UNIVERSITAT DE BARCELONA

Facultat de Matemàtiques i Informàtica

GRADO DE MATEMÁTICAS

Trabajo fin de grado

Cadena de Markov y el problema de la ruina del jugador

Autor: Pablo Ji

Director: Dra. Maria Carmen Florit Selma

Realizado en: Departamento de Matemáticas e Informática

Barcelona, 13 de junio de 2023

Abstract

In this project we will study the Markov chains. First, we will see all importants definitions and theorems; then we will show it's applications, mainly the Gambler's ruin.

Abstracto

En este proyecto estudiaremos las cadenas de Markov en el tiempo discreto. Primero veremos todas las definiciones y teoremas importantes; luego mostraremos algunas aplicaciones de las cadenas de Markov, principalmente sobre el problema de la ruina del jugador.

2020 Mathematics Subject Classification. 60G05, 60G07, 60J10

Índice

1. Preliminares

1.1. Probabilidad

Antes de estudiar las cadenas de Markov veremos algunos conceptos básicos y propiedades sobre la probabilidad, que nos ayuda a tener una mejor comprensión de las definiciones, teoremas y demostraciones que veremos posteriormente.

Definición 1.1. *Un espacio de probabilidad está compuesto por* $(\Omega, \mathcal{A}, \mathbb{P})$.

Ω *es el espacio mostral, un conjunto de todos los resultados posibles de una experiencia aleatoria.*

$\mathcal{A}$ *es una familia de partes de* Ω *que tiene estructura de* σ*-algebra, es decir:*

1. $\Omega \in \mathcal{A}$
2. *Si* $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$, *donde* A^c *es el conjunto complementario de* A.
3. *Si* $\{A_n, n \geq 1\} \subset \mathcal{A} \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{A}$

$\mathbb{P}$ *es la probabilidad. Es una aplicación* $\mathbb{P} : \mathcal{A} \longrightarrow [0,1]$ *tales que:*

1. $\mathbb{P}(\Omega) = 1$.
2. *(*$\sigma -$ *additividad) Si* $\{A_n, n \geq 1\}$ *es una sucesión de conjuntos de* $\mathcal{A}$ *disjuntos dos a dos, entonces*
$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) \tag{1.1}$$

Durante todo el trabajo, vamos a suponer que estaremos en el espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$.

Propiedades 1.2. *Las propiedades de la probabilidad.*

1. $\mathbb{P}(\emptyset)$ *= 0.*
2. *La propiedad* $\sigma -$ *additividad implica la additividad finita.*
3. *Para todo* $A \in \mathcal{A}$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
4. *Si* $A, B \in \mathcal{A}$ *y* $A \subset B$, *entonces* $\mathbb{P}(A) \leq \mathbb{P}(B)$.
5. *Sean* $A, B \in \mathcal{A}$, *entonces*
$$\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) \tag{1.2}$$
6. *Toda las probabilidades* P *son subadditivas, es decir,*
$$\mathbb{P}(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \mathbb{P}(A_i) \tag{1.3}$$

7. Sean $\{A_i\}_{1\le i\le n} \subset \mathcal{A}$ *entonces,*

$$\mathbb{P}\big(\cup_{i=1}^n A_i\big) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i<j,1\le i,j\le n} \mathbb{P}(A_i\cap A_j) + \tag{1.4}$$

$$\sum_{i<j<k,1\le i,j,k\le n} \mathbb{P}(A_i\cap A_j\cap A_k) - \cdots + (-1)^{n-1}\mathbb{P}\big(\cap_{i=1}^n A_i\big).$$

Definición 1.3. *(Probabilidad condicionada) Sea* $(\Omega,\mathcal{A},\mathbb{P})$ *un espacio de probabilidad asociado a un experimento aleatorio. Sea* $B\in\mathcal{A}$ *no nulo.*
La probabilidad de un conjunto $A\in\mathcal{A}$ *condicionado por* B *es*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} \tag{1.5}$$

Proposición 1.4. *(Fórmula de las probabilidades compuestas) Sean* $A_1,\dots,A_n$ *elementos de* $\mathcal{A}$ *tales que* $\mathbb{P}(A_1\cap\cdots\cap A_n) > 0$*, entonces,*

$$\mathbb{P}(A_1\cap\cdots\cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1\cap A_2)\cdots\mathbb{P}(A_n|A_1\cap\cdots\cap A_{n-1}) \tag{1.6}$$

Proposición 1.5. *(Fórmula de Bayes) Sean* $\mathcal{P}_1 = \{A_1,\dots,A_n\}$ *y* $\mathcal{P}_2 = \{B_1,\dots,B_m\}$ *particiones de* Ω *y* $\{A_i\}_{1\le i\le n}, \{B_j\}_{1\le j\le m}$*, conjuntos de* $\mathcal{A}$ *de probabilidad no nula. Entonces,*

$$\mathbb{P}(A_i|B_j) = \frac{\mathbb{P}(A_i)\mathbb{P}(B_j|A_i)}{\mathbb{P}(B_j)} = \frac{\mathbb{P}(A_i)\mathbb{P}(B_j|A_i)}{\sum_{k=1}^n \mathbb{P}(B_j|A_k)\mathbb{P}(A_k)} \tag{1.7}$$

Definición 1.6. $A,B\in\mathcal{A}$ *son independientes si*

$$\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B) \tag{1.8}$$

1.2. Relación de recurrencia

En esta sección veremos dos tipos de recurrencias lineales que nos será útil en el desarrollo de los cálculos que se verá posteriormente.

La primera que veremos es una recurrencia lineal de primer grado,

$$x_{n+1} = ax_n + b.$$

Suponemos que la recurrencia tiene una solución constante, $x_n = x$. entonces

$$x = ax + b \Rightarrow x = \frac{b}{1-a}, \text{ si } a \neq 1$$

Consideremos

$$y_n = x_n - \frac{b}{1-a} \tag{1.9}$$

que satisface

$$y_{n+1} = ay_n$$

ya que,

$$\begin{cases} y_{n+1} = ay_n = a(x_n - \frac{b}{1-a}) = ax_n - \frac{ab}{1-a} \\ y_{n+1} = x_{n+1} - \frac{b}{1-a} = ax_n + b - \frac{b}{1-a} = ax_n - \frac{ab}{1-a} \end{cases}$$

por tanto,

$$y_{n+1} = ay_n = a^2 y_{n-1} = \cdots = a^n y_0.$$

Teniendo esto en cuenta, lo aplicamos en (1.9) y obtenemos una solución general

$$x_n = Aa^n + \frac{b}{1-a}, \text{ si } a \neq 1.$$

donde $A = y_0$ es un constante. Si $a = 1$, tenemos

$$x_n = x_{n-1} + b = x_{n-2} + 2b = \cdots = x_0 + nb.$$

La otra recurrencia que veremos son las relaciones de recurrencias de segundo grado,

$$ax_{n+1} + bx_n + cx_{n-1} = 0, \text{ donde } a, b, c \neq 0$$

Supongamos que la solución se del tipo $x_n = \lambda^n$, entonces

$$a\lambda^{n+1} + b\lambda^n + c\lambda^{n-1} = 0 \Rightarrow a\lambda^2 + b\lambda + c = 0.$$

Supongamos que α y β son las soluciones de la ecuación de segundo grado. Entonces

$$x_n = A\alpha^n + B\beta^n$$

es una solución de la recurrencia. Si $\alpha \neq \beta$, entonces A, B son las soluciones del sistema de ecuación

$$\begin{cases} x_0 = A + B \\ x_1 = A\alpha + B\beta \end{cases}$$

Si $\alpha = \beta \neq 0$, entonces

$$x_n = (A + nB)\alpha^n$$

donde A, B son las soluciones del sistema de ecuación

$$\begin{cases} x_0 = A\alpha^n \\ x_1 = (A + B)\alpha^n \end{cases}$$

Por tanto, la solución general de la relación de recurrencia es

$$x_n = \begin{cases} A\alpha^n + B\beta^n & \text{si } \alpha \neq \beta \\ (A + nB)\alpha^n & \text{si } \alpha = \beta \end{cases} \tag{1.10}$$

2. Introducción

En la teoría de la probabilidad, la cadena de Markov es un tipo proceso estocástico donde los elementos (estados) de la cadena son numerables o finitos, y la probabilidad de que suceda un evento depende únicamente del evento anterior. Estos eventos pueden transcurrir a tiempo discreto o continuo.

El trabajo está estructurado de la siguiente manera:

- Primero veremos las definiciones, notaciones y propiedades para un mejor entendimiento sobre las cadenas de Markov.
- Luego, veremos las comunicaciones que hay entre los estados de la cadena para luego clasificarlos en clases.
- Después, miramos la probabilidad y el tiempo medio que tardaría en llegar a un estado.
- Seguidamente, se muestra algunas características sobre cada estado de la cadena, como por ejemplo, las veces que visita al un estado es infinito o finito.
- A continuación, veremos la existencia de las distribuciones invariantes y la convergencia al equilibrio.
- Posteriormente, estudiamos sobre las cadenas de Markov invirtiendo el sentido del tiempo.
- Más tarde, estudiamos el teorema ergódico sobre las cadenas de Markov.
- Por último, nos centramos en aplicar las cadenas de Markov al problema de la ruina del jugador.

3. Definiciones y propiedades

Sea I un conjunto numerable. Llamaremos I el espacio de los estados, donde sus elementos $i \in I$ son los **estados**.

Consideremos también el conjunto T como un espacio de parámetros, donde sus elementos t∈T representan el tiempo, que puede ser discreto o continuo. En este proyecto solo trabajaremos en el tiempo discreto.

Definición 3.1. *Un proceso estocástico con el espacio de estados I es una familia de variables aleatorias* $\{\mathcal{X}_t\}_{t\in T}$, $\mathcal{X}_t : \Omega \to I$ *indexados por un conjunto T. Llamaremos* $\mathcal{X}_t$ *el estado del proceso en el tiempo t.*

Definición 3.2. *Llamaremos* $\lambda = (\lambda_i : i \in I)$ *la medida en I si* $0 \leq \lambda_i < \infty$, $\forall i \in I$. *Definiremos*

$$\lambda_i = \mathbb{P}(\mathcal{X} = i) = \mathbb{P}(\{\omega : \mathcal{X}(\omega) = i\})$$

y λ *es una* ***distribución*** *si* $\sum_{i\in I} \lambda_i = 1$.

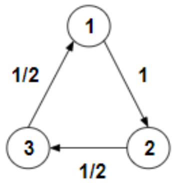
Definición 3.3. *Una matriz* $P = (p_{ij} : i, j \in I)$ *es estocástica si para cada fila i* $(p_{ij} : j \in I)$ *es una distribución, es decir,*

$$\sum_{j\in I} p_{ij} = 1, \text{ para todo } i \in I.$$

Llamaremos la matriz estocástica P como la matriz de transición si $p_{ij} \geq 0$, *para todo i,j* $\in I$.

Hay una correspondencia biyectiva entre una matriz de transición P con un diagrama de los estados, donde van de un estado a otro según la probabilidad que tengan.

Ejemplos 3.4.



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

La fila 1 y columna 1 de la matriz representan el estado 1; la fila 2 y columna 2 representan el estado 2 y la fila 3 y columna 3 representan el estado 3. Sea p_{ij} el elemento de la matriz P de la fila i y columna j. Definamos p_{ij} como la probabilidad que va del estado i a j. En nuestro ejemplo, el estado 1 va al estado 2 con probabilidad 1; el estado 2 va al estado 3 con probabilidad 1/2, como la matriz P es estocástica y el estado 2 va al estado 3, el otro 1/2 tiene que ir a sí mismo, es decir, $p_{22} = 1/2$. Por último, el estado 3 va al estado 1 con probabilidad 1/2 y también va a sí mismo con probabilidad 1/2.

Definición 3.5. *La sucesión de variables aleatorias* $(\mathcal{X}_n)_{n\geq 0}$ *es una cadena de Markov, Markov*(λ, P)*, con la distribución incial* λ *y la matriz de transición* P *si:*

I. $\mathcal{X}_0$ *tiene distribución* λ*;*

II. *para cada* $n \geq 0$*, condicionado en* $\mathcal{X}_n = i$*,* $\mathcal{X}_{n+1}$ *tiene distribución* $(p_{ij} : j \in I)$ *y es independiente de* $\mathcal{X}_0, \ldots, \mathcal{X}_{n-1}$.

Es decir, para cada $n \geq 0$*, con* $i_0, \ldots, i_{n+1} \in I$

I. $\mathbb{P}(\mathcal{X}_0 = i_0) = \lambda_{i_0}$.

II. $P(\mathcal{X}_{n+1} = i_{n+1} | \mathcal{X}_0 = i_0, \ldots, \mathcal{X}_n = i_n) = p_{i_n i_{n+1}}$.

Recordemos que p_{ij} es la probabilidad de ir el estado i al estado j.

Observación 3.6. Si $(\mathcal{X}_n)_{0\leq n\leq N}$ es una sucesión finita de variables aleatorias satisfaciendo las propiedades I e II para n = 0,$\ldots, N-1$, entonces, $(\mathcal{X}_n)_{0\leq n\leq N}$ también es Markov(λ, P).

Notación 3.7. *Sean* $A_0, \ldots, A_n$ *conjuntos, entonces* $\mathbb{P}(A_0, \ldots, A_n) = \mathbb{P}(A_0 \cap \cdots \cap A_n)$.

Teorema 3.8. *En un experimento aleatorio en tiempo discreto,* $(\mathcal{X}_n)_{0\leq n\leq N}$ *es Markov*(λ, P) *si y solo si,*

$$\mathbb{P}(\mathcal{X}_0 = i_0, \mathcal{X}_1 = i_1, \ldots, \mathcal{X}_N = i_N) = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{N-1} i_N}, \text{ para todo } i_0, \ldots, i_N \in I \ . \tag{3.1}$$

Demostración. $\Rightarrow$) Si $(\mathcal{X}_n)_{0\leq n\leq N}$ es Markov(λ, P), entonces, por proposición 1.4 (fórmula de las probabilidades compuestas) y las propiedades I e II de la definición 3.5

$$\mathbb{P}(\mathcal{X}_0 = i_0, \mathcal{X}_1 = i_1, \ldots, \mathcal{X}_N = i_N)$$
$$= \mathbb{P}(\mathcal{X}_0 = i_0)\mathbb{P}(\mathcal{X}_1 = i_1 | \mathcal{X}_0 = i_0) \cdots \mathbb{P}(\mathcal{X}_N = i_N | \mathcal{X}_0 = i_0, \mathcal{X}_1 = i_1, \ldots, \mathcal{X}_{N-1} = i_{N-1})$$
$$= \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{N-1} i_N}$$

$\Leftarrow$) Suponemos (3.1) para algún N.

Sea $i_N \in I$ y $\sum_{j\in I} p_{ij} = 1$. Suponemos que es cierto para N y veremos si es cierto para $N-1$.

Sea

$$\mathbb{P}(\mathcal{X}_0 = i_0, \mathcal{X}_1 = i_1, \ldots, \mathcal{X}_n = i_n) = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}, \text{ para todo } n = 0, \ldots, N$$

Si $n = 0$, $\mathbb{P}(\mathcal{X}_0 = i_0) = \lambda_{i0}$. Entonces para $n = 0, \ldots, N-1$ tenemos

$$\mathbb{P}(\mathcal{X}_{n+1} = i_{n+1} | \mathcal{X}_0 = i_0, \ldots, \mathcal{X}_n = i_n)$$
$$= \mathbb{P}(\mathcal{X}_0 = i_0, \ldots, \mathcal{X}_n = i_n, \mathcal{X}_{n+1} = i_{n+1}) / \mathbb{P}(\mathcal{X}_0 = i_0, \mathcal{X}_1 = i_1, \ldots, \mathcal{X}_n = i_n)$$
$$= \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_n i_{n+1}} / \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} = p_{i_n i_{n+1}}.$$

Aplicando la definición de la probabilidad condicionada.

Por tanto, $(\mathcal{X}_n)_{0\leq n\leq N}$ es Markov(λ, P). $\square$

Definición 3.9. *Sea $\delta_i = (\delta_{ij} : j \in I)$ la masa unitaria de i, donde*

$$\delta_{ij} = \begin{cases} 1 & si\ i = j \\ 0 & si\ no \end{cases} \tag{3.2}$$

El teorema que veremos a continuación nos muestra que, aunque empecemos las cadenas de Markov en el paso m, sigue siendo una cadena de Markov.

Teorema 3.10. *(Propiedad de Markov) Sea $(\mathcal{X}_n)_{0\leq n\leq N}$ de Markov(λ, P), entonces, condicionado en $X_m = i$, $(\mathcal{X}_{m+n})_{n\geq 0}$ es Markov(δ_i, P) y es independiente de las variables aleatorias $\mathcal{X}_0, \ldots, \mathcal{X}_m$.*

Demostración. Sea A un evento determinado por $\mathcal{X}_0, \ldots, \mathcal{X}_m$, entonces queremos ver si

$$\mathbb{P}(\{\mathcal{X}_m = i_m, \ldots, \mathcal{X}_{m+n} = i_{m+n}\} \cap A|\mathcal{X}_m = i) = \delta_{ii_m} p_{i_m i_{m+1}} \cdots p_{i_{m+n-1} i_{m+n}} \mathbb{P}(A|\mathcal{X}_m = i) \tag{3.3}$$

teniendo en cuenta de que $\{\mathcal{X}_m = i_m, \ldots, \mathcal{X}_{m+n} = i_{m+n}\}$ es independiente de A.

Suponemos que $A = \{\mathcal{X}_0 = i_0, \ldots, \mathcal{X}_m = i_m\}$, entonces

$$\mathbb{P}(\{\mathcal{X}_m = i_m, \ldots, \mathcal{X}_{m+n} = i_{m+n}\} \cap A|\mathcal{X}_m = i) = \mathbb{P}(\mathcal{X}_0 = i_0, \ldots, \mathcal{X}_{m+n} = i_{m+n} \text{ y } i = i_m)/\mathbb{P}(\mathcal{X}_m = i)$$

$$= \delta_{ii_m} p_{i_m i_{m+1}} \cdots p_{i_{m+n-1} i_{m+n}} \mathbb{P}(\mathcal{X}_0 = i_0, \ldots, \mathcal{X}_{m+n} = i_{m+n} \text{ y } i = i_m)/\mathbb{P}(\mathcal{X}_m = i)$$

que es cierto por el Teorema 3.8. En general, para cada evento A determinado por $\mathcal{X}_0, \ldots, \mathcal{X}_m$ sea la unión disjunta de eventos elementales

$$A = \bigcup_{k=1}^{\infty} A_k.$$

Entonces obtenemos la identidad (3.3) sumando las identidades correspondientes de A_k. □

Queremos averiguar cuál es la probabilidad de un estado después de n pasos. Este problema se reducirá en calcular la potencia n-ésima de la matriz de transición, P.

Supongamos $I = \{1, 2, \ldots, N\}$ un espacio de estados finito de N elementos. Entonces la distribución y medida λ es un vector con elementos indexado por I y la matriz de transición P también está indexado por $I \times I$. Por tanto, λ será un N-vector y P un $N \times N$ - matriz.

Definición 3.11. *Sea λ la medida y P la matriz de transición, entonces (λP) es una medida y P^2 es una matriz tales que*

$$(\lambda P)_j = \sum_{i\in I} \lambda_i p_{ij},\ (P^2)_{ik} = \sum_{j\in I} p_{ij} p_{jk}.$$

Definimos la matriz P^n para cualquier $n \geq 0$, donde P^0 es la matriz identidad, $(I_d)_{ij} = \delta_{ij}$. Y escribiremos $p_{ij}^{(n)} = (P^n)_{ij}$ como la probabilidad de ir el estado i a j después de n pasos.

Definición 3.12. *Sea $\lambda_i \geq 0$, entonces $\mathbb{P}_i(A) = \mathbb{P}(A|\mathcal{X}_0 = i)$ es la probabilidad condicionada.*

Entonces, por la Propiedad de Markov con $m = 0$, bajo $\mathbb{P}_i$, $(\mathcal{X}_n)_{n\geq 0}$ es Markov(δ_i, P). Por tanto, $(\mathcal{X}_m)_{n\geq 0}$ bajo $\mathbb{P}_i$ no depende de λ.

Teorema 3.13. *Sea $(\mathcal{X}_n)_{n\geq 0}$ una cadena de markov Markov(λ, P), entonces, $\forall n, m \geq 0$*

I. $\mathbb{P}(\mathcal{X}_n = j) = (\lambda P^n)_j$;

II. $\mathbb{P}_i(\mathcal{X}_n = j) = \mathbb{P}(\mathcal{X}_{n+m} = j | \mathcal{X}_m = i) = p_{ij}^{(n)}$.

Demostración. I. Por el Teorema 3.8,

$$\mathbb{P}(\mathcal{X}_n = j) = \sum_{i_0 \in I} \cdots \sum_{i_{n-1} \in I} \mathbb{P}(\mathcal{X}_0 = i_0, \ldots, \mathcal{X}_{n-1} = i_{n-1}, \mathcal{X}_n = j)$$

$$= \sum_{i_0 \in I} \cdots \sum_{i_{n-1} \in I} \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} = (\lambda P^n)_j$$

II. Por la Propiedad de Markov, condicionado en $\mathcal{X}_m = i$, $(\mathcal{X}_{m+n})_{n\geq 0}$ es Markov(δ_i, P). Por tanto, escogemos $\lambda = \delta_i$ y lo aplicamos en I.

□

4. Estructura de clases

En esta sección veremos cómo clasificaremos los estados según sus movimientos.

Definición 4.1. *Decimos que el estado i se dirige al estado j, $i \to j$ si*

$$\mathbb{P}_i(\mathcal{X}_n = j, \textit{ para alguna } n \geq 0) = \mathbb{P}(\mathcal{X}_n = j, \textit{ para alguna } n \geq 0 | \mathcal{X}_0 = i) > 0.$$

Decimos que el estado i comunica con el estado j, $i \leftrightarrow j$, si $i \to j$ y $j \to i$.

Teorema 4.2. *Sean los estados i, $j \in I$ cualesquieras, entonces son equivalentes*

I. $i \to j$;

II. $p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} > 0$ *para algunos* $i_0, \ldots, i_n$ *con* $i_0 = i$ *e* $i_n = j$;

III. $p_{ij}^{(n)} > 0$ *para algún* $n \geq 0$.

Demostración. Observamos que

$$p_{ij}^{(n)} \leq \mathbb{P}_i(\mathcal{X}_n = j, n \geq 0) \leq \sum_{n=0}^{\infty} p_{ij}^{(n)}$$

Por tanto, $I \iff III$. También tenemos,

$$p_{ij}^{(n)} = \sum_{i_1, \ldots, i_{n-1}} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}$$

que prueba la equivalencia $II \iff III$. □

Proposición 4.3. *La comunicación entre estados, $\leftrightarrow$, es una relación de equivalencia.*

Demostración. I. Refrexiva: $\forall i \in I$, $i \leftrightarrow i$.

Como $p_{ii}^{(0)} = \mathbb{P}(\mathcal{X}_0 = i | \mathcal{X}_0 = 1) = \mathbb{P}(\mathcal{X}_0 = i) = \lambda_i > 0$, entonces por el Teorema 4.2., $i \leftrightarrow i$.

II. Simétrica: $\forall i, j \in I$, $i \leftrightarrow j \Rightarrow j \leftrightarrow i$.

Si $i \leftrightarrow j$, entonces por el Teorema 4.2., $p_{ij}^{(n)} > 0$ y $p_{ji}^{(n)} > 0$. Por tanto, $j \leftrightarrow i$.

III. Transitiva: $\forall i, j, k \in I$, si $i \leftrightarrow j$ y $j \leftrightarrow k \Rightarrow i \leftrightarrow k$.

Si $i \leftrightarrow j$ y $j \leftrightarrow k$, entonces existen $i_0, \ldots, i_n$ y $j_0, \ldots, j_m$ tales que $p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} > 0$ y $p_{j_0 j_1} p_{j_1 j_2} \cdots p_{j_{n-1} j_m} > 0$ con $i_0 = i$, $i_m = j_0 = j$ y $j_m = k$.

Por tanto, $p_{ii_1} p_{i_1 i_2} \cdots p_{i_{m-1} j} p_{jj_1} \cdots p_{j_{n-1} k} > 0$ y $i \to k$, por el Teorema 4.2. Por último, aplicamos la propiedad simétrica y obtenemos $i \leftarrow k$

□

Por tanto, las particiones de I, $\mathcal{P}(I)$, son clases de comunicación.

Definición 4.4. *Una clase de comunicación $C \in \mathcal{P}(I)$ es cerrado si*

$$i \in C, i \to j \Rightarrow j \in C.$$

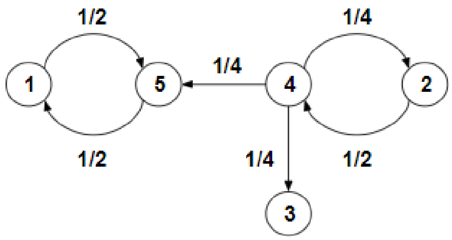
Definición 4.5. *Un estado $i \in I$ es* ***absorbente*** *si $\{i\}$ es una clase de comunicación cerrado.*

Definición 4.6. *Una cadena o una matriz de transición P es irreducible si I es la única clase de comunicación cerrada.*

Ejemplos 4.7. Veremos las clases de la siguiente matriz de transición:

$$P = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

Primero lo transcribimos en diagrama de estados para detectar las clases de comunicación con más facilidad.



Las clases son $\{1,5\}$, $\{2,4\}$ y $\{3\}$. Observamos que $\{1,5\}$ y $\{3\}$ clases son cerradas y $\{3\}$ es una clase absorbente. Pero la clase $\{2,4\}$ no es una clase cerrada, ya que si estamos en el estado 4, puede ir también al estado 3 o 5.

5. Probabilidad de absorción y tiempo medio de absorción

Sea $(\mathcal{X}_n)_{n\geq 0}$ una cadena de Markov con la matriz de transición P.

Definición 5.1. *El **tiempo de llegada** de $A \subset I$ es una variable aleatoria $H^A : \Omega \to \{0, 1, \dots\} \cup \{\infty\}$ tales que*

$$H^A(\omega) = \inf\{n \geq 0 : \mathcal{X}_n(\omega) \in A\}$$

es decir, es el mínimo pasos que se necesita para llegar al estado o conjunto A. Para el caso $\emptyset$, tenemos $H^A(\emptyset) = \infty$.

Definición 5.2. *La probabilidad de que $(\mathcal{X}_n)_{n\geq 0}$ vaya A empezando por el estado i es*

$$h_i^A = \mathbb{P}_i(H^A < \infty) = \mathbb{P}(H^A < \infty | X_0 = i).$$

Si A es una clase cerrada, entonces h_i^A es la **probabilidad de absorción**.

Observación 5.3. Si $i \in A$, entonces $h_i^A = 1$. Y si $i \in B$ tales que B es una clase cerrada y $B \cap A = \emptyset$, entonces $h_i^A = 0$.

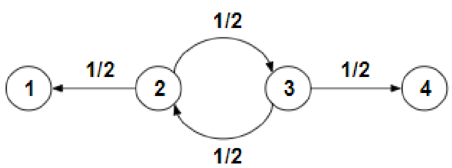
Definición 5.4. *El **tiempo medio** de $(\mathcal{X}_n)_{n\geq 0}$ para llegar a A es*

$$k_i^A = \mathbb{E}_i(H^A) = \mathbb{E}(H^A | \mathcal{X}_0 = i) = \sum_{n<\infty} n\mathbb{P}(H^A = n | \mathcal{X}_0 = i) + \infty\mathbb{P}(H^A = \infty | \mathcal{X}_0 = i)$$

Notación 5.5. *Escribiremos $h_i^A = \mathbb{P}_i$(ir a A) y $k_i^A = \mathbb{E}_i$(tiempo en llegar a A).*

Ahora veremos un ejemplo para entender mejor los conceptos mencionados anteriormente.

Ejemplos 5.6. Consideremos una cadena con el siguiente diagrama:



Vamos a averiguar cuál es la probabilidad de ir el estado 4 empezando por el estado 2 y cuánto tiempo tarda en llegar en las clases absorbentes 1 y 4.

Sean $h_i = \mathbb{P}_i(\text{ir a } 4)$ y $k_i = \mathbb{E}_i(\text{tiempo de llegar a } \{1, 4\})$.

Es trivial de que $h_1 = \mathbb{P}_1(\text{ir a } 4) = 0$, $h_4 = \mathbb{P}_4(\text{ir a } 4) = 1$ y $k_1 = k_4 = 0$. Supongamos que empezamos por el estado 2 y realizamos un paso. Tenemos una probabilidad de $1/2$ de ir al estado 1 y una probabilidad de $1/2$ de ir al estado 3, entonces,

$$h_2 = \frac{1}{2}h_1 + \frac{1}{2}h_3,\ k_2 = 1 + \frac{1}{2}k_1 + \frac{1}{2}k_3.$$

Hay un 1 en la ecuación del k_2 porque contamos el primer paso. Análogamente, si empezamos por el estado 3, tenemos

$$h_3 = \frac{1}{2}h_2 + \frac{1}{2}h_4,\ k_3 = 1 + \frac{1}{2}k_2 + \frac{1}{2}k_4.$$

Resolviendo estos sistemas de ecuaciones, obtenemos $h_2 = 1/3$ y $k_2 = 2$. Es decir, la probabilidad de llegar al estado 4 empezando por el estado 2 es $1/3$ y el tiempo medio en llegar al estado 4 es de 2 pasos.

A continuación veremos los resultados generales de estos dos conceptos.

Teorema 5.7. *El vector de probabilidades $h^A = (h_i^A : i \in I)$ es minimal no negativa y es solución del sistema de ecuaciones lineales*

$$\begin{cases} h_i^A = 1 & si\ i \in A \\ h_i^A = \sum_{j\in I} p_{ij} h_j^A & si\ i \notin A. \end{cases} \tag{5.1}$$

(h^A es minimal si para todo i, $x_i \geq h_i$, donde $x = (x_i : x_i \geq 0, i \in I)$ sea otra solución).

Demostración. Veremos primero si h^A cumple (5.1).

Si $\mathcal{X}_0 = i \in A$ entonces $H^A = \text{ínf}\{n \geq 0 : \mathcal{X}_0 = i \in A\} = 0$. Por tanto, $h_i^A = 1$.

Si $\mathcal{X}_0 = i \notin A$ entonces $H^A \geq 1$. Aplicando la propiedad de Markov, siendo $m = 1$

$$\mathbb{P}_i(H^A < \infty | \mathcal{X}_1 = j) = \mathbb{P}_j(H^A < \infty) = h_j^A.$$

$$h_i^A = \mathbb{P}_i(H^A < \infty) = \sum_{j\in I} \mathbb{P}_i(H^A < \infty, \mathcal{X}_1 = j) = \sum_{j\in I} \mathbb{P}_i(H^A < \infty | \mathcal{X}_1 = j)\mathbb{P}_i(\mathcal{X}_1 = j) = \sum_{j\in I} h_j^A p_{ij}.$$

Supongamos que $x = (x_i : i \in I)$ es otra solución del (5.1). Si $i \in A$, $h_i^A = x_i = 1$. Si $i \notin A$, entonces

$$x_i = \sum_{j\in I} p_{ij} x_j = \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij} x_j.$$

Substituyendo $x_j = \sum_{k\in A} p_{jk} + \sum_{k\notin A} p_{jk} x_k$

$$x_i = \sum_{j\in I} p_{ij} x_j = \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij} \Big(\sum_{k\in A} p_{jk} + \sum_{k\notin A} p_{jk} x_k \Big) = \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij} \sum_{k\in A} p_{jk} + \sum_{j\notin A} p_{ij} \sum_{k\notin A} p_{jk} x_k$$

$$= \mathbb{P}_i(\mathcal{X}_1 \in A) + \mathbb{P}_i(\mathcal{X}_1 \notin A, \mathcal{X}_2 \in A) + \sum_{j\notin A} \sum_{k\notin A} p_{ij} p_{jk} x_k.$$

Repetimos el proceso n veces y obtenemos,

$$x_i = \mathbb{P}_i(\mathcal{X}_1 \in A) + \cdots + \mathbb{P}_i(\mathcal{X}_1 \notin A, \ldots, \mathcal{X}_{n-1} \in A, \mathcal{X}_n \in A) +$$

$$+ \sum_{j_1\notin A} \cdots \sum_{j_n\notin A} p_{ij_1} p_{j_1 j_2} \cdots p_{j_{n-1} j_n} x_{j_n}.$$

Si x es no negativa, entonces el término del sumatorio también lo es. Como el resto de los términos suman $\mathbb{P}_i(H^A \leq n)$, $x_i \geq \mathbb{P}_i(H^A \leq n)$ para todo n y

$$x_i \geq \lim_{n\to\infty} \mathbb{P}_i(H^A \leq n) = \mathbb{P}_i(H^A < \infty) = h_i.$$

□

A continuación veremos la forma general del tiempo medio de llegada. Recordemos que $k_i^A = \mathbb{E}_i(H^A)$, donde H^A es la primera vez que $(\mathcal{X}_n)_{n\geq 0}$ llega a A.

Notación 5.8. $\mathbf{1}_{\mathcal{X}_1=j}$ *es una variable aleatoria tales que*

$$\mathbf{1}_{\mathcal{X}_1=j} = \begin{cases} 1 & si\ \mathcal{X}_1 = j \\ 0 & si\ no. \end{cases}$$

Teorema 5.9. *El vector del tiempo medio de llegada* $k^A = (k_i^A : i \in A)$ *es la solución minimal no negativa del sistema de ecuaciones lineales*

$$\begin{cases} k_i^A = 0 & si\ i \in A \\ k_i^A = 1 + \sum_{j\notin I} p_{ij} k_j^A & si\ i \notin A \end{cases} \tag{5.2}$$

Demostración. Veremos primero si k^A cumple (5.2).

Si $\mathcal{X}_0 = i \in A$ entonces $H^A = 0$. Por tanto, $k_i^A = 0$.

Si $\mathcal{X}_0 = i \notin A$ entonces $H^A \geq 1$. Aplicando la propiedad de Markov,

$$\mathbb{E}_i(H^A | X_1 = j) = 1 + \mathbb{E}_j(H^A).$$

Entonces,

$$k_i^A = \mathbb{E}_i(H^A) = \sum_{j\in I} \mathbb{E}_i(H^A \mathbf{1}_{\mathcal{X}_1=j}) = \sum_{j\in I} \mathbb{E}_i(H^A | \mathcal{X}_1 = j)\mathbb{P}_i(\mathcal{X}_1 = j) = 1 + \sum_{j\notin A} p_{ij} k_j^A.$$

Supongamos que $y = (y_i : i \in I)$ también es solución de (5.2). Si $i \in A$, $k_i^A = y_i = 0$. Si $i \notin A$,

$$y_i = 1 + \sum_{j\notin A} p_{ij} y_i = 1 + \sum_{j\notin A} p_{ij} \Big(1 + \sum_{k\notin A} p_{jk} y_k\Big) = \mathbb{P}_i(H^A \geq 1) + \mathbb{P}_i(H^A \geq 2) + \sum_{j\notin A}\sum_{k\notin A} p_{ij} p_{jk} y_k.$$

Repitiendo el proceso n veces,

$$y_i = \mathbb{P}_i(H^A \geq 1) + \cdots + \mathbb{P}_i(H^A \geq n) + \sum_{j_1\notin A} \cdots \sum_{j_n\notin A} p_{ij_1} p_{j_1 j_2} \cdots p_{j_{n-1} j_n} y_{j_n}.$$

Si y es no negativa,

$$y_i \geq \mathbb{P}_i(H^A \geq 1) + \cdots + \mathbb{P}_i(H^A \geq n).$$

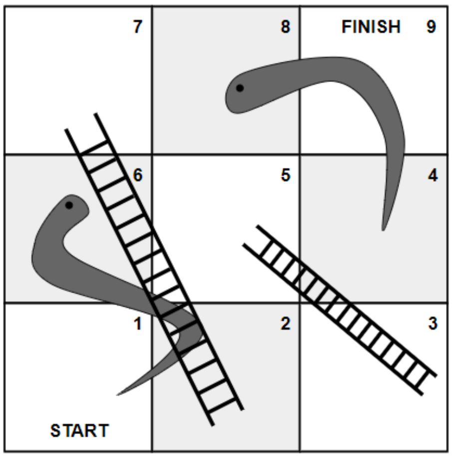
Entonces, cuando $n \to \infty$

$$y_i \geq \sum_{n=1}^{\infty} \mathbb{P}_i(H^A \geq n) = \mathbb{E}_i(H^A) = k_i^A.$$

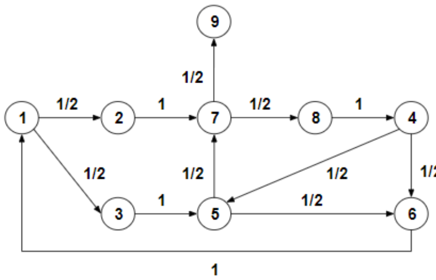
□

Ejemplos 5.10. Consideremos el juego de serpientes y escaleras con 9 casillas.

En cada turno, los jugadores tiran una moneda. Avanzas una casilla si es cara y avanzas dos si es cruz. Si estás en la cabeza de la serpiente, te deslizas hacia abajo, y si estás en el pie de la escalera, subes. ¿En cuántos turnos acabaría el juego? ¿Cuál es la probabilidad de acabar el juego?



Primero dibujamos el diagrama de la cadena, donde cada estado representa el número de la casilla:



Queremos saber en cuántos turnos llegamos en la casilla 9, es decir, calcular el $k_1^{\{9\}}$. Aplicando el teorema anterior,

$$\begin{cases} k_9 = 0 \\ k_1 = 1 + \frac{1}{2}k_2 + \frac{1}{2}k_3 \\ k_2 = 1 + k_7 \\ k_3 = 1 + k_5 \\ k_4 = 1 + \frac{1}{2}k_5 + \frac{1}{2}k_6 \\ k_5 = 1 + \frac{1}{2}k_6 + \frac{1}{2}k_7 \\ k_6 = 1 + k_1 \\ k_7 = 1 + \frac{1}{2}k_8 + \frac{1}{2}k_9 \\ k_8 = 1 + k_4 \end{cases} \Longrightarrow \begin{cases} k_9 = 0 \\ k_1 = 1 + \frac{1}{2}k_2 + \frac{1}{2}k_3 = 5 + \frac{4}{7}k_1 \Rightarrow k_1 = \mathbf{11{,}67} \\ k_2 = 1 + k_7 = 4 + \frac{3}{7}k_1 \\ k_3 = 1 + k_5 = 4 + \frac{5}{7}k_1 \\ k_4 = 1 + \frac{1}{2}k_5 + \frac{1}{2}k_6 = \frac{21}{8} + \frac{3}{4}k_1 + \frac{1}{8}k_4 \Rightarrow k_4 = 3 + \frac{6}{7}k_1 \\ k_5 = 1 + \frac{1}{2}(1 + k_1) + \frac{1}{2}(\frac{3}{2} + \frac{1}{2}k_4) = \frac{9}{4} + \frac{1}{2}k_1 + \frac{1}{4}k_4 \\ k_6 = 1 + k_1 \\ k_7 = 1 + \frac{1}{2}(1 + k_4) = \frac{3}{2} + \frac{1}{2}k_4 \\ k_8 = 1 + k_4 \end{cases}$$

Por tanto, el tiempo medio de un jugador en acabar la partida son 12 turnos.

A continuación vamos a averiguar cuál es la probabilidad de terminar la partida. Observamos que $\{9\}$ es la única clase cerrada de la cadena y es una clase absorbente. Por tanto, $h_1^{\{9\}}$ es la probabilidad de absorción.

$$\begin{cases} h_9 = 1 \\ h_8 = h_4 \\ h_7 = \frac{1}{2}h_8 + \frac{1}{2}h_9 \\ h_6 = h_1 \\ h_5 = \frac{1}{2}h_6 + \frac{1}{2}h_7 \\ h_4 = \frac{1}{2}h_5 + \frac{1}{2}h_6 \\ h_3 = h_5 \\ h_2 = h_7 \\ h_1 = \frac{1}{2}h_2 + \frac{1}{2}h_3 \end{cases} \implies \begin{cases} h_2 = \frac{1}{2}h_4 + \frac{1}{2} \\ h_3 = \frac{1}{2}h_1 + \frac{1}{2}h_2 \\ h_4 = \frac{1}{2}h_3 + \frac{1}{2}h_1 \\ h_1 = \frac{1}{2}h_2 + \frac{1}{2}h_3 \end{cases} \implies \begin{cases} h_2 = \frac{1}{2}(\frac{1}{2}h_3 + \frac{1}{2}h_1) + \frac{1}{2} \\ = \frac{1}{4}(\frac{1}{2}h_1 + \frac{1}{2}h_2) + \frac{1}{4}h_1 + \frac{1}{2} \\ h_1 = \frac{1}{2}h_2 + \frac{1}{2}(\frac{1}{2}h_1 + \frac{1}{2}h_2) \end{cases}$$

$$\implies \begin{cases} \frac{7}{8}h_2 = \frac{3}{8}h_1 + \frac{1}{2} \\ h_1 = h_2 \end{cases}$$

Por tanto, $h_1^{\{9\}} = 1$, es decir, el juego seguro que acabará.

6. Propiedad fuerte de Markov

Definición 6.1. *Una variable aleatoria $T : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$ es **tiempo de parada** si el evento $\{T = n\}$ depende únicamente de $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_n$ para $n = 0, 1, \dots$.*

La propiedad fuerte de Markov está vinculado con el tiempo de parada. Si T es un tiempo de parada y $B \subseteq \Omega$ determinado por $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_T$, entonces $B \cap \{T = m\}$ está determinado por $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_m$ para todo $m = 0, 1, 2, \dots$.

Teorema 6.2 (Propiedad fuerte de Markov)**.** *Sea la cadena $(\mathcal{X}_n)_{n\geq 0}$ de Markov(λ, P) y sea T el tiempo de parada de $(\mathcal{X}_n)_{n\geq 0}$, entonces, condicionado en $T < \infty$ y $\mathcal{X}_T = i$, $(\mathcal{X}_{T+n})_{n\geq 0}$ es Markov(δ_i, P) e independiente de $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_T$.*

Demostración. Sea B un evento determinado por $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_T$, entonces $B \cap \{T = m\}$ está determinado por $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_m$. Por la propiedad de Markov en el paso m

$$\mathbb{P}(\{\mathcal{X}_T = j_0, \mathcal{X}_{T+1} = j_1, \dots, \mathcal{X}_{T+n} = j_n\} \cap B \cap \{T = m\} \cap \{\mathcal{X}_T = i\}) =$$

$$= \mathbb{P}_i(\mathcal{X}_0 = j_0, \mathcal{X}_1 = j_1, \dots, \mathcal{X}_n = j_n)\mathbb{P}(B | T < \infty, \mathcal{X}_T = i).$$

$\square$

7. Recurrencia y transitoriedad

Definición 7.1. *Sea una cadena* $(\mathcal{X}_n)_{n\geq 0}$ *de Markov con la matriz de transición* P.

I. *El estado* i *es recurrente si* $\mathbb{P}_i(\mathcal{X}_n = i$ *por infinitos* $n) = 1$*, es decir, vuelve al estado* i *infinitamente.*

II. *El estado* i *es transitorio si* $\mathbb{P}_i(\mathcal{X}_n = i$ *por infinitos* $n) = 0$*, es decir, pasa por el estado* i *finitamente.*

Definición 7.2. *Llamaremos el* ***primer tiempo de llegada*** *del estado* i *a la variable aleatoria* T_i, $T_i(\omega) = \text{ínf}\{n \geq 1 : \mathcal{X}_n(\omega) = i\}$.

Observación 7.3. Si no existe alguna n tales que $\mathcal{X}_n(\omega) = i$, es decir, $\{n \geq 1 : \mathcal{X}_n(\omega) = i\} = \emptyset$ entonces $\text{ínf}\,\emptyset = \infty$.

Definición 7.4. *La r-ésima tiempo de llegada del estado* i*,* $T_i^{(r)}$*, es*

$$T_i^{(r)}(\omega) = \begin{cases} 0 & \text{si } r = 0 \\ T_i(\omega) & \text{si } r = 1 \\ \text{ínf}\{n \geq T_i^{(r)}(\omega) + 1 : \mathcal{X}_n(\omega) = i\} & \text{si } r > 1. \end{cases}$$

Definición 7.5. *Llamaremos la* ***longitud*** *de la r-ésima excursión al estado* i *como*

$$S_i^{(r)} = \begin{cases} T_i^{(r)} - T_i^{(r-1)} & \text{si } T_i^{(r-1)} < \infty \\ 0 & \text{si } T_i^{(r-1)} = \infty. \end{cases}$$

Nos enfocamos en el estudio de esta longitud cuando analizamos la recurrencia y la transitoriedad.

Lema 7.6. *Sea* $r > 1$ *condicionado en* $T_i^{(r-1)} < \infty$*, entonces* $S_i^{(r)}$ *es independiente de los* $\mathcal{X}_m$ *cuando* $m \leq T_i^{(r-1)}$ *y*

$$\mathbb{P}(S_i^{(r)} = n | T_i^{(r-1)} < \infty) = \mathbb{P}_i(T_i = n).$$

Demostración. Aplicamos el Teorema 6.2, la propiedad fuerte de Markov, con el tiempo de llegada $T = T_i^{r-1}$. Es automatico de que $\mathcal{X}_T = i$ cuando $T < \infty$. Entonces, condicionado en $T < \infty$, $(\mathcal{X}_{T+n})_{n\geq 0}$ es Markov(δ_i, P) e independiente de $\mathcal{X}_0, \mathcal{X}_1, \ldots, \mathcal{X}_T$. Pero como

$$S_i^{(r)} = \text{ínf}\{n \geq 1 : \mathcal{X}_{T+n} = i\}$$

entonces $S_i^{(r)}$ es el primer tiempo de llegada de $(\mathcal{X}_{T+n})_{n\geq 0}$ al estado i. $\square$

Definición 7.7. *Llamaremos el número de visitas al estado* i*,* V_i*, si*

$$V_i = \sum_{n=0}^{\infty} \mathbf{1}_{\{\mathcal{X}_n = i\}}.$$

La esperanza del número de visitas es

$$\mathbb{E}_i(V_i) = \mathbb{E}_i\Big(\sum_{n=0}^{\infty} \mathbf{1}_{\{\mathcal{X}_n = i\}}\Big) = \sum_{n=0}^{\infty} \mathbb{E}_i(\mathbf{1}_{\{\mathcal{X}_n = i\}}) = \sum_{n=0}^{\infty} \mathbb{P}_i(\mathcal{X}_n = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)}.$$

Se suele calcular la distribución de V_i con $\mathbb{P}_i$ en término de la ***probabilidad de retorno****,*

$$f_i = \mathbb{P}_i(T_i < \infty).$$

Lema 7.8. $\mathbb{P}_i(V_i > r) = f_i^r$ *para* $r = 0, 1, 2, \ldots$.

Demostración. Observamos que si $\mathcal{X}_0 = i$ entonces $\{V_i > r\} = \{T_i^{(r)} < \infty\}$.

Si r = 0, entonces,

$$f_i^0 = \mathbb{P}_i(V_i > 0) = \mathbb{P}(T_i^{(0)} < \infty)$$

que es cierto por definición. Suponemos inductivamente que es cierto para r, entonces veremos si es cierto para $r+1$

$$\begin{aligned}\mathbb{P}_i(V_i > r+1) &= \mathbb{P}_i(T_i^{(r+1)} < \infty)\\ &= \mathbb{P}_i(T_i^{(r)} < \infty \text{ y } S_i^{(r+1)} < \infty)\\ &= \mathbb{P}_i(S_i^{(r+1)} < \infty | T_i^{(r)} < \infty)\mathbb{P}_i(T_i^{(r)} < \infty)\\ &= f_i f_i^r = f_i^{r+1}\end{aligned}$$

por el Lema 7.6. Por tanto, por inducción, $\mathbb{P}_i(V_i > r) = f_i^r$ es cierto para todo $r \geq 0$. □

En el siguiente teorema obtenemos otra manera de determinar la recurrencia y la transitoriedad de un estado.

Teorema 7.9. *Se cumplen,*

I. *Si* $\mathbb{P}_i(T_i < \infty) = 1$ *entonces el estado i es recurrente y* $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.

II. *Si* $\mathbb{P}_i(T_i < \infty) < 1$ *entonces el estado i es transitorio y* $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$.

En particular, cada estado son tanto recurrentes como transitorios.

Demostración. I) Si $f_i = \mathbb{P}_i(T_i < \infty) = 1$, entonces por el lema 7.8

$$\mathbb{P}_i(V_i = \infty) = \lim_{r\to\infty} \mathbb{P}_i(V_i > r) = 1$$

ja que $\mathbb{P}_i(V_i > r) = \mathbb{P}_i(T_i < \infty) = 1$. Entonces, la probabilidad de visitar al estado i es 1. Por tanto, el estado i es recurrente y

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i(V_i) = \infty.$$

II) Si $f_i = \mathbb{P}_i(T_i < \infty) < 1$, entonces por el lema 7.8

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i(V_i) = \sum_{r=0}^{\infty} \mathbb{P}_i(V_i > r) = \sum_{r=0}^{\infty} f_i^r = \lim_{r\to\infty} \frac{1-f_i^r}{1-f_i} = \frac{1}{1-f_i} < \infty.$$

Ya que $f_i < 1$. Entonces, como $\mathbb{E}_i(V_i) < \infty$, $\mathbb{P}_i(V_i = \infty) = 0$ y, por tanto, i es transitorio. □

A continuación veremos que la recurrencia y la transitoriedad tienen propiedades sobre las clases de comunicación.

Teorema 7.10. *Sea C una clase de comunicación, entonces todos los estados de C son recurrentes o todos son transitorios.*

Demostración. Sean los estados $i, j \in C$ cualesquiera y suponemos que el estado i es transitorio. Entonces existen $n, m \geq 0$ con $p_{ij}^{(n)} > 0$ y $p_{ji}^{(m)} > 0$. Teniendo esto en cuenta,

$$p_{ii}^{(n+r+m)} \geq p_{ij}^{(n)} p_{jj}^{(r)} p_{ji}^{(m)}$$

por tanto,

$$\sum_{r=0}^{\infty} p_{jj}^{(r)} \leq \frac{1}{p_{ij}^{(n)} p_{ji}^{(m)}} \sum_{r=0}^{\infty} p_{ii}^{(n+r+m)} < \infty$$

aplicando el Teorema 7.9 con el estado i transitorio. Por tanto, por el Teorema 7.9, j también es transitorio. $\square$

Gracias al Teorema 7.10 podemos hablar de clases recurrentes o transitorios.

Teorema 7.11. *Todas las clases recurrentes son cerradas.*

Demostración. Supongamos que la clase C no es cerrada, entonces existe $i \in C$, $j \notin C$ y $m \geq 1$ con

$$\mathbb{P}_i(\mathcal{X}_m = j) > 0.$$

Como tenemos

$$\mathbb{P}_i(\{\mathcal{X}_m = j\} \cap \{\mathcal{X}_n = i \text{ para infinitos n}\}) = 0$$

implica que

$$\mathbb{P}_i(\mathcal{X}_n = i \text{ para infinitos n}) < 1$$

entonces i no es recurrente, y por el Teorema 7.10, C tampoco es recurrente. $\square$

Teorema 7.12. *Todas las clases cerrada finita es recurrente.*

Demostración. Supongamos la clase C es cerrada y finita y la cadena $(\mathcal{X}_n)_{n\geq 0}$ empieza en C. Entonces para algún $i \in C$ tenemos

$$0 < \mathbb{P}_i(\mathcal{X}_n = i \text{ para infinitos n}) = \mathbb{P}_i(\mathcal{X}_n = i \text{ para algún n})\mathbb{P}_i(\mathcal{X}_n = i \text{ para infinitos n})$$

aplicando Teorema 6.2, la propiedad fuerte de Markov. Esto implica que

$$0 < \mathbb{P}_i(\mathcal{X}_n = i \text{ para infinitos n})$$

entonces i no es transitorio. Por tanto, aplicando el Teorema 7.9 y el Teorema 7.10, C es recurrente. $\square$

Gracias al teorema anterior, podemos identificar rápidamente si una clase finita es recurrente o transitorio.

Teorema 7.13. *Supongamos P es irreductible y recurrente, entonces para todo $j \in I$ tenemos $\mathbb{P}(T_j < \infty) = 1$.*

Demostración. Por la propiedad de Markov tenemos

$$\mathbb{P}(T_j < \infty) = \sum_{i \in I} \mathbb{P}(\mathcal{X}_0 = i)\mathbb{P}_i(T_j < \infty)$$

entonces es suficiente ver si $\mathbb{P}_i(T_j < \infty) = 1$ para todo $i \in I$.

Escogemos una m tales que $p_{ij}^{(m)} > 0$, entonces, por el Teorema 7.9 tenemos

$$1 = \mathbb{P}_j(\mathcal{X}_n = j \text{ para infinitos n}) = \mathbb{P}_j(\mathcal{X}_n = j \text{ para algunos } n \geq m+1)$$

$$= \sum_{k \in I} \mathbb{P}_j(\mathcal{X}_n = j \text{ para algunos } n \geq m+1 | \mathcal{X}_m = k)\mathbb{P}_j(\mathcal{X}_m = k) = \sum_{k \in I} \mathbb{P}_k(T_j < \infty) p_{jk}^{(m)}$$

aplicando la propiedad de Markov en la última igualdad. Por tanto, $\mathbb{P}_i(T_j < \infty) = 1$, ya que $\sum_{k \in I} p_{jk}^{(m)} = 1$. □

8. Distribución invariante

Durante el estudio de las cadenas de Markov podemos encontrar la noción de la distribución invariante o medida invariante. Recordemos que una medida λ es un vector $(\lambda_i : i \in I)$ no negativas y decimos que λ es invariante si $\lambda P = \lambda$.

A continuación veremos los términos de equilibrio y estacionario.

Teorema 8.1. *Sea $(\mathcal{X}_n)_{n\geq 0})$ una cadena de Markov(λ, P) y suponemos que λ es invariante para P. Entonces, $(\mathcal{X}_{m+n})_{n\geq 0})$ también es de Markov(λ, P).*

Demostración. Por el Teorema 3.13 tenemos

$$\mathbb{P}(\mathcal{X}_m = i) = (\lambda P^m))_i = \lambda_i, \forall i.$$

Entonces, condicionado en $\mathcal{X}_{m+n} = i$, $\mathcal{X}_{m+n+1}$ es independiente de $\mathcal{X}_m, \mathcal{X}_{m+1}, \ldots, \mathcal{X}_{m+n}$ y tienen distribución $(p_{ij} : j \in I)$. $\square$

Este teorema nos explica los términos estacionario. El siguiente resultado nos explica los términos de equilibrio.

Teorema 8.2. *Sea I finito. Supongamos que para alguna $i \in I$ tales que $p_{ij}^{(n)} \to \pi_j$ cuando $n \to \infty$, $\forall j \in I$. Entonces $\pi = (\pi_j : j \in I)$ es una distribución invariante.*

Demostración. Tenemos

$$\sum_{j\in I} \pi_j = \sum_{j\in I} \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{j\in I} p_{ij}^{(n)} = 1$$

y

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{k\in I} p_{ik}^{(n)} p_{kj} = \sum_{k\in I} \lim_{n\to\infty} p_{ik}^{(n)} p_{kj} = \sum_{k\in I} \pi_k p_{kj}$$

donde usamos la finitud de I para justificar el cambio de la operación límite con el sumatorio. Por lo tanto, π es una distribución invariante. $\square$

Observación 8.3. Este teorema no nos sirve para indicar la relación entre las distribuciones invariantes con la probabilidad de la n-ésima transición.

Ejemplos 8.4. Consideremos una cadena de Markov de dos estados con la siguiente matriz de transición.

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

Ignorando los casos triviales $\alpha = \beta = 0$ y $\alpha = \beta = 1$. Queremos ver a qué tiende P^n cuand $n \to \infty$.

Sea $P^{n+1} = P^n P$, entonces $p_{11}^{(n+1)} = p_{12}^{(n)}\beta + p_{11}^{(n)}(1-\alpha)$. Sabiendo que

$$p_{12}^{(n)} + p_{11}^{(n)} = \mathbb{P}_1(\mathcal{X}_n = 1 \text{ o } 2) = 1.$$

Aislamos $p_{12}^{(n)}$ y substituimos en la relación anterior y obtenemos la siguiente relación de recurrencia para $p_{11}^{(n)}$

$$p_{11}^{(n+1)} = (1-p_{11}^{(n)})\beta + p_{11}^{(n)}(1-\alpha) = (1-\alpha-\beta)p_{11}^{(n)} + \beta$$

donde $p_{11}^{(0)} = 1$. Tiene una solución única (ver preliminar sobre relación de recurrencia de primer orden)

$$p_{11}^{(n)} = \begin{cases} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \text{si } \alpha+\beta>0 \\ 1 & \text{si } \alpha+\beta=0. \end{cases}$$

Como $p_{11}^{(n)} + p_{12}^{(n)} = 1$

$$p_{12}^{(n)} = \begin{cases} 1-\frac{\beta}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \text{si } \alpha+\beta>0 \\ 0 & \text{si } \alpha+\beta=0. \end{cases}$$

Análogamente para calcular $p_{22}^{(n)}$ y obtenemos

$$p_{22}^{(n)} = \begin{cases} \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n & \text{si } \alpha+\beta>0 \\ 1 & \text{si } \alpha+\beta=0. \end{cases}$$

Como $p_{21}^{(n)} + p_{22}^{(n)} = 1$

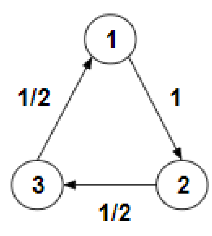
$$p_{21}^{(n)} = \begin{cases} 1-\frac{\alpha}{\alpha+\beta} - \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n & \text{si } \alpha+\beta>0 \\ 0 & \text{si } \alpha+\beta=0. \end{cases}$$

Por tanto, cuando $n \longrightarrow \infty$

$$P^n = \begin{pmatrix} p_{11}^{(n)} & p_{12}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{\beta}{\alpha+\beta} & 1-\frac{\beta}{\alpha+\beta} \\ 1-\frac{\alpha}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix}$$

ya que $(1-\alpha-\beta)^n \longrightarrow 0$ porque $(1-\alpha-\beta) < 1$. Entonces, por el Teorema 8.2, la distribución $(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta})$ tiene que ser invariante.

Ejemplos 8.5. Consideremos la cadena de Markov $(\mathcal{X}_n)_{n\geq 0}$ con el diagrama



y su matriz de transición P

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

Para encontrar la distribución invariante, π, tiene que cumplir $\pi P = \pi$. Es decir, resolver el sistema de ecuaciones

$$\begin{cases} \pi_1 = \frac{1}{2}\pi_3 \\ \pi_2 = \pi_1 + \frac{1}{2}\pi_2 \\ \pi_3 = \frac{1}{2}\pi_2 + \frac{1}{2}\pi_3 \end{cases} \Rightarrow \begin{cases} \pi_3 = t \\ \pi_2 = \pi_3 = t \\ \pi_1 = \frac{1}{2}t \end{cases}, \text{ donde } t \in \mathbb{R}.$$

El sistema es compatible indeterminado. Pero, como π es una distribución, tiene que satisfacer

$$\pi_1 + \pi_2 + \pi_3 = 1.$$

Por tanto, $t = \frac{2}{5}$, y $\pi = (\frac{1}{5}, \frac{2}{5}, \frac{2}{5})$.

A continuación veremos dos resultados que nos muestra que todas las matrices estocásticas irreducibles y recurrentes, P, tienen una única medida invariante positiva. Como cada fila de la matriz P suman 1, entonces, por la existencia de un vector fila invariante visto en álgebra lineal, un vector de unos es un vector propio con valor propio 1. Por tanto, la matriz P debe tener una fila de vector propio con valor propio 1.

Definición 8.6. *Para un estado k fijado, para cada estado i, el* ***tiempo esperado pasado*** *de las visitas de i a k es*

$$\gamma_i^k = \mathbb{E}_k \sum_{n=0}^{T_k-1} \mathbf{1}_{\{\mathcal{X}_n = i\}}.$$

Estamos contando las veces que pasan $\mathcal{X}_n = i$ antes del primer tiempo de llegada T_k.

Teorema 8.7. *Sea P irreducible y recurrente. Entonces*

I. $\gamma_k^k = 1$*;*

II. $\gamma_k^k = (\gamma_i^k : i \in I)$ *tales que* $\gamma^k P = \gamma^k$*;*

III. $0 < \gamma_i^k < \infty$ *para todo* $i \in I$*.*

Demostración. I. Es trivial por la definición del tiempo esperado pasado.

II. Sean los eventos $\{n \leq T_k\}$ que dependen únicamente de $\mathcal{X}_0, \mathcal{X}_1, \ldots, \mathcal{X}_{n-1}$, donde $n = 1, 2, \ldots$. Entonces, por la propiedad de Markov para $n-1$ tenemos

$$\mathbb{P}_k(\mathcal{X}_{n-1} = i, \mathcal{X}_n = j \text{ y } n \leq T_k) = \mathbb{P}_k(\mathcal{X}_{n-1} = i \text{ y } n \leq T_k) p_{ij}.$$

Como P es recurrente, tenemos $T_k < \infty$ y $\mathcal{X}_0 = \mathcal{X}_{T_k} = k$ con probabilidad 1 en $\mathbb{P}_k$. Por tanto,

$$\gamma_j^k = \mathbb{E}_k \sum_{n=1}^{T} {}_k \mathbf{1}_{\{\mathcal{X}_n = j\}} = \mathbb{E}_k \sum_{n=1}^{\infty} \mathbf{1}_{\{\mathcal{X}_n = j \text{ y } n \leq T_k\}} = \sum_{n=1}^{\infty} \mathbb{P}_k(\mathcal{X}_n = j \text{ y } n \leq T_k)$$

$$= \sum_{i \in I} \sum_{n=1}^{\infty} \mathbb{P}_k(\mathcal{X}_{n-1} = i, \mathcal{X}_n = j \text{ y } n \leq T_k) = \sum_{i \in I} \sum_{n=1}^{\infty} \mathbb{P}_k(\mathcal{X}_{n-1} = i \text{ y } n \leq T_k) p_{ij}$$

$$= \sum_{i \in I} p_{ij} \mathbb{E}_k \sum_{m=0}^{\infty} \mathbf{1}_{\{\mathcal{X}_m = i \text{ y } m \leq T_k - 1\}} = \sum_{i \in I} p_{ij} \mathbb{E}_k \sum_{m=0}^{T_k - 1} \mathbf{1}_{\{\mathcal{X}_m = i\}} = \sum_{i \in I} \gamma_i^k p_{ij}.$$

III. Como P es irreducible, por cada estado i existen $n, m \geq 0$ tales que $p_{ik}^{(n)}, p_{ki}^{(m)} > 0$. Entonces, $\gamma_i^k \geq \gamma_k^k p_{ki}^{(m)} > 0$ y $\gamma_i^k p_{ik}^{(n)} \leq \gamma_k^k = 1$ por I) y II).

$\square$

Teorema 8.8. *Sean P irreducible y λ una mesura invariante de P con $\lambda_k = 1$. Entonces $\lambda \geq \gamma^k$. Si P es irreducible y recurrente, entonces obtenemos la igualdad $\lambda = \gamma^k$.*

Demostración. Para cada $j \in I$ tenemos

$$\lambda_j = \sum_{i_1 \in I} \lambda_{i_1} p_{i_1 j} = \sum_{i_1 \neq k} \lambda_{i_1} p_{i_1 j} + p_{kj}$$

$$= \sum_{i_1, i_2 \neq k} \lambda_{i_2} p_{i_2 i_1} p_{i_1 j} + \Big(p_{kj} + \sum_{i_1 \neq k} p_{k i_1} p_{i_1 j}\Big)$$

$$\vdots$$

$$= \sum_{i_1, \ldots, i_n \neq k} \lambda_{i_n} p_{i_n i_{n-1}} \cdots p_{i_1 j} + \Big(p_{kj} + \sum_{i_1 \neq k} p_{k i_1} p_{i_1 j} + \cdots + \sum_{i_1, \ldots, i_{n-1} \neq k} p_{k i_{n-1}} \cdots p_{i_2 i_1} p_{i_1 j}\Big).$$

Entonces cuando $j \neq k$ tenemos

$$\lambda_j \geq \mathbb{P}_k(\mathcal{X}_1 = j \text{ y } T_k \geq 1) + \mathbb{P}_k(\mathcal{X}_2 = j \text{ y } T_k \geq 2) + \cdots + \mathbb{P}_k(\mathcal{X}_n = j \text{ y } T_k \geq n) \to \gamma_j^k$$

cuando $n \to \infty$. Por tanto, $\lambda \geq \gamma^k$.

Si P es recurrente, entonces γ^k es invariante por el Teorema 8.7. Por tanto, $\mu = \lambda - \gamma^k$ también es invariante y $\mu > 0$. Como P también es irreducible, dado $i \in I$, tenemos $p_{ik}^{(n)} > 0$ para algún n, y $0 = \mu_k = \sum_{j \in I} \mu_j p_{jk}^{(n)} \geq \mu_i p_{ik}^{(n)}$. Por tanto, $\mu_i = 0$. $\square$

Recoredemos que un estado i es recurrente si

$$\mathbb{P}_i(\mathcal{X}_n = i, \text{ por infinitos n}) = 1$$

o

$$\mathbb{P}_i(T_i < \infty) = 1$$

por el Teorema 7.9.

Definición 8.9. *Un estado i es **recurrente positiva** si el tiempo de retorno esperado, $m_i = \mathbb{E}_i(T_i)$ es finito.*

*Un estado recurrente que no cumple esta propiedad se les conocen como **recurrente nulo**.*

Teorema 8.10. *Sea P irreducible. Entonces son equivalente,*

I. *Todos los estados son recurrentes positivas;*

II. *algún estado i es recurrente positiva;*

III. *P tiene una distribución invariante, π. Además, cuando cumple III) tenemos $m_i = \frac{1}{\pi_i}$, $\forall i$.*

Demostración. I) $\Rightarrow$ II) es trivial.

II) $\Rightarrow$ III) Si el estado i es recurrente positiva, i también es recurrente, entonces P es recurrente. Por el Teorema 8.7, γ^i es invariante. Pero

$$\sum_{j \in I} \gamma_j^i = m_i < \infty$$

por tanto, $\pi_j = \frac{\gamma_j^i}{m_i}$ se define como una distribución invariante.

III) $\Rightarrow$ I) Escojamos un estado k cualesquiera. Como P es irreducible y $\sum_{i \in I} \pi_i = 1$ tenemos

$$\pi_k = \sum_{i \in I} \pi_i p_{ik}^{(n)} > 0$$

para alguna n. Definamos $\lambda_i = \frac{\pi_i}{\pi_k}$. Entonces λ es una mesura invariante con $\lambda_k = 1$. Por tanto, por el Teorema 8.8, $\lambda \geq \gamma^k$. Por ende,

$$m_k = \sum_{i \in I} \gamma_i^k \leq \sum_{i \in I} \frac{\pi_i}{\pi_k} = \frac{1}{\pi_k} < \infty \tag{8.1}$$

y el estado k es recurrente positiva.

Si P también es recurrente, entonces $\lambda = \gamma^k$ y la desigualdad de 8.1 sería una igualdad. $\square$

9. Convergencia al equilibrio

A continuación veremos cómo es el comportamiento de los límites de la probabilidad transitoria, $p_{ij}^{(n)}$, cuando $n \to \infty$. Por el teorema 8.2, si hay el espacio de estados es finito y para algún estado i existe el límite para todo estado j, entonces este debe ser una distribución invariante. Pero el límite no siempre existe.

Ejemplos 9.1. Consideremos una cadena de dos estados con la siguiente matriz de transición.

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

Observamos que

$$P^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_d.$$

Entonces $P^{2n} = I_d$ y $P^{2n+1} = P$ para todo n. Por lo tanto, $p_{ij}^{(n)}$ no converge para todo i, j.

Definición 9.2. *Un estado i es* ***aperiódico*** *si $p_{ii}^{(n)} > 0$ para todo n suficientemente grande.*

Lema 9.3. *Supogamos que P es irreducible y tiene un estado i aperiódico. Entonces, para todo los estados j y k, $p_{ik}^{(n)} > 0$ para todo n suficientemente grande. En particular, todos los estados son aperiódicos.*

Demostración. Existen $r, s \geq 0$ con $p_{ji}^{(r)} > 0$ y $p_{ik}^{(s)} > 0$. Entonces

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

para todo n suficientemente grande. □

A continuación nos encontraremos con dos teoremas tales que usaremos el argumento de aparejamiento para demostrarlos.

Teorema 9.4 (Convergencia al equilibrio)**.** *Sea P irreducible, aperiódico y tiene una distribución invariante π. Sea λ una distribución cualesquiera.*

Supongamos que $(\mathcal{X}_n)_{n\geq 0}$ es Markov(λ, P), entonces

$$\mathbb{P}(\mathcal{X}_n = j) \to \pi_j, \text{ cuando } n \to \infty \text{ para todo } j.$$

En particular, $p_{ij}^{((n)} \to \pi_j$ cuando $n \to \infty$ para todo i, j.

Demostración. Aplicaremos el argumento de pareja.

Sea $(\mathcal{Y}_n)_{n\geq 0}$ Markov(π, P) e independiente de $(\mathcal{X}_n)_{n\geq 0}$. Fijando un estado b de referencia y considerando el conjunto

$$T = \text{ínf}\{n \geq 1 : \mathcal{X}_n = \mathcal{Y}_n = b\}$$

Paso 1. Veremos si $\mathbb{P}(T < \infty) = 1$.

Sea $\mathcal{W}_n = (\mathcal{X}_n, \mathcal{Y}_n)$ una cadena de Markov en IxI con probabilidad de transición

$$\tilde{p}_{(i,k)(j,l)} = p_{ij} p_{kl}$$

y con distribución inicial

$$\mu_{(i,k)} = \lambda_i \pi_k.$$

Como P es aperiódica, para todo i, j, k estados tenemos

$$\tilde{p}^{(n)}_{(i,k)(j,l)} = p^{(n)}_{ik} p^{(n)}_{kl} > 0$$

para todo n suficientemente grande. Entonces, $\tilde{P}$ es irreducible y también tiene una distribución invariante

$$\tilde{\pi}_{(i,k)} = \pi_i \pi_k.$$

Por lo tanto, por el Teorema 8.10, $\tilde{P}$ es recurrente positiva. Pero T es el primer tiempo de llegada de $\mathcal{W}_n$ a (b, b), entonces, por el Teorema 7.13, $\mathbb{P}(T < \infty) = 1$.

Paso 2. Consideremos el proceso

$$\mathcal{Z}_n = \begin{cases} \mathcal{X}_n & \text{si } n < T \\ \mathcal{Y}_n & \text{si } n \geq T \end{cases}$$

Veremos si $(\mathcal{Z}_n)_{n\geq 0}$ es Markov(λ, P).

Aplicando la propiedad fuerte de Markov a $(\mathcal{W}_n)_{n\geq 0}$ en el tiempo T, tenemos que la cadena $(\mathcal{X}_{T+n}, \mathcal{Y}_{T+n})_{n\geq 0}$ es Markov$(\delta_{(b,b)}, \tilde{P})$ e independiente de $(\mathcal{X}_0, \mathcal{Y}_0), (\mathcal{X}_1, \mathcal{Y}_1), \ldots, (\mathcal{X}_T, \mathcal{Y}_T)$. Por simetría, podemos reeemplazar el proceso $(\mathcal{X}_{T+n}, \mathcal{Y}_{T+n})_{n\geq 0}$ por $(\mathcal{Y}_{T+n}, \mathcal{X}_{T+n})_{n\geq 0}$ tales que ambos son de Markov$(\delta_{(b,b)}, \tilde{P})$ y mantienen la independencia con $(\mathcal{X}_0, \mathcal{Y}_0), (\mathcal{X}_1, \mathcal{Y}_1), \ldots,$ $(\mathcal{X}_T, \mathcal{Y}_T)$.

Por lo tanto, $\mathcal{W}'_n = (\mathcal{Z}_n, \mathcal{Z}'_n)$ es Markov$(\mu, \tilde{P})$ tales que

$$\mathcal{Z}'_n = \begin{cases} \mathcal{Y}_n & \text{si } n < T \\ \mathcal{X}_n & \text{si } n \geq T. \end{cases}$$

En particular, $(\mathcal{Z}_n)_{n\geq 0}$ es Markov$(\lambda.P)$.

Paso 3. Tenemos

$$\mathbb{P}(\mathcal{Z}_n = j) = \mathbb{P}(\mathcal{X}_n = j \text{ y } n < T) + \mathbb{P}(\mathcal{Y}_n = j \text{ y } n \geq T)$$

entonces

$$|\mathbb{P}(\mathcal{X}_n = j) - \pi_j| = |\mathbb{P}(\mathcal{Z}_n = j) - \mathbb{P}(\mathcal{Y}_n = j)|$$

$$= |\mathbb{P}(\mathcal{X}_n = j \text{ y } n < T) - \mathbb{P}(\mathcal{Y}_n = j \text{ y } n < T)| \leq \mathbb{P}(n < T)$$

y $\mathbb{P}(n < T) \to 0$ cuando $n \to \infty$.

$\square$

Ahora veremos casos que excluimos en el último teorema, es decir, cuando $(\mathcal{X}_n)_{n\geq 0}$ es periódico, transitorio o recurrente nulo.

Teorema 9.5. *Sea P irreducible, consideremos un entero $d \geq 1$ y una partición*

$$I = C_0 \cup C_1 \cup \cdots \cup C_{d-1}$$

tales que (teniendo en cuenta que $C_{nd+r} = C_r$)

I. $p_{ij}^{(n)} > 0$ *solo si $i \in C_r$ y $j \in C_{r+n}$ para alguna r;*

II. $p_{ij}^{(nd)} > 0$ *para todo n suficientemente grande, para todo $i, j \in C_r$ y para todo r.*

Demostración. Fijamos el estado k y consideramos el conjunto $S = \{n \geq 0 : p_{kk}^{(n)} > 0\}$. Escogemos $n_1, n_2 \in S$ con $n_1 < n_2$ tales que $d := n_2 - n_1$ sea lo más pequeño posible.

Notación 9.6. $:=$ *es una igualdad por definición.*

Definamos

$$C_r = \{i \in I : p_{ki}^{(nd+r)} > 0 \text{ para alguna } n > 0\}$$

para $r = 0, 1, \ldots, d-1$. Entonces $C_0 \cup C_1 \cup \cdots \cup C_{d-1} = I$, por irreductibilidad.

Además, si $p_{ki}^{(nd+r)} > 0$ y $p_{ki}^{(nd+r)} > 0$ para alguna $r, s \in \{0, 1, \ldots, d-1\}$, entonces, escogiendo un $m \geq 0$ tales que $p_{ik}^{(m)} > 0$ tenemos $p_{kk}^{(md+r+n)} > 0$ y $p_{kk}^{(md+s+n)} > 0$. Esto implica que $r = s$ por la minimalidad de d. Por tanto, obtenemos la partición.

I. Supongamos que $p_{ij}^{(n)}$ e $i \in C$. Escogemos una m tales que $p_{ki}^{(md+r)} > 0$, entonces $p_{ki}^{(md+r+n)} > 0$. Esto implica que $j \in C_{r+n}$, como queríamos.

   Si $i = j = k$, veremos que d debe dividir a cada elemento de S. En particular, a n_1.

   Sea $nd \geq n_1^2$, podemos escribir $nd = qn_1 + r$ para los enteros $q > n_1$ y $0 \leq r \leq n_1 - 1$. Como d divide a n_1, tendremos $r = md$ para algún entero m. Entonces, $nd = (q-m)n_1 + mn_2$. Por tanto,

$$p_{kk}^{(nd)} \geq (p_{kk}^{(n_1)})^{(q-m)} (p_{kk}^{(n_2)})^m > 0$$

   y por lo tanto, $nd \in S$.

II. Sean $i, j \in C_r$, escogemos m_1 y m_2 tales que $p_{ik}^{(m_1)} > 0$ y $p_{kj}^{(m_2)} > 0$, entonces

$$p_{ij}^{(m_1+nd+m_2)} \geq p_{ik}^{(m_1)} p_{kk}^{(nd)} p_{kj}^{(m_2)} > 0$$

   cuando $nd \geq n_1^2$. Como $m_1 + m_2$ tiene que ser un múltiplo de d, obtenemos el resultado.

□

Llamaremos d como el período de la matriz de transición P. El teorema anterior nos muestra que para todo $i \in I$ tales que d es el máximo común divisor del conjunto $\{n \geq 0 : p_{ii}^{(n)} > 0\}$. Esto nos será útil para identificar la d.

Por último, tenemos una descripción completa del comportamiento del límite para cadenas irreducibles. Esto generaliza el teorema 9.4 en dos aspectos, ya que no requiere ni aperiocidad ni existencia de una distribución invariante. Para el caso de la recurrencia nula usamos el argumento descubierto por B. Fristedt y L. Gray.

Teorema 9.7. *Sea P irreducible de período d y sean $C_0, C_1, \dots, C_{d-1}$ particiones obtenidas por el teorema 9.5. Sea λ una distribución tales que $\sum_{i \in C_0} \lambda_i = 1$. Supongamos que $(\mathcal{X}_n)_{n \geq 0}$ es Markov(λ, P). Entonces para $r = 0, 1, \dots, d-1$ tenemos*

$$\mathbb{P}(\mathcal{X}_{nd+r} = j) \to \frac{d}{m_j}, \textit{ si } n \to \infty$$

donde m_j es el tiempo de retorno esperado al estado j. En particular, para $i \in C_0$ y $j \in C_r$ tenemos

$$p_{ij}^{(nd+r)} \to \frac{d}{m_j}, \textit{ si } n \to \infty$$

Demostración. **Paso 1**. Solo nos centramos en los casos aperiódicos.

Sea $\nu = \lambda P^r$, entonces por el teorema 9.5 tenemos

$$\sum_{i \in C_r} \nu_i = 1.$$

Sea $\mathcal{Y}_n = \mathcal{X}_{nd+r}$ entonces $(\mathcal{Y}_n)_{n \geq 0}$ es Markov(λ, P^d) y, por el teorema 9.5, P^d es irreducible y aperiódico en C_r. Para $j \in C_r$, el tiempo de retorno esperado de $(\mathcal{Y}_n)_{n \geq 0}$ al estado j es $\frac{m_j}{d}$. Entonces, si el teorema se cumple para los casos aperiódicos, esto implica que

$$\mathbb{P}(\mathcal{X}_{nd+r} = j) = \mathbb{P}(\mathcal{Y}_n = j) \to \frac{d}{m_j}, \text{ si } n \to \infty$$

por lo tanto, el teorema se cumpliría en general.

Paso 2. Supongamos que P es aperiódico. Si P es recurrente positiva, entonces $\frac{1}{m_j} = \pi_j$, donde π es la única distribución invariante. En consecuencia, se obtiene el resultado por el teorema 9.4. De lo contrario, $m_j = \infty$ y tendremos que demostrar que

$$\mathbb{P}(\mathcal{X}_n = j) \to 0, \text{ si } n \to \infty.$$

Si P fuera transitorio sería fácil y lo dejaríamos en el caso de recurrencia nula.

Paso 3. Supongamos que P es aperiódico y de recurrencia nula. Entonces

$$\sum_{k=0}^{\infty} \mathbb{P}_j(T_j > k) = \mathbb{E}_j(T_j) = \infty.$$

Dado $\epsilon > 0$, escogemos un K tales que

$$\sum_{k=0}^{K-1} \mathbb{P}_j(T_j > k) \geq \frac{2}{\epsilon}.$$

Entonces, para $n \geq K-1$

$$1 \geq \sum_{k=n-K+1}^{n} \mathbb{P}(\mathcal{X}_k = j \text{ y } \mathcal{X}_m \neq j \text{ para } m = k+1, \ldots, n)$$

$$= \sum_{k=n-K+1}^{n} \mathbb{P}(\mathcal{X}_k = j)\mathbb{P}_j(T_j > n-k)$$

$$= \sum_{k=0}^{K-1} \mathbb{P}(\mathcal{X}_{n-k} = j)\mathbb{P}_j(T_j > k)$$

entonces tenemos $\mathbb{P}(\mathcal{X}_{n-k} = j) \leq \frac{\epsilon}{2}$ para alguna $k \in \{0, 1, \ldots, K-1\}$.

Volveremos a usar el argumento de aparejamiento que habíamos aplicado en la demostración del teorema 9.4. Solo que ahora $(\mathcal{Y}_n)_{n\geq 0}$ es Markov(μ, P), donde μ es escogido más adelante. Sea $\mathcal{W}_n = (\mathcal{X}_n, \mathcal{Y}_n)$. Como antes, la aperiocidad de $(\mathcal{X}_n)_{n\geq 0}$ asegura la irreductibilidad de $(\mathcal{W}_n)_{n\geq 0}$.

Si $(\mathcal{X}_n)_{n\geq 0}$ es transitorio, entonces, tomando $\mu = \lambda$, obtenemos

$$\mathbb{P}(\mathcal{X}_n = j)^2 = \mathbb{P}(\mathcal{W}_n = (j, j)) \to 0, \text{ como queríamos.}$$

Supongamos que $(\mathcal{W}_n)_{n\geq 0}$ es recurrente. Entonces, por el teorema 9.4, tenemos $\mathbb{P}(T < \infty) = 1$ y con el argumento de aparejamiento tenemos

$$|\mathbb{P}(\mathcal{X}_n = j) - \mathbb{P}(\mathcal{Y}_n = j)| \to 0, \text{ si } n \to \infty.$$

Obtenemos esta convergencia tomando $\mu = \lambda P^k$ para $k = 1, \ldots, K-1$, tales que $\mathbb{P}(\mathcal{Y}_n = j) = \mathbb{P}(\mathcal{X}_{n+k} = j)$. Podemos encontrar N tales que

$$|\mathbb{P}(\mathcal{X}_n = j) - \mathbb{P}(\mathcal{X}_{n+k} = j)| \leq \frac{\epsilon}{2}$$

para $n > N$ y $k = 1, \ldots, K-1$.

Pero para cualquier n podemos encontrar una $k \in \{0, 1, \ldots, K-1\}$ tales que $\mathbb{P}(\mathcal{X}_{n+k} = j) \leq \frac{\epsilon}{2}$. Por lo tanto, para $n \geq N$

$$\mathbb{P}(\mathcal{X}_n = j) \leq \epsilon.$$

Como $\epsilon > 0$ era arbitrario,

$$\mathbb{P}(\mathcal{X}_n = j) \to 0, \text{ si } n \to \infty$$

como queríamos.

$\square$

10. Inversión del tiempo

Para las cadenas de Markov, tanto el pasado como el futuro son independientes del presente. Esto también lo podemos apreciar en el tiempo. A continuación veremos cadenas de Markov con el tiempo corriendo hacia atrás.

Teorema 10.1. *Sea P irreducible y tiene una distribución invariante* π. *Supongamos que* $(\mathcal{X}_n)_{0\leq n\leq N}$ *es Markov*(π, P) *y consideremos* $\mathcal{Y}_n = \mathcal{X}_{N-n}$. *Entonces,* $(\mathcal{Y}_n)_{0\leq n\leq N}$ *es Markov*$(\pi, \widehat{P})$, *donde* $\widehat{P} = (\widehat{p}_{ij})$ *está definido por*

$$\pi_j \widehat{p}_{ji} = \pi_i p_{ij}, \text{ para todo } i, j.$$

y $\widehat{P}$ *también es irreducible con distribución invariante* π.

Demostración. Primero veremos si $\widehat{P}$ es una matriz estocástica.

$$\sum_{j\in I} \widehat{p}_{ji} = \frac{1}{\pi_j} \sum_{j\in I} \pi_i p_{ij} = 1$$

ya que π es invariante para P. Por tanto, $\widehat{P}$ es estocástica. Ahora miramos si π es invariante para $\widehat{P}$.

$$\sum_{j\in I} \pi_j \widehat{p}_{ji} = \sum_{j\in I} \pi_i p_{ij} = \pi_i \sum_{j\in I} p_{ij} = \pi_i$$

ya que P es una matriz estocástica. Por tanto, π es invariante para $\widehat{P}$.

Después, veremos si $(\mathcal{Y}_n)_{0\leq n\leq N}$ es Markov$(\pi, \widehat{P})$.

Sea

$$\begin{aligned}\mathbb{P}(\mathcal{Y}_0 = i_0, \mathcal{Y}_1 = i_1, \ldots, \mathcal{Y}_N = i_N) &= \mathbb{P}(\mathcal{X}_0 = i_N, \mathcal{X}_1 = i_{N-1}, \ldots, \mathcal{X}_N = i_0)\\ &= \pi_{i_N} p_{i_N i_{N-1}} \cdots p_{i_1 i_0} = \pi_{i_0} \widehat{p}_{i_N i_{N-1}} \cdots \widehat{p}_{i_1 i_0}.\end{aligned}$$

Por el teorema 3.8, $(\mathcal{Y}_n)_{0\leq n\leq N}$ es Markov$(\pi, \widehat{P})$.

Por último, como P es irreducible, para cada par de estado i, j hay una cadena de estados $i_0 = i, i_1, \ldots, i_{n-1}, i_n = j$ con $p_{i_0 i_1} \cdots p_{i_{n-1} i_n} > 0$. Entonces

$$\widehat{p}_{i_n i_{n-1}} \cdots \widehat{p}_{i_1 i_0} = \frac{\pi_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}}{\pi_{i_n}} > 0.$$

Por lo tanto, $\widehat{P}$ también es irreducible. □

Definición 10.2. *Una matriz estocástica P y una medida* λ *están en equilibrio detallado si*

$$\lambda_i p_{ij} = \lambda_j p_{ji}, \text{ para todo } i, j.$$

Lema 10.3. *Si P y* λ *están en equilibrio detallado, entonces* λ *es invariante para P.*

Demostración. Tenemos

$$(\lambda P)_i = \sum_{j\in I} \lambda_j p_{ji} = \sum_{j\in I} \lambda p_{ij} = \lambda_i$$

□

Definición 10.4. *Sea $(\mathcal{X}_n)_{n\geq 0}$ Markov(λ, P), con P irreducible. $(\mathcal{X}_n)_{n\geq 0}$ es reversible si para todo $N \geq 1$, $(\mathcal{X}_{N-n})_{0\leq n\leq N}$ también es Markov(λ, P).*

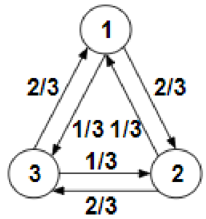
Teorema 10.5. *Sea P una matriz estocástica y irreducible y sea λ una distribución. Supongamos que $(\mathcal{X}_n)_{n\geq 0}$ es Markov(λ, P). Entoces son equivalentes,*

I. *$(\mathcal{X}_n)_{n\geq 0}$ es reversible;*

II. *P y λ están en equilibrio detallado.*

Demostración. Tanto I como II implica que λ sea invariante para P. Entonces I y II equivalen a que $\widehat{P} = P$ por el teorema 10.1. □

Ejemplos 10.6. Consideremos una cadena de Markov con el siguiente diagrama



Su matriz de transición es

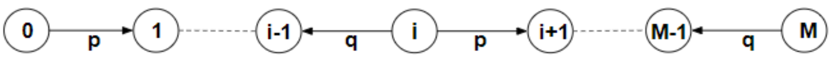
$$P = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{pmatrix}$$

y $\pi = (1/3, 1/3, 1/3)$ es invariante. Observamos que la matriz de transición P es irreducible porque la única clase cerrada es I. Como

$$\widehat{p}_{ji} = \frac{\pi_i p_{ij}}{\pi_j} = p_{ij}, \text{ para todo i, j}$$

ya que $\pi_i = 1/3$ para $i = 1, 2, 3$. Entonces, $\widehat{P} = P^T$, la transpuesta de P. Pero P no es simétrica. Por lo tanto, $P \neq \widehat{P}$ y esta cadena no es reversible.

Ejemplos 10.7. Consideremos una cadena de Markov con el siguiente diagrama



donde $0 < p = 1 - q < 1$. Observamos que, como en el ejemplo anterior, la matriz de transición P es irreducible, ya que la única clase cerrada es I.

Sea λ una medida que está en equilibrio detallado con P y satisface

$$\lambda_i p_{i,i+1} = \lambda_{i+1} p_{i+1,i}, \text{ para } i = 0, 1, \ldots, M-1. \tag{10.1}$$

La ecuación (10.1) es una relación de recurrencia de primer grado, donde su solución es

$$\lambda_{i+1} = \frac{p_{i,i+1}}{p_{i+1,i}} \lambda_i = \frac{p}{q} \lambda_i = \left(\frac{p}{q}\right)^2 \lambda_{i-1} = \cdots = \left(\frac{p}{q}\right)^{i+1} \lambda_0.$$

Por tanto,
$$\lambda = \left(\left(\frac{p}{q}\right)^i : i = 0, 1, \ldots, M\right), \text{ donde } \lambda_0 = 1.$$
Y podemos normalizar esta medida para que sea una distribución en equilibrio detallado con P. Por tanto, esta cadena es reversible.

11. Teorema ergódico

El teorema ergódico está relacionado con los comportamientos de los límites de los promedios a lo largo del tiempo. A continuación veremos cómo el teorema identifica, para cadenas de Makorv, la proporción a largo plazo del tiempo pasado en cada estado. También veremos una herramienta sobre el teorema para las variables aleatorias independientes, que es una versión de la Ley fuerte de los grandes números.

Teorema 11.1 (Ley fuerte de los grandes números)**.** *Sea $\mathcal{Y}_1, \mathcal{Y}_2, \dots$ una sucesión de variables aleatorias independientes, idénticamente distribuido y no negativos con $\mathbb{E}(\mathcal{Y}_1) = \mu$. Entonces*

$$\mathbb{P}\Big(\frac{\mathcal{Y}_1 + \cdots + \mathcal{Y}_n}{n} \to \mu, \textit{ si } n \to \infty\Big) = 1.$$

Demostración. Para el caso $\mu < \infty$, tenemos la demostración en Probability with martigales de David Willians (Cambridg Univerity Press, 1991).

Si $\mu = \infty$. Fijamos una $N < \infty$ y sea $\mathcal{Y}_n^{(N)} = \mathcal{Y}_n \wedge N$, entonces

$$\frac{\mathcal{Y}_1 + \cdots + \mathcal{Y}_n}{n} \geq \frac{\mathcal{Y}_1^{(N)} + \cdots + \mathcal{Y}_n^{(N)}}{n} \to \mathbb{E}(\mathcal{Y}_1 \wedge N), \text{ si } n \to \infty$$

con probabilidad 1. Si

$$N \to \infty \Rightarrow \mathbb{E}(\mathcal{Y}_1 \wedge N) \to \mu$$

por la convergencia monótona (ver sección 6.4 del [1]). Por lo tanto, obtenemos

$$\frac{\mathcal{Y}_1 + \cdots + \mathcal{Y}_n}{n} \to \infty, \text{ si } n \to \infty$$

con probabilidad 1. □

Definición 11.2. *Sea $V_i(n)$ el número de las veces que visita al estado i antes del paso n y*

$$V_i(n) = \sum_{k=0}^{n-1} \mathbf{1}_{\{\mathcal{X}_k = i\}}.$$

Entonces $\frac{v_i(n)}{n}$ es la proporción del tiempo gastado antes del paso n en el estado i.

A continuación, veremos los resultados obtenidos de la proporción del tiempo gastados a largo plazo por una cadena de Markov en cada estado.

Teorema 11.3 (Teorema ergódico)**.** *Sea P irreducible y λ una distribución cualesquiera. Si $(\mathcal{X}_n)_{n\geq 0}$ es Markov(λ, P) entonces*

$$\mathbb{P}\Big(\frac{V_i(n)}{n} \to \frac{1}{m_i}, \textit{ si } n \to \infty\Big) = 1.$$

Donde $m_i = \mathbb{E}_i(T_i)$ es el tiempo de retorno esperado en el estado i. Además, en el caso de la recurrencia positiva, cualquiera función acotada $f : I \to \mathbb{R}$ tendremos

$$\mathbb{P}\Big(\frac{1}{n}\sum_{k=0}^{n-1} f(\mathcal{X}_k) \to \overline{f}, \textit{ si } n \to \infty\Big) = 1.$$

Donde
$$\overline{f} = \sum_{i \in I} \pi_i f_i.$$
Y $(\pi_i : i \in I)$ *es la única distribución invariante.*

Demostración. Si P fuera transitorio, entonces, teniendo probabilidad 1, el número total V_i de visitas al estado i es finito. Por tanto,
$$\frac{V_i(n)}{n} \geq \frac{V_i}{n} \to 0 = \frac{1}{m_i}.$$
Recordemos que $V_i = \sum_{n=0}^{\infty} \mathbf{1}_{\{\mathcal{X}_n = i\}}$.

Supongamos ahora que P es recurrente. Fijamos un estado i. Para $T = T_i$, tenemos $\mathbb{P}(T < \infty) = 1$, por el teorema 7.13. También tenemos que $(\mathcal{X}_n)_{n \geq 0}$ es Markov(δ_i, P) e independiente de $\mathcal{X}_0, \ldots, \mathcal{X}_T$ por la propiedad fuerte de Markov. La proporción a largo plazo del tiempo pasado en el estado i es la misma tanto para $(\mathcal{X}_{T+n})_{n \geq 0}$ como para $(\mathcal{X}_n)_{n \geq 0}$, entonces es suficiente considerar el caso $\lambda = \delta_i$.

Recordemos que $S_i^{(r)}$ es la distancia de la r - ésima excursión al estado i. Por el Lema 7.6, las variables aleatorias no negativa $S_i^{(1)}, S_i^{(2)}, \ldots$ son independientes e idénticamente distribuidos con $\mathbb{E}_i(S_i^{(r)}) = m_i$. Ahora,
$$S_i^{(n)} + \cdots + S_i^{(V_i(n)-1)} \leq n - 1.$$
Donde la parte izquierda de la desigualdad es el tiempo de la última visita al estado i antes del paso n. También tenemos
$$S_i^{(n)} + \cdots + S_i^{(V_i(n))} \geq n.$$
Donde la parte izquierda de la desigualdad es la última visita al estado i antes del paso $n - 1$. Por lo tanto,
$$\frac{S_i^{(1)} + \cdots + S_i^{(V_i(n)-1)}}{V_i(n)} \leq \frac{n-1}{V_i(n)} \leq \frac{n}{V_i(n)} \leq \frac{S_i^{(1)} + \cdots + S_i^{(V_i(n))}}{V_i(n)}. \tag{11.1}$$
Entonces, por el teorema 11.1,
$$\mathbb{P}\Big(\frac{S_i^{(1)} + \cdots + S_i^{(n)}}{n} \to \infty, \text{ si } n \to \infty\Big) = 1$$
y, como P es recurrente,
$$\mathbb{P}(V_i(n) \to \infty, \text{ si } n \to \infty) = 1.$$
Entonces, cuando $n \to \infty$ en 11.1, tenemos
$$\mathbb{P}\Big(\frac{n}{V_i(n)} \to m_i, \text{ si } n \to \infty\Big) = 1 \Rightarrow \mathbb{P}\Big(\frac{V_i(n)}{n} \to \frac{1}{m_i}, \text{ si } n \to \infty\Big) = 1.$$

Supongamos ahora que $(\mathcal{X}_n)_{n \geq 0}$ tiene una distribución invariante $(\pi_i : i \in I)$. Sea $f : I \to \mathbb{R}$ una función acotada y supongamos que, sin pérdida de generalidad, $|f| \geq 1$. Entonces, para cualquier $J \subseteq I$ tenemos,
$$\left|\frac{1}{n}\sum_{k=0}^{n-1} f(\mathcal{X}_k) - \overline{f}\right| = \left|\frac{1}{n}\sum_{k=0}^{n-1} f(\mathcal{X}_k) - \sum_{i \in I} \pi_i f_i\right| = \left|\sum_{i \in I}\left(\frac{V_i(n)}{n} - \pi_i\right) f_i\right| \leq$$

$$\le \sum_{i\in I}\left|\left(\frac{V_i(n)}{n}-\pi_i\right)\right||f_i| \le \sum_{i\in I}\left|\frac{V_i(n)}{n}-\pi_i\right| \le$$

ya que $|f|\le 1$,

$$\le \sum_{i\in J}\left|\frac{V_i(n)}{n}-\pi_i\right| + \sum_{i\notin J}\left|\frac{V_i(n)}{n}-\pi_i\right| \le \sum_{i\in J}\left|\frac{V_i(n)}{n}-\pi_i\right| + \sum_{i\notin J}\left(\frac{V_i(n)}{n}-\pi_i\right)$$

$$\le 2\sum_{i\in J}\left|\frac{V_i(n)}{n}-\pi_i\right| + 2\sum_{i\notin J}\pi_i.$$

Como habíamos demostrado que

$$\mathbb{P}\left(\frac{V_i(n)}{n}\to\pi_i,\text{ si } n\to\infty \text{ para todo i}\right)=1.$$

Entonces, dado $\epsilon>0$ y escogemos un conjunto J finito tales que

$$\sum_{i\notin J}\pi_i<\frac{\epsilon}{4}$$

y sea $N=N(\omega)$ tales que, para $n\ge N(\omega)$,

$$\sum_{i\in J}\left|\frac{V_i(n)}{n}-\pi_i\right|<\frac{\epsilon}{4}.$$

Por lo tanto, para $n\ge N(\omega)$, tenemos

$$\left|\frac{1}{n}\sum_{k=0}^{n-1}f(\mathcal{X}_k)-\overline{f}\right| \le 2\sum_{i\in J}\left|\frac{V_i(n)}{n}-\pi_i\right| + 2\sum_{i\notin J}\pi_i < 2\frac{\epsilon}{4}+2\frac{\epsilon}{4}=\epsilon.$$

que establece la condición de convergencia. □

Consideremos ahora el problema estadístico de estimación de una matriz transitiva P desconocida en base de las observaciones de la correspondiente cadena de Markov.

Consideremos, para empezar, el caso donde tenemos $N+1$ observaciones $(\mathcal{X}_n)_{0\le n\le N}$. La función log - verosimilitud dado por

$$l(P)=\log(\lambda_{\mathcal{X}_0}p_{\mathcal{X}_0\mathcal{X}_1}\cdots p_{\mathcal{X}_{N-1}\mathcal{X}_N})=\sum_{i,j\in I}N_{ij}\log(p_{ij}).$$

Donde N_{ij} es el número de transiciones del estado i a j.

Un procedimiento estadístico estándar para encontrar el máximo verosimilitud estimado $\widehat{P}$ es la elección de P maximizando $l(P)$. Como P debe cumplir la restricción lineal

$$\sum_j p_{ij}=1,\text{ para cada i.}$$

Primero maximizamos
$$l(P) + \sum_{i,j \in I} \mu_i p_{ij}$$
y luego escogemos $(\mu_i : i \in I)$ para que cumpla las restricciones. Esto es el método de multiplicadores de Lagrange.

Entonces obtendremos
$$\widehat{p}_{ij} = \sum_{n=0}^{N-1} \mathbf{1}_{\{\mathcal{X}_n = i, \mathcal{X}_{n+1} = j\}} \Big/ \sum_{n=0}^{N-1} \mathbf{1}_{\{\mathcal{X}_n = i\}}$$
que es la proporción de saltos del estado i que va al estado j.

Ahora pasamos a ver la consistencia de este tipo de estimación. Esto es para decir que
$$\widehat{p}_{ij} \to p_{ij}$$
con probabilidad 1 si $N \to \infty$. Como esto no se cumple cuando i es transitorio, modificaremos ligeramente el enfoque.

Tengamos en cuenta que para buscar $\widehat{p}_{ij}$ solo hay que maximizar
$$\sum_{j \in I} N_{ij} \log(p_{ij})$$
sujeto a
$$\sum_j p_{ij} = 1.$$

Los demás términos y restricciones son irrelevantes.

Supongamos, entonces, que en lugar de $N+1$ observaciones, hacemos todas las observaciones necesarias para asegurar que la cadena sale del estado i en un total de N veces. En el caso transitorio, esto puede implicar a reiniciar la cadena varias veces.

Para maximizar la verosimilitud del $(p_{ij} : jinI)$, también maximizamos
$$\sum_{j \in I} N_{ij} \log(p_{ij})$$
sujeto a
$$\sum_j p_{ij} = 1.$$
Lo que implica a maximizar la verosimilitud estimada
$$\widehat{p}_{ij} = \frac{N_{ij}}{N}.$$
Pero $N_{ij} = \mathcal{Y}_1 + \cdots + \mathcal{Y}_N$, donde
$$\mathcal{Y}_n = \begin{cases} 1 & \text{si la n-ésima transición va del estado i a j;} \\ 0 & \text{si no.} \end{cases}$$
Por la propiedad fuerte de Markovm $\mathcal{Y}_1, \cdots, \mathcal{Y}_N$ son variables aleatorias independientes e idénticamente distribuidos con media igual a p_{ij}. Entonces, por el teorema 11.1,
$$\mathbb{P}(\widehat{p}_{ij} \to p_{ij}, \text{ si } N \to \infty = 1.$$
Esto demuestra que $\widehat{p}_{ij}$ es **consistente**.

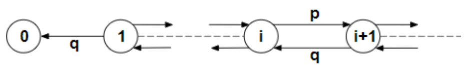
12. Aplicaciones

En esta sección nos concentramos en ver el problema de la ruina del jugador con diversas estrategias. Primero lo haremos a nivel general, luego lo aplicamos en un ejemplo real.

12.1. El problema de la ruina del jugador

Supongamos que un jugador entra en un casino con i euros. En cada apuesta, se apuesta 1 euro. Si gana, obtiene 1 euro y recupera el euro apostado. Definamos I como el espacio de los estados, que representa la fortuna del jugador. En cada apuesta o ronda tiene p probabilidad de ganar y q probabilidad de perder.

Por tanto, tendremos una cadena de Markov con el siguiente diagrama:



donde $0 < p = 1 - q < 1$. Las probabilidades de transición son:

$$\begin{cases} p_{00} = 1; \\ p_{i,i-1} = q, p_{i,i+1} = p, \text{ para todo } i > 0 \text{ .} \end{cases}$$

Observamos que la clase $\{0\}$ es la única clase cerrada y es absorbente.

Entonces, el juego terminará si el jugador se queda sin dinero. Pero, ¿Cuál es la probabilidad de que ocurra?

Sea

$$h_i^{\{0\}} = \mathbb{P}_i(\text{ir a } 0) \text{ y } k_i^{\{0\}} = \mathbb{E}_i(\text{tiempo en llegar a } 0)$$

Entonces, por el teorema 5.7, h es la solución minimal no negativa de

$$\begin{cases} h_0 = 1; \\ h_i = ph_{i+1} + qh_{i-1}, \text{ para todo } i > 0. \end{cases}$$

Vamos a resolver la relación de recurrencia $h_i = ph_{i+1} + qh_{i-1}$. Visto en el preliminar, teníamos resuelto las relaciones de recurrencia de segundo grado a nivel general. Entonces, tomando $a = p$, $b = -1$ y $c = q$

$$p\lambda^2 - \lambda + q = 0$$

$$\lambda = \frac{1 \pm \sqrt{1 - 4pq}}{2p} = \frac{1 \pm \sqrt{1 - 4(1-p)p}}{2p} = \frac{1 \pm \sqrt{1 - 4p + p^2}}{2p} = \frac{1 \pm \sqrt{(2p-1)^2}}{2p}$$

$$= \frac{1 \pm (2p - 1)}{2p} = \begin{cases} \lambda = 1 \\ \lambda = \frac{1-p}{p} = \frac{q}{p} \end{cases}$$

Entonces,

$$h_i = A1^i + B\left(\frac{q}{p}\right)^i = A + B\left(\frac{q}{p}\right)^i \tag{12.1}$$

es la solución de la recurrencia, donde A, B se obtiene resolviendo el sistema de ecuación

$$\begin{cases} h_0 = A + B \\ h_1 = A + \frac{q}{p} B \end{cases} \Rightarrow \begin{cases} 1 = A + B \\ h_1 = A + \frac{q}{p} B \end{cases}$$

Veremos los casos que depende de los valores de p, q.

Si $p < q$, entonces $q/p > 1$. Como tenemos la restricción $0 \leq h_i \leq 1$, esto obliga a $B = 0$. Por tanto $h_i = 1$ para todo i.

Si $p > q$, entonces tenemos una familia de soluciones

$$h_i = A + B\Big(\frac{q}{p}\Big)^i = A + (1 - A)\Big(\frac{q}{p}\Big)^i = \Big(\frac{q}{p}\Big)^i + A\left(1 - B\Big(\frac{q}{p}\Big)^i\right)$$

teniendo en cuenta que $h_0 = 1 = A + B$. Como la solución h es no negativa, $A \geq 0$. Y por la condición de la mínima, A tiene que ser 0. Por tanto,

$$h_i = \Big(\frac{q}{p}\Big)^i, \text{ para todo } i \geq 0.$$

Por último, si $p = q$, entonces tenemos la relación de recurrencia

$$h_i = A + Bi.$$

Como tenemos la restricción $0 \leq h_i \leq 1$, esto obliga a $B = 0$. Por tanto $h_i = 1$ para todo i.

Como resultado, el jugador siempre acabará arruinado, a no ser que encuentre un juego que haya más probabilidad de ganar que de perder. Si no acaba arruinado, entonces la cadena nunca acabaria.

Ahora vamos a calcular el tiempo medio de acabar arruinado. Por el teorema 5.9, tenemos

$$\begin{cases} k_0 = 0; \\ k_i = 1 + pk_{i+1} + qk_{i-1} \quad \text{para todo } i > 0. \end{cases}$$

Tenemos una relación de recurrencia lineal de orden 2 no homogénea. La solución de esta recurrencia es la suma de la solución de la recurrencia homogénea más el caso particular,

$$k_i = k_i^{homogénea} + k_i^{Particular}, \text{ para todo } i \geq 0.$$

Por la ecuación general 12.1, la solución homogénea es

$$k_i^{homogénea} = A + B\Big(\frac{q}{p}\Big)^i.$$

La solución particular debe ser un constante, ya que 1 es un constante. Supongamos que $k_i = C$ es la solución particular, donde C es un constante cualquiera, entonces

$$C = pC + qC + 1 \Rightarrow 0 = -C + (p + q)C + 1 \Rightarrow 0 = -C + C + 1 \Rightarrow 0 = 1$$

ya que $p = 1 - q$. Como no funciona, supongamos ahora que $k_i = Ci$, entonces

$$Ci = p(C(i+1)) + q(C(i-1)) + 1 \Rightarrow Ci(1-p-q) + C(q-p) = 1 \Rightarrow C = \frac{1}{q-p}.$$

Por tanto, la solución general es

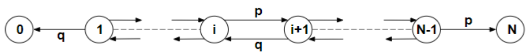
$$k_i = \frac{i}{q-p} + A + B\Big(\frac{q}{p}\Big)^i, \text{ para todo } i \geq 0, \text{ con parámetros A, B} \tag{12.2}$$

Como $k_0 = 0$, tenemos $k_0 = 0 = A + B \Rightarrow A = -B$. Por tanto, tenemos una familia de soluciones

$$k_i = \frac{i}{q-p} + B\Big(\Big(\frac{q}{p}\Big)^i - 1\Big), \text{ para todo } i \geq 0.$$

Por tanto, no podemos calcular con exactitud el tiempo medio de llegar al estado 0.

Ahora supongamos que el jugador se irá del casino cuando obtiene N euros o cuando se arruine. Entonces tenemos la siguiente cadena de Markov con su respectivo diagrama:



Las probabilidades de transición son:

$$\begin{cases} p_{00} = p_{NN} = 1; \\ p_{i,i+1} = p; p_{i,i-1} = q \end{cases}$$

para $i = 1, 2, 3, \ldots, N-1, N$.

I. La distribución inicial es $(0, \ldots, 0, 1, 0, \ldots, 0)$ en la i-ésima posición, que representa el dinero que tiene el jugador.

II. $\{0\}$ y $\{N\}$ son las únicas clases cerradas. Además, también son estados absorbentes.

III. $I - \{0, 1\}$ es una clase no cerrada.

IV. La cadena no es irreducible, ya que I no es la única clase cerrada.

V. Como $\mathbb{P}(T_i = \infty | X_0 = i) < 1$, el estado i es transitorio. Por tanto, la clase $I - \{0, 1\}$ es transitoria, puesto que, $i \in I - \{0, 1\}$.

VI. La cadena tiene estados transitorios y absorbentes, entonces no puede ser ergótica.

VII. Distribución invariante: veremos si existe una distribución v tales que $vP = v$, entonces tenemos el siguiente sistema de ecuaciones:

$$\begin{cases} v_0 + qv_1 = v_0 \\ qv_2 = v_1 \\ pv_1 + qv_3 = v_2 \\ pv_2 + qv_4 = v_3 \\ \vdots \\ pv_{i-1} + qv_{i+2} = v_i \quad \text{para todo } i = 2, 3, \ldots, N-2 \\ \vdots \\ pv_{N-2} = v_{N-1} \\ pv_{N-1} + v_N = v_N \\ \sum_{n=0}^{N} v_n = 1 \end{cases}$$

Notamos que $v_1 = v_2 = \cdots = v_{N-1} = 0$. Por tanto, tenemos

$$v_0 + v_N = 1.$$

Entonces, las distribuciones $(v_0, 0, \ldots, 0, v_N = 1 - v_0)$ son invariantes, con $0 \leq v_0 \leq 1$. Por tanto, tenemos infinitas distribuciones invariantes. Esto es consecuencia de tener dos estados absorbentes.

VIII. Probabilidad de absorción: Primero calculamos la probabilidad de llegar al estado de absorción 0 empezando por el estado i. Aplicaremos el teorema 5.7,

$$\begin{cases} h_0 = 1 \\ h_N = 0 \\ h_i = ph_{i+1} + qh_{i-1} \quad \text{si } 0 < i < N-1. \end{cases}$$

Como $p = 1 - q \Rightarrow p + q = 1$, entonces

$$h_i = (p+q)h_i = ph_i + qh_i = ph_{i+1} + qh_{i-1} \Rightarrow p(h_{i+1} - h_i) = q(h_i - h_{i-1})$$

$$\Rightarrow h_{i+1} - h_i = \frac{q}{p}(h_i - h_{i-1}).$$

Supongamos que $p \neq q$, entonces teniendo en cuenta la condición anterior

$$h_i - h_{i-1} = \left(\frac{q}{p}\right)^{i-1}(h_1 - h_0).$$

Por tanto

$$h_i - h_1 = \sum_{j=2}^{i} h_j - h_{j-1} = \sum_{j=2}^{i} \left(\frac{q}{p}\right)^{j-1}(h_1 - h_0) = (h_1 - h_0)\frac{\frac{q}{p} - (\frac{q}{p})^i}{1 - \frac{q}{p}}. \tag{12.3}$$

Como $h_N = 0$ y $h_0 = 1$, tenemos

$$h_N - h_1 = -h_1 = (h_1 - 1)\frac{\frac{q}{p} - (\frac{q}{p})^N}{1 - \frac{q}{p}}.$$

Es una ecuación de primer grado, con h_1 de incógnita. Por tanto,

$$h_1 = \frac{\frac{\frac{q}{p}-(\frac{q}{p})^N}{1-\frac{q}{p}}}{1+\frac{\frac{q}{p}-(\frac{q}{p})^N}{1-\frac{q}{p}}} = \frac{\frac{q}{p}-(\frac{q}{p})^N}{1-(\frac{q}{p})^N}.$$

Substituyendo este valor en la ecuación 12.3, tenemos

$$\begin{aligned}
h_i = h_1 + (h_1 - 1)\frac{\frac{q}{p}-(\frac{q}{p})^i}{1-\frac{q}{p}} &= \frac{\frac{q}{p}-(\frac{q}{p})^N}{1-(\frac{q}{p})^N} + \left(\frac{\frac{q}{p}-(\frac{q}{p})^N}{1-(\frac{q}{p})^N} - 1\right)\frac{\frac{q}{p}-(\frac{q}{p})^i}{1-\frac{q}{p}} \\
&= \frac{\frac{q}{p}-(\frac{q}{p})^N}{1-(\frac{q}{p})^N} + \left(\frac{\frac{q}{p}-(\frac{q}{p})^N - 1 + (\frac{q}{p})^N}{1-(\frac{q}{p})^N}\right)\frac{\frac{q}{p}-(\frac{q}{p})^i}{1-\frac{q}{p}} \\
&= \frac{\frac{q}{p}-(\frac{q}{p})^N}{1-(\frac{q}{p})^N} - \frac{\frac{q}{p}-(\frac{q}{p})^i}{1-(\frac{q}{p})^N} = \frac{(\frac{q}{p})^i-(\frac{q}{p})^N}{1-(\frac{q}{p})^N}.
\end{aligned}$$

Por tanto, la probabilidad de arruinarse empezando con i euros es

$$h_i = \frac{(\frac{q}{p})^i-(\frac{q}{p})^N}{1-(\frac{q}{p})^N}. \tag{12.4}$$

Ahora solo nos falta ver la probabilidad de obtener N euros empezando con i euros. La resolución del problema es análoga del caso anterior, pero con $h_N = 1$ y $h_0 = 0$. Por tanto, solo hay que cambiar los papeles de q y p

$$h_i' = \frac{(\frac{p}{q})^{N-i}-(\frac{p}{q})^N}{1-(\frac{p}{q})^N}.$$

Se puede comprobar que

$$\frac{(\frac{q}{p})^i-(\frac{q}{p})^N}{1-(\frac{q}{p})^N} + \frac{(\frac{p}{q})^{N-i}-(\frac{p}{q})^N}{1-(\frac{p}{q})^N} = 1.$$

Por tanto, el juego tiene un fin y acabará en uno de los dos estados absorbentes. Estos resultados lo podemos ver con más detalle en [2] de las referencias.

Supongamos ahora que $p = q$. Entonces $p = q = 1/2$. Como antes, vamos a calcular primero la probabilidad de arruinarse. Teniendo en cuenta que

$$h_i - h_1 = \sum_{n=2}^{i} \left(\frac{q}{p}\right)^{n-1}(h_1 - h_0) = (i-1)(h_1 - h_0).$$

Entonces, como $h_N = 0$ y $h_0 = 1$,

$$h_N - h_1 = (N-1)(h_1 - h_0) \Rightarrow -h_1 = (N-1)(h_1 - 1) \Rightarrow h_1 = \frac{N-1}{N}.$$

Por tanto,

$$h_i = h_1 + (i-1)(h_1 - 1) = \frac{N-1}{N} + (i-1)(\frac{N-1}{N} - 1) = \frac{N-1+(i-1)(-1)}{N} = \frac{N-i}{N} = 1 - \frac{i}{N}.$$

Y obtenemos, como antes, la probabilidad de llegar a tener N euros cambiando i por $N-i$

$$h_i = 1 - \frac{N-i}{N} = \frac{i}{N}.$$

Volvemos a observar que

$$1 - \frac{i}{N} + \frac{i}{N} = 1.$$

Por tanto, el juego también acabará en uno de los estados absorbentes.

IX. Tiempo medio de absorción: Vamos a averiguar cuántas rondas necesita la cadena para llegar al estado de absorción 0 o N. Aplicaremos el teorema 5.9.

$$\begin{cases} k_0 = k_N = 0; \\ k_i = 1 + pk_{i+1} + qk_{i-1} \quad \text{si } 0 < i \leq N-1. \end{cases}$$

Esta recurrencia lo habíamos visto antes, pero esta vez tenemos dos condiciones iniciales $k_0 = k_N = 0$, entonces aplicando estas condiciones en la ecuación 12.2 y suponiendo que $p \neq q$

$$\begin{cases} k_0 = 0 = A + B \\ k_N = 0 = \frac{N}{q-p} + A + B(\frac{q}{p})^N. \end{cases} \Rightarrow B = \frac{N}{(q-p)(1-(\frac{q}{p})^N)} \text{ y } A = -\frac{N}{(q-p)(1-(\frac{q}{p})^N)}$$

entonces,

$$\begin{aligned} k_i &= \frac{i}{q-p} - \frac{N}{(q-p)(1-(\frac{q}{p})^N)} + \frac{N}{(q-p)(1-(\frac{q}{p})^N)}\left(\frac{q}{p}\right)^i \\ &= \frac{i}{q-p} - \frac{N}{(q-p)}\frac{1-(\frac{q}{p})^i}{1-(\frac{q}{p})^N}. \end{aligned}$$

Por tanto,

$$k_i = \frac{i}{q-p} - \frac{N}{(q-p)}\frac{1-(\frac{q}{p})^i}{1-(\frac{q}{p})^N}$$

es el tiempo medio de acabar el juego empezando con i euros. Ahora supongamos que $p = q$, entonces $p = q = \frac{1}{2}$ y tenemos la siguiente recurrencia

$$\begin{cases} k_0 = k_N = 0; \\ k_i = 1 + \frac{1}{2}k_{i+1} + \frac{1}{2}k_{i-1} \quad \text{si } 0 < i \leq N-1. \end{cases}$$

Sabemos que $k_i^{Homogénea} = A + iB$, por la solución general 1.10. Como $k_i = C$ y $k_i = Ci$ no soluciones, supongamos que la solución particular es $k_i = Ci^2$.

$$Ci^2 = 1 + \frac{1}{2}C(i+1)^2 + \frac{1}{2}C(i-1)^2 \Rightarrow Ci^2 = 1 + C(i^2+1) \Rightarrow C = -1.$$

Por tanto,

$$k_i = A + iB - i^2$$

como $k_0 = k_N = 0$

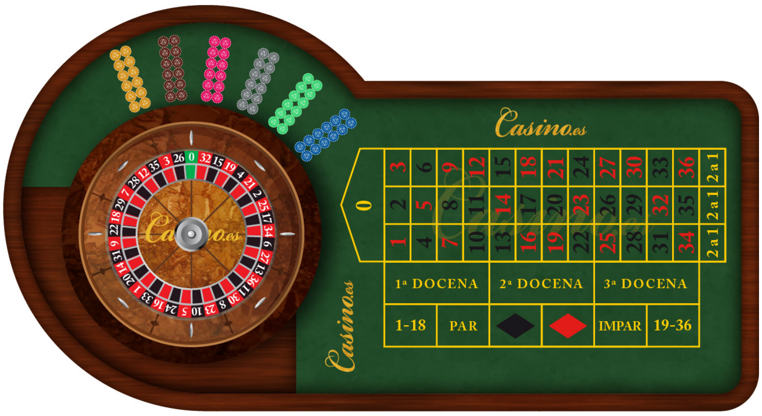
$$\begin{cases} k_0 = 0 = A \\ k_N = 0 = A + NB - N^2 \end{cases} \Rightarrow \begin{cases} A = 0 \\ B = N. \end{cases}$$

Entonces el tiempo medio de terminar el juego empezando con i euros es

$$k_i = iN - i^2, \text{ para todo } 0 \leq i \leq N.$$

12.2. El problema de la ruina del jugador sobre la ruleta

La ruleta es uno de los juegos de azar más comunes en los casinos. En la ruleta hay 37 números, de 0 hasta 36. El 0 es de color verde y hay 18 números que son rojos y otros 18 son negros. En esta sección únicamente tenemos en cuenta en las apuestas por el color, la paridad (el 0 no es ni par ni impar) y si el número está entre $1-18$ o $19-36$. Todas ellas tienen $\frac{19}{37}$ probabilidad de perder y $\frac{18}{37}$ de ganar obteniendo un premio de 1 x 1. Es decir, por cada euro apostado, gana un euro y recupera el dinero apostado.



Supongamos que un jugador juega a la ruleta apostando un euro cada ronda y se irá cuando gane 9 euros. Sea $q = \frac{19}{37}$ y $p = \frac{18}{37}$ las probabilidades de perder y de ganar, entonces, por la ecuación 12.4, la probabilidad de arruinarse es

$$h_i = \frac{(\frac{19}{18})^i - (\frac{19}{18})^9}{1 - (\frac{19}{18})^9}.$$

Vemos que si empieza a partir de 6 euros, la probabilidad de arruinarse es menor que 50 % y el tiempo medio de acabar arruinado u obteniendo 10 euros son 18 rondas (ver anexo A.2).

Observación 12.1. I. h_i decrece cuando aumentamos i, es decir, la probabilidad de arruinarse es menor cuando el jugador tiene mucho dinero. Sobre todo cuando se acerca a N, ya que la probabilidad de obtener N euros es elevada y se iría del casino antes de arruinarse.

II. $k_i = \frac{i}{q-p} - \frac{N}{(q-p)} \frac{1-(\frac{q}{p})^i}{1-(\frac{q}{p})^N} = 37i - 37N \frac{1-(\frac{19}{18})^i}{1-(\frac{19}{18})^N}$, para $i \in [0, N]$. Notamos que si aumentamos la fortuna del jugador, aumenta el tiempo medio de acabar el juego, pero llega un momento donde empieza a decrecer, ya que se acerca a N (ver anexo A.2).

III. La matriz de transición para $N = 9$ es

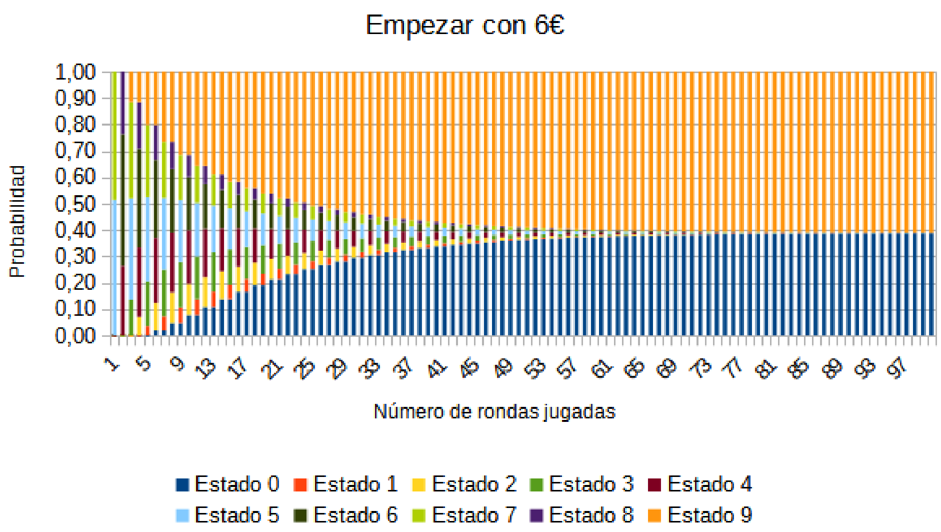
$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{19}{37} & 0 & \frac{18}{37} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{19}{37} & 0 & \frac{18}{37} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{19}{37} & 0 & \frac{18}{37} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{19}{37} & 0 & \frac{18}{37} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{19}{37} & 0 & \frac{18}{37} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{19}{37} & 0 & \frac{18}{37} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{19}{37} & 0 & \frac{18}{37} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{19}{37} & 0 & \frac{18}{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Supongamos que empezamos con 6 euros, entonces la distribución inicial sería

$$\lambda = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Vamos a calcular las distribuciones λP^n para saber qué probabilidad hay de estar en los estados después de n rondas. Para realizar estos cálculos, usaremos el segundo código del anexo A.1.

La gráfica que veremos a continuación es el resultado de las distribuciones λP^n, para $1 \leq n \leq 100$.



. (12.5)

Observación 12.2.

I. La probabilidad de ganar 9 euros es cada vez mayor cuando aumenta el número de rondas jugadas.

II. Si el número de rondas jugadas tiende a infinito, independientemente de la distribución inicial, acabaría en el estado 0 o 9, ya que son estados absorbentes (ver las graficas A.3 A.4 A.5 A.6 A.7 A.8 A.9 del anexo).

Referencias

[1] Norris, J. R. (1997) Markov Chains. Vol. 2. [Online]. West Nyack: Cambridge University Press.

[2] Solé, Josep Lluís. ≪Incertesa i probabilitat : un passeig per algunes paradoxes i problemes clàssics de la teoria de la probabilitat≫. Noubiaix: revista de la FEEMCAT i la SCM, 2011, Núm. 30, p. 62-86, https://raco.cat/index.php/Noubiaix/article/view/264381.

A. Anexo

A.1. Códigos en C++

Código para calcular las probabilidades y el tiempo medio de absorción

```
#include <math.h>
#include <fstream>
#include <stdlib.h>
#include <stdio.h>
#include <iostream>
#include <iomanip>
using namespace std;

int main(void){
          float k, p, q, h, a, b, g;
          int i, N;
          cout <<"Cuanto dinero quieres ganar en la ruleta? \n";
          cin >> N;
          k= 0.0;
          p = 0.48649;
          q = 0.51351;
          a = 1.05556; /* a = q/p */
          b = pow(a,N);
          g = N/(q-p);
          cout << "i    h              k" << endl;
          for(i = 0; i <= N; i++){
                    h = (pow(a,i) - b)/(1-b);
                    k = i/(q-p) - g*(1- pow(a,i))/(1-pow(a,N));
                    printf("%d     %f  %f \n", i, h, k);
       }
}
```

Código para calcular λP^n, para $n \geq 1$;

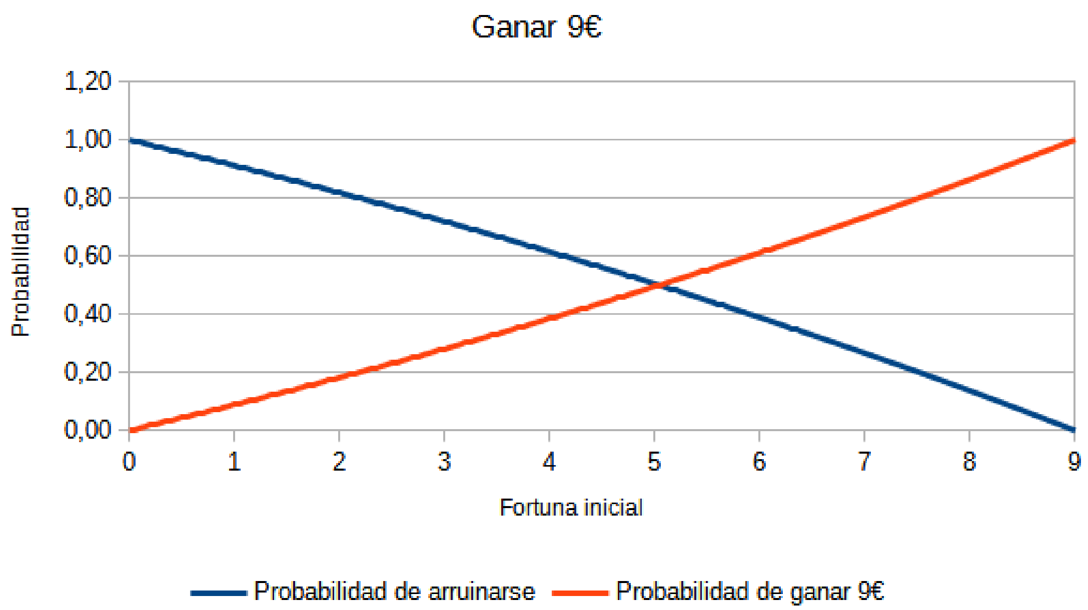
```
#include <math.h>
#include <fstream>
#include <stdlib.h>
#include <stdio.h>
#include <iostream>
#include <iomanip>
using namespace std;

int main(void){
       int i,j,k,l;
       int z = 0;
       int N = 10;
       long double A[10][10], B[10][10], C[10][10];
       long double b[10], c[10];
       long  double q = 0.51351351;
       long double  p  = 0.48648649;

       cout << "Empezar por el estado " << endl;
       cin >> l;

       for (i=0;i<10;i++)
              {for (j=0;j<10;j++){
                     A[i][j] = B[i][j] = C[i][j] =0;
```

```
                }
                b[i] = c[i] = 0;
        }

    b[1]= C[0][0] = A[0][0] = C[9][9] = A[9][9]=1;

    for(i = 1; i < 9; i++){
          for(j = 0; j <=9; j++){
                    if(i == j+1){
                            C[i][j] =A[i][j] = q;
                            }else if (i == j-1){
                                  C[i][j] =A[i][j] = p;
                                }else {
                                      C[i][j] =A[i][j] = 0;
                                }
                }
        }

    for (i=0;i<10;i++)
        {for (j=0;j<10;j++){
                c[i] += b[j]*A[j][i];
                }
        }

    for (j=0;j<10;j++){
        cout << c[j] << " ";
        c[j] = 0;
        }
      cout << endl;

    while (z < 99){
          for (i=0;i<10;i++)
                {for (j=0;j<10;j++)
                  { B[i][j]=0;
                            for (k=0;k<10;k++)
                                {B[i][j] += A[i][k] * C[k][j];
                        }
                            }
                        }
                for (i=0;i<10;i++)
                {for (j=0;j<10;j++){
                        C[i][j] = B[i][j];
                        }
                        }

            for (i=0;i<10;i++)
                {for (j=0;j<10;j++){
                            c[i] += b[j]*B[j][i];
                            }
                        }
            for (j=0;j<10;j++){
                        cout << c[j] << " ";
                        c[j] = 0;
                        }
            cout << endl;
                z += 1;
         }
        return 1;}
```

A.2. Tabla de probabilidad de absorción y el tiempo medio de absorción

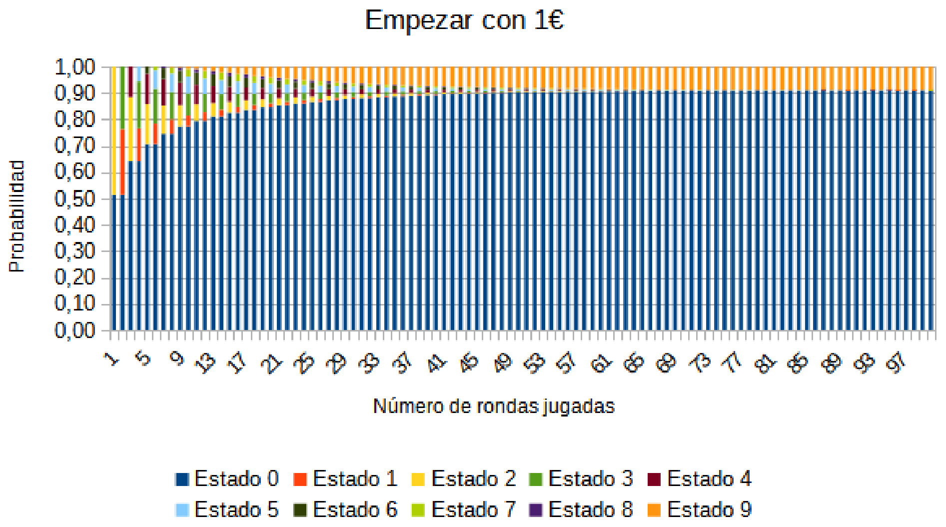
Fortuna inicial	Probabilidad de arruinarse	Probabilidad de ganar 9€	Rondas esperadas
0	1,00	0,00	0
1	0,91	0,09	7
2	0,82	0,18	13
3	0,72	0,28	17
4	0,61	0,39	20
5	0,50	0,50	20
6	0,39	0,61	18
7	0,27	0,73	15
8	0,14	0,86	8
9	0,00	1,00	0

(A.1)

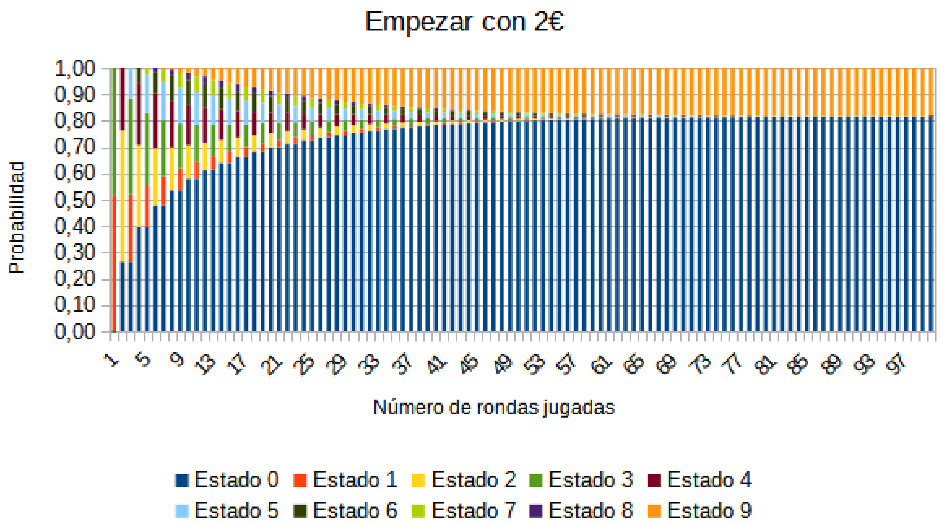


(A.2)

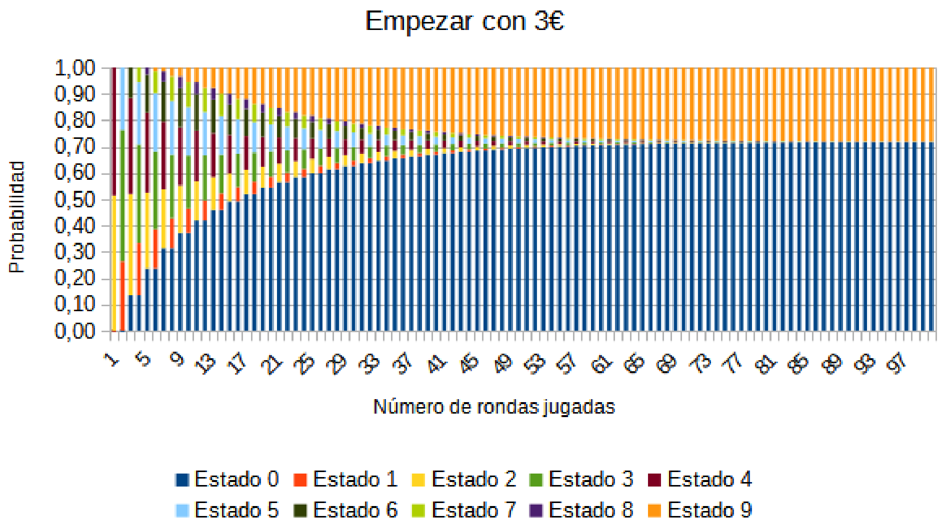
Las probabilidades de los estados tras n rondas jugadas según la distribución inicial.



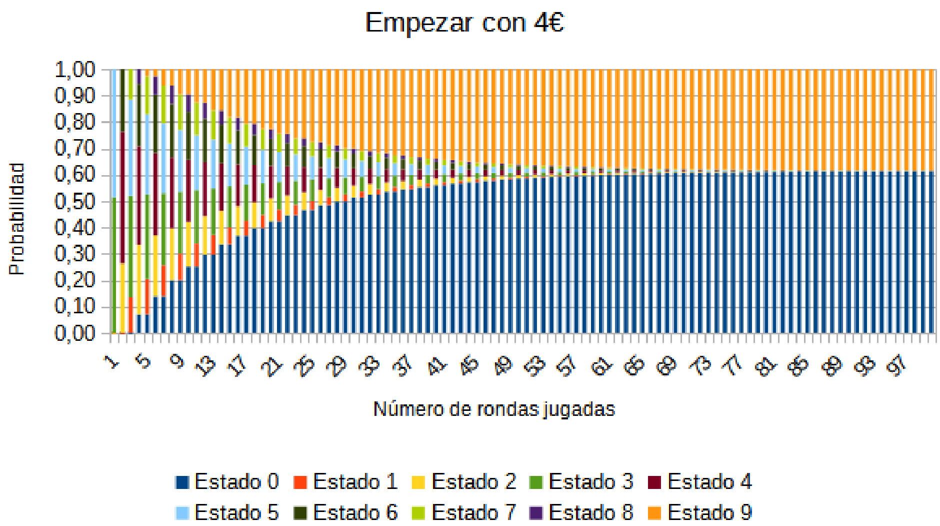
(A.3)



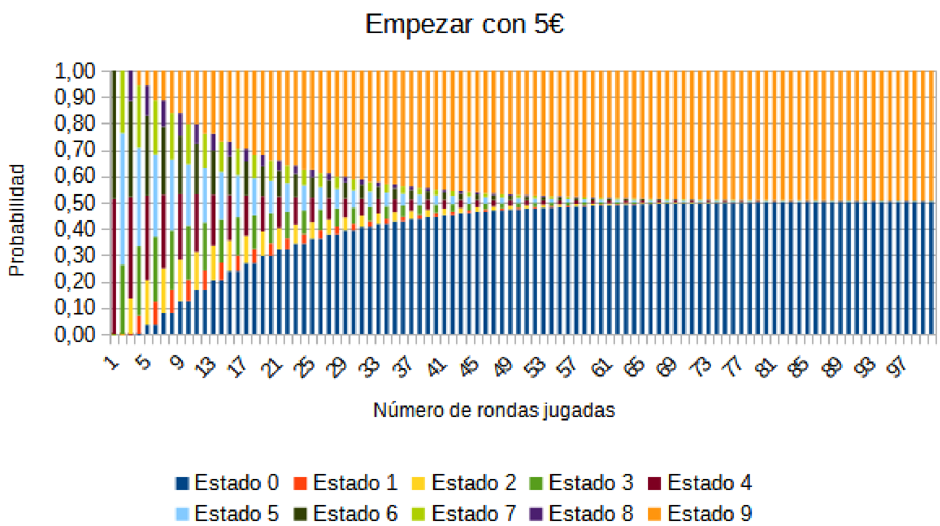
(A.4)


Empezar con 3€
Probabilidad
Número de rondas jugadas
Estado 0 Estado 1 Estado 2 Estado 3 Estado 4
Estado 5 Estado 6 Estado 7 Estado 8 Estado 9

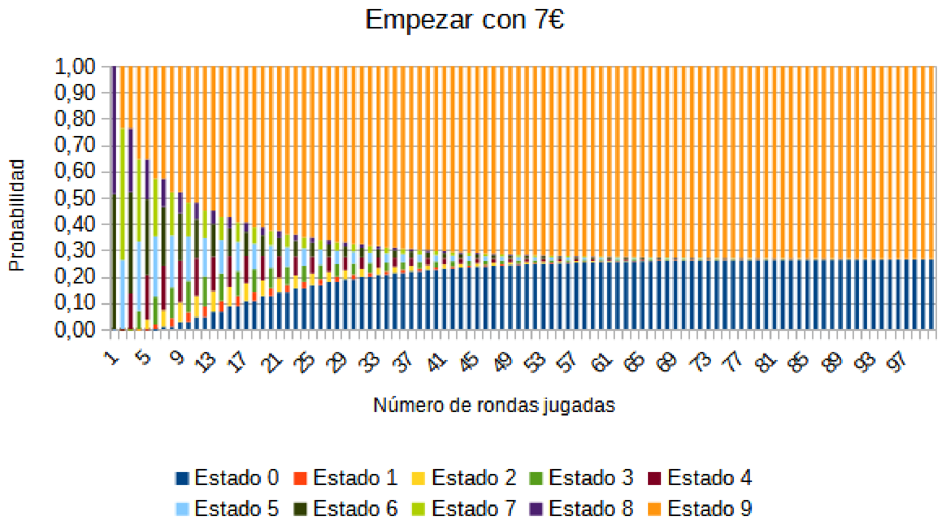
(A.5)


Empezar con 4€
Probabilidad
Número de rondas jugadas
Estado 0 Estado 1 Estado 2 Estado 3 Estado 4
Estado 5 Estado 6 Estado 7 Estado 8 Estado 9

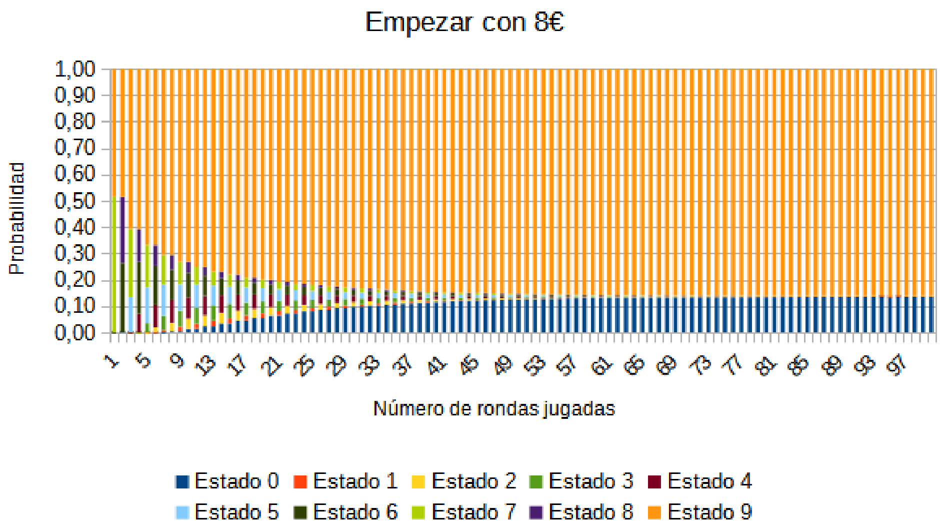
(A.6)


Empezar con 5€
Probabilidad
Número de rondas jugadas
Estado 0 Estado 1 Estado 2 Estado 3 Estado 4
Estado 5 Estado 6 Estado 7 Estado 8 Estado 9

(A.7)


Empezar con 7€
Probabilidad
1,00
0,90
0,80
0,70
0,60
0,50
0,40
0,30
0,20
0,10
0,00
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97
Número de rondas jugadas
Estado 0 Estado 1 Estado 2 Estado 3 Estado 4
Estado 5 Estado 6 Estado 7 Estado 8 Estado 9

(A.8)


Empezar con 8€
Probabilidad
1,00
0,90
0,80
0,70
0,60
0,50
0,40
0,30
0,20
0,10
0,00
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97
Número de rondas jugadas
Estado 0 Estado 1 Estado 2 Estado 3 Estado 4
Estado 5 Estado 6 Estado 7 Estado 8 Estado 9

(A.9)