B = \int_B G(x, dy) > 0\}$.

Como $\psi(B) = 0$, entonces,

$$0 = \psi K(B) = \int_E \psi(dx) K(x, B),$$

por el Lema 2.12, pág 37

$$0 = \psi G(B) = \int_E \psi(dx) G(x, B),$$

es decir $F = \{x \in E : x \rightarrow B\} = \emptyset$. Entonces

$$\begin{aligned} \psi G(B^+) &= \psi G(B) + \psi G(F) \\ &= 0 + \int_E \psi(dx) G(x, F) \\ &= \int_E \psi(dx) G(x, \emptyset) \\ &= 0. \end{aligned}$$

□

Definición 2.18. (*Conjunto completo*). Si K es irreducible con medida máxima de irreducibilidad ψ , se dice que un conjunto $B \in \mathcal{E}$ es **completo** si $\psi(B^c) = 0$.

Proposición 2.19. *Suponga que K es irreducible con medida máxima de irreducibilidad ψ , entonces:*

1. *Todo conjunto cerrado F es completo.*
2. *Si B es un conjunto completo entonces existe un conjunto cerrado $F \subseteq B$.*
3. *Si $F_i, i \geq 1$, son conjuntos cerrados entonces también su intersección $\bigcap_{n=0}^{\infty} F_i$.*

Demostración.

(1). Supongamos que $\psi(F^c) > 0$.

Como K es irreducible y F^c es ψ -positivo, se tiene que, para toda $x \in E$, $K^n(x, F^c) > 0$, para algún $n \geq 1$; en particular cuando $x \in F$, también $K^n(x, F^c) > 0$, pero esto es una contradicción ya que por hipótesis F es cerrado.

Por lo tanto $\psi(F^c) = 0$.

(2). Sea $F = (B^c)^0 = \{x \in B : x \not\rightarrow B^c\}$.

Por la Proposición 2.9, pág 35, $F = (B^c)^0$ es vacío o cerrado, entonces sólo queda por demostrar que F es no vacío.

Como B es completo, entonces $\psi(B^c) = 0$, por la Proposición 2.17, pág 40, también $\psi((B^c)^+) = 0$, donde $(B^c)^+ = B^c \cup \{x \in B : x \rightarrow B^c\}$; entonces todos los elementos $x \in E$ que se comunican con B^c tienen medida 0.

Entonces $\psi(F) > 0$, es decir F es no vacío, por lo tanto F es cerrado.

(3). En general la intersección de un número contable de conjuntos cerrados es cerrado. Pero ahora bajo el hecho de irreducibilidad como todo F_i es completo, entonces su intersección también lo es, veamos este hecho. Como cada F_i es cerrado, por (1.), también cada F_i es completo, con $i \geq 1$.

Notando que

$$\begin{aligned} \psi \left[\left(\bigcap_{i=1}^{\infty} F_i \right)^c \right] &= \psi \left[\bigcup_{i=1}^{\infty} F_i^c \right] \\ &\leq \sum_{i=1}^{\infty} \psi[F_i^c] = 0, \end{aligned}$$

Por lo tanto, como $\psi[(\bigcap_{i=1}^{\infty} F_i)^c] = 0$, entonces, $\bigcap_{i=1}^{\infty} F_i$ es completa.

□

2.3. Funciones pequeñas

En esta sección se definirá una clase importante de funciones, llamadas funciones pequeñas, estas funciones jugarán un papel muy importante en la separación de espacios de estados de una cadena de Markov.

Se definirá que para toda subclase $\mathcal{A} \subseteq \mathcal{G}_+$ de funciones medibles no negativas en (E, \mathcal{E}) , \mathcal{A}^+ denota las subclases de elementos ψ -positivos en \mathcal{A} , es decir:

$$\mathcal{A}^+ = \{f \in \mathcal{A} : \psi(f) > 0\},$$

donde $\psi(f) = \int f(x)\psi(dx) = \int_E f(x)\psi(dx)$. Cuando $1_A \in \mathcal{A}^+$, es decir, $\psi(A) > 0$, sólo se escribirá $A \in \mathcal{A}^+$.

Se dice que K *satisface la condición de minorización* $M(m_0, \beta, s, \nu)$, si

$$K^{m_0}(x, A) \geq \beta s(x)\nu(A) \quad \forall x \in E, A \in \mathcal{E},$$

donde $m_0 \geq 1$ es un *entero*, $\beta > 0$ es una *constante*, $s \in \mathcal{G}^+$ una *función* y $\nu \in \mathcal{M}^+(\mathcal{E})$ una *medida*, una manera abreviada para denotar esta condición es $K^{m_0} \geq \beta s \otimes \nu$. Esta cota inferior para el kernel será fundamental para la construcción de la separación del espacio de estados de una cadena de Markov.

Definición 2.20. (*Función, medida y conjunto pequeño*). Se denotará con \mathcal{S} la clase de funciones pequeñas y se define:

- Una función $s \in \mathcal{G}^+$ es llamada una **función pequeña** si K satisface $M(m_0, \beta, s, \nu)$, para algún $m_0 \geq 1$, $\beta > 0$ y $\nu \in \mathcal{M}^+(\mathcal{E})$.
- Una medida $\nu \in \mathcal{M}^+(\mathcal{E})$ es llamada una **medida pequeña** si K satisface $M(m_0, \beta, s, \nu)$, para algún $m_0 \geq 1$, $\beta > 0$ y $s \in \mathcal{G}^+$.
- Un conjunto $C \in \mathcal{G}^+$ es llamado **conjunto pequeño** si su función indicadora 1_C es una función pequeña.

De aquí en adelante se usarán los símbolos s y ν para denotar una función y una medida pequeña respectivamente.

Observación 2.21. 1. *Una medida pequeña ν es siempre una medida de irreducibilidad para K y es absolutamente continua con respecto a ψ .*

2. *Si se toma $K^{m_0}(x, A) \geq \beta s(x)\nu(A)$ y a esta desigualdad le multiplicamos por ambos lados a $K^m(x, A)$, es decir*

$$K^m K^{m_0}(x, A) \geq \beta s(x) K^m \nu(A)$$

se puede ver que $K^m s$ es una función pequeña y también νK^m es una medida pequeña, para toda $m \geq 0$, por irreducibilidad y sin perder generalidad se supone que $s(x)\nu(A) > 0$.

3. *Si $K = P$, donde P es la probabilidad de transición de una cadena de Markov, entonces ν es finita y s es acotado por $(\beta\nu(E))^{-1}$, por lo que se puede suponer sin pérdida de generalidad que*

$$P^{m_0}(x, A) \geq s(x)\nu(A), \quad \forall x \in E, A \in \mathcal{E},$$

donde $0 \leq s(x) \leq 1$ y ν es una medida de probabilidad.

La siguiente proposición será utilizada más adelante.

Proposición 2.22. 1. *Para toda $f \in \mathcal{G}^+$ y toda función pequeña $s \in \mathcal{S}$, existe un entero $m \geq 1$ y una constante $\gamma > 0$ tales que*

$$K^m f \geq \gamma s.$$

2.3 Funciones pequeñas

2. Para toda $f \in \mathcal{G}^+$ y $s \in \mathcal{S}$, existe una constante $\gamma \geq 0$ tal que

$$Gf \geq \gamma Gs.$$

Demostración.

(1). Tenemos

$$K^m f \geq \beta(\nu K^{m-m_0} f)s \quad \forall m \geq m_0.$$

Elijiendo $m \geq m_0$ tal que $\gamma = \beta\nu K^{m-m_0} f > 0$, entonces

$$K^m f \geq \gamma s.$$

(2). Del inciso anterior, tenemos

$$K^m f \geq \gamma s.$$

multiplicando por K^n

$$K^n K^m f \geq \gamma K^n s,$$

sumando sobre $n \in \mathbb{N}$

$$\sum_{n=0}^{\infty} K^n K^m f \geq \sum_{n=0}^{\infty} \gamma K^n s$$

por el Teorema de convergencia monótona

$$K^m \sum_{n=0}^{\infty} K^n f \geq \gamma Gs,$$

$$K^m Gf \geq \gamma Gs,$$

se tiene que

$$K^m Gf \leq Gf$$

por lo tanto

$$Gf \geq \gamma Gs.$$

□

2.4. Recurrencia y transitoriedad

Con la teoría revisada anteriormente, ahora se podrán analizar los conceptos de recurrencia y transitoriedad. Comenzando por un kernel general irreducible, después con un kernel potencial y por último cuando $K = P$ donde P es una probabilidad de transición de una cadena de Markov, y P no es necesariamente irreducible. En el caso particular, en el que la cadena de Markov es recurrente e irreducible se tendrá el concepto de Harris-recurrencia.

2.4.1. Elementos básicos de teoría potencial

A lo largo de esta sección se utilizarán nociones básicas de teoría potencial aplicadas a los kernels no negativos.

Definición 2.23. (*Función armónica, superarmónica y potencial*). Sea $h, p \in \mathcal{G}_+$ funciones, con $h \neq \infty$ y $p \neq \infty$, entonces:

- h es llamada *armónica* para el kernel K , si:

$$h = Kh.$$

- h es llamada *superarmónica* para el kernel K , si:

$$h \geq Kh.$$

- p es llamado *potencial*, si existe una función $g \in \mathcal{G}_+$ llamada la *carga* de p tal que:

$$p = Gg,$$

donde $G = \sum_{n=0}^{\infty} K^n$ es el kernel potencial.

Proposición 2.24. *Todo potencial p es una función superarmónica.*

Demostración. Por la Proposición 2.8, pág 35

$$\begin{aligned}
 p = Gg &= (I + KG)g \\
 &= (Ig + KGg) \\
 &= g + Kp \\
 &\geq Kp,
 \end{aligned}$$

por lo tanto p es superarmónica.

□

Proposición 2.25. *Suponga que h es una función superarmónica, entonces*

1. *El conjunto $\{x \in E : h(x) < \infty\}$ es cerrado para K .*
2. *Para toda $x \in E$, $h(x) > 0$ ó bien el conjunto $\{x \in E : h(x) = 0\}$ es cerrado para K .*

Demostración.

(1). Sea $F = \{x \in E : h(x) < \infty\}$ y $F^c = \{x \in E : h(x) = \infty\}$. Cuando $x \in F$ se observa que

$$\begin{aligned}
 \infty > h(x) \geq Kh(x) &= \int_E K(x, dy)h(y) \\
 &\geq \int_{F^c} K(x, dy)h(y) \\
 &\geq \infty \int_{F^c} K(x, dy) \\
 &\geq \infty \cdot K(x, F^c)
 \end{aligned}$$

entonces $K(x, F^c) = 0$ para toda $x \in F$, por lo tanto $F = \{x \in E : h(x) < \infty\}$ es cerrado.

(2). Sea $F = \{x \in E : h(x) = 0\}$ y $F^c = \{x \in E : h(x) > 0\}$.

Si $x \in F$

$$0 = h(x) \geq Kh(x) = \int_E K(x, dy)h(y) \geq \int_{F^c} K(x, dy)h(y)$$

si $K(x, F^c) > 0$, entonces

$$\int_{F^c} K(x, dy)h(y) > 0,$$

entonces

$$0 = h(x) \geq Kh(x) \geq \int_{F^c} K(x, dy)h(y) > 0,$$

lo cual no puede ser, por lo tanto F es cerrado, es decir

$$K(x, F^c) = 0, \quad \forall x \in F.$$

Ahora se enuncia un resultado importante, el *teorema de descomposición de Riesz*.

Teorema 2.26. 1. Si $h \in \mathcal{G}_+$, con $h \not\equiv \infty$, es la suma de un potencial $p = Gg$ y una función armónica h^∞ , es decir, $h = p + h^\infty$, entonces h es superarmónica. En el conjunto cerrado $F = \{x \in E : h(x) < \infty\}$ se tiene

- $g = h - Kh$ y

- $h^\infty = \lim_{n \rightarrow \infty} \downarrow K^n h$

(donde h^∞ representa una sucesión decreciente cuando $n \rightarrow \infty$).

2. Si h es superarmónica, entonces h es la suma de un potencial $p = Gg$ y una función armónica h^∞ en el conjunto $F = \{x \in E : h(x) < \infty\}$, donde $g = h - Kh$ y $h^\infty = \lim_{n \rightarrow \infty} \downarrow K^n h$.

3. Si $h \in \mathcal{G}_+$ es superarmónica, y $g \in \mathcal{G}_+$ es tal que $h \geq g + Kh$, entonces

$$h \geq Gg.$$

La pareja (p, h^∞) es llamada la descomposición de Riesz de la función superarmónica h en el conjunto $F = \{x \in E : h(x) < \infty\}$ y la parte 3 es llamado el teorema de Balayage.

Demostración.

(1). Tenemos que $p = Gg$, donde G es el kernel potencial, $g \in \mathcal{G}_+$ y $h^\infty = Kh^\infty$ ya que h^∞ es

armónica y por la Proposición 2.8, pág 35, tenemos

$$\begin{aligned}
 h &= p + h^\infty \\
 &= p + Kh^\infty \quad (h^\infty \text{ armónica}) \\
 &\geq Kp + Kh^\infty \quad (p \text{ superarmónica}) \\
 &= K(p + h^\infty) \\
 &= Kh,
 \end{aligned}$$

entonces h es superarmónica.

Si $x \in F$, entonces

$$\begin{aligned}
 h(x) &\geq Kh(x) \\
 \infty &> h(x) - Kh(x) \geq 0.
 \end{aligned}$$

entonces llamemos $g = h - Kh$. Por lo anterior $g \in \mathcal{G}_+$, y éste es finito en F .

(2). Supongamos h superarmónica, sabemos que $g = h - Kh > 0$, valido en F despejando h de la ecuación anterior se obtiene $h = g + Kh$, entonces realizando la siguiente iteración

$$\begin{aligned}
 h &= g + Kh \\
 &= g + K(g + Kh) = g + Kg + K^2h \\
 &= g + K(g + K(g + Kh)) = g + Kg + K^2g + K^3h \\
 &= g + K(g + K(g + K(g + Kh))) = g + Kg + K^2g + K^3g + K^4h \\
 &\vdots \\
 &= \sum_{m=0}^{n-1} K^m g + K^n h, \\
 &\vdots \\
 &= \lim_{n \rightarrow \infty} \left[\sum_{m=0}^{n-1} K^m g + K^n h \right] = p + h^\infty
 \end{aligned}$$

cuando $n \rightarrow \infty$ el primer término del lado derecho crece a $Gg = p$ y el segundo término decrece a la función h^∞ y es armónica en el conjunto F por el Teorema de convergencia monótona.

(3). Como $h \in \mathcal{G}_+$, es superarmónica y $g \in \mathcal{G}_+$ es tal que $h \geq g + Kh$, entonces

$$\begin{aligned}
 h &\geq g + Kh \\
 &\geq g + K(g + Kh) = g + Kg + K^2h \\
 &\geq g + K(g + K(g + Kh)) = g + Kg + K^2g + K^3h \\
 &\geq g + K(g + K(g + K(g + Kh))) = g + Kg + K^2g + K^3g + K^4h \\
 &\vdots \\
 &\geq \lim_{n \rightarrow \infty} \left[\sum_{m=0}^{n-1} K^m g + K^n h \right] \\
 &= Gg + h^\infty \\
 &\geq Gg.
 \end{aligned}$$

□

2.4.2. R-transitoriedad y R-recurrencia

En esta parte se define R-transitoriedad y R-recurrencia del kernel K .

Para toda $0 \leq r < \infty$, se denota por $G^{(r)}$ al kernel potencial del kernel rK , es decir $G^{(r)} = \sum_{n=0}^{\infty} r^n K^n$.

Definición 2.27. (*Parámetro de convergencia*). Un número real $0 \leq R < \infty$ es llamado *Parámetro de convergencia del kernel K* , cuando existe un conjunto cerrado F tal que:

1. $G^{(r)}s < \infty$ en F , para toda $0 \leq r < R$, $s \in \mathcal{S}$, y
2. $G^{(r)}f \equiv \infty$, en F , para toda $r > R$, $f \in \mathcal{G}^+$.

El kernel K es llamado **R-transitorio** si sucede (1.) cuando $r = R$. Y es llamado **R-recurrente** si sucede (2.) cuando $r = R$.

El siguiente teorema nos dice que cuando tenemos un kernel irreducible existe un parámetro de convergencia y podemos hablar de R-recurrencia y R-transitoriedad.

Teorema 2.28. *Sea K un kernel irreducible, entonces:*

1. *Existe un parámetro de convergencia $0 \leq R < \infty$.*
2. *El kernel K es R -transitorio o R -recurrente.*

Demostración.

(2.) Fijemos una función pequeña $s \in \mathcal{S}$. Sea

$$R = \sup\{r \geq 0 : G^{(r)}s(x) < \infty, \text{ para algún } x \in E\}.$$

$$F^{(r)} = \{x \in E : G^{(r)}s(x) < \infty\}, \quad 0 \leq r < \infty.$$

Observemos que si $r > R$, entonces $G^{(r)}s(x) = \infty$.

Sabemos que $G^{(r)}s(x)$ es un potencial en $F^{(r)}$ y por la Proposición 2.24, pág 46, también es una función superarmónica para el kernel rK .

Por la Proposición 2.25 (1), pág 47, los conjuntos $F^{(r)}$, $0 \leq r < R$, son cerrados y por la Proposición 2.19 (3), pág 42, la intersección de los conjuntos $F^{(r)}$ es cerrada

$$F = \bigcap_{0 \leq r < R} F^{(r)} = \lim_{r_m \uparrow R} \downarrow F^{(r_m)}.$$

(En el caso especial cuando $R = 0$, el conjunto $F = F^{(0)} = E$).

Por otro lado sea $s' \in \mathcal{S}$ otra función pequeña, por la Proposición 2.22 (2), pág 44, existe una constante $\gamma > 0$ tal que

$$G^{(r)}s \geq \gamma G^{(r)}s';$$

además cuando estamos en el conjunto F

$$\infty > G^{(r)}s(x) \geq \gamma G^{(r)}s'(x), \quad \forall 0 \leq r < R,$$

por lo tanto, para cualquier función pequeña s' ,

$$G^{(r)}s'(x) < \infty,$$

es decir K es R -transitorio.

Por la misma Proposición 2.22, se puede ver que

$$G^{(r)}f \geq \gamma G^{(r)}s,$$

y cuando $r > R$

$$G^{(r)}f \geq \gamma G^{(r)}s = \infty,$$

entonces $G^{(r)}f = \infty$ cuando $r > R$; por lo tanto K es R -recurrente.

(1.) Para ver que el parámetro de convergencia R es finito; supongamos que K , satisface la condición de minorización $M(m_0, \beta, s, \nu)$, con $s \otimes \nu > 0$ como se vio en la Observación 2.21, pág 44, entonces

$$\begin{aligned} K^{m_0} &\geq \beta s \otimes \nu \\ K^{nm_0} &\geq \beta^n (s \otimes \nu)^n \\ K^{nm_0} s &\geq (\beta s \otimes \nu)^n s = (\beta \nu(s))^n s. \end{aligned}$$

Se sigue que $G^{(r)}s \geq \sum_{n=0}^{\infty} r^{nm_0} K^{nm_0} s = \infty$ en el conjunto $\{x \in E : s(x) > 0\}$, donde $r^{m_0} K^{m_0} \geq 1$ para que la suma sea infinita; pero también por la condición de minorización $r^{m_0} K^{m_0} \geq r^{m_0} (\beta s \otimes \nu)$, entonces podemos hacer

$$r^{m_0} (\beta s \otimes \nu) \geq 1$$

$$r \geq (\beta \nu(s))^{-\frac{1}{m_0}}.$$

Por lo tanto, para que $G^{(r)} < \infty$, se necesita que

$$R \leq (\beta \nu(s))^{-\frac{1}{m_0}}.$$

Por lo tanto existe un parámetro de convergencia $R < \infty$.

□

2.4.3. Tiempos de paro para una cadena de Markov

De aquí en adelante se supondrá que $K = P$ donde P es una probabilidad de transición para una cadena de Markov (no necesariamente irreducible), y definiremos los tiempos de paro de una cadena de Markov. Estas definiciones serán muy útiles para la separación de estados de la cadena de Markov.

Definición 2.29. Sea (\mathcal{F}_n) una historia. Una variable aleatoria $\bar{\mathbb{N}}$ -valuada T , es llamado un *tiempo aleatorio* si:

$$\{T = n\} \in \mathcal{F}_n, \quad \forall n \geq 0,$$

entonces el tiempo aleatorio T es llamado *tiempo de paro relativo a la historia* (\mathcal{F}_n) . Si $\{X_n\}_{n \geq 0}$ es una cadena de Markov con respecto a su historia (\mathcal{F}_n) entonces T es llamado *tiempo de paro para la cadena de Markov* (X_n, \mathcal{F}_n) . Si T es un tiempo de paro relativo a la historia interna (\mathcal{F}_n^X) entonces esta es llamada simplemente *tiempo de paro para la cadena de Markov* $\{X_n\}_{n \geq 0}$.

Definición 2.30. Un tiempo de paro T es llamado *tiempo de paro aleatorio* (para la cadena $\{X_n\}_{n \geq 0}$), si para toda $n \geq 0$, el evento $\{T = n\}$ y las post-cadena (X_n, X_{n+1}, \dots) son condicionalmente independientes, dada la pre-n-cadena (X_0, \dots, X_n) , es decir:

$$\mathbb{P}\{T = n | \mathcal{F}^X\} = \mathbb{P}\{T = n | \mathcal{F}_n^X\} = f_n(X_0, \dots, X_n),$$

para alguna función $f_n, \mathcal{E}^{\otimes(n+1)}$ medible.

Las siguientes definiciones son ejemplos importantes de un tiempo de paro y a su vez serán muy importante para el desarrollo central de este trabajo.

Definición 2.31. Sea $B \in \mathcal{E}$ un conjunto arbitrario.

- Se define por T_B al *tiempo de primera visita* al conjunto B , es decir:

$$T_B = \inf\{n \geq 0 : X_n \in B\},$$

- Se define por S_B al *tiempo del primer regreso* al conjunto B , es decir:

$$S_B = \inf\{n \geq 1 : X_n \in B\}.$$

Por convención $\inf \emptyset = \infty$. La iteración $T_B(i)$, $i \geq 0$, es definida por: $T_B(0) = T_B$, y

$$T_B(i) = \inf\{n > T_B(i-1) : X_n \in B\},$$

es decir $T_B(i)$ es el tiempo de la $(i+1)$ -ésima visita al conjunto B . Se define $S_B(i)$ con $i \geq 1$ por: $S_B(1) = S_B$ y

$$S_B(i) = \inf\{n > S_B(i-1) : X_n \in B\}.$$

Entonces los tiempos aleatorios $T_B(i)$ y $S_B(i)$ son *tiempos de paro* para la cadena de Markov (X_n) .

Definición 2.32. Se define el kernel $I_B(x, A)$ como:

$$I_B(x, A) = 1_{A \cap B}(x), \quad x \in E, A \in \mathcal{E}.$$

Se definen los kernels G_B , G'_B y U_B por

$$\begin{aligned} G_B(x, A) &= \mathbb{E}_x \sum_{n=0}^{T_B} 1_A(X_n) = \sum_{n=0}^{\infty} \mathbb{P}_x\{X_n \in A, T_B \geq n\} \\ &= \sum_{n=0}^{\infty} (I_{B^c}P)^n(x, A), \\ G'_B(x, A) &= \mathbb{E}_x \sum_{n=0}^{S_B-1} 1_A(X_n) = \sum_{n=0}^{\infty} \mathbb{P}_x\{X_n \in A, S_B > n\} \\ &= \sum_{n=0}^{\infty} (PI_{B^c})^n(x, A) \\ U_B(x, A) &= \mathbb{E}_x \sum_{n=1}^{S_B} 1_A(X_n) = \sum_{n=0}^{\infty} P(I_{B^c}P)^n(x, A), \quad x \in E, A \in \mathcal{E}. \end{aligned}$$

Donde $G_B(x, A)$ representa “el valor esperado del número de visitas al conjunto $A \in \mathcal{E}$ desde un estado inicial $x \in E$ sin pasar por el conjunto $B \in \mathcal{E}$, terminando cuando se llegue por primera vez a B ”.

Y también se tiene

$$\begin{aligned} G_B I_B(x, A) &= \mathbb{P}_x\{X_{T_B} \in A, T_B < \infty\}, \\ U_B I_B(x, A) &= \mathbb{P}_x\{X_{S_B} \in A, S_B < \infty\}, \end{aligned}$$

La siguiente definición no es un tiempo de paro, pero será utilizado más adelante.

Definición 2.33. Sea $B \in \mathcal{E}$ un conjunto arbitrario; se define por L_B como el *último tiempo de salida* desde el conjunto B , es decir:

$$L_B = \sup\{n \geq 0 : X_n \in B\}.$$

L_B es definido sólo en el conjunto $\{T_B < \infty\} = \{X_n \in B \text{ para algún } n \geq 0\}$. L_E es llamado el *tiempo de vida* de $\{X_n\}_{n \geq 0}$, definido como:

$$L_E = \sup\{n \geq 0 : X_n \in E\} = \inf\{n \geq 0 : X_{n+1} = \Delta\}.$$

Se definen las siguientes funciones:

$$\begin{aligned} h_B(x) &= \mathbb{P}_x\{T_B < \infty\} = G_B(x, B), \\ p_B(x) &= \mathbb{P}_x\{0 \leq L_B < \infty\}, \\ g_B(x) &= \mathbb{P}_x\{L_B = 0\} = 1_B(x)\mathbb{P}_x\{S_B = \infty\} \\ h_B^\infty(x) &= \mathbb{P}_x\{X_n \in B \text{ i.o.}\} \end{aligned}$$

(donde i.o. significa: *ocurre infinidad de veces*).

Teorema 2.34. La función h_B es superarmónica, p_B es el potencial con carga g_B , h_B^∞ es armónica y (p_B, h_B^∞) es la descomposición de Riesz de h_B . Además, h_B es la función superarmónica h más pequeña que satisface $h \geq 1_B$.

Demostración. Ver [6, Teorema 3.4. Pág. 36].

□

Se denotará por B^∞ :

$$B^\infty = \{h_B^\infty = 1\} = \{x \in E : \mathbb{P}_x\{X_n \in B \text{ i.o.}\} = 1\}.$$

Proposición 2.35. Para todo conjunto $B \in \mathcal{E}$:

1. Siempre $h_B^\infty > 0$ ó el conjunto $\{x \in E : h_B^\infty(x) = 0\}$ es cerrado.
2. Si B es absorbente entonces $p_B \equiv 0$ y $h_B = h_B^\infty$.

Demostración.

(1). Sea $F = \{x \in E : h_B^\infty(x) = 0\}$ y $C = \{x \in E : h_B^\infty(x) > 0\}$.

Si $x \in F$

$$0 = h_B^\infty(x) \geq K h_B^\infty(x) = \int_E K(x, dy) h_B^\infty(y) \geq \int_{F^c} K(x, dy) h_B^\infty(y)$$

si $K(x, F^c) > 0$, entonces

$$\int_{F^c} K(x, dy) h_B^\infty(y) > 0,$$

entonces

$$0 = h_B^\infty(x) \geq K h_B^\infty(x) \geq \int_{F^c} K(x, dy) h_B^\infty(y) > 0,$$

lo cual no puede ser, por lo tanto F es cerrado, es decir

$$K(x, F^c) = 0, \quad \forall x \in F.$$

(2). Se sabe que $p_B = \mathbb{P}_x\{0 \leq L_B < \infty\}$, como B es absorbente, entonces $K(x, B) = 1$ para toda $x \in B$, es decir, x nunca sale del conjunto B . Por lo tanto $p_B = 0$.

Por otro lado $h_B^\infty = \mathbb{P}_x\{X_n \in B \text{ i.o.}\} = 1$, para toda $x \in B$ ya que B es absorbente, en particular la probabilidad de partir de un estado $x \in B$, y llegar por primera vez a un estado $y \in B$ en un tiempo finito es 1, es decir, $h_B = 1$. Por lo tanto

$$h_B^\infty = h_B = 1.$$

2.4.4. Las partes disipativa y conservativa

Aquí se verá que el espacio de estados de una cadena de Markov puede ser separado en dos partes. La parte disipativa E_d que es la unión numerable de conjuntos B_i , tal que, la cadena de Markov es transitoria en cada B_i , y la parte conservativa E_c , donde la cadena de Markov es recurrente. El espacio de estados queda separado en $E = E_c \cup E_d$.

Definición 2.36. ■ Un conjunto $B \in \mathcal{E}$ es llamado *transitorio*, si:

$$\mathbb{P}_x\{L_B < \infty\} = 1, \quad \forall x \in B,$$

ó equivalentemente, $h_B^\infty = 0$ en B .

- El conjunto $B \in \mathcal{E}$ es llamado **disipativo**, si existe una función $g \in \mathcal{G}_+$ tal que

$$g > 0 \text{ y } Gg < \infty, \quad \text{en } B.$$

donde G es el kernel potencial.

Si todo el espacio de estados E es transitorio, se dice que la cadena de Markov $\{X_n\}_{n \geq 0}$ es *transitoria*. De la misma manera si E es disipativa entonces toda la cadena de Markov $\{X_n\}_{n \geq 0}$ es *disipativa*.

Nótese que $\{X_n\}_{n \geq 0}$ es transitoria, si y sólo si, el tiempo de vida L_E de $\{X_n\}_{n \geq 0}$ es finito.

Definición 2.37. Un conjunto $B \in \mathcal{E}$ φ -positivo es llamado **φ -conservativo**, si para todo conjunto $A \subseteq B$, con A φ -positivo, se tiene

$$h_A^\infty = 1 \quad \varphi - c.s. \text{ en } A.$$

Si todo el espacio E es φ -conservativo decimos que la cadena de Markov $\{X_n\}_{n \geq 0}$ es φ -conservativa.

Se puede decir que los conjuntos φ -conservativos pueden ser caracterizados como el complemento de todos los conjuntos disipativos φ -positivos.

Ahora se enuncia el teorema de la descomposición Hopf, el cual dice que el espacio de estados puede ser dividido en dos partes, en la parte disipativa y en la parte conservativa. Este teorema no se demuestra pero se puede consultar en [6, pág 41].

Teorema 2.38. (Descomposición de Hopf). *Sea $\varphi \in \mathcal{M}^+(\mathcal{E})$ una medida. El espacio de estados puede ser dividido en dos partes:*

$$E = E_d(\varphi) + E_c(\varphi),$$

donde $E_d(\varphi)$ es la parte disipativa y $E_c(\varphi)$ es la parte conservativa.

□

2.5. Recurrencia

Todo lo visto a lo largo de este Capítulo fue para poder definir los conceptos de recurrencia y Harris recurrencia de una cadena de Markov; se verán sólo los conceptos necesarios para poder analizar la separación de estados de la cadena.

Seguiremos denotando a ψ como la medida de máxima irreducibilidad.

Definición 2.39. Una cadena de Markov irreducible (X_n) es llamada *recurrente*, sí:

- $h_B^\infty > 0$ y
- $h_B^\infty = 1$ ψ -c.s., para todo conjunto B ψ -positivo,

Definición 2.40. Una cadena de Markov irreducible $\{X_n\}_{n \geq 0}$ es llamada *Harris-recurrente* o ψ -*recurrente*, sí:

- $h_B^\infty \equiv 1$ para todo conjunto B ψ -positivo.

Notemos que si $\{X_n\}_{n \geq 0}$ es recurrente entonces también es ψ -conservativo. El conjunto $\{x \in E : \mathbb{P}_x\{X_n \in B \text{ i.o.}\} = 1\}$ es absorbente para todo B ψ -positivo; en particular si $\{X_n\}_{n \geq 0}$ es Harris-recurrente entonces P es estocástica.

Definición 2.41. (*conjunto Harris para $\{X_n\}_{n \geq 0}$*). Sea $H \in \mathcal{E}$ un conjunto absorbente, tal que la cadena de Markov $\{X_n\}_{n \geq 0}$ restringida a H es Harris-recurrente, se dice que H es un **conjunto Harris para $\{X_n\}_{n \geq 0}$** .

Teorema 2.42. *Si la cadena de Markov $\{X_n\}_{n \geq 0}$ es Harris-recurrente entonces para toda función superarmónica h existe una constante $0 \leq c < \infty$ tal que $h = c$ ψ -c.s., y $h \geq c$.*

Demostración. Supongamos que $h \in \mathcal{G}_+$ es superarmónica; sea $c = \text{esssup } h$, ψ -c.s.¹ Entonces, para toda $\varepsilon > 0$ el conjunto $B = \{x : h(x) \geq c - \varepsilon\}$ es ψ -positivo, es decir, $\psi(B) > 0$.

Por hipótesis se sabe que

$$h(x) \geq (c - \varepsilon)1_B(x),$$

¹ $\text{esssup } |h| = \inf\{c : c \geq 0, |h| \leq c, \psi\text{-c.s.}\}$, [7, pág 70].

entonces

$$\left(\frac{1}{c-\epsilon}h\right)(x) \geq 1_B(x),$$

por la minimalidad de h_B (Teorema 2.34, pág 55):

$$\frac{1}{c-\epsilon}h(x) \geq h_B(x) \geq 1_B(x),$$

$$h(x) \geq (c-\epsilon)h_B(x) = c-\epsilon$$

esto sucede por que h_B es la función superarmónica más pequeña que cumple que $h_B \geq 1_B$ y $h_B = 1$ ya que la cadena de Markov es Harris recurrente. Por lo tanto

$$h \geq c.$$

Ahora sea $\{x : h(x) = c \text{ } \psi\text{-c.s.}\}$, entonces $\psi(A) > 0$. Se sabe que

$$h(x) = c1_A(x)$$

entonces

$$\frac{h(x)}{c} = 1_A(x) \Leftrightarrow h(x) = c \text{ } \psi\text{-c.s.}$$

□

Proposición 2.43. *Suponga que la cadena $\{X_n\}$ es recurrente, entonces existe una función armónica \underline{h} , con $\underline{h} = 1$ ψ -c.s., $h \leq 1$, la cual es mínima entre las funciones h superarmónicas que satisfacen $h = 1$ ψ -c.s.; \underline{h} es dada por:*

$$\underline{h} = h_H = h_H^\infty \quad \forall \text{ conjunto Harris } H \in \mathcal{E}.$$

Demostración. Sea $H \in \mathcal{E}$ un conjunto Harris; por la Proposición 2.35 (2), pág 55, $h_H = h_H^\infty$ ya que H es absorbente, además por el Teorema 2.34, pág 55 es armónica.

Si h es alguna función superarmónica que satisface $h = 1$ ψ -c.s. Por el Teorema 2.42, pág 58, entonces $h \geq 1$ en H , una vez más por el Teorema 2.34, implica que $h \geq h_H^\infty \geq 1_H$, además $h_H^\infty = 1$. Por lo tanto h_H^∞ es mínima.

□

Teoría de Renovación

En la sección 3.1 se analizarán los elementos generales de la teoría de renovación, tales como el proceso de conteo de renovación, la función de renovación, la ecuación de renovación y algunas propiedades básicas. Posteriormente, de la sección 3.2 en adelante se aplicará esta teoría para el estudio de kernels irreducibles descritos en el Capítulo 2 y se describirá el método de regeneración, que en general habla de cómo la cadena se regenera cada vez que visita a un conjunto $\{x_0\} \in \mathcal{E}$, usando como hipótesis la condición de minorización $M(m_0, \beta, s, \nu)$, con $m_0 = 1$. La bibliografía principal para este Capítulo fue: [8, Capítulo 3] y [6, Capítulo 4].

3.1. Elementos generales de Teoría de renovación

Supongamos que $\{Y_n; n \geq 0\}$ es una sucesión de variables aleatorias independientes, las cuales toman sólo valores no negativos, además que la sucesión $\{Y_n; n \geq 1\}$ es idénticamente distribuida con distribución común $F(x)$. Y suponemos además que

$$F(0-) = 0, \quad F(0) < 1,$$

o equivalentemente: para cada $n \geq 1$,

$$\mathbb{P}(Y_n < 0) = 0, \quad \mathbb{P}(Y_n = 0) < 1.$$

Para $n \geq 0$, definimos:

$$S_n = Y_0 + \cdots + Y_n.$$

La sucesión $\{S_n\}_{n \in \mathbb{N}}$ es llamada *sucesión de renovación*. Las cantidades S_n son pensadas como tiempos de ocurrencia y son llamadas *tiempos de renovación ó épocas*; también $\{S_n\}_{n \in \mathbb{N}}$ mide el tiempo total del proceso observado.

Definición 3.1. $\{S_n\}_{n \in \mathbb{N}}$ es llamado *Proceso con retraso* si

- $S_0 = Y_0$ y
- $Y_0 \neq 0$

y $\{S_n\}_{n \in \mathbb{N}}$ es llamado *Proceso puro* si

- $S_0 = 0$ y
- $Y_0 \equiv 0$

F es llamada la distribución interarribo, si F es propia *i.e.* ($F(\infty) = 1$), entonces el proceso es llamado *Propio*. Si F es *defectuosa i.e.* ($F(\infty) < 1$), entonces $\{S_n\}_{n \in \mathbb{N}}$ es *terminante o transitorio*. $F(t)$ es la función de distribución del tiempo de vida.

3.1.1. Proceso de conteo de renovación

Definición 3.2. La *función de conteo o proceso de conteo de renovación* se define como:

$$N(t) = \sum_{n=0}^{\infty} 1_{[0,t]}(S_n).$$

Esta función nos proporciona el número de renovaciones en el intervalo $[0, t]$.

Definición 3.3. La *función de renovación* es definida como la esperanza de la función de conteo:

$$U(t) = E[N(t)].$$

Las ecuaciones de renovación son una herramienta básica para resolver probabilidades de interés en teoría de renovación.

3.1 Elementos generales de Teoría de renovación

Si tenemos un *proceso puro*, entonces la función de renovación queda expresada de la siguiente manera:

$$\begin{aligned}
 U(t) &= E \left[\sum_{n=0}^{\infty} 1_{[0,t]}(S_n) \right] \\
 &= \sum_{n=0}^{\infty} E [1_{[0,t]}(S_n)] \quad (\text{por Teor. de conv. monótona}) \\
 &= \sum_{n=0}^{\infty} [1 \cdot \mathbb{P}(S_n \leq t) + 0 \cdot \mathbb{P}(S_n > t)] \\
 &= \sum_{n=0}^{\infty} \mathbb{P}(Y_1 + \dots + Y_n \leq t) \\
 &= \sum_{n=0}^{\infty} F^{n*}(t).
 \end{aligned}$$

Por otro lado, si tenemos un *Proceso con retraso*, si $S_0 = Y_0$ y Y_0 tiene una distribución G , entonces la función de renovación queda expresada como:

$$\begin{aligned}
 V(t) &= E \left[\sum_{n=0}^{\infty} 1_{[0,t]}(S_n) \right] \\
 &= \sum_{n=0}^{\infty} E [1_{[0,t]}(S_n)] \quad (\text{por Teor. de conv. monótona}) \\
 &= \sum_{n=0}^{\infty} [1 \cdot \mathbb{P}(S_n \leq t) + 0 \cdot \mathbb{P}(S_n > t)] \\
 &= \sum_{n=0}^{\infty} \mathbb{P}(Y_0 + Y_1 + \dots + Y_n \leq t) \\
 &= \sum_{n=1}^{\infty} G * F^{(n-1)*}(t) \\
 &= \sum_{n=1}^{\infty} F^{(n-1)*}(t) * G \\
 &= G * U(t).
 \end{aligned}$$

Las siguientes relaciones entre $\{S_n\}_{n \in \mathbb{N}}$ y $N(t)$ serán útiles para demostrar las propiedades límite de $N(t)$:

1. $\{N(t) \leq n\} = \{S_n > t\}$, $n \geq 0$,
2. $S_{N(t)-1} \leq t < S_{N(t)}$ en $\{N(t) \geq 1\}$

3. $\{N(t) = n\} = \{S_{n-1} \leq t < S_n\}, \quad n \geq 1.$

Teorema 3.4. *Supongamos $\mu = E(Y_1) = \int_0^\infty xF(dx) < \infty$ que es la Media del tiempo Interarribo.*

1. *Si $\mathbb{P}(Y_0 < \infty) = 1$, entonces,*

$$\frac{N(t)}{t} \rightarrow \frac{1}{\mu} \text{ c.s., cuando } t \rightarrow \infty,$$

2. *Si $\sigma^2 = \text{Var}(Y_1) < \infty$, entonces, $N(t)$ es $N(\mu t^{-1}, t\sigma^2\mu^{-3})$; i.e.,*

$$\lim_{t \rightarrow \infty} \mathbb{P}[(N(t) - t\mu^{-1})/(t\sigma^2\mu^{-3})^{1/2} \leq x] = N(0, 1, x).$$

donde $N(0, 1, x)$ es la función de distribución normal estándar con media cero y varianza uno evaluada en x y $N(\mu t^{-1}, t\sigma^2\mu^{-3})$ es la función de distribución normal con media μt^{-1} y varianza $t\sigma^2\mu^{-3}$.

Demostración.

(1). Sea $S_n = Y_0 + Y_1 + \dots + Y_n$ una sucesión de renovación con $Y_i, i \in \mathbb{N}$ v.a.i.i.d.

Usando la *Ley fuerte de los grandes números* tenemos:

$$\lim_{n \rightarrow \infty} \frac{1}{n} S_n = \lim_{n \rightarrow \infty} \left[\frac{1}{n} Y_0 + \frac{1}{n} \sum_{i=1}^n Y_i \right]$$

como cada límite del lado derecho de la igualdad existe, tenemos

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} S_n &= \lim_{n \rightarrow \infty} Y_0 \frac{1}{n} + \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_i \\ &= 0 + \mu \\ &= \mu \end{aligned}$$

Observamos que $N(t) \rightarrow \infty$, cuando $n \rightarrow \infty$ c.s. Entonces,

$$\lim_{t \rightarrow \infty} \frac{S_{N(t)}}{N(t)} = \lim_{t \rightarrow \infty} \left[\frac{Y_0}{N(t)} + \frac{\sum_{i=1}^{N(t)} Y_i}{N(t)} \right],$$

3.1 Elementos generales de Teoría de renovación

usando la ley fuerte de los grandes números una vez más, obtenemos

$$\lim_{t \rightarrow \infty} \frac{S_{N(t)}}{N(t)} = \mu \quad \text{con probabilidad 1.}$$

Por otro lado tenemos que

$$S_{N(t)-1} \leq t < S_{N(t)}$$

entonces

$$\frac{S_{N(t)-1}}{N(t)-1} \frac{N(t)-1}{N(t)} \leq \frac{t}{N(t)} \leq \frac{S_{N(t)}}{N(t)}$$

cuando $t \rightarrow \infty$, ambos extremos convergen a μ , y por lo tanto

$$\frac{N(t)}{t} = \frac{1}{\mu}, \quad \text{cuando } n \rightarrow \infty.$$

(2). Sea $[x]$ el mayor entero menor que o igual a x .

Usaremos la identidad $\{N(t) \leq n\} = \{S_n > t\}$ y el *Teorema de límite central*:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{(S_n - n\mu)}{\sigma\sqrt{n}} \leq x \right] = N(0, 1, x).$$

Tenemos:

$$\mathbb{P} \left[\frac{N(t) - t\mu^{-1}}{(\sigma^2 t \mu^{-3})^{1/2}} \leq x \right] = \mathbb{P} \left[N(t) \leq [x(\sigma^2 t \mu^{-3})^{1/2} + t\mu^{-1}] \right]$$

llamemos $h(t) = [x(\sigma^2 t \mu^{-3})^{1/2} + t\mu^{-1}]$.

Como $\{N(t) \leq n\} = \{S_n > t\}$

$$\mathbb{P}[S_{h(t)} > t] = \mathbb{P} \left[\frac{S_{h(t)} - \mu h(t)}{\sigma h^{1/2}(t)} > \frac{t - \mu h(t)}{\sigma h^{1/2}(t)} \right].$$

Llamemos $z(t) = [t - \mu h(t)] / \sigma h^{1/2}(t)$.

Es suficiente mostrar que $h(t) \rightarrow \infty$ (ya que $h(t)$ juega el papel de n en el Teorema de límite central) y que $z(t) \rightarrow -x$, lo que, por el teorema del límite central implicaría:

$$\mathbb{P} \left[\frac{S_{h(t)} - \mu h(t)}{\sigma h^{1/2}(t)} > z(t) \right] \rightarrow 1 - N(0, 1, -x) = N(0, 1, x).$$

Primero veamos qué pasa con $h(t)$; cuando $t \rightarrow \infty$

$$\begin{aligned} \lim_{t \rightarrow \infty} h(t) &= \lim_{t \rightarrow \infty} [x(\sigma^2 t \mu^{-3})^{1/2} + t\mu^{-1}] \\ &= \infty. \end{aligned}$$

Además $h(t) \sim t\mu^{-1}$, ya que son absolutamente continuas una con respecto a la otra, y cuando $t \rightarrow \infty$ también $t\mu^{-1} \rightarrow \infty$.

Para ver que $z(t) \rightarrow -x$, tenemos:

$$h(t) = x(\sigma^2 t \mu^{-3})^{1/2} + t\mu^{-1} + \varepsilon(t),$$

donde $|\varepsilon(t)| \leq t$, y de aquí:

$$\begin{aligned} z(t) &= \frac{t - \mu[x(\sigma^2 t \mu^{-3})^{1/2} + t\mu^{-1} + \varepsilon(t)]}{\sigma h^{1/2}(t)} \\ &= \frac{t - \mu x(\sigma^2 t \mu^{-3})^{1/2} - t - \mu\varepsilon(t)}{\sigma h^{1/2}(t)} \\ &\sim \frac{-\mu x(\sigma^2 t \mu^{-3})^{1/2}}{\sigma(t\mu^{-1})^{1/2}} \\ &= \frac{-x(\sigma^2 t \mu^2 \mu^{-3})^{1/2}}{(\sigma^2 t \mu^{-1})^{1/2}} \\ &= -x. \end{aligned}$$

El término $\varepsilon(t)$ representa la diferencia en $h(t)$ provocada por poner x en lugar de $[x]$. También se utilizó la equivalencia $h(t) \sim t\mu^{-1}$ para completar la última parte de la prueba.

□

Teorema 3.5. *Teorema elemental de Renovación.* Sea $\mu = E[Y_1] \leq \infty$. Entonces

$$\lim_{t \rightarrow \infty} \frac{V(t)}{t} = \lim_{t \rightarrow \infty} \frac{U(t)}{t} = \frac{1}{\mu}$$

dado que $Y_0 < \infty$ c.s.

Demostración. Del lema de Fatou (ver apéndice) y el Teorema 3.4 (1), implica que:

$$\begin{aligned} \mu^{-1} &= E \left[\liminf_{t \rightarrow \infty} t^{-1} N(t) \right] \\ &\leq \liminf_{t \rightarrow \infty} E \left[\frac{N(t)}{t} \right] \\ &= \liminf_{n \rightarrow \infty} \frac{1}{t} E[N(t)] \\ &= \liminf_{t \rightarrow \infty} \frac{1}{t} V(t). \end{aligned}$$

3.1 Elementos generales de Teoría de renovación

Ahora, sea $Y_0^* = 0$, $Y_i^* = Y_i \wedge b$, $S_0^* = 0$, $S_n^* = Y_1^* + \cdots + Y_n^*$, $N^*(t) = \sum_{n=0}^{\infty} 1_{[0,t]}(S_n^*)$ y $V^*(t) = E(N^*(t))$. Entonces, $S_n \geq S_n^*$, ya que S_n es un proceso con retraso y $N^*(t) \geq N(t)$ ya que elegimos $Y_i^* = y_i \wedge b$.

Usando la identidad de Wald (ver apéndice), tenemos

$$\begin{aligned} E[S_{N^*(t)}^*] &= E \left[\sum_{i=1}^{N^*(t)} Y_i^* \right] \\ &= E[Y_1^*] E[N^*(t)] \\ &= E[Y_1^*] V^*(t) \end{aligned}$$

como

$$N^*(t) \geq N(t),$$

$$V^*(t) = E[N^*(t)] \geq E[N(t)] = V(t).$$

Por lo tanto

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{V(t)}{t} &\leq \limsup_{t \rightarrow \infty} \frac{V^*(t)}{t} \\ &= \limsup_{t \rightarrow \infty} \frac{E[S_{N^*(t)}^*]}{t E[Y_1^*]} \\ &= \limsup_{t \rightarrow \infty} \frac{1}{t} \left[\frac{E[S_{N^*(t)-1}^* + Y_{N^*(t)}^*]}{E[Y_1^*]} \right] \\ &\leq \limsup_{t \rightarrow \infty} \frac{1}{t} \frac{(t+b)}{E[Y_1^*]} \\ &= \frac{1}{E[Y_1^*]} \limsup_{t \rightarrow \infty} \frac{t+b}{t} \\ &= \frac{1}{E[Y_1 \wedge b]}. \end{aligned}$$

Hagamos que $b \rightarrow \infty$, sabemos que $E[Y_1] < \infty$, entonces, usando el Teorema de convergencia monótona se obtiene

$$\begin{aligned} \lim_{b \rightarrow \infty} E[Y_1 \wedge b] &= E \left[\lim_{b \rightarrow \infty} (Y_1 \wedge b) \right] \\ &= E[Y_1] \\ &= \mu. \end{aligned}$$

Entonces tenemos:

$$\limsup_{t \rightarrow \infty} \frac{V(t)}{t} \leq \frac{1}{\mu} \quad y \quad \liminf_{t \rightarrow \infty} \frac{V(t)}{t} \geq \frac{1}{\mu}.$$

Por lo tanto

$$\lim_{t \rightarrow \infty} \frac{V(t)}{t} = \frac{1}{\mu}.$$

□

3.1.2. Ecuación de renovación

Definición 3.6. Sean $F(t)$, $z(t)$ y $Z(t)$ funciones definidas para $t \geq 0$. Suponga que $F(t)$ y $z(t)$ son conocidas y $Z(t)$ desconocida. Se dice que $Z(t)$ satisface una ecuación de renovación si cumple la ecuación integral

$$Z(t) = z(t) + \int_0^t Z(t-y)F(dy).$$

Expresado de otra manera, la ecuación de renovación es la ecuación convolución de la forma:

$$Z = z + F * Z.$$

F es una distribución definida en $[0, \infty)$. Si $F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1$, la ecuación de renovación es llamada *ecuación propia*.

Ejemplo. Consideremos la función de renovación $U(t)$. Tenemos.

$$\begin{aligned} U(t) &= \sum_{n=0}^{\infty} F^{n*}(t) \\ &= F^{0*}(t) + \sum_{n=1}^{\infty} F^{n*}(t) \\ &= F^{0*}(t) + F * \sum_{n=1}^{\infty} F^{(n-1)*}(t) \\ &= F^{0*}(t) + F * U(t) \end{aligned}$$

y obtenemos la ecuación de renovación con $Z = U$ y $z = F^{0*}$.

□

3.1 Elementos generales de Teoría de renovación

A una función g se le llama **localmente acotada** si g es acotada en intervalos finitos.

Lema 3.7. Para una función localmente acotada g y una distribución F , definidas en $\mathbb{R}_+ = [0, \infty)$, entonces $F * g$ es localmente acotada; de hecho

$$\sup_{0 \leq s \leq t} |F * g(s)| \leq \left[\sup_{0 \leq s \leq t} |g(s)| \right] F(t).$$

Demostración. Definamos $\|g\| =: \sup_{0 \leq s \leq t} |g(s)|$, el cual es finito para todo t por hipótesis. Entonces, para toda $s \leq t$,

$$\begin{aligned} |F * g(s)| &= \left| \int_0^s g(s-x)F(dx) \right| \\ &\leq \int_0^s |g(s-x)||F(dx)| \\ &\leq \|g\| \int_0^s |F(dx)| \\ &\leq \|g\|F(s) < \infty. \end{aligned}$$

Este resultado se obtiene del Corolario 4.12, pág 106.

□

Teorema 3.8. Supongamos $z(t) = 0$ para $t < 0$, z localmente acotada y $F(0) < 1$.

1. Una solución localmente acotada de la ecuación de renovación

$$Z(t) = z(t) + \int_0^t Z(t-y)F(dy) \quad \text{es} \quad U * z(t) = \int_0^t z(t-u)U(du).$$

2. No hay otra solución localmente acotada.

Demostración.

(1). Primero veamos que $U * z$ es localmente acotada.

Para algún $T > 0$ y utilizando el Lema 3.7, pág 69 obtenemos:

$$\begin{aligned} \sup_{0 \leq t \leq T} |U * z(t)| &= \sup_{0 \leq t \leq T} \left| \int_0^t z(t-u)U(du) \right| \\ &\leq \left[\sup_{0 \leq t \leq T} |z(t)| \right] U(T) < \infty. \end{aligned}$$

Para verificar que $U * z$ es una solución, observemos que por la asociatividad de la convolución

$$F * (U * z) = (F * U) * z.$$

Recordando la ecuación de renovación para U del ejemplo anterior tenemos:

$$U(t) = F^{0*}(t) + F * U(t),$$

de donde

$$U(t) - F^{0*}(t) = F * U(t).$$

Entonces

$$\begin{aligned} F * (U * z) &= (F * U) * z \\ &= (U - F^{0*}) * z \\ &= U * z - z \end{aligned}$$

ya que $F^{0*} = 1_{[0, \infty)}(x)$.

Por lo tanto

$$U * z = z + F * (U * z)$$

se concluye que $U * z$ es una solución de la ecuación de renovación.

(2). Sean Z_1, Z_2 dos soluciones localmente acotadas de la ecuación de renovación tales que

$$Z_i = z + F * Z_i, \quad i = 1, 2.$$

Sea $H = Z_1 - Z_2$. H también es localmente acotada, además si hacemos la diferencia de las dos ecuaciones obtenemos

$$\begin{aligned} H &= Z_1 - Z_2 \\ &= [z + F * Z_1] - [z + F * Z_2] \\ &= F * Z_1 - F * Z_2 \\ &= F * [Z_1 - Z_2] = F * H. \end{aligned}$$

Iterando, para $n \geq 1$, obtenemos

$$H = F^{n*} * H.$$

3.2 Sucesiones y procesos de renovación.

Por lo tanto, para $T > 0$, y utilizando una vez más el Lema 3.7, pág 69 se obtiene:

$$\begin{aligned} \sup_{0 \leq t \leq T} |H(t)| &= \sup_{0 \leq t \leq T} \left| \int_0^t (Z_1(t-y) - Z_2(t-y)) F^{n*}(dy) \right| \\ &\leq \sup_{0 \leq t \leq T} H(t) F^{n*}(T) \rightarrow 0 \end{aligned}$$

cuando $n \rightarrow \infty$, puesto que H es localmente acotada y $U(T) < \infty$ implica que $F^{n*}(t) \rightarrow \infty$. De este modo $H \equiv 0$ y $Z_1 = Z_2$.

□

3.2. Sucesiones y procesos de renovación.

En esta sección realizamos un cambio en la notación, con la finalidad de abarcar un enfoque más general. Nos enfocaremos en los elementos necesarios para poder estudiar kernels irreducibles.

Suponga que X y Y son variables aleatorias no negativas independientes con

$$\mathbb{P}(X = k) = a_k, \quad \mathbb{P}(Y = k) = b_k, \quad k = 0, 1, \dots$$

para $n \geq 0$

$$\begin{aligned} [X + Y = n] &= \bigcup_{i=0}^n [X = i, Y = n - i] \\ \mathbb{P}(X + Y = n) &= \mathbb{P} \left\{ \bigcup_{i=0}^n [X = i, Y = n - i] \right\} \\ &= \sum_{i=0}^n \mathbb{P}(X = i, Y = n - i) \\ &= \sum_{i=0}^n a_i b_{n-i} =: p_n. \end{aligned}$$

Definición 3.9. La convolución de dos sucesiones $\{a_n\}_{n \geq 0}$ y $\{b_n\}_{n \geq 0}$, denotado como $\{c_n\}_{n \geq 0} = \{a_n\}_{n \geq 0} * \{b_n\}_{n \geq 0}$ resulta una nueva sucesión $\{c_n\}_{n \geq 0}$ definida como

$$c_n = \sum_{i=0}^n a_i b_{n-i}.$$

- La convolución es una operación conmutativa

$$\{a_n\}_{n \geq 0} * \{b_n\}_{n \geq 0} = \{b_n\}_{n \geq 0} * \{a_n\}_{n \geq 0}$$

- La convolución es una operación asociativa

$$\{a_n\}_{n \geq 0} * (\{b_n\}_{n \geq 0} * \{c_n\}_{n \geq 0}) = (\{a_n\}_{n \geq 0} * \{b_n\}_{n \geq 0}) * \{c_n\}_{n \geq 0}.$$

Sea $b = \{b_n, n \geq 0\}$ una sucesión que satisface

$$b_0 = 0, \quad 0 \leq b_n < \infty \quad \forall n \geq 0, \quad b_n > 0 \text{ para algún } n \geq 1.$$

Se define una *sucesión de renovación puro* $u = \{u_n, n \geq 0\}$, correspondiente a la sucesión de incrementos b , como

$$u = \sum_{i=0}^{\infty} b^{i*}.$$

Se define una *sucesión de renovación con retraso*, si $a = \{a_n, n \geq 0\}$ es una sucesión arbitraria no negativa y es tal que

$$v = a * u = a * \sum_{i=0}^{\infty} b^{i*},$$

se dice que a es la sucesión de retraso.

A continuación se enuncian proposiciones acerca de estas sucesiones de renovación.

Proposición 3.10. *La sucesión de renovación $u = \{u_n : n \geq 0\}$ es la única sucesión no negativa que satisface la ecuación de renovación*

$$u = \delta + b * u.$$

Donde

$$\delta = b^{0*} = (1, 0, 0, \dots)$$

3.2 Sucesiones y procesos de renovación.

Más generalmente, la sucesión con retraso $v = \{v_n : n \geq 0\} = a * u$ es la única sucesión que satisface la ecuación

$$v = a + b * u.$$

Demostración. Primero veamos que pasa con la sucesión u

$$\begin{aligned} u = \sum_{i=0}^{\infty} b^{i*} &= b^{0*} + \sum_{i=1}^{\infty} b^{i*} \\ &= b^{0*} + b * \sum_{i=1}^{\infty} b^{(i-1)*} \\ &= \delta + b * u. \end{aligned}$$

Ahora para la sucesión v

$$\begin{aligned} v = a * u &= a * \sum_{i=0}^{\infty} b^{i*} \\ &= a * (b^{0*} + \sum_{i=1}^{\infty} b^{i*}) \\ &= a * (b^{0*} + b * \sum_{i=1}^{\infty} b^{(i-1)*}) \\ &= a * b^{0*} + a * b * \sum_{i=1}^{\infty} b^{(i-1)*} \\ &= a + b * v. \end{aligned}$$

□

Ahora veamos qué pasa en el caso probabilístico.

Si $M_b^{(0)} = \sum_{n=1}^{\infty} b_n \leq 1$, entonces la sucesión de renovación u que le corresponde es *probabilística* y se dará una interpretación igual a como se vio en la primera sección de este Capítulo.

Sea $t(i)$, $1 \leq t(i) \leq \infty$, c.s., con $i = 1, 2, \dots$, tiempos aleatorios independientes e idénticamente distribuidos, con distribución común b . Como notación se tiene que $t(1) = t$, entonces para toda $i \geq 1$, se tiene

$$\blacksquare \quad \mathbb{P}\{t(i) = n\} = \mathbb{P}\{t = n\} = b_n, \quad n \geq 1$$

- $\mathbb{P}\{t(i) = \infty\} = \mathbb{P}\{t = \infty\} = 1 - M_b^{(0)}.$

Sea $T(0)$, $0 \leq T(0) \leq \infty$ c.s., tiempos aleatorios, independientes de $\{t(i) : i \geq 1\}$. Se define a $T(0)$ como el **retraso**, con distribución $a = \{a_n : n \geq 0\}$, es decir

- $\mathbb{P}\{T(0) = n\} = a_n, \quad n \geq 0,$
- $\mathbb{P}\{T(0) = \infty\} = 1 - \sum_{n=0}^{\infty} a_n$

Definición 3.11. Sea $T(i)$, $i \geq 1$, definido como

$$T(i) = T(0) + \sum_{j=0}^i t(j)$$

entonces $\{T(i) : i \geq 0\}$ es llamado **proceso de renovación con distribución de retraso a y distribución de incrementos b** .

La sucesión $\{Y_n : n \geq 0\}$ definida por

$$Y_n = 1_{\{T(i)=n \text{ para algún } i \geq 1\}}$$

es llamado el **proceso de incidencia** del proceso de renovación $\{T(i) : i \geq 0\}$.

Para especificar una distribución de retraso a , se escribirá como \mathbb{P}_a en lugar de \mathbb{P} .

Se puede observar que $v_n = a * u_n = \mathbb{P}_a\{Y_n = 1\}$; en particular para una sucesión de renovación sin retraso o puro u , se tiene que $u_n = \mathbb{P}_\delta\{Y_n = 1\}$ y en el caso probabilístico $0 \leq u_n \leq 1$ para toda $n \geq 0$.

Definición 3.12. Un proceso de renovación puro $\{T(i) : i \geq 0\}$ es llamado **recurrente** si $T(i) < \infty$ c.s. para toda $i \geq 0$ y en otro caso es llamado **transitorio**.

A continuación se define la función generadora de b y u

Definición 3.13. Se define la **función generadora** de $\hat{b}(r)$ y $\hat{u}(r)$ como

$$\hat{b}(r) = \sum_{n=1}^{\infty} r^n b_n, \quad \hat{u}(r) = \sum_{n=0}^{\infty} r^n u_n, \quad 0 \leq r < \infty,$$

donde $\hat{b}(0) = 0$ y $\hat{u}(0) = 1$.

3.3. Un átomo para kernels y cadenas de Markov

Antes de pasar al esquema general de regeneración se analizará cuando un kernel o una cadena de Markov tienen un átomo propio, el cual es esencial para la técnica de separación.

Definición 3.14. (*Átomo propio*). Un conjunto no vacío $\alpha \in \mathcal{E}$ con $\alpha \neq E$ es llamado un *átomo propio* para el kernel K , si

1. $K(x, \cdot) = K(y, \cdot)$ para todo $x, y \in \alpha$, y
2. $x \rightarrow \alpha$ para algún $x \in \alpha$.

Si tenemos un átomo propio $\alpha \in \mathcal{E}$, entonces para toda función $f \in \mathcal{G}$, si f es constante en α , por convención se escribe $f(\alpha)$ para el valor común de $f(x)$, $x \in \alpha$, y entonces podemos escribir $K(\alpha, \cdot)$, del inciso 1.

Se observa también que si x_0 es un estado comunicante, es decir, satisface

$$\{x_0\} \in \mathcal{E}, \quad x_0 \rightarrow x_0.$$

entonces $\{x_0\}$ es un átomo propio.

Cuando el kernel K tiene un átomo propio $\alpha \in \mathcal{E}$, con el átomo α siempre hay asociada una sucesión de renovación, como se verá en la siguiente Proposición.

Para la demostración de la siguiente proposición, se necesitará el siguiente Lema.

Lema 3.15. Sean γ y δ dos elementos de un anillo (una clase no vacía de subconjuntos de \mathcal{E} cerrado bajo uniones y diferencias conjuntistas), entonces para toda $n \geq 1$

$$\begin{aligned} (\gamma + \delta)^n &= \gamma^n + \sum_{m=1}^n \gamma^{m-1} \delta (\gamma + \delta)^{n-m} \\ &= \gamma^n + \sum_{m=1}^n (\gamma + \delta)^{m-1} \delta \gamma^{n-m}. \end{aligned}$$

Demostración. Ver [6, Lema 4.1. pág 53]

Proposición 3.16. Sea $\nu = K(\alpha, \cdot)$.

1. La sucesión definida por

$$u_n = K^n(\alpha, \alpha) = \begin{cases} 1 & \text{si } n = 0, \\ \nu K^{n-1}(\alpha) & \text{si } n \geq 1, \end{cases}$$

es una sucesión de renovación. Su sucesión de incrementos b esta determinada por

$$b_n = K(I_{\alpha^c}K)^{n-1}(\alpha, \alpha) = \nu(I_{\alpha^c}K)^{n-1}(\alpha), \quad n \geq 1.$$

2. Para toda $x \in E$, la sucesión $v(x)$ definida por

$$v(x) = K^n(x, \alpha), \quad n \geq 0,$$

es una sucesión de renovación con retraso

$$v(x) = a(x) * u,$$

donde la sucesión de retraso $a(x)$ esta dada por

$$a_n(x) = (I_{\alpha^c}K)^n(x, \alpha), \quad n \geq 0.$$

3. Para toda $A \in \mathcal{E}$, la sucesión $w(A)$ definida por

$$w_n(A) = K^{n+1}(\alpha, A) = \nu K^n(A), \quad n \geq 0,$$

tiene una sucesión de renovación con retraso:

$$w(A) = u * \sigma(A),$$

donde la sucesión de retraso $\sigma(A)$ está dada por

$$\sigma_n(A) = K(I_{\alpha^c}K)^n(\alpha, A) = \nu(I_{\alpha^c}K)^n(A), \quad n \geq 0.$$

Demostración. Sea $\gamma = I_{\alpha^c}K$ y $\delta = I_{\alpha}K$, entonces

$$\begin{aligned} \gamma + \delta &= I_{\alpha^c}K + I_{\alpha}K \\ &= (I_{\alpha^c} + I_{\alpha})K \\ &= K. \end{aligned}$$

(1) y (2). Por el Lema anterior se tiene que para la sucesión $v_n(x) = K^n(x, \alpha)$, $n \geq 0$:

$$\begin{aligned}
 v_n(x) &= (\gamma + \delta)^n(x, \alpha) \\
 &= (I_{\alpha^c}K)^n(x, \alpha) + \sum_{m=1}^n (I_{\alpha^c}K)^{m-1}I_{\alpha}K(I_{\alpha^c}K + I_{\alpha}K)^{n-m}(x, \alpha) \\
 &= (I_{\alpha^c}K)^n(x, \alpha) + \sum_{m=1}^n (I_{\alpha^c}K)^{m-1}I_{\alpha}K^{n-m+1}(x, \alpha) \\
 &= a_n(x) + (I_{\alpha^c}K)^0I_{\alpha}K^n(x, \alpha) + \cdots + (I_{\alpha^c}K)^{n-1}I_{\alpha}K^1(x, \alpha) \\
 &= a_n(x) + a_0(x)K^n(\alpha, \alpha) + \cdots + a_{n-1}(x)K^1(\alpha, \alpha) \\
 &= \sum_{m=0}^n a_m(x)K^{n-m}(\alpha, \alpha) \\
 &= a(x) * u_n,
 \end{aligned}$$

donde $u_n = K^n(\alpha, \alpha) = Kv_{n-1}(\alpha)$ y $b_n = K(I_{\alpha^c}K)^{n-1}(\alpha, \alpha) = Ka_{n-1}(\alpha)$, $n \geq 1$. Entonces u satisface la ecuación de renovación $u_n = \delta + b * u_n$, y por la Proposición 3.10, pág 72 es una sucesión de renovación.

(3). Una vez más, por el Lema anterior, se tiene que para la sucesión $v_n(x) = K^n(x, \alpha)$, $n \geq 0$:

$$\begin{aligned}
 w_n(A) &= (\gamma + \delta)^{n+1}(\alpha, A) \\
 &= (\gamma + \delta)(\gamma + \delta)^n(\alpha, A) \\
 &= K \left[(I_{\alpha^c}K)^n(\alpha, A) + \sum_{m=1}^n (I_{\alpha^c}K)^{m-1}I_{\alpha}K^{n-m+1}(\alpha, A) \right] \\
 &= K \left[\sum_{m=0}^n (I_{\alpha^c}K)^m I_{\alpha}K^{n-m}(\alpha, A) \right] \\
 &= \sum_{m=0}^n K(I_{\alpha^c}K)^m I_{\alpha}K^{n-m}(\alpha, A) \\
 &= \sum_{m=0}^n \sigma_n(A)K^{n-m}(\alpha, \alpha) \\
 &= u * \sigma_n(A),
 \end{aligned}$$

donde $u_n = K^n(\alpha, \alpha) = w_{n-1}(A)$ y $b_n = K(I_{\alpha^c}K)^{n-1}(\alpha, \alpha) = \sigma_{n-1}(\alpha)$, $n \geq 1$.

□

La sucesión definida por $u = \{K^n(\alpha, \alpha) : n \geq 0\}$ es llamada **sucesión de renovación embebida, asociada con el átomo propio α** .

Ahora pasemos a un caso particular, este es cuando el kernel K es una probabilidad de transición de una cadena de Markov $\{X_n\}_{n \geq 0}$, es decir, $K = P$ y P tiene un átomo propio $\alpha \in \mathcal{E}$.

Supongamos que $\{X_n\}_{n \geq 0}$ tiene un estado inicial $X_0 = x \in E$. Los tiempos $\{T_\alpha(i) : i \geq 0\}$ de visitas sucesivas al átomo propio α forman un proceso de renovación que corresponde al proceso de incidencia $Y_n = 1_\alpha(X_n)$, $n \geq 0$. Su retraso es $T_\alpha(0) = T_\alpha$. La distribución de retraso $a(x)$ está dada por

$$a_n(x) = \mathbb{P}_x\{T_\alpha = n\} = (I_{\alpha^c}P)^n(x, \alpha), \quad n \geq 0.$$

Los incrementos $t_\alpha(i) = T_\alpha(i) - T_\alpha(i-1)$, $i \geq 1$, tienen distribución común

$$b_n = \mathbb{P}_\alpha\{S_\alpha = n\} = P(I_{\alpha^c}P)^{n-1}(\alpha, \alpha), \quad n \geq 1.$$

La sucesión de renovación sin retraso o puro u es

$$u_n = \mathbb{P}_\alpha\{X_n \in \alpha\} = P^n(\alpha, \alpha), \quad n \geq 0,$$

la sucesión de renovación con retraso $v(x)$ está dada por

$$v_n(x) = \mathbb{P}_x\{X_n \in \alpha\} = P^n(x, \alpha) = a(x) * u_n, \quad n \geq 0 \tag{3.3.1}$$

La identidad (3.3.1) es llamada **descomposición de la primera entrada** de la probabilidad de transición $P^n(x, \alpha)$.

3.4. Esquema general de regeneración.

En esta sección se mostrará que un esquema de regeneración, se puede adaptar para un kernel irreducible que satisface la condición de minorización $M(m_0, 1, s, \nu)$, con $m_0 = 1$.

3.4 Esquema general de regeneración.

Definición 3.17. (*Átomo*). Una pareja (s, ν) , $s \in \mathcal{G}^+$, $\nu \in \mathcal{M}^+(\mathcal{E})$, es llamado un **átomo** si el kernel K satisface la condición de minorización $M(1, 1, s, \nu)$, es decir,

$$K \geq s \otimes \nu.$$

que es igual a decir $K(x, A) \geq s(x)\nu(A)$, con $x \in E$ y $A \in \mathcal{E}$.

Se dice que $s \in \mathcal{G}^+$ es accesible desde $\nu \in \mathcal{M}^+(\mathcal{E})$ si

$$\nu K^n s > 0 \quad \text{para algún } n \geq 0.$$

Teorema 3.18. *Sea (s, ν) un átomo, entonces*

1. *La sucesión $u = \{u_n : n \geq 0\}$ definida por*

$$u_0 = 0, \quad u_n = \nu K^{n-1} s, \quad n \geq 1,$$

es una sucesión de renovación. Su sucesión de incrementos b está dada por

$$b_0 = 0, \quad b_n = \nu(K - s \otimes \nu)^{n-1} s, \quad n \geq 1.$$

2. *Para todo $x \in E$, la sucesión $\{K^n s(x) : n \geq 0\}$ es una sucesión de renovación con retraso. Su sucesión de retraso $a(x) = \{a_n(x) : n \geq 0\}$ está dada por*

$$a_n(x) = (K - s \otimes \nu)^n s(x), \quad n \geq 0;$$

es decir,

$$K^n s(x) = a(x) * u_n, \quad n \geq 0.$$

3. *Para toda $A \in \mathcal{E}$, la sucesión $\{\nu K^n(A) : n \geq 0\}$ es una sucesión de renovación con retraso. Su sucesión de retraso $\sigma(A) = \{\sigma_n(A) : n \geq 0\}$ está dada por*

$$\sigma_n(A) = \nu(K - s \otimes \nu)^n(A), \quad n \geq 0;$$

es decir,

$$\nu K^n(A) = u * \sigma(A)_n, \quad n \geq 0.$$

4. Para todo $n \geq 1$, $x \in E$, $A \in \mathcal{E}$

$$K^n(x, A) = (K - s \otimes \nu)^n(x, A) + a(x) * u * \sigma(A)_{n-1}.$$

Demostración.

(1) y (2). Sea $\gamma = K - s \otimes \nu$, y $\delta = s \otimes \nu$, usando el Lema 3.15, tenemos que para toda $n \geq 1$:

$$K^n = (\gamma + \delta)^n$$

$$K^n = (K - s \otimes \nu)^n + \sum_{m=1}^n (K - s \otimes \nu)^{m-1} s \otimes \nu K^{n-m}. \quad (3.4.2)$$

Si hacemos $K^n s(x)$, y reordenamos la suma, obtenemos

$$\begin{aligned} K^n s(x) &= \sum_{m=0}^n (K - s \otimes \nu)^m s(x) s K^{n-m-1} \nu \\ &= \sum_{m=0}^n (K - s \otimes \nu)^m s(x) u_{n-m} \\ &= \sum_{m=0}^n a_m(x) u_{n-m} \\ &= a(x) * u_n. \end{aligned}$$

De donde $b_n = \nu(a_{n-1}) = \nu(K - s \otimes \nu)^{n-1} s$ para toda $n \geq 1$. Se sigue que u satisface la ecuación de renovación $u = \delta + b * u$ y que es una sucesión de renovación.

(3). Usando (3.4.2)

$$\begin{aligned} \nu K^n(A) &= \sum_{m=0}^n \nu(K - s \otimes \nu)^m(A) s K^{n-m-1} \nu \\ &= \sum_{m=0}^n \sigma_m(A) u_{n-m} \\ &= u * \sigma(A)_n \end{aligned}$$

(4). Por (3.4.2) y (3)

$$\begin{aligned}
 K^n(x, A) &= (K - s \otimes \nu)^n(x, A) + \sum_{m=1}^n (K - s \otimes \nu)^{m-1} s \otimes \nu K^{n-m}(x, A) \\
 &= (K - s \otimes \nu)^n(x, A) + \sum_{m=1}^n [(K - s \otimes \nu)^{m-1} s] [\nu K^{n-m}](x, A) \\
 &= (K - s \otimes \nu)^n(x, A) + \sum_{m=1}^n [a_{m-1}(x)] [u * \sigma_{n-m}(A)] \\
 &= (K - s \otimes \nu)^n(x, A) + a(x) * u * \sigma(A)_{n-1}.
 \end{aligned}$$

□

Proposición 3.19. *Suponga que K es un kernel irreducible y que tiene un átomo (s, ν) , entonces*

1. *El parámetro de convergencia R del kernel K y de la sucesión de renovación encajada u coinciden, y*

$$R = \sup\{r \geq 0 : \hat{u}(r) < \infty\} = \sup\{r \geq 0 : \hat{b}(r) < 1\}.$$

2. $\hat{b}(R) \leq 1$.

3. K es: R -recurrente $\Leftrightarrow \hat{u}(R) = \infty \Leftrightarrow \hat{b}(R) = 1$.

Demostración. Ver [6, Proposición 4.7. Pág 60]

□

Capítulo 4

Técnica de separación.

En este capítulo se describe la técnica que permite separar el espacio de estados de una cadena de Markov. La bibliografía principal para este Capítulo fue: [6, Capítulo 4], [5] y [10, Capítulo 1]

Supongamos ahora que el kernel $K = P$, es una probabilidad de transición de una cadena de Markov irreducible $\{X_n\}_{n \geq 0}$ y además P satisface la condición de minorización $M(m_0, 1, s, \nu)$, con $m_0 = 1$, es decir

$$P \geq s \otimes \nu,$$

$$i.e. \quad P(x, A) \geq s(x)\nu(A), \quad \forall x \in E, A \in \mathcal{E}.$$

Sin pérdida de generalidad se puede suponer que ν es una medida de probabilidad como se mencionó en la observación 2.21 (3), pág 44 y por consecuencia también $0 \leq s \leq 1$.

Ahora se define un kernel que será fundamental para la separación del espacio de estados.

Proposición 4.1. *Sean $s, \nu \in \mathcal{S}$, una función y una medida pequeña respectivamente y $P(x, A)$ con $x \in E$ y $A \in \mathcal{E}$ una probabilidad de transición de una cadena de Markov, entonces la función Q en (E, \mathcal{E}) definida por*

$$\begin{aligned} Q(x, A) &= \frac{P(x, A) - s(x)\nu(A)}{1 - s(x)}, & si \ s(x) < 1 \\ &= 1_A(x), & si \ s(x) = 1, \end{aligned}$$

es un kernel subestocástico.

Demostración. Primero veamos que para toda $x \in E$ fija, $Q(x, \cdot)$ es una medida.

$$\begin{aligned} Q(x, \emptyset) &= \frac{P(x, \emptyset) - s(x)\nu(\emptyset)}{1 - s(x)} \\ &= \frac{0 - 0}{1 - s(x)} \\ &= 0, \end{aligned}$$

ya que P es un kernel y ν es una medida de probabilidad. Sean $A_i \in \mathcal{E}$, con $i \in \mathbb{N}$, entonces

$$\begin{aligned} Q(x, \cup_{n \in \mathbb{N}} A_n) &= \frac{P(x, \cup_{n \in \mathbb{N}} A_n) - s(x)\nu(\cup_{n \in \mathbb{N}} A_n)}{1 - s(x)} \\ &\leq \frac{\sum_{n=0}^{\infty} P(x, A_i) - s(x) \sum_{n=0}^{\infty} \nu(A_i)}{1 - s(x)} \\ &= \sum_{n=0}^{\infty} \frac{P(x, A_i) - s(x)\nu(A_i)}{1 - s(x)} \\ &= \sum_{n=0}^{\infty} Q(x, A_i), \end{aligned}$$

Por lo tanto $Q(x, \cdot)$ es una medida.

Debido al Teorema [7, Pág. 42, (a),(b),(c)], si f y g son funciones medibles y $c \in \mathbb{R}$ entonces $f + g$, fg , f/g y cf también son funciones medibles. Entonces

$$Q(\cdot, A) = \frac{P(\cdot, A) - s(\cdot)\nu(A)}{1 - s(\cdot)}$$

es una función medible.

Cuando $s(x) = 1$ se tiene que

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

Por lo tanto $Q(x, A)$ es un kernel subestocástico.

□

A partir de la expresión del kernel Q podemos ver que la probabilidad de transición $P(x, A)$ está dada por

$$P(x, A) = s(x)\nu(A) + (1 - s(x))Q(x, A);$$

entonces la probabilidad de transición se separa en 2 partes. Así, una probabilidad de transición que comienza desde un estado inicial $x \in E$, puede ser interpretada en dos etapas. Primero, la función $s(x)$ se comporta como una función Bernoulli. Es decir, se lanza una moneda con probabilidad de obtener águila igual a $s(x)$, y sol igual a $1 - s(x)$. Si se obtiene como resultado *águila*, entonces la cadena de Markov se mueve de acuerdo a la medida de probabilidad $\nu(\cdot)$; si se obtiene como resultado *sol*, entonces se mueve de acuerdo a $Q(x, \cdot)$. Un hecho importante de esto es que al obtener *águila* en el experimento, la transición se mueve de acuerdo a una ley de probabilidad $\nu(\cdot)$, la cual es independiente del estado x . Entonces la cadena de Markov $\{X_n\}_{n \geq 0}$ tendrá asociado un proceso estocástico $\{Y_n\}_{n \geq 0}$ $\{0, 1\}$ -valuado, que representa los resultados del experimento de obtener *sol* o *águila* respectivamente, a partir del cuál se obtendrá una nueva cadena $\{(X_n, Y_n) : n \geq 0\}$ en el espacio de estados $E \times \{0, 1\}$ y veremos que el subconjunto $E \times \{1\}$ es un átomo propio para la nueva cadena $\{X_n, Y_n\}_{n \geq 0}$.

Para ver lo dicho anteriormente, tomemos una cadena de Markov $\{X_n\}_{n \geq 0}$ con respecto a una historia arbitraria fija $\{\mathcal{F}_n\}$. Si la probabilidad de transición P no es estocástica se puede completar a un kernel estocástico con una extensión $(E_\Delta, \mathcal{E}_\Delta)$ como la descrita en la sección 2.1.2. También al espacio de probabilidad (Ω, \mathcal{F}) hay que anexarle los resultados del experimento del proceso $\{Y_n\}_{n \geq 0}$; el espacio resultante es $(\Omega \times \{0, 1\}^{\times \infty}, \mathcal{F} \otimes \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}^{\otimes \infty})$.

Para construir la cadena separada se tendrá el nuevo espacio de estados $E^* = E \times \{0, 1\}$, donde $A_0 = A \times \{0\}$ y $A_1 = A \times \{1\}$, con $A \in \mathcal{E}$. Se denota por \mathcal{E}^* la σ -álgebra en E^* generada por los conjuntos A_i ($A \in \mathcal{E}, i = 0, 1$). Y por $x_i = (x, i), i = 0, 1$ a los elementos de E^* .

Si λ es una medida en $\mathcal{M}_+(\mathcal{E})$, entonces el siguiente paso en esta construcción es la separación de la medida λ en dos medidas en $E \times \{0\}$ y $E \times \{1\}$ como sigue:

$$\begin{aligned} \lambda^*(A_0) &= \lambda I_{1-s}(A) = \int_A \lambda(dx)(1 - s(x)) \\ \lambda^*(A_1) &= \lambda I_s(A) = \int_A \lambda(dx)s(x) \end{aligned}$$

A la medida $\lambda^* \in \mathcal{M}_+(\mathcal{E}^*)$ se le llama la *separación* de $\lambda \in \mathcal{M}_+(\mathcal{E})$. Como $A_0 \cup A_1 = A$ y

$A_0 \cap A_1 = \emptyset$ la separación está bien definida.

Sea $\{Y_n\}_{n \geq 0}$ un proceso estocástico $\{0, 1\}$ -valuado que depende de la cadena de Markov (X_n, \mathcal{F}_n) a través de las fórmulas:

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in A, Y_n = 1 | \mathcal{F}_n \vee \mathcal{F}_{n-1}^Y\} &= \mathbb{P}\{X_{n+1} \in A, Y_n = 1 | X_n\} \\ &= s(X_n)\nu(A), \end{aligned} \quad (4.0.1)$$

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in A, Y_n = 0 | \mathcal{F}_n \vee \mathcal{F}_{n-1}^Y\} &= \mathbb{P}\{X_{n+1} \in A, Y_n = 0 | X_n\} \\ &= [1 - s(X_n)]Q(X_n, A) \\ &= [1 - s(X_n)] \left[\frac{P(X_n, A) - s(X_n)\nu(A)}{(1 - s(X_n))} \right] \\ &= P(X_n, A) - s(X_n)\nu(A), \end{aligned} \quad (4.0.2)$$

$A \in \mathcal{E}_\Delta$, $n \geq 0$. Por convención se define $\mathcal{F}_{-1}^Y = \{\emptyset, \Omega\}$ y $\nu(\{\Delta\}) = 0$. Las ecuaciones anteriores también son equivalentes a las siguientes condiciones

$$\mathbb{P}\{X_{n+1} \in A | \mathcal{F}_n \vee \mathcal{F}_{n-1}^Y\} = \mathbb{P}\{X_{n+1} \in A | X_n\} = P(X_n, A), \quad (4.0.3)$$

$$\mathbb{P}\{Y_n = 1 | \mathcal{F}_n \vee \mathcal{F}_{n-1}^Y\} = \mathbb{P}\{Y_n = 1 | X_n\} = s(X_n), \quad (4.0.4)$$

$$\mathbb{P}\{X_{n+1} \in A | \mathcal{F}_n \vee \mathcal{F}_{n-1}^Y; Y_n = 1\} = \mathbb{P}\{X_{n+1} \in A | Y_n = 1\} = \nu(A), \quad (4.0.5)$$

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in A | \mathcal{F}_n \vee \mathcal{F}_{n-1}^Y; Y_n = 0\} &= \mathbb{P}\{X_{n+1} \in A | X_n; Y_n = 0\} \\ &= \mathbb{P}\{X_{n+1} \in A | Y_n = 0\} \\ &= Q(x, A), \end{aligned} \quad (4.0.6)$$

$A \in \mathcal{E}$, $n \geq 0$. La condición (4.0.3), nos dice que $\{X_n\}_{n \geq 0}$ es una cadena de Markov con respecto a la historia $(\mathcal{F}_n \vee \mathcal{F}_{n-1}^Y) = \sigma(\mathcal{F}_n, \mathcal{F}_{n-1}^Y)$. La condición (4.0.4) significa que la probabilidad de obtener 1 en el n -ésimo lanzamiento es igual a $s(X_n)$, que es independiente de la historia previa \mathcal{F}_n de la cadena $\{X_n\}_{n \geq 0}$ y de los $n - 1$ experimentos anteriores. La condición (4.0.5) nos dice que si en el n -ésimo experimento se obtiene como resultado 1, la siguiente transición obedece a la medida de probabilidad $\nu(\cdot)$ independientemente de la historia pasada \mathcal{F}_n de la cadena y de los $n - 1$ experimentos Bernoulli realizados anteriormente. Finalmente

la condición (4.0.6) nos dice que si en el n -ésimo experimento se obtiene como resultado 0, la siguiente transición obedece a la medida $Q(x, \cdot)$ independientemente de la historia pasada \mathcal{F}_n de la cadena y de los $n - 1$ experimentos Bernoulli realizados anteriormente.

A la cadena de Markov bivariada $\{X_n, Y_n : n \geq 0\}$ se le llamará la *separación de la cadena* de $\{X_n\}_{n \geq 0}$.

Ahora se enunciará un Teorema clave de este trabajo, en el cual se hace el enlace entre la cadena de Markov $\{X_n\}_{n \geq 0}$ y la nueva cadena $\{X_n, Y_n\}_{n \geq 0}$.

Teorema 4.2. *La distribución de la cadena $\{X_n\}_{n \geq 0}$ es la distribución marginal de la cadena $\{X_n, Y_n\}_{n \geq 0}$, esto es, para toda distribución inicial $\lambda \in \mathcal{M}_+(\mathcal{E})$ y para toda $A \in \mathcal{E}$, tenemos*

$$\int_E \lambda(dx) P^n(x, A) = \sum_{i=0}^1 \int_{E^*} \lambda^*(dy, i) \check{P}^n((y, i), A_0 \cup A_1),$$

donde $\check{P}^n((y, i), A_0 \cup A_1)$, es la probabilidad de transición de ir de un estado en $x_i \in \mathcal{E}^*$, con $i = 0, 1$ a un conjunto $A = A_0 \cup A_1$.

Demostración. Lo que se quiere demostrar es lo siguiente:

$$\lambda P^n(A) = \lambda^* \check{P}^n(A_0 \cup A_1). \quad \forall n \in \mathbb{N} \tag{4.0.7}$$

Pues integrando sobre el espacio de estados y calculando la distribución marginal de la cadena $\{X_n, Y_n\}_{n \geq 0}$ se obtiene

$$\int_E \lambda(dx) P^n(x, A) = \sum_{i=0}^1 \int_{E^*} \lambda^*(dy, i) \check{P}^n((y, i), A_0 \cup A_1).$$

Probémoslo para el caso $n = 1$.

$$\begin{aligned}
& \sum_{i=0}^1 \int_{E^*} \lambda^*(dy, i) \check{P}((y, i), A_0 \cup A_1) \\
&= \int_{E^*} \lambda^*(dy, 0) \check{P}((y, 0), A_0 \cup A_1) + \int_{E^*} \lambda^*(dy, 1) \check{P}((y, 1), A_0 \cup A_1) \\
&= \int_{E^*} [\lambda I_{1-s}(E)] Q(y, A) + \int_{E^*} [\lambda I_s(E)] \nu(A) \\
&= \int_E \lambda(dy) (1-s(y)) Q(y, A) + \int_E \lambda(dy) s(y) \nu(A) \\
&= \int_E \lambda(dy) (1-s(y)) Q(y, A) + \lambda(dy) s(y) \nu(A) \\
&= \int_E \lambda(dy) (1-s(y)) \frac{P(y, A) - s(y) \nu(A)}{1-s(y)} + \lambda(dy) s(y) \nu(A) \\
&= \int_E \lambda(dy) [P(y, A) - s(y) \nu(A)] + \lambda(dy) s(y) \nu(A) \\
&= \int_E \lambda(dy) P(y, A) \\
&= \lambda P(A).
\end{aligned}$$

Suponemos válido para n .

Por demostrar para el caso $n + 1$.

$$\begin{aligned}
& \sum_{i=0}^1 \int_{E^*} \lambda^*(dy, i) \check{P}^{n+1}((y, i), A_0 \cup A_1) \\
&= \int_{E^*} \lambda^*(dy, 0) \check{P}^{n+1}((y, 0), A_0 \cup A_1) + \int_{E^*} \lambda^*(dy, 1) \check{P}^{n+1}((y, 1), A_0 \cup A_1) \\
&= \int_{E^*} \lambda^*(dy, 0) \int_{E^*} \check{P}((y, 0), dz) P^n(z, A) + \int_{E^*} \lambda^*(dy, 1) \int_{E^*} \check{P}((y, 1), dz) P^n(z, A) \\
&= \int_{E^*} [\lambda I_{1-s}(E)] \int_{E^*} Q(y, dz) P^n(z, A) + \int_{E^*} [\lambda I_s(E)] \int_{E^*} \nu(dz) P^n(z, A) \\
&= \int_E \lambda(dy) (1-s(y)) \int_E Q(y, dz) P^n(z, A) + \int_E \lambda(dy) s(y) \int_E \nu(dz) P^n(z, A) \\
&= \int_E \int_E \lambda(dy) (1-s(y)) \frac{P(y, dz) - s(y) \nu(dz)}{1-s(y)} P^n(z, A) + \lambda(dy) s(y) \nu(dz) P^n(z, A) \\
&= \int_E \int_E \lambda(dy) [P(y, dz) - s(y) \nu(dz)] P^n(z, A) + \lambda(dy) s(y) \nu(dz) P^n(z, A) \\
&= \int_E \int_E \lambda(dy) P(y, dz) P^n(z, A) \\
&= \int_E \lambda(dy) P^{n+1}(y, A) \\
&= \lambda P^{n+1}(A)
\end{aligned}$$

□

Este Teorema es de gran importancia ya que la *cadena separada* $\{X_n, Y_n\}_{n \geq 0}$ le hereda a la cadena original $\{X_n\}_{n \geq 0}$ propiedades importantes con la ventaja de que la cadena separada ya contiene un átomo.

Teorema 4.3. 1. *El proceso estocástico $\{X_n, Y_n : n \geq 0\}$ es una cadena de Markov con respecto a su historia $\{\mathcal{F}_n \vee \mathcal{F}_n^Y : n \geq 0\}$.*

2. *La cadena de Markov $\{X_n, Y_n : n \geq 0\}$ es irreducible.*

3. *El conjunto definido como $\alpha = E \times \{1\}$ es un átomo propio para la cadena $\{X_n, Y_n : n \geq 0\}$.*

4. *La sucesión de renovación*

$$u_0 = 1, \quad u_n = \nu P^{n-1} s, \quad n \geq 1,$$

asociada con el átomo (s, ν) del kernel P es también la sucesión de renovación asociada con el átomo propio α de la cadena de Markov $\{X_n, Y_n\}_{n \geq 0}$.

Demostración.

(1.) De las fórmulas (4.0.1 a la 4.0.6) tenemos que

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in dy, Y_{n+1} = 1 | \mathcal{F}_n \vee \mathcal{F}_n^Y\} &= \mathbb{P}\{X_{n+1} \in dy, Y_{n+1} = 1 | X_n, Y_n\} \\ &= \begin{cases} \nu(dy)s(y) & \text{si } Y_n = 1, \\ Q(x, dy)s(y) & \text{si } Y_n = 0, \end{cases} \end{aligned}$$

y también

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in dy, Y_{n+1} = 0 | \mathcal{F}_n \vee \mathcal{F}_n^Y\} &= \mathbb{P}\{X_{n+1} \in dy, Y_{n+1} = 0 | X_n, Y_n\} \\ &= \begin{cases} \nu(dy)(1 - s(y)) & \text{si } Y_n = 1, \\ Q(x, dy)(1 - s(y)) & \text{si } Y_n = 0. \end{cases} \end{aligned}$$

De estas condiciones $\{X_n, Y_n\}_{n \geq 0}$ es una cadena de Markov con respecto a la historia $(\mathcal{F}_n \vee \mathcal{F}_n^Y)$. También se puede observar que las transiciones que comienzan desde $\alpha = E \times \{1\}$ son

idénticas en distribución.

(2.) Para mostrar que la cadena separada $\{X_n, Y_n\}_{n \geq 0}$ es irreducible, basta con mostrar que

$$\mathbb{P}\{Y_n = 1 \text{ para algún } n \geq 1 \mid X_0 = x, Y_0 = i\} > 0$$

para todo estado inicial $(x, i) \in E \times \{0, 1\}$ de la cadena separada $\{X_n, Y_n\}_{n \geq 0}$.

Sabemos que

$$\begin{aligned} \varphi(s) &= \int_E \varphi(dx) s(x) = \varphi^*(E \times \{1\}) \\ \varphi^*(\alpha) &= \varphi^*(E \times \{1\}) = \varphi I_s(E) = \int_E \varphi(dx) s(x) \end{aligned}$$

entonces

$$\varphi^*(\alpha) = \varphi(s) > 0,$$

por lo tanto α es φ -positivo, además $\alpha \subset E \times \{0, 1\}$.

Recordemos el Lema 2.12, pág 37, que nos dice que si ψ es una medida y $\psi P \ll \psi$, entonces $\psi G \sim \psi$.

Supongamos que

$$\psi^*(\alpha) = \psi^*(E \times \{1\}) = \psi I_s(E) = \int_E \psi(dx) s(x) = 0.$$

Por demostrar que $\psi^* \check{P}(\alpha) = 0$

$$\psi^* \check{P}(\alpha) = \int_{E^*} \psi^*(dx_i) \check{P}(x_i, \alpha).$$

Si $i = 1$,

$$\begin{aligned} \int_{E^*} \psi^*(dx, 1) \int_E \nu(dy) s(y) &= \int_E \psi(dx) s(x) \int_E \nu(dy) s(y) \\ &= 0. \end{aligned}$$

Si $i = 0$

$$\begin{aligned}
 \int_{E^*} \psi^*(dx, 0) \check{P}((x, 0), \alpha) &= \int_{E^*} \psi^*(dx, 0) \int_E Q(x, dy) s(y) \\
 &= \int_{E^*} \psi I_{1-s}(E) \int_E Q(x, dy) s(y) \\
 &= \int_E \int_E \psi(dx) [P(x, dy) - s(x) \nu(dy)] s(y) \\
 &= \int_E \int_E \psi(dx) P(x, dy) - \psi(dx) s(x) \nu(dy) s(y) \\
 &= 0 - 0,
 \end{aligned}$$

como la cadena $\{X_n\}_{n \geq 0}$ es irreducible el primer término es 0 y el segundo es 0 por el caso anterior.

Por lo tanto $\psi^* \check{P} \ll \psi^*$.

Por el Lema 2.12 se tiene que $\psi^* \check{G} \sim \psi^*$.

Es decir, si

$$\psi^* \check{G}(\alpha) = \int_{E^*} \psi^*(dx_i) \check{G}(x_i, \alpha) = \int_{E^*} \psi^*(dx_i) \sum_{n=0}^{\infty} \check{P}^n(x_i, \alpha) = 0,$$

entonces

$$\psi^*(\alpha) = \psi I_s(E) = 0,$$

pero ya probamos que $\psi^*(\alpha) > 0$, entonces

$$\int_{E^*} \sum_{n=0}^{\infty} \psi(dx_i) \check{P}^n(x_i, \alpha) > 0,$$

por lo tanto

$$\mathbb{P}\{Y_n = 1 \text{ para algún } n \geq 1 \mid X_0 = x, Y_0 = i\} > 0.$$

Por lo que podemos decir que la cadena separada $\{X_n, Y_n\}_{n \geq 0}$ es irreducible.

(3.) Como dijimos en (1.), todas las transiciones que comienzan en $\alpha = E \times \{1\}$ son idénticas en distribución, es decir, cumplen con

$$\check{P}(x_1, \cdot) = \check{P}(y_1, \cdot) \quad \forall x_1, y_1 \in \alpha.$$

Ahora, para mostrar que $x_1 \rightarrow \alpha$ para algún $x_1 \in \alpha$ usemos (4.0.3 a la 4.0.6), de donde se obtiene

$$\begin{aligned}\mathbb{P}\{Y_1 = 1 \mid X_0 = x, Y_0 = 1\} &= \int_E \nu(dy) s(y) = \int_E \nu(dy) P^0 s(y) \\ \mathbb{P}\{Y_2 = 1 \mid X_0 = x, Y_0 = 1\} &= \int_E \nu(dy) P s(y) \\ \mathbb{P}\{Y_n = 1 \mid X_0 = x, Y_0 = 1\} &= \int_E \nu(dy) P^{n-1} s(y), \quad \forall n \geq 1.\end{aligned}\tag{4.0.8}$$

Por irreducibilidad $\nu P^{n-1} s > 0$ para algún $n > 1$, entonces $x_1 \rightarrow \alpha$ para algún $x_1 \in \alpha$. Por lo tanto α es un *átomo propio*.

De manera similar se tiene que

$$\begin{aligned}\mathbb{P}\{Y_1 = 1 \mid X_0 = x, Y_0 = 0\} &= \int_E Q(x, dy) s(y) = \int_E Q(x, dy) P^0 s(y) \\ \mathbb{P}\{Y_2 = 1 \mid X_0 = x, Y_0 = 0\} &= \int_E Q(x, dy) P s(y) \\ \mathbb{P}\{Y_n = 1 \mid X_0 = x, Y_0 = 0\} &= \int_E Q(x, dy) P^{n-1} s(y) > 0, \quad \text{para algún } n \geq 1.\end{aligned}\tag{4.0.9}$$

(4.) De (4.0.8) se dedujo que

$$\mathbb{P}\{Y_n = 1 \mid X_0 = x, Y_0 = 1\} = \nu P^{n-1} s,$$

donde

$$u_0 = \mathbb{P}\{Y_0 = 1 \mid X_0 = x, Y_0 = 1\} = 1 \quad y \quad u_n = \nu P^{n-1} s.$$

donde u_n es la sucesión de renovación asociada con el átomo propio α de la cadena de Markov separada $\{X_n, Y_n\}_{n \geq 0}$.

□

Se denota por \mathbb{P}_α a la medida de probabilidad definida en la σ -álgebra $\sigma(X_n, n \geq 1 : Y_n, n \geq 0)$ correspondiente al estado inicial $Y_0 = 1$ y $X_0 = x$, donde x es un estado arbitrario, es decir

$$\mathbb{P}_\alpha = \mathcal{L}(X_n, n \geq 1 : Y_n, n \geq 0 \mid Y_0 = 1).$$

Para diferenciar se usa el símbolo P para denotar una probabilidad de transición de la cadena separada $\{X_n, Y_n\}_{n \geq 0}$ y el símbolo \mathbb{P} para una medida de probabilidad, por ejemplo

$$\begin{aligned} \check{P}^n(\alpha, A) &= \mathbb{P}_\alpha\{X_n \in A\} = \nu P^{n-1}(A) = \int_A P^n(x, dy) s(y), \quad A \in \mathcal{E}_\Delta, n \geq 1, \\ \check{P}^n(x, \alpha) &= \mathbb{P}_x\{Y_n = 1\} = \mathbb{E}_x s(X_n) = P^n s(x), \quad x \in E, n \geq 0, \\ \check{P}^n(\alpha, \alpha) &= \mathbb{P}_\alpha\{Y_n = 1\} = u_n = \begin{cases} 1 & \text{si } n = 0, \\ \nu P^{n-1} s & \text{si } n \geq 1. \end{cases} \end{aligned} \quad (4.0.10)$$

Ahora se definen tiempos de paro para la cadena separada $\{X_n, Y_n\}_{n \geq 0}$, de una manera similar al Capítulo 2. T_α, S_α que son la primer visita y primer regreso al átomo propio α , así como $T_\alpha(i)$ que es la sucesión de retornos a α .

$$\begin{aligned} T_\alpha &= T_\alpha(0) = \inf\{n \geq 0 : Y_n = 1\}, \\ S_\alpha &= \inf\{n \geq 1 : Y_n = 1\}, \end{aligned}$$

e inductivamente para $i \geq 1$

$$T_\alpha(i) = \inf\{n > T_\alpha(i-1) : Y_n = 1\}.$$

Entonces la sucesión $\{T_\alpha(i); i \geq 0\}$ es el proceso de renovación embebido asociado con el átomo propio α de la cadena separada. Su distribución de incrementos $b = \{b_n; n \geq 1\}$ está dada por

$$b_n = \mathbb{P}_\alpha\{S_\alpha = n\} = \nu(P - s \otimes \nu)^{n-1} s, \quad n \geq 1, \quad (4.0.11)$$

y la distribución retraso $a(x) = \{a_n(x); n \geq 0\}$ que corresponde a $X_0 = x$, está dada por

$$a_n(x) = \mathbb{P}_x\{T_\alpha = n\} = (P - s \otimes \nu)^n s(x), \quad n \geq 0. \quad (4.0.12)$$

Ahora por el Teorema 3.18 (2), pág 79, cuando $K = P$ puede ser interpretado como la *descomposición de primera entrada*, (también vista en la fórmula (3.3.1), pág 78), entonces se tiene

$$\begin{aligned} P^n s(x) &= \mathbb{P}_x\{Y_n = 1\} \\ &= \sum_{m=0}^n \mathbb{P}_x\{T_\alpha = m\} \mathbb{P}_\alpha\{Y_{n-m} = 1\} \\ &= a(x) * u_n, \quad n \geq 0, \end{aligned}$$

De manera similar se tiene

$$\begin{aligned}\sigma_n(A) &= \nu(P - s \otimes \nu)^n(A) \\ &= \mathbb{P}_\alpha\{X_{n+1} \in A, S_\alpha \geq n + 1\}, \quad n \geq 0,\end{aligned}\tag{4.0.13}$$

y

$$L_\alpha(n) = \text{máx}\{m : Y_m = 1, 0 \leq m \leq n\}, \quad n \geq 0.$$

También se puede interpretar la parte (3) del Teorema 3.18, pág 79, como la *última salida de descomposición*, es decir

$$\begin{aligned}\nu P^{n-1}(A) &= \mathbb{P}_\alpha\{X_n \in A\} \\ &= \sum_{m=0}^{n-1} \mathbb{P}_\alpha\{L_\alpha(n-1) = m, X_n \in A\} \\ &= \sum_{m=0}^{n-1} \mathbb{P}_\alpha\{Y_m = 1\} \mathbb{P}_\alpha\{X_{n-m} \in A, S_\alpha \geq n - m\} \\ &= u * \sigma(A)_{n-1}, \quad n \geq 1.\end{aligned}\tag{4.0.14}$$

Y la parte (4) del mismo Teorema como la *descomposición de la primera entrada y última salida*

$$\begin{aligned}P^n(x, A) &= \mathbb{P}_x\{X_n \in A\} \\ &= \mathbb{P}_x\{X_n \in A, T_\alpha \geq n\} + \sum_{m=0}^{n-1} \sum_{p=m}^{n-1} \mathbb{P}_x\{T_\alpha = m, L_\alpha(n-1) = p, X_n \in A\} \\ &= (P - s \otimes \nu)^n(x, A) + a(x) * u * \sigma(A)_{n-1}, \quad n \geq 1.\end{aligned}\tag{4.0.15}$$

Sea $G_\alpha = G_{s,\nu}^{(1)} = \sum_{n=0}^{\infty} (P - s \otimes \nu)^n$, entonces por (4.0.10) y (4.0.11)

$$\begin{aligned}\sum_{n=0}^{\infty} u_n &= 1 + \nu G s \leq \infty, \\ \sum_{n=1}^{\infty} b_n &= \mathbb{P}_\alpha\{S_\alpha < \infty\} = \nu G_\alpha s \leq 1.\end{aligned}$$

Ahora se supondremos que la cadena de Markov $\{X_n\}_{n \geq 0}$ es Harris-recurrente, y demostraremos que la cadena separada $\{X_n, Y_n\}_{n \geq 0}$ es Harris-recurrente; además que α es un átomo recurrente.

Teorema 4.4. *Sea $\{X_n\}_{n \geq 0}$ una cadena de Markov Harris-recurrente, entonces*

1. *La cadena separada $\{X_n, Y_n\}_{n \geq 0}$ es Harris-Recurrente.*

2. El conjunto $\alpha = E \times \{1\}$ es un átomo recurrente.

Demostración.

(1.) Por la definición de la cadena separada $\{X_n, Y_n\}_{n \geq 0}$, y de que $\{X_n\}_{n \geq 0}$ es Harris-recurrente, se sabe que para toda $z_i \in E^* = E \times \{0, 1\}$, $A \in \mathcal{E}$, con $\varphi(A) > 0$ se tiene

$$\mathbb{P}_{z_i}(S_A < \infty) = \check{P}(z_i, A) + \int_{A^c} \check{P}(z_i, dy) \mathbb{P}_y(S_A < \infty) = 1. \quad (4.0.16)$$

Fijemos $z_i \in E^*$ y $A \in \mathcal{E}$ con $\varphi(A_1) = \varphi I_s(A) > 0$, entonces existe $\beta > 0$ y $C \subset A$, tal que $\varphi(C) > 0$ y $s \geq \beta$ en C .

Sea S_C^n el instante de la n -ésima visita de la cadena $\{X_n\}_{n \geq 0}$ al conjunto C y se define para $n \geq 1$

$$Z_n = Y_{S_C^n} 1_{\{S_C^n < \infty\}}.$$

Entonces

$$\mathbb{P}_{z_i}\{Z_n = 1 \mid Z_1, Z_2, \dots, Z_{n-1}\} \geq \beta$$

por (4.0.16)

$$\mathbb{P}_{z_i}\{S_{C_1} < \infty\} = \mathbb{P}_{z_i}\{Z_n = 1 \text{ para algún } n\} = 1,$$

como la cadena $\{X_n, Y_n\}_{n \geq 0}$ es irreducible entonces $\varphi(C_1) = \varphi I_s(C) > 0$, por lo tanto

$$\mathbb{P}_{z_i}\{S_{A_1} < \infty\} = 1.$$

De manera análoga $\mathbb{P}_{z_i}(S_{A_0} < \infty) = 1$ para toda $z_i \in E^*$ y $A \in \mathcal{E}$ con $\varphi(A_0) > 0$.

(2.) Se sabe que $\varphi^*(\alpha) = \varphi(s) > 0$, además $\mathbb{P}_{z_i}\{S_{A_1} < \infty\} = 1$, con $A_1 \in \alpha$, $z \in E \times \{0, 1\}$ y $\varphi(A_1) > 0$ y también que la cadena es irreducible, entonces

$$(z, 1) \rightarrow A \times \{1\}, \quad \forall (z, 1) \in E \times \{1\} = \alpha.$$

Por lo tanto

$$\mathbb{P}_{z_1}\{S_\alpha < \infty\} = 1,$$

es decir α es un átomo recurrente.

□

Conclusiones

A lo largo de este trabajo vimos el comportamiento de algunas propiedades de las cadenas de Markov en un espacio de estados discreto, con la finalidad de poder compararlas con una cadena de Markov definida en un espacio de estados general; donde se lograron ver similitudes al momento de manipular dichas cadenas con la técnica de separación.

Por otra parte, pudimos demostrar que es posible aplicar la técnica de separación, obteniendo una cadena bivariada a partir de una cadena definida en un espacio de estados general; como resultado se observó que la distribución marginal de la cadena separada es idéntica en distribución a la cadena original. Por lo tanto, las propiedades de la cadena separada son automáticamente heredadas a la cadena original y esto es válido para toda cadena de Markov irreducible definida en un espacio de estados general, por ende se cumple para toda cadena recurrente o Harris-Recurrente.

Finalmente, se logró el segundo objetivo; donde usando la teoría de renovación se pudo contar el número de regresos que realiza la cadena separada al átomo y observamos que la sucesión de renovación para la cadena original es idéntica a la sucesión de la cadena separada.

Esta técnica es muy útil en la teoría de probabilidad ya que existen problemas en donde se puede utilizar esta técnica y resulte más fácil su estudio. Un ejemplo de ello es una prueba constructiva para la existencia de medidas invariantes; otro más se puede ver en el estudio de

convergencia de sumas de transiciones de probabilidad [5, Pág. 313-317].

Apéndice

Teorema de convergencia monótona

Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias no negativas. Si $X_n \nearrow X$ cuando $n \rightarrow \infty$, entonces

$$E[X_n] \nearrow E[X], \quad \text{cuando } n \rightarrow \infty.$$

Demostración. Primero definamos una sucesión de variables aleatorias $\{Y_{k,n}, n \geq 1\}$ para toda k , tales que

$$Y_{k,n} \nearrow X_k, \quad \text{cuando } n \rightarrow \infty.$$

tales sucesiones existen por definición y consistencia.

Sea

$$Z_n = \max_{1 \leq k \leq n} Y_{k,n}, \quad n \geq 1.$$

Por construcción y dominación

$$Y_{k,n} \leq Z_n \leq X_n \quad y \quad E[Y_{k,n}] \leq E[Z_n] \leq E[X_n]. \quad (4.0.17)$$

Cuando $n \rightarrow \infty$,

$$X_k \leq \lim_{n \rightarrow \infty} Z_n \leq \lim_{n \rightarrow \infty} X_n = X$$

y haciendo $k \rightarrow \infty$

$$X \leq \lim_{n \rightarrow \infty} Z_n \leq X,$$

por lo que

$$\lim_{n \rightarrow \infty} E[Z_n] = E[X] = E[\lim_{n \rightarrow \infty} Z_n], \quad (4.0.18)$$

donde la primera igualdad se obtiene por definición (y consistencia) y la segunda por equivalencia [2, Teorema 4.4.(g), Pág 52], que nos dice que si X, Y son variables aleatorias no negativas y $X = Y$, c.s., entonces $E[X] = E[Y]$.

El mismo procedimiento en la desigualdad entre las esperanzas en (4.0.17), nos lleva a

$$E[X_k] \leq \lim_{n \rightarrow \infty} E[Z_n] \leq \lim_{n \rightarrow \infty} E[X_n],$$

y entonces

$$\lim_{n \rightarrow \infty} E[X_k] \leq \lim_{n \rightarrow \infty} E[Z_n] \leq \lim_{n \rightarrow \infty} E[X_n].$$

Combinando la última ecuación con (4.0.18) finalmente mostramos que

$$\lim_{n \rightarrow \infty} E[X_n] = \lim_{n \rightarrow \infty} E[Z_n] = E[\lim_{n \rightarrow \infty} Z_n] = E[X].$$

□

Teorema 4.5. Sean X y Y variables aleatorias no negativas y $Y \leq X$ c.s., entonces

$$E[Y] \leq E[X].$$

Demostración. Ver [2, Teorema 4.4.(h), pág 52].

□

Lema de Fatou

1. Si $\{X_n, n \geq 1\}$ es una sucesión de variables aleatorias no negativas, entonces

$$E \left[\liminf_{n \rightarrow \infty} X_n \right] \leq \liminf_{n \rightarrow \infty} E[X_n].$$

2. Si, Y y Z son variables aleatorias integrables, tal que $Y \leq X_n \leq Z$ casi seguramente para toda n , entonces

$$E \left[\liminf_{n \rightarrow \infty} X_n \right] \leq \liminf_{n \rightarrow \infty} E[X_n] \leq \limsup_{n \rightarrow \infty} E[X_n] \leq E \left[\limsup_{n \rightarrow \infty} X_n \right].$$

Demostración.

(1.) Sea $Y_n = \inf_{k \geq n} X_k$, $n \geq 1$. Donde

$$Y_n = \inf_{k \geq n} X_k \nearrow \liminf_{n \rightarrow \infty} X_n, \quad \text{cuando } n \rightarrow \infty,$$

el Teorema de convergencia monótona nos lleva

$$E[Y_n] \nearrow E[\liminf_{n \rightarrow \infty} X_n].$$

Donde, además $Y_n \leq X_n$, y por el Teorema 4.5, tenemos que

$$E[Y_n] \leq E[X_n], \quad \forall n.$$

(2.) Notemos que

$$\liminf_{n \rightarrow \infty} (X_n - Y) = \liminf_{n \rightarrow \infty} X_n - Y \quad \text{y} \quad \liminf_{n \rightarrow \infty} (Z - X_n) = Z - \liminf_{n \rightarrow \infty} X_n,$$

por lo que (2.) se sigue de (1.) y de la aditividad, donde $\{X_n - Y, n \geq 1\}$ y $\{Z - X_n, n \geq 1\}$ son variables aleatorias no negativas.

□

Teorema de convergencia dominada

Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias no negativas tales que $|X_n| \leq Y$, para toda n , donde $E[Y] < \infty$, y que convergen a una variable aleatoria X , casi seguramente cuando $n \rightarrow \infty$. Entonces

$$E|X_n - X| \rightarrow 0 \quad \text{cuando } n \rightarrow \infty,$$

en particular,

$$E[X_n] \rightarrow E[X] \quad \text{cuando } n \rightarrow \infty.$$

□

Teorema de convergencia acotada

Sean $a(x)$, $x \in E$ números no negativos que tienen suma finita, y sea $\{b_n(x)\}_{n \in \mathbb{N}}$ una sucesión de reales, $x \in E$ y $n \geq 1$, tal que $|b_n(x)| \leq 1$, $x \in E$ y $n \geq 1$, y

$$\lim_{n \rightarrow \infty} b_n(x) = b(x), \quad x \in E.$$

Entonces

$$\lim_{n \rightarrow \infty} \sum_x a(x)b_n(x) = \sum_x a(x)b(x).$$

□

Ecuación de Wald

Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias no negativas independientes idénticamente distribuidas y con esperanza finita. Sea N una variable aleatoria con valores en el conjunto $\{1, 2, \dots\}$ con esperanza finita e independiente de la sucesión. Entonces

$$E \left[\sum_{k=1}^N X_k \right] = E[X]E[N].$$

Demostración.

$$E \left[\sum_{k=1}^N X_k \right] = E \left[E \left[\sum_{i=1}^N X_i / N \right] \right]$$

pero

$$\begin{aligned} E \left[\sum_{i=1}^N X_i / N = n \right] &= E \left[\sum_{i=1}^n X_i / N = n \right] \\ &= E \left[\sum_{i=1}^n X_i \right] \\ &= nE[X] \quad (\text{por independencia de } X_i \text{ y } N) \end{aligned}$$

donde $E[X] = E[X_i]$

lo cual cumple que

$$E \left[\sum_{i=1}^N X_i / N \right] = NE[X]$$

entonces

$$E \left[\sum_{i=1}^N X_i \right] = E[N E[X]] = E[N]E[X].$$

□

Teorema de las tres series de Kolmogorov

Sea $A > 0$. Suponga que X_1, X_2, \dots son variables independientes y sea para $k \geq 1$,

$$Y_k = \begin{cases} X_k & \text{cuando } |X_k| \leq A, \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces

$$\sum_{k=1}^{\infty} X_k \text{ converge c.s., cuando } n \rightarrow \infty,$$

si y sólo si

1. $\sum_{k=1}^{\infty} \mathbb{P}(X_k \neq Y_k) = \sum_{k=1}^{\infty} \mathbb{P}(|X_k| > A) < \infty$;
2. $\sum_{k=1}^{\infty} E[Y_k]$ converge;
3. $\sum_{k=1}^{\infty} \text{var}[Y_k] < \infty$.

Demostración. Ver [2, Teorema 5.5. Pág. 289].

□

Proposición 4.6. *Suponga que X, X_1, X_2, \dots son variables aleatorias independientes e idénticamente distribuidas, tales que $E|X|^r < \infty$ para alguna $r > 0$, y sea*

$$Y_n = X_n I\{|X_n| \leq n^{1/r}\}, \quad n \geq 1.$$

Entonces X_1, X_2, \dots y Y_1, Y_2, \dots son equivalentes en convergencia

Demostración Ver [2, Proposición 1.2. Pág. 267].

□

Lema 4.7. Sean $0 < r < 2$ y $\{X_i\}_{i \geq 1}$ v.a.i.i.d. Definimos $Y_n = X_n I\{|X_n| \leq n^{1/r}\}$ para $n \geq 1$.

Si $E[|X|^r] < \infty$, entonces

$$\sum_{n=1}^{\infty} \text{var} \left[\frac{Y_n}{n^{1/r}} \right] = \sum_{n=1}^{\infty} \frac{\text{var}[Y_n]}{n^{2/r}} < \infty.$$

Demostración. Ver [2, Lema 6.1, pág. 294].

□

Lema 4.8. Si $\{X_i\}_{i \geq 1}$ son v.a., sea $a_0 = 0$, y sea $\{a_n\}_{n \geq 1}$, una sucesión creciente a infinito de números positivos. Entonces

$$\sum_{k=1}^n \frac{X_k}{a_k} \text{ converge c.s.} \Rightarrow \frac{1}{a_n} \sum_{k=1}^n X_k \rightarrow 0 \text{ cuando } n \rightarrow \infty$$

Demostración. Ver [2, Lema 5.1. Pág. 284].

□

Lema 4.9. Suponga que $a_n \in \mathbb{R}$, $n \geq 1$. Si $a_n \rightarrow a$ cuando $n \rightarrow \infty$, entonces

$$\frac{1}{n} \sum_{k=1}^n a_k \rightarrow a, \quad \text{cuando } n \rightarrow \infty.$$

Demostración. Ver [2, Lema A.6.1, pág 564].

□

Lema 4.10. (Lema de Kronecker). Supongamos que $\{X_n\}_{n \geq 1}$ es una sucesión de variables aleatorias, sea $a_0 = 0$ y $\{a_n\}_{n \geq 1}$ una sucesión creciente a $+\infty$ de números positivos. Entonces si

$$\sum_{k=1}^{\infty} \frac{X_k}{a_k} \text{ converge c.s.} \Rightarrow \frac{1}{a_n} \sum_{k=1}^n X_k \rightarrow 0, \text{ c.s., cuando } n \rightarrow \infty.$$

Demostración. Ver [2, Lema de Kronecker, pág. 284].

□

Teorema 4.11. (Criterio de convergencia de Kolmogorov). Sean X_1, X_2, \dots variables aleatorias con sumas parciales S_n , $n \geq 1$. Entonces

$$\sum_{n=1}^{\infty} \text{Var } X_n < \infty \Rightarrow \sum_{n=1}^{\infty} (X_n - E[X_n]) \text{ converge c.s.}$$

Demostración. Ver [2, Teorema 5.2. Pág. 286]

□

Ley fuerte de los grandes números

Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas. Si estas variables tienen media finita μ , entonces

$$\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu.$$

Demostración. Sea

$$Y_n = X_n I\{|X_n| \leq n\} = \begin{cases} X_n & \text{si } |X_n| \leq n, \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces por la Proposición 4.6, pág. 103, X_1, X_2, \dots y Y_1, Y_2, \dots son equivalentes en convergencia, es decir $\sum_{n=1}^{\infty} \mathbb{P}(X_n \neq Y_n) < \infty$. Entonces

$$\sum_{n=1}^{\infty} \mathbb{P} \left[\left| \frac{X_n}{n} \right| > 1 \right] < \infty,$$

pues debido al Teorema de las tres series de Kolmogorov (1.) es convergente (con $A = 1$).

Ahora, por el Lema 4.7 (con $r = 1$) se tiene que

$$\sum_{n=1}^{\infty} \text{var} \left[\frac{Y_n}{n} \right] < \infty,$$

tal que la tercera suma en el Teorema de las tres series de Kolmogorov es convergente. Por el Teorema 4.11, pág 104 se tiene que

$$\sum_{n=1}^{\infty} \frac{Y_n - E[Y_n]}{n} \text{ converge c.s.,}$$

tal que por el Lema 4.8, se tiene

$$\frac{1}{n} \sum_{k=1}^n (Y_k - E[Y_k]) \rightarrow 0, \quad \text{c.s.}$$

Finalmente,

$$E[Y_n] = E[X_n I\{|X_n| \leq n\}] = E[XI\{|X| \leq n\}] \rightarrow E[X] = \mu \quad \text{cuando } n \rightarrow \infty,$$

tal que por el Lema 4.9 se tiene que

$$\frac{1}{n} \sum_{k=1}^n E[Y_k] \rightarrow \mu \quad \text{cuando } n \rightarrow \infty,$$

lo cual implica que

$$\frac{1}{n} \sum_{k=1}^n Y_k \rightarrow \mu \quad \text{c.s. cuando } n \rightarrow \infty.$$

debido a la equivalencia en convergencia, el Teorema queda demostrado. □

Corolario 4.12. *Sea $I =: [a, b]$ y sea $f : I \rightarrow \mathbb{R}$ integrable en I , entonces, la función valor absoluto $|f|$ es integrable en I , y*

$$\left| \int_a^b f \right| \leq \int_a^b |f| \leq K(b-a),$$

donde $|f(x)| \leq K$ para toda $x \in I$.

Demostración. Ver [1, Corolario 7.2.6 (a), pág. 271] □

Proposición 4.13. *Sea:*

$$\psi = \sum_{n=0}^{\infty} 2^{-(n+1)} \tilde{\varphi} \tilde{K}^n,$$

donde $\tilde{\varphi}$ es una medida de probabilidad equivalente a φ y \tilde{K} es un kernel subestocástico equivalente a K . Entonces ψ es una medida finita que cumple $\psi \ll \psi K$.

Demostración. Primero veamos que ψ es una medida. Sabemos que $\tilde{K}(x, \cdot)$, para toda $x \in E$ es una medida,

$$\begin{aligned} \psi(\emptyset) &= \sum_{n=0}^{\infty} 2^{-(n+1)} \tilde{\varphi} \tilde{K}^n(\emptyset) \\ &= \sum_{n=0}^{\infty} 2^{-(n+1)} \int_E \tilde{\varphi}(dx) \tilde{K}^n(x, \emptyset) \\ &= 0, \end{aligned}$$

ya que $\tilde{K}^n(x, \emptyset) = 0$.

Ahora, sea $A_i \in \mathcal{E}$, $i \in \mathbb{N}$, con $A_i \cap A_j = \emptyset$, $i \neq j$, entonces

$$\begin{aligned}
 \psi(\cup_{i=0}^{\infty} A_i) &= \sum_{n=0}^{\infty} 2^{-(n+1)} \tilde{\varphi} \tilde{K}^n(\cup_{i=0}^{\infty} A_i) \\
 &= \sum_{n=0}^{\infty} 2^{-(n+1)} \int_E \tilde{\varphi}(dx) \tilde{K}^n(x, \cup_{i=0}^{\infty} A_i) \\
 &= \sum_{n=0}^{\infty} 2^{-(n+1)} \int_E \sum_{i=0}^{\infty} \tilde{\varphi}(dx) \tilde{K}^n(x, A_i) \\
 &= \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} 2^{-(n+1)} \int_E \tilde{\varphi}(dx) \tilde{K}^n(x, A_i) \quad (\text{por Teo. conver. monótona}) \\
 &= \sum_{i=0}^{\infty} \psi(A_i).
 \end{aligned}$$

La medida ψ es finita ya que cada término de ψ es menor a 1.

Ahora, sea \tilde{K} un kernel subestocástico equivalente al kernel K . Si $\psi(A) = 0$,

$$\begin{aligned}
 \psi \tilde{K}(A) &= \int_E \psi(dx) \tilde{K}(x, A) \\
 &= \int_E \left[\sum_{n=0}^{\infty} 2^{-(n+1)} \tilde{\varphi} \tilde{K}^n \right] (dx) \tilde{K}(x, A) \\
 &= \int_E \sum_{n=0}^{\infty} 2^{-(n+1)} \int_E \tilde{\varphi}(dy) \tilde{K}^n(y, dx) \tilde{K}(x, A) \\
 &= \sum_{n=0}^{\infty} 2^{-(n+1)} \int_E \tilde{\varphi}(dy) \tilde{K}^{n+1}(y, A) \quad (\text{por Teo. Conv. monótona}) \\
 &= 0,
 \end{aligned}$$

por lo tanto $\psi \tilde{K} \ll \psi$.

□

Bibliografía

- [1] Bartle, R. (2005). *Introducción al análisis matemático de una variable*. Limusa Wiley.
- [2] Gut, A. (2005). *Probability: A Graduate Course*. Springer.
- [3] Hoel, P. G. (1972). *Introduction to stochastic Processes*. University of California, Los Angeles.
- [4] Mein, S. P. (1996). *Markov Chains and Stochastic Stability*. Springer verlag.
- [5] Nummelin, E. (1978). *A Splitting Technique for Harris Recurrent Markov chains*. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 43, 309-318.
- [6] Nummelin, E. (1984). *General irreducible Markov chains and non-negative operators*. Cambridge university press.
- [7] Pap, E. (2002). *Handbook of Measure Theory, Volume I*. Elsevier.
- [8] Resnick, S. (1994). *Adventures in Stochastic Processes*. Birkäuser, Boston.
- [9] Revuz, D. (1975). *Markov Chains*. Amsterdam: North Holland.
- [10] Chen, X. (1999). *Limits Theorems for Functionals of Ergodic Markov Chain with General State Space*. Memoirs of the American Mathematical Society no. 664.