## § 1. Distribuciones estacionarias

En estas notas se pretenden dar algunos de los resultados principales de las distribuciones *estacionarias*, *invariantes* o *de equilibrios*. Estas distribuciones nos permitirán estudiar comportamientos asintóticos de las cadenas de Markov,[1] en particular si la cadena tiende a equilibrarse estocásticamente en una distribución dada, aunque también sabremos cosas de la cadena si esto último no sucede. Las notas están basadas esencialmente en [Hoe72] y [Nor97].

Para motivar un poco el estudio de las distribuciones estacionarias, pensemos en la cadena de Markov $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ más simple, aquella con espacio de estados $\mathcal{E} = \{0, 1\}$ y matriz de probabilidades de transición

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

Supongamos que $\pi_0 := (\mathbb{P}(X_0 = 0), \mathbb{P}(X_0 = 1)) = (r, 1-r)$. En este caso podemos calcular $\pi_n := (\mathbb{P}(X_n = 0), \mathbb{P}(X_n = 1))$ para toda $n \in \mathbb{N}$ de forma recursiva, pues por probabilidad total tendremos que

$$\begin{aligned}
\pi_{n+1}(0) &= \mathbb{P}(X_{n+1} = 0 \,|\, X_n = 0)\pi_n(0) + \mathbb{P}(X_{n+1} = 0 \,|\, X_n = 1)\pi_n(1) \\
&= (1-p)\pi_n(0) + q\pi_n(1) \\
&= q + (1-p-q)\pi_n(0).
\end{aligned}$$

Así las cosas, $\pi_1(0) = q + (1-p-q)r$, $\pi_2(0) = (1-p-q)^2 r + q(1 + (1-p-q))$ y en general

$$\pi_n(0) = (1-p-q)^n r + q\sum_{k=0}^{n-1}(1-p-q)^k. \tag{1}$$

Notemos que si $p + q = 0$ entonces es claro que $\pi_n = \pi_0$ para toda $n \in \mathbb{N}$, pues la suma que aparece en (1) será cero y entonces $\pi_n(0) = r$. Por otra parte, si $0 < p + q < 2$,

$$\begin{aligned}
\pi_n(0) &= (1-p-q)^n r + q\frac{1-(1-p-q)^n}{p+q} \\
&= \frac{q}{p+q} + (1-p-q)^n\left(r - \frac{q}{q+p}\right) \xrightarrow{n\to\infty} \frac{q}{p+q}.
\end{aligned}$$

Además, observando que $1 - r - p/(p+q) = q/(p+q) - r$,

$$\pi_n(1) = 1 - \pi_n(0) = \frac{p}{p+q} + (1-p-q)^n\left(1 - r - \frac{p}{p+q}\right) \xrightarrow{n\to\infty} \frac{p}{p+q}.$$

En este caso, $0 < p + q < 2$,vemos que si $n$ es suficientemente grande entonces $X_n$ tendrá un distribución que será aproximadamente Bernoulli $\big(p/(p+q)\big)$, de forma sorprendente, sin importar el valor de $r$, lo que particularmente implica que

$$\lim_{n\to\infty} \mathbb{P}(X_n = 0 \,|\, X_0 = 0) = \lim_{n\to\infty} P^n_{00} = \frac{q}{q+n},$$

de donde la influencia de $X_0$ se va desvaneciendo mientras más pasos se den. El caso $p+q = 1$ será estudiado posteriormente, cuando las herramientas pertinentes se hayan desarrollado.

[1] Recordemos que todas las cadenas con las que se trabajan en este curso son a tiempo discreto y homogéneas.

## §1.1. Propiedades básicas de distribuciones estacionarias

Habiendo visto un pequeño ejemplo en el que parece que una cadena de Markov se estabiliza o equilibra asintóticamente, procedemos a dar algunas propiedades de las distribuciones estacionarias, pero para poder hacerlo debemos saber qué es una distribución estacionaria.

**Definición 1.1.** Sea $\{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de probabilidades de transición $P$. Diremos que $\pi : \mathcal{E} \to \mathbb{R}$ es una *medida invariante* si satisface las siguientes condiciones:

1) $\pi(i) \geq 0$ para toda $i \in \mathcal{E}$.
2) Para toda $i \in \mathcal{E}$, $\pi(i) = \sum_{j\in\mathcal{E}} \pi(j)P_{ji}$.

Si además $\sum_{i\in\mathcal{E}} \pi(i) = 1$, diremos que $\pi$ es una *distribución estacionaria* o *distribución invariante*.

Si se representa a $\pi$ mediante un vector renglón entonces la condición 2 de la definición 1.1 es equivalente a pedir a $\pi = \pi P$.

**Ejemplo 1.1.** Consideremos la cadena de Markov $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ con espacio de estados $\mathcal{E} = \{1, 2, 3\}$ y matriz de probabilidades de transición

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

Si queremos encontrar una distribución estacionaria para esta cadena, debemos encontrar un vector $\pi$ con entradas no negativas, tal que satisfaga $\sum_{i\in\mathcal{E}} \pi(i) = 1$ y $\pi = \pi P$, lo que nos da el siguiente sistema de ecuaciones:

$$\begin{aligned} \pi(1) &= \frac{1}{2}\pi(3), \\ \pi(2) &= \pi(1) + \frac{1}{2}\pi(2), \\ \pi(3) &= \frac{1}{2}\pi(2) + \frac{1}{2}\pi(3), \\ 1 &= \pi(1) + \pi(2) + \pi(3). \end{aligned}$$

Resolviendo el sistema de ecuaciones planteado se obtiene que $\pi = (1/5, 2/5, 2/5)$ es una distribución estacionaria.

Ahora veremos un resultado que explica el término *distribución estacionaria*.

**Proposición 1.2.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de probabilidades de transición $P$. Si la distribución $\pi$ de $X_0$ es invariante entonces para toda $n \in \mathbb{N}$ la distribución de $X_n$ es $\pi$.*

*Demostración.* Denotemos por $\pi_n$ la distribución de $X_n$. Por ser $\pi_0 = \pi$ una distribución estacionaria,

$$\pi_1 = \pi_0 P = \pi P = \pi = \pi_0.$$

El resultado se sigue de forma inductiva, usando que $\pi_{n+1} = \pi_n P$ para toda $n \in \mathbb{N}$. ⁕

Los siguientes resultados ayudan a explicar, parcialmente, el concepto de distribución de *equilibrio*.

**Proposición 1.3.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ finito y matriz de probabilidades de transición $P$. Si existe $i \in \mathcal{E}$ tal que*

$$\lim_{n\to\infty} P_{ij}^n = \pi(j) \tag{2}$$

*para toda $j \in \mathcal{E}$, entonces $\pi = \{\pi(j) : j \in \mathcal{E}\}$ es una distribución estacionaria.*

*Demostración.* Observemos que $\pi(j) \geq 0$ para toda $j \in \mathcal{E}$ por ser límite de números no negativos. Por ser $\mathcal{E}$ finito y de (2) tendremos

$$\sum_{j\in\mathcal{E}} \pi(j) = \sum_{j\in\mathcal{E}} \lim_{n\to\infty} P_{ij}^n = \lim_{n\to\infty} \sum_{j\in\mathcal{E}} P_{ij}^n = 1.$$

Por Chapman–Kolmogorov, (2) y la finitud de $\mathcal{E}$,

$$\pi(j) = \lim_{n\to\infty} P_{ij}^{n+1} = \lim_{n\to\infty} \sum_{k\in\mathcal{E}} P_{ik}^n P_{kj} = \sum_{k\in\mathcal{E}} \lim_{n\to\infty} P_{ik}^n P_{kj} = \sum_{n\in\mathcal{E}} \pi(k) P_{kj}.$$

Por lo tanto, $\pi$ resulta ser una distribución invariante. ⋇

Notemos que en la demostración anterior la condición de que $\mathcal{E}$ es finito se requiere para cambiar la suma y el límite, además de la condición (2) es para una $i \in \mathcal{E}$ fija y no para toda $i \in \mathcal{E}$. Un ejemplo en el que se puede observar que el límite $\pi$ que se obtiene en la proposición 1.3 depende de $i \in \mathcal{E}$ es el siguiente.

**Ejemplo 1.2.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ la cadena de Markov con espacio de estados $\mathcal{E} = \{0, 1\}$ y probabilidades de transición $P_{00} = P_{11} = 1$. Entonces $P_{00}^n = P_{11}^n = 1$ para toda $n \in \mathbb{N}$ y por la proposición 1.3, tanto $\pi^0 = (1, 0)$ como $\pi^1 = (0, 1)$ resultan ser distribuciones estacionarias.

Para obtener un resultado asintótico sobre la distribución de $X_n$ en el que no influya la distribución inicial, supondremos que el límite en (2) es uniforme en $i \in \mathcal{E}$, por lo que podremos relajar la condición de que $\mathcal{E}$ sea finito. Empero, a diferencia de la proposición 1.3, no podremos deducir que $\pi$ es una distribución estacionaria, sino que tendrá que ser parte de los supuestos.

**Proposición 1.4.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$, matriz de probabilidades de transición $P$. Sea $\pi$ es una distribución estacionaria. Si para cualesquiera $i, j \in \mathcal{E}$ se cumple*

$$\lim_{n\to\infty} P_{ij}^n = \pi(j) \tag{3}$$

*entonces*

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \pi(j)$$

*se satisfará para toda $j \in \mathcal{E}$.*

*Demostración.* Supongamos que $\pi$ es una distribución estacionaria para $\boldsymbol{X}$, mientras que $\pi_0$ es la distribución inicial para la cadena, y que se satisface (3). Fijando $j \in \mathcal{E}$, tendremos

$$\mathbb{P}(X_n = j) = \sum_{i\in\mathcal{E}} \pi_0(i) P_{ij}^n = \mathbb{E}[Y_n^j],$$

donde $Y_n^j$ es una variable aleatoria que toma el valor $P_{ij}^n$ con probabilidad $\pi_0(i)$. Por (3), $Y_n^j \to \pi(j)$ cuando $n \to \infty$ y además $|Y_n^j| \leq 1$ para toda $n \in \mathbb{N}$, entonces por el teorema A.3,

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \lim_{n\to\infty} \mathbb{E}[Y_n^j] = \mathbb{E}[\pi(j)] = \pi(j). \quad ⋇$$

Un resultado que se obtiene directamente de la proposición anterior es el siguiente.

**Corolario 1.5.** *Considerando las mismas hipótesis que en la proposición 1.4 la distribución estacionaria de $\boldsymbol{X}$ es única.*

*Demostración.* Sea $\pi$ la distribución estacionaria del enunciado. Supongamos que la distribución inicial de $\boldsymbol{X}$, $\tilde{\pi}$, es estacionaria. Entonces $\mathbb{P}(X_n = j) = \mathbb{P}(X_0 = j) = \tilde{\pi}(j)$ para toda $n \in \mathbb{N}$ y para toda $j \in \mathcal{E}$, por lo que $\lim_{n\to\infty} \mathbb{P}(X_n = j) = \tilde{\pi}(j)$ para toda $j \in \mathcal{E}$. Luego, por la proposición 1.4 concluimos que $\tilde{\pi}(j) = \pi(j)$ para toda $j \in \mathcal{E}$. ⁕

Los resultados que aparecen en esta sección, si bien nos dan relaciones entre las probabilidades de transición a $n$ pasos y las distribuciones estacionarias, no son útiles en el sentido de que no dan información de qué cadenas tienen distribuciones estacionarias y cuándo se cumplen (2) y (3). Estas preguntas se responderán en secciones posteriores.

## §1.2. Visitas a estados recurrentes y la existencia de distribuciones estacionarias

Para poder estudiar el comportamiento asintótico de $P_{ij}^n$ para ciertas cadenas de Markov, primero estudiaremos el comportamiento asintótico de

$$\frac{1}{n}\sum_{m=1}^{n} P_{ij}^m$$

para cadenas de Markov arbitrarias. Esto nos pondrá en una posición favorable para después estudiar la existencia de distribuciones estacionarias, comenzando con cadenas irreducibles, para después pasar a cadenas de Markov generales. Las pruebas de los resultados que se presentan en esta sección se basan en interpretaciones probabilistas.

Con este fin en mente, considerando una cadena de Markov $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ con espacio estados $\mathcal{E}$ definamos, para $j \in \mathcal{E}$ fija,

$$N_j^n = \sum_{m=1}^{n} \mathbb{1}_{\{X_m = j\}} \quad \text{y} \quad N_j = \sum_{m=1}^{\infty} \mathbb{1}_{\{X_m = j\}},$$

que cuentan el número de veces que la cadena $\boldsymbol{X}$ pasa por el estado $j$ en los primeros $n$ pasos y en total de forma respectiva. A continuación se enuncia un resultado, que usaremos más adelante, sobre estas variables cuando $j \in \mathcal{E}$ es un estado transitorio.

**Lema 1.6.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $j \in \mathcal{E}$ es un estado transitorio, entonces $N_j < \infty$ con probabilidad uno y $\mathbb{E}[N_j \,|\, X_0 = i] < \infty$ para cualquier $i \in \mathcal{E}$.*

*Demostración.* Recordemos que $T_j = \inf\{n \geq 1 : X_n = j\}$ es la primera vez que $\boldsymbol{X}$ está en el estado $j$, con $\inf \varnothing = \infty$ por convención, y que $\rho_{jj} = \mathbb{P}(T_j < \infty \,|\, X_0 = j) < 1$ por ser $j$ un estado transitorio. Es claro que $T_j$ es un tiempo de paro —ver definición A.5—. Usaremos esto para probar que

$$\mathbb{P}(N_j = n \,|\, X_0 = i) = \begin{cases} 1 - \rho_{ij} & \text{si } n = 0, \\ \rho_{ij}\rho_{jj}^{n-1}(1 - \rho_{jj}) & \text{si } n \geq 1. \end{cases} \tag{4}$$

Así las cosas, notemos que $\{N_j = 0\} = \{T_j = \infty\}$, por lo que

$$\mathbb{P}(N_j = 0 \,|\, X_0 = i) = \mathbb{P}(T_j = \infty \,|\, X_0 = i) = 1 - \mathbb{P}(T_j < \infty \,|\, X_0 = i) = 1 - \rho_{ij}.$$

Para $n \geq 1$ tendremos que

$$\begin{aligned}
\mathbb{P}(N_j = n \,|\, X_0 = i) &= \mathbb{P}(N_j = n \,|\, T_j < \infty, X_0 = i)\mathbb{P}(T_j < \infty \,|\, X_0 = i) \\
&\quad + \mathbb{P}(N_j = n \,|\, T_j = \infty, X_0 = i)\mathbb{P}(T_j = \infty \,|\, X_0 = i) \\
&= \mathbb{P}\left( \sum_{m=1}^{T_j} \mathbb{1}_{\{X_m = j\}} + \sum_{m=T_j+1}^{\infty} \mathbb{1}_{\{X_m = 1\}} = n \,\middle|\, T_j < \infty, X_0 = i \right) \rho_{ij} \\
&= \mathbb{P}\left( \sum_{m=T_j+1}^{\infty} \mathbb{1}_{\{X_m = j\}} = n - 1 \,\middle|\, T_j < \infty, X_0 = i \right) \rho_{ij} \\
&= \mathbb{P}\left( \sum_{m=1}^{\infty} \mathbb{1}_{\{X_m = j\}} = n - 1 \,\middle|\, X_0 = j \right) \rho_{ij} \\
&= \mathbb{P}(N_j = n - 1 \,|\, X_0 = j)\rho_{ij},
\end{aligned}$$

donde se usó la propiedad fuerte de Markov —ver teorema A.6— en la cuarta igualdad. Un argumento inductivo nos permite obtener (4). Teniendo esto, para $i \in \mathcal{E}$,

$$\mathbb{E}[N_j \,|\, X_0 = i] = \sum_{n=1}^{\infty} n\mathbb{P}(N_j = n \,|\, X_0 = i) = \sum_{n=1}^{\infty} n\rho_{ij}\rho_{jj}^{n-1}(1 - \rho_{jj}) = \frac{\rho_{ij}}{1 - \rho_{jj}} < \infty.$$

Entonces para toda $i \in \mathcal{E}$, $\mathbb{P}(N_j < \infty \,|\, X_0 = i) = 1$ y, por un argumento de probabilidad total, $\mathbb{P}(N_j < \infty) = 1$, lo cual concluye la prueba. ⁕

¿Cómo es que aplicamos esto? Si $j \in \mathcal{E}$ es transitorio, observando que $N_j^n \leq N_j < \infty$ con probabilidad uno, se sigue que

$$\lim_{n \to \infty} \frac{N_j^n}{n} = 0$$

con probabilidad uno. Además tendremos que

$$\begin{aligned}
\lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m &= \lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} \mathbb{E}[\mathbb{1}_{\{X_m = j\}} \,|\, X_0 = i] \\
&= \lim_{n \to \infty} \mathbb{E}\left[ \frac{N_j^n}{n} \,\middle|\, X_0 = i \right] \\
&\leq \lim_{n \to \infty} \frac{1}{n} \mathbb{E}[N_j \,|\, X_0 = i] \\
&= 0.
\end{aligned}$$

Antes de proseguir, consideremos un estado cualquiera $j \in \mathcal{E}$ y definamos el *tiempo medio de retorno* a $j$ de una cadena que comience en $j$, $m_j = \mathbb{E}[T_j \,|\, X_0 = j]$ que puede tomar el valor $\infty$. Definamos además $\mathbb{1}_{\{T_j < \infty\}}$, la variable aleatoria que es 1 si $T_j < \infty$ y cero en otro caso. El siguiente resultado es uno de los resultados principales de la sección, del cual hemos demostrado ya una parte.

**Teorema 1.7.** *Sea $\boldsymbol{X} = \{X_n\}_{n \in \mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $j \in \mathcal{E}$ entonces*

$$\lim_{n \to \infty} \frac{N_j^n}{n} = \frac{1}{m_j} \mathbb{1}_{\{T_j < \infty\}} \tag{5}$$

*con probabilidad uno. Además*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} P_{ij}^{m} = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[N_j^n \mid X_0 = i] = \frac{\rho_{ij}}{m_j} \tag{6}$$

*para toda $i \in \mathcal{E}$.*

Si $m_j = \infty$, el lado derecho de (5) y (6) se define como cero por convención, para que se satisfaga en el caso transitorio.[2]

*Demostración.* El caso en el que $j \in \mathcal{E}$ es transitorio se probó antes de enunciar el teorema, por lo que podremos suponer que $j \in \mathcal{E}$ es recurrente, *i.e.* $\rho_{jj} = 1$. Supongamos además que $X_0 = j$. Definamos $W_0^j = 0$ y, para $n \geq 1$, $W_n^j = \inf\{m \geq 1 : N_j^m = n\}$, el tiempo en el que $\boldsymbol{X}$ visita $j$ por $n$-ésima vez. Definamos además $T_n^j = W_n^j - W_{n-1}^j$, los tiempos entre visitas a $j$, para $n \geq 1$. De la propiedad de Markov y homogeneidad se deduce fácilmente que

$$\mathbb{P}(T_n^j = m_n, \ldots, T_1^j = m_1 \mid X_0 = j) = \prod_{k=1}^{n} \mathbb{P}(T_j = m_k \mid X_0 = j).$$

Por lo tanto $\{T_n^j\}_{n\geq 1}$, condicional en $X_0 = j$, es una sucesión de variables aleatorias independientes e idénticamente distribuidas con media $m_j = \mathbb{E}[T_j \mid X_0 = j]$. Además $W_n^j = \sum_{m=1}^{n} T_m^j$ y así, por el teorema A.4,

$$\lim_{n\to\infty} \frac{W_n^j}{n} = m_j \tag{7}$$

con probabilidad uno.

Ahora notemos que $W_{N_j^n}^j \leq n < W_{N_j^n+1}^j$, pues la $N_j^n$-ésima visita a $j$ tiene que suceder a lo más al tiempo $n$ y $N_j^n + 1$-ésima visita debe suceder después del tiempo $n$. Por lo tanto,

$$\frac{W_{N_j^n}^j}{N_j^n} \leq \frac{n}{N_j^n} < \frac{W_{N_j^n+1}^j}{N_j^n},$$

con $n$ suficientemente grande para que $N_j^n \geq 1$. Como $j$ es recurrente, $N_j^n \to N_j = \infty$ con probabilidad uno. Luego, por las desigualdades anteriores y (7),

$$\lim_{n\to\infty} \frac{n}{N_j^n} = m_j$$

con probabilidad uno, o de forma equivalente, $N_j^n/n \to 1/m_j$ con probabilidad uno.

Ahora dejemos que $X_0$ tenga una distribución arbitraria, si $T_j < \infty$, entonces el argumento anterior es válido por la propiedad fuerte de Markov, y si $T_j = \infty$ entonces $N_j^n/n \to 0$. Así las cosas, hemos obtenido que con probabilidad uno se satisfará

$$\lim_{n\to\infty} \frac{N_j^n}{n} = \frac{1}{m_j}\mathbb{1}_{\{T_j<\infty\}}.$$

Para concluir notemos que para cada $n \geq 1$, $0 \leq N_j^n/n \leq 1$. Por el teorema A.3,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} P_{ij}^{m} = \lim_{n\to\infty} \mathbb{E}\left[\frac{N_j^n}{n} \,\middle|\, X_0 = i\right] = \mathbb{E}\left[\frac{1}{m_j}\mathbb{1}_{\{T_j<\infty\}} \,\middle|\, X_0 = i\right] = \frac{\rho_{ij}}{m_j}. \qquad ⁂$$

[2]Y ciertos casos recurrentes.

Recoredemos que $\mathcal{C} \subset \mathcal{E}$ es una clase cerrada si para cualquier $i \in \mathcal{C}$ y cualquier $j \notin \mathcal{C}$, $P_{ij} = 0$. Además se dice que $\mathcal{C}$ es irreducible si $\rho_{ij} > 0$ para cualesquiera $i, j \in \mathcal{C}$. Un resultado de la primera parte de cadenas de Markov nos dice que si $\mathcal{C}$ es cerrado e irreducible, entonces todos los estados son transitorios o son recurrentes. De cualquier forma, tendremos el siguiente corolario directo del teorema 1.7.

**Corolario 1.8.** *Sean $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$, y $\mathcal{C} \subset \mathcal{E}$ una clase cerrada. Entonces*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = \frac{\rho_{ij}}{m_j}$$

*para cualesquiera $i, j \in \mathcal{C}$. Si además se tiene que $\mathbb{P}(X_0 \in \mathcal{C}) = 1$ entonces*

$$\lim_{n\to\infty} \frac{N_j^n}{n} = \frac{1}{m_j},$$

*para toda $j \in \mathcal{C}$. En particular, si $\mathcal{C}$ es una clase cerrada de estados recurrentes, la conclusión se satisface con $\rho_{ij} = 1$ para cualquier par $i, j \in \mathcal{C}$.*

Hasta ahora hemos clasificado los estados en recurrentes y transitorios. Es momento de hacer una subdivisión de los estados recurrentes, en los que son *recurrentes positivos* y los que son *recurrentes nulos*.

**Definición 1.9.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Se dice que $j \in \mathcal{E}$ es *recurrente nulo* si $m_j = \infty$ y que es *recurrente positivo* si $m_j < \infty$.

Observemos que esta definición junto al teorema 1.7 nos dicen que si $j \in \mathcal{E}$ es un estado recurrente nulo, entonces para cualquier $i \in \mathcal{E}$ se satisfará

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = 0.$$

Esto que se acaba de deducir, así como el siguiente resultado, nos dan la pauta para fijarnos únicamente en los estados recurrentes positivos.

**Proposición 1.10.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $\pi$ es una medida invariante, $\pi(j) = 0$ para todo estado $j \in \mathcal{E}$ transitorio o recurrente nulo.*

*Demostración.* Por ser $\pi$ una medida invariante, $\pi = \pi P^m$ para toda $m \in \mathbb{N}$, en particular

$$\pi(j) = \sum_{i\in\mathcal{E}} \pi(i) P_{ij}^m.$$

Sumando sobre $m \in \{1, \ldots, n\}$ y dividiento entre $n$ obtenemos

$$\pi(j) = \sum_{i\in\mathcal{E}} \pi(i) \left( \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m \right).$$

Cuando $j \in \mathcal{E}$ es transitorio o recurrente nulo, por el teorema A.3, usando un argumento similar a la proposición 1.4,

$$\pi(j) = \lim_{n\to\infty} \sum_{i\in\mathcal{E}} \pi(i) \left( \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m \right) = \sum_{i\in\mathcal{E}} \pi(i) \left( \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m \right) = 0. \qquad ⁂$$

0.20787957635076193864189444561570819682757565328522246119872598426 2...

Ha llegado el momento de probar la existencia de distribuciones estacionarias para las cadenas de Markov irreducibles con estados recurrentes (positivos). Para poder probar este resultado, deberemos probar un par de lemas en el que usaremos las variables $\gamma_i^j$, definidas por

$$\gamma_i^j = \mathbb{E}\left[\sum_{m=0}^{T_j-1} \mathbb{1}_{\{X_m=i\}} \,\middle|\, X_0 = j\right],$$

donde $i, j \in \mathcal{E}$. Notemos que $\gamma_i^j$ cuenta la cantidad esperada de veces que la cadena $\boldsymbol{X}$ está en el estado $i$ antes de regresar por primera vez al estado $j$.

**Lema 1.11.** *Sea $\boldsymbol{X}$ una cadena de Markov irreducible y recurrente, con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Entonces se cumplirán:*

*1) $\gamma_j^j = 1$ para toda $j \in \mathcal{E}$.*

*2) Para cada $j \in \mathcal{E}$, $\gamma^j = \{\gamma_i^j : i \in \mathcal{E}\}$, representado como vector renglón, satisface $\gamma^j = \gamma^j P$.*

*3) $0 < \gamma_i^j < \infty$ para cualesquiera $i, j \in \mathcal{E}$.*

*Demostración.* Notemos que para $1 \leq n < T_j$, $\mathbb{1}_{\{X_n=0\}}$, por lo que

$$\gamma_j^j = \mathbb{E}[\mathbb{1}_{\{X_0=j\}} \,|\, X_0 = j] = 1,$$

probando (1).

Supongamos que (2) es verdadero y probemos (3). Como $\boldsymbol{X}$ es irreducible, para cualquier par de estados $i, j \in \mathcal{E}$ existen $n, m \in \mathbb{N}$ tales que $P_{ij}^n, P_{ji}^m > 0$. De (2), $\gamma_i^j \geq \gamma_j^j P_{ji}^m$ y $\gamma_i^j P_{ij}^n \leq \gamma_j^j$. Por (1), $\gamma_j^j = 1$ y por lo tanto

$$0 < P_{ji}^m \leq \gamma_i^j \leq \frac{1}{P_{ij}^n} < \infty,$$

lo que prueba (3). Así las cosas, concluiremos la demostración si logramos probar (2).

Ahora observemos que para $n \geq 1$, $n \leq T_j$ si y sólo si $X_1, \ldots, X_{n-1} \neq j$, por lo que para cada $k \in \mathcal{E}$ y cada $i \in \mathcal{E}$, por la propiedad de Markov y homogeneidad,

$$\mathbb{P}(X_{n-1} = k, X_n = i, n \leq T_j \,|\, X_0 = j) = \mathbb{P}(X_{n-1} = k, n \leq T_j \,|\, X_0 = j) P_{ki},$$

donde esto último es igual a cero si $k = j$. Como $j$ es recurrente, condicionado a $X_0 = j$, $T_j < \infty$ y $X_0 = X_{T_j} = j$ con probabilidad uno, lo que en particular implica, condicional en $X_0 = j$, que

$$\sum_{n=0}^{T_j-1} \mathbb{1}_{\{X_n=i\}} = \sum_{n=1}^{T_j} \mathbb{1}_{\{X_n=i\}}$$

se satisface para cualquier $i \in \mathcal{E}$. Notemos además que

$$\sum_{n=1}^{T_j} \mathbb{1}_{\{X_n=i\}} = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=i, n \leq T_j\}}$$

mientras que

$$\sum_{n=0}^{T_j-1} \mathbb{1}_{\{X_n=i\}} = \sum_{n=0}^{\infty} \mathbb{1}_{\{X_n=i, n \leq T_j-1\}}$$

para cualesquiera $i, j \in \mathcal{E}$. Por lo tanto,

$$\begin{aligned}
\gamma_i^j &= \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=i, n\leq T_j\}} \,\middle|\, X_0 = j\right] \\
&= \sum_{n=1}^{\infty} \mathbb{P}(X_n = i, n \leq T_j \,|\, X_0 = j) \\
&= \sum_{n=1}^{\infty} \sum_{k\in\mathcal{E}} \mathbb{P}(X_{n-1} = k, X_n = i, n \leq T_j \,|\, X_0 = j) \\
&= \sum_{n=1}^{\infty} \sum_{k\in\mathcal{E}} \mathbb{P}(X_{n-1} = k, n \leq T_j \,|\, X_0 = j) P_{ki} \\
&= \sum_{k\in\mathcal{E}} P_{ki} \sum_{n=1}^{\infty} \mathbb{P}(X_{n-1} = k, n \leq T_j \,|\, X_0 = j) \\
&= \sum_{k\in\mathcal{E}} P_{ki} \sum_{m=0}^{\infty} \mathbb{P}(X_m = k, m \leq T_j - 1 \,|\, X_0 = j) \\
&= \sum_{k\in\mathcal{E}} P_{ki} \mathbb{E}\left[\sum_{m=0}^{\infty} \mathbb{1}_{\{X_m=i, m\leq T_j-1\}} \,\middle|\, X_0 = j\right] \\
&= \sum_{k\in\mathcal{E}} P_{ki} \gamma_k^i,
\end{aligned}$$

donde los intercambios entre sumas y sumas y esperanzas son consecuencia del teorema A.1. Como eso se satisface para toda $i \in \mathcal{E}$, se sigue que $\gamma^j = \gamma^j P$, probando (2). ⁂

**Lema 1.12.** *Sea $\boldsymbol{X}$ una cadena de Markov irreducible, con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $\pi$ es una medida invariante tal que $\pi(j) = 1$ para alguna $j \in \mathcal{E}$, entonces $\pi \geq \gamma^j$, donde la desigualdad se entiende entrada a entrada. Si además $\boldsymbol{X}$ es recurrente, entonces $\pi = \gamma^j$.*

*Demostración.* Para cada $i \in \mathcal{E}$ y cada $n \geq 1$, por ser $\pi$ una medida invariante tal que $\pi(j) = 1$,

$$\begin{aligned}
&\sum_{k\neq j} \pi(k) \mathbb{P}(X_n = i, n \leq T_j \,|\, X_0 = k) \\
&\quad= \sum_{k\neq j} \left(\sum_{l\in\mathcal{E}} \pi(j) P_{lk}\right) \mathbb{P}(X_n = i, n \leq T_j \,|\, X_0 = k) \\
&\quad= \sum_{k\neq j} P_{jk} \mathbb{P}(X_n = i, n \leq T_j \,|\, X_0 = k) + \sum_{k\neq j} \sum_{l\neq j} \pi(l) P_{lk} \mathbb{P}(X_n = i, n \leq T_j \,|\, X_0 = k) \\
&\quad= \sum_{k\neq j} P_{jk} \mathbb{P}(X_{n+1} = i, n+1 \leq T_j \,|\, X_1 = k) \\
&\qquad+ \sum_{l\neq j} \pi(l) \sum_{k\neq j} P_{lk} \mathbb{P}(X_{n+1} = i, n+1 \leq T_j \,|\, X_1 = k) \\
&\quad= \mathbb{P}(X_{n+1} = i, n+1 \leq T_j \,|\, X_0 = j) + \sum_{l\neq j} \pi(l) \mathbb{P}(X_{n+1} = i, n+1 \leq T_j \,|\, X_0 = l),
\end{aligned}$$

donde se ha usado la homogeneidad de $\boldsymbol{X}$ en la tercera igualdad.

0.20787957635076193864189444561570819682757565328522246119872598426 2...

Ahora observemos que para $i \in \mathcal{E}$ fijo, por ser $\pi$ una medida invariante con $\pi(j) = 0$,

$$\begin{aligned}\pi(i) &= \sum_{k\in\mathcal{E}} \pi(k)P_{ki} = P_{ji} + \sum_{k\neq j} \pi(k)P_{ki} \\ &= \mathbb{P}(X_1 = i, 1 \leq T_j \mid X_0 = j) + \sum_{k\neq j} \pi(k)\mathbb{P}(X_1 = i, 1 \leq T_j \mid X_0 = k).\end{aligned}$$

Luego entonces, juntando las igualdades se obtiene fácilmente que para toda $n \geq 1$,

$$\begin{aligned}\pi(i) &= \sum_{m=1}^{n} \mathbb{P}(X_m = i, m \leq T_j \mid X_0 = j) + \sum_{k\neq j} \pi(k)\mathbb{P}(X_{n+1} = i, n+1 \leq T_j \mid X_0 = k) \\ &\geq \sum_{m=1}^{n} \mathbb{P}(X_m = i, m \leq T_j \mid X_0 = j) = \mathbb{E}\left[\sum_{m=1}^{n} \mathbb{1}_{\{X_m=i, m\leq T_i\}} \,\middle|\, X_0 = j\right].\end{aligned}$$

Así pues, por el teorema A.1 tendremos que

$$\pi(i) \geq \lim_{n\to\infty} \mathbb{E}\left[\sum_{m=1}^{n} \mathbb{1}_{\{X_m=i, m\leq T_j\}} \,\middle|\, X_0 = j\right] = \mathbb{E}\left[\sum_{m=1}^{\infty} \mathbb{1}_{\{X_m=i, m\leq T_j\}} \,\middle|\, X_0 = j\right] = \gamma_i^j.$$

Por lo tanto $\pi \geq \gamma^j$.

Ahora supongamos que $\boldsymbol{X}$ es recurrente, por el lema 1.11 $\gamma^j$ es una medida invariante. Por lo que acabamos de probar, $\mu = \pi - \gamma^j$ es también una medida invariante. En efecto, $\mu \geq 0$ pues $\pi \geq \gamma^j$ y $\mu P = (\pi - \gamma^j)P = \pi P - \gamma^j P = \pi - \gamma^j = \mu$. Por ser $\boldsymbol{X}$ irreducible, para cada $i \in \mathcal{E}$ existe $n \in \mathbb{N}$ tal que $P_{ij}^n > 0$ y así

$$\mu(j) = \sum_{k\in\mathcal{E}} \mu(k)P_{kj}^n \geq \mu(i)P_{ij}^n.$$

Como $\mu(j) = 0$ se sigue que $\mu(i) = 0$, o de forma equivalente $\pi(i) = \gamma_i^j$. ⁜

Ahora sí, el teoremón de la sección, en el que probamos la existencia —y unicidad— para cadenas de Markov irreducibles con estados recurrentes positivos.

**Teorema 1.13.** *Sea $\boldsymbol{X}$ una cadena de Markov irreducible, con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Los siguientes enunciados son equivalentes:*

*1) Todo estado de $\mathcal{E}$ es recurrente positivo.*

*2) Existe $j \in \mathcal{E}$ recurrente positivo.*

*3) $\boldsymbol{X}$ tiene una distribución estacionaria $\pi$.*

*Más aún, $\pi(j) = 1/m_j$ para toda $j \in \mathcal{E}$.*

*Demostración.* Que (1) implica (2) es evidente. Supongamos entonces que (2) se cumple y probemos (3). Como $j \in \mathcal{E}$ es recurrente y $\boldsymbol{X}$ es irreducible, $\boldsymbol{X}$ es recurrente y por el lema 1.11, $\gamma^j$ es invariante y

$$\begin{aligned}\sum_{i\in\mathcal{E}} \gamma_i^j &= \sum_{i\in\mathcal{E}} \mathbb{E}\left[\sum_{n=1}^{T_j} \mathbb{1}_{\{X_n=i\}} \,\middle|\, X_0 = j\right] = \mathbb{E}\left[\sum_{n=1}^{T_j}\sum_{i\in\mathcal{E}} \mathbb{1}_{\{X_n=i\}} \,\middle|\, X_0 = j\right] \\ &= \mathbb{E}\left[\sum_{n=1}^{T_j} \mathbb{1}_{\{X_n\in\mathcal{E}\}} \,\middle|\, X_0 = j\right] = \mathbb{E}[T_j \mid X_0 = j] = m_j,\end{aligned}$$

de donde $\sum_{i\in\mathcal{E}} \gamma_i^j = m_j < \infty$, pues $j \in \mathcal{E}$ es recurrente positivo. Entonces $\pi = \gamma^j/m_j$ es claramente una distribución estacionaria.

Finalmente, supongamos que (3) es verdadera y demostremos (1). Notemos que existe $j \in \mathcal{E}$ tal que $\pi(j) > 0$, pues $\sum_{i\in\mathcal{E}} \pi(i) = 1$. Fijemos $i \in \mathcal{E}$, por ser $\boldsymbol{X}$ irreducible, para existe $n \in \mathbb{N}$ tal que $P_{ji}^n > 0$ y por ser $\pi$ estacionaria, $\pi(i) = \sum_{k\in\mathcal{E}} \pi(k)P_{ki}^n \leq \pi(j)P_{ji}^n > 0$. Definamos una nueva medida estacionaria $\mu = \pi/\pi(i)$ y notemos qeu $\mu(i) = 1$, por lo que $\mu \geq \gamma^i$ por el lema 1.12. Por lo tanto,

$$m_i = \sum_{k\in\mathcal{E}} \gamma_k^i \leq \sum_{k\in\mathcal{E}} \frac{\pi(k)}{\pi(i)} = \frac{1}{\pi(i)} < \infty, \tag{8}$$

donde la última igualdad se da por ser $\pi$ distribución estacionaria.

Con esto hemos probado la equivalencia de (1), (2) y (3). Para finalizar la demostración, notemos que bajo cualquiera de los enunciados, $\boldsymbol{X}$ es recurrente y por el lema 1.12, la desigualdad en (8) es de hecho una igualdad. ⋇

Después de un teorema de tal importancia, ahora nos podemos concentrar en dar consecuencias de éste. Recordemos que si $\mathcal{C} \subset \mathcal{E}$ es una clase cerrada e irreducible, entonces se dice que es transitoria si todos sus estaos son transitorios y es recurrente si todos sus estados son recurrentes. Ahora refinaremos esta última clasificación y diremos que $\mathcal{C}$ es recurrente nula si todos sus estados son recurrentes nulos y que es recurrente positiva si todo $j \in \mathcal{C}$ es recurrente positivo.

**Corolario 1.14.** *Sea $\boldsymbol{X}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $\mathcal{C} \subset \mathcal{E}$ es cerrada e irreducible, entonces $\mathcal{C}$ es transitoria, recurrente nula o recurrente positiva.*

*Demostración.* Sabemos $\mathcal{C}$ es o transitoria o recurrente. Supongamos que $\mathcal{C}$ es recurrente y que $X_0 \in \mathcal{C}$ con probabilidad 1, por lo que podemos pensar en $\boldsymbol{X}$ como una cadena de Markov con espacio de estados $\mathcal{C}$. Así, por el teorema 1.13, si existe un estado $j \in \mathcal{C}$ recurrente positivo, $\mathcal{C}$ es recurrente positiva y en caso contrario $\mathcal{C}$ es recurrente nula. ⋇

Asimismo, si $\boldsymbol{X}$ es una cadena de Markov irreducible, se dirá que es transitoria, recurrente nula o recurrente positiva si $\mathcal{E}$ es transitoria, recurrente nula o recurrente positiva. Algo que resulta sorprendente es que solamente es necesario ver la transitoriedad o tipo de recurrencia de un único estado en $j \in \mathcal{E}$ para clasificar la cadena.

**Corolario 1.15.** *Sea $\boldsymbol{X}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $\mathcal{C} \subset \mathcal{E}$ es cerrada, irreducible y finita, $\mathcal{C}$ es recurrente positiva.*

*Demostración.* Notemos que por ser una clase cerrada y finita, para cada $i \in \mathcal{C}$,

$$\sum_{j\in\mathcal{C}} P_{ij}^m = 1.$$

Sumando sobre $m \in \{1, \ldots, n\}$ y dividiendo entre $n$, para cada $i \in \mathcal{C}$,

$$\sum_{j\in\mathcal{C}} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = 1.$$

Por el teorema 1.7, cuando $n \to \infty$,

$$\sum_{j\in\mathcal{C}} \frac{1}{m_j} = 1,$$

de donde debe existir $k \in \mathcal{C}$ tal que $1/m_k > 0$ o bien $m_k < \infty$. Por el teorema 1.13 $\mathcal{C}$ es recurrente positiva. ⁜

Ahora damos una serie de resultados inmediatios —por lo que se omiten las pruebas— que serán de utilidad más adelante.

**Corolario 1.16.** *Sea $\boldsymbol{X}$ una cadena de Markov irreducible con espacio de estados $\mathcal{E}$ finito y matriz de transición $P$, entonces $\boldsymbol{X}$ es recurrente positiva.*

**Corolario 1.17.** *Una cadena de Markov $\boldsymbol{X}$ con una cantidad finita de estados no tiene estados recurrentes nulos.*

**Corolario 1.18.** *Una cadena de Markov $\boldsymbol{X}$ irreducible es recurrente positiva si y sólo si tiene distribución estacionaria.*

**Corolario 1.19.** *Una cadena de Markov $\boldsymbol{X}$ irreducible, con espacio de estados $\mathcal{E}$ finito, tiene distribución estacionaria única.*

**Corolario 1.20.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov irreducible y recurrente positiva, con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Entonces*

$$\lim_{n\to\infty} \frac{N_j^n}{n} = \pi(j)$$

*con probabilidad uno para cada $j \in \mathcal{E}$.*

El último corolario resulta de importante a la hora de simular cadenas de Markov irreducibles y recurrentes positivas puesto que se pueden simular los primeros $n$ pasos de la cadena y, si $n$ es suficientemente grande, podremos aproximar $\pi(j)$ para toda $j \in \mathcal{E}$ por medio de $N_j^n/n$, incluso si $\mathcal{E}$ no es finito. ¿Y qué pasa si una cadena no es irreducible? Para poder responder esto, consideremos una clase $\mathcal{C} \subset \mathcal{E}$ y una distribución inicial $\pi$ para una cadena de Markov $\boldsymbol{X}$. Diremos que $\pi$ es concentra en $\mathcal{C}$ si

$$\pi(j) = 0$$

para toda $j \notin \mathcal{C}$. Si además $\mathcal{C}$ es cerrada e irreducible, por el mismo argumento que el teorema 1.13 tendremos el siguiente resultado.

**Teorema 1.21.** *Sea $\boldsymbol{X}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $\mathcal{C} \subset \mathcal{E}$ es cerrada, irreducible y positiva recurrente, entonces $\boldsymbol{X}$ tiene una única distribución estacionaria $\pi$, concentrada en $\mathcal{C}$, dada por*

$$\pi(j) = \frac{1}{m_j}\mathbb{1}_{\mathcal{C}}(j) = \begin{cases} 1/m_j & \text{si } j \in \mathcal{C}, \\ 0 & \text{en otro caso.} \end{cases}$$

Así pues, si $\mathcal{C}_1$ y $\mathcal{C}_2$ son dos clases cerradas, irreducibles y recurrente positivas ajenas, por el teorema 1.21 existirán dos distribuciones estacionarias $\pi^1$ y $\pi^2$ concentradas en $\mathcal{C}_1$ y $\mathcal{C}_2$. Entonces

$$\pi^\alpha = \alpha\pi^1 + (1-\alpha)\pi^2$$

es una distribución estacionaria para cada $\alpha \in [0,1]$, por lo que en este caso, no tendremos una distribución estacionaria única.

**Ejemplo 1.3.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ la cadena de Markov con espacio de estados $\mathcal{E} = \{0,1\}$ y probabilidades de transición $P_{00} = P_{11} = 1$. Entonces $\mathcal{C}_0 = \{0\}$ y $\mathcal{C}_1 = \{1\}$ son clases cerradas irreducibles y, por ser finitas, recurrentes positivas. En el ejemplo 1.2 deducimos que $\pi^0 = (1,0)$ y $\pi^1 = (0,1)$ eran distribuciones estacionarias para $\boldsymbol{X}$, las cuales podemos observar que están concentradas en $\mathcal{C}_0$ y $\mathcal{C}_1$ de forma respectiva. Ahora podemos asegurar que son las únicas que satisfacen esta condición y por lo tanto, todas las distribuciones estacionarias de la cadena están dadas por $\pi^\alpha = (\alpha, 1-\alpha)$ para $\alpha \in [0,1]$.

Para finalizar la sección, daremos un resultado en el cual resumimos la existencia y unicidad de distribuciones estacionarias para cadenas de Markov.

**Proposición 1.22.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Sea $\mathcal{E}_P \subset \mathcal{E}$ el conjunto de estados recurrentes positivos. Entonces se cumple uno y sólo uno de los siguientes incisos:*

1) *Si $\mathcal{E}_P = \varnothing$ entonces la cadena $\boldsymbol{X}$ no tiene distribuciones estacionarias.*
2) *Si $\mathcal{E}_P \neq \varnothing$ es irreducible entonces existe una única distribución estacionaria para $\boldsymbol{X}$.*
3) *Si $\mathcal{E}_P \neq \varnothing$ no es irreducible entonces existe una colección infinita de distribuciones estacionarias para $\boldsymbol{X}$.*

Queda pendiente el estudiar el comportamiento asintótico de $P_{ij}^n$ para ciertas cadenas de Markov, antes de estudiar obtendremos las distribuciones estacionarias para algunas cadenas de Markov, especificando cuándo son únicas y cuándo no lo son.

## §1.3. Ejemplos

El objetivo de esta sección es ejemplificar el uso de los resultados de la sección anterior para obtener distribuciones estacionarias de algunas cadenas de Markov y así estudiar el comportamiento asintótico de las cadenas.

**Ejemplo 1.4.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E} = \{0,1,2,3,4\}$ y matriz de transición

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

Observemos que los estados $\mathcal{E}_T = \{1,3\}$ son transitorios mientras que los estados $\mathcal{E}_R = \{0,2,4\}$ son recurrentes; más aún, por ser $\mathcal{E}$ finito son recurrentes positivos. Además, $\mathcal{C}_1 = \{0,4\}$ y $\mathcal{C}_2 = \{2\}$ son clases cerradas e irreducibles. Entonces habrá una distribución estacionaria $\pi^1$ concentrada en $\mathcal{C}_1$ y otra $\pi^2$ concentrada en $\mathcal{C}_2$. Para obtener $\pi^1$, consideremos la submatriz de transición para $\mathcal{C}_1$, dada por

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

La única distribución estacionaria para esta matriz está dada por $\tilde{\pi}^1 = (1/2, 1/2)$ y, por lo tanto, $\pi^1 = (1/2, 0, 0, 0, 1/2)$. De forma similar, se obtenemos $\pi^2 = (0,0,1,0,0)$, por lo que las distribuciones estacionarias de $\boldsymbol{X}$ están caracterizadas por $\pi^\alpha = (\alpha/2, 0, 1-\alpha, 0, \alpha/2)$ para $\alpha \in [0,1]$.

Recordemos por otra parte que para cualesquiera $i, j \in \mathcal{E}$ se cumplen

$$\lim_{n\to\infty} \frac{N_j^n}{n} = \frac{1}{m_j} \mathbb{1}_{\{T_j<\infty\}}$$

con probabilidad uno y

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = \frac{\rho_{ij}}{m_j}.$$

Esto lo podemos usar para estudiar la proporción de pasos que estará una cadena en el estado $j \in \mathcal{E}$ si $X_0 = i$, pues en este caso

$$\lim_{n\to\infty} \frac{N_j^n}{n} = \frac{1}{m_j} \mathbb{1}_{\{T_j<\infty, X_0=i\}}$$

con probabilidad uno. Esto último lo podemos escribir para toda $i \in \mathcal{E}$ y toda $j \in \mathcal{E}$ mediante el arreglo matricial

$$\begin{pmatrix} 1/m_0 & 0 & 0 & 0 & 1/m_4 \\ \mathbb{1}_{\{T_0<\infty, X_0=1\}}/m_0 & 0 & \mathbb{1}_{\{T_2<\infty, X_0=1\}}/m_2 & 0 & \mathbb{1}_{\{T_4<\infty, X_0=1\}}/m_4 \\ 0 & 0 & 1/m_2 & 0 & 0 \\ \mathbb{1}_{\{T_0<\infty, X_0=3\}}/m_0 & 0 & \mathbb{1}_{\{T_3<\infty, X_0=3\}}/m_2 & 0 & \mathbb{1}_{\{T_4<\infty, X_0=3\}}/m_4 \\ 1/m_0 & 0 & 0 & 0 & 1/m_4 \end{pmatrix}$$

y por el teorema 1.21, $m_0 = 1/\pi^1(0)$, $m_2 = 1/\pi^2(2)$ y $m_4 = 1/\pi^1(4)$, de donde la matriz anterior es igual a

$$\begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ \mathbb{1}_{\{T_0<\infty, X_0=1\}}/2 & 0 & \mathbb{1}_{\{T_2<\infty, X_0=1\}} & 0 & \mathbb{1}_{\{T_4<\infty, X_0=1\}}/2 \\ 0 & 0 & 1 & 0 & 0 \\ \mathbb{1}_{\{T_0<\infty, X_0=3\}}/2 & 0 & \mathbb{1}_{\{T_3<\infty, X_0=3\}} & 0 & \mathbb{1}_{\{T_4<\infty, X_0=3\}}/2 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

Por otra parte, podemos resumir la información que nos da

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = \frac{\rho_{ij}}{m_j}$$

para cualesquiera $i, j \in \mathcal{E}$ en la siguiente matriz

$$\begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ \rho_{10}/2 & 0 & \rho_{12} & 0 & \rho_{14}/2 \\ 0 & 0 & 1 & 0 & 0 \\ \rho_{30}/2 & 0 & \rho_{32} & 0 & \rho_{34}/2 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix},$$

por lo que basta encontrar $\rho_{10}, \rho_{12}, \rho_{14}, \rho_{30}, \rho_{32}$ y $\rho_{34}$ para obtener todos los límites explícitamente. Para obtener estos valores, podemos obtener $\rho_{\mathcal{C}_1}(1), \rho_{\mathcal{C}_1}(3), \rho_{\mathcal{C}_2}(1)$ y $\rho_{\mathcal{C}_2}(3)$ pues si entramos a una clase cerrada, casi seguramente visitaremos todos los estados en la clase.

0.207879576350761938641894445615708196827575653285222461198725984262...

Más aún, como solamente hay dos clases cerradas, $\rho_{\mathcal{C}_1}(1) = 1 - \rho_{\mathcal{C}_2}(1)$ y $\rho_{\mathcal{C}_1}(3) = 1 - \rho_{\mathcal{C}_2}(3)$, entonces bastará resolver el sistema de ecuaiones

$$\begin{aligned} \rho_{\mathcal{C}_1}(1) &= \frac{1}{2}\rho_{\mathcal{C}_1}(1) + \frac{1}{2}\rho_{\mathcal{C}_1}(3), \\ \rho_{\mathcal{C}_1}(3) &= \frac{1}{4}\rho_{\mathcal{C}_1}(1) + \frac{1}{4}\rho_{\mathcal{C}_1}(3) + \frac{1}{4}, \end{aligned}$$

de donde $\rho_{\mathcal{C}_1}(1) = \rho_{\mathcal{C}_1}(3) = 1/2$ y, por lo tanto,

$$\begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ \rho_{10}/2 & 0 & \rho_{12} & 0 & \rho_{14}/2 \\ 0 & 0 & 1 & 0 & 0 \\ \rho_{30}/2 & 0 & \rho_{32} & 0 & \rho_{34}/2 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 1/4 & 0 & 1/2 & 0 & 1/4 \\ 0 & 0 & 1 & 0 & 0 \\ 1/4 & 0 & 1/2 & 0 & 1/4 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

De esta forma hemos obtenido el comportamiento asintótico de la cadena dependiendo del estado en el que comience la cadena $\boldsymbol{X}$.

**Ejemplo 1.5.** Consiremos que $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ es una cadena de nacimiento y muerte en $\mathbb{N}$ irreducible, es decir que $p_i := P_{i,i+1} > 0$ para $i \in \mathbb{N}$ y $q_i := P_{i,i-1}$ para $i \geq 1$. Denotemos además $r_i = P_{ii}$ para toda $i \in \mathbb{N}$. Para obtener una distribución estacionaria $\pi$, se debe satisfacer $\pi = \pi P$, que nos da el sistema de ecuaciones

$$\begin{aligned} \pi(0) &= \pi(0)r_0 + \pi(1)q_1, \\ \pi(i) &= \pi(i-1)p_{i-1} + \pi(i)r_i + \pi(i+1)q_{i+1} \qquad \text{para } i \geq 1. \end{aligned}$$

De $1 - r_i = p_i + q_i$, el sistema de ecuaciones anterior se reduce a

$$\begin{aligned} q_1\pi(1) - p_0\pi(0) &= 0, \\ q_{i+1}\pi(i+1) - p_i\pi(i) &= q_i\pi(i) - p_{i-1}\pi(i-1) \qquad \text{para } i \geq 1. \end{aligned}$$

De un argumento inductivo se sigue que

$$q_{i+1}\pi(i+1) - p_i\pi(i) = 0$$

para toda $i \in \mathbb{N}$ y entonces

$$\pi(i+1) = \frac{p_i}{q_{i+1}}\pi(i)$$

es cierto para toda $i \in \mathbb{N}$. Al definir

$$\eta_i = \frac{\prod_{j=0}^{i-1} p_j}{\prod_{j=1}^{i} q_j},$$

donde el producto vacío es uno, tendremos de forma recursiva que $\pi(i) = \eta_i\pi(0)$ para toda $i \in \mathbb{N}$. Si $\sum_{i=0}^{\infty} \eta_i < \infty$, la cadena de nacimiento y muerte tiene una distribución estacionaria $\pi$ dada por

$$\pi(i) = \frac{\eta_i}{\sum_{j=0}^{\infty} \eta_j}$$

para toda $i \in \mathbb{N}$, que además es única por ser una $\boldsymbol{X}$ irreducible. Como la cadena es irreducible y existe una distribución estacionaria cuando $\sum_{i=0}^{\infty} \eta_i < \infty$, vemos que una condición necesaria y suficiente para que la cadena de nacimiento y muerte sea recurrente positiva es que

$$\sum_{i=0}^{\infty} \frac{\prod_{j=0}^{i-1} p_j}{\prod_{j=1}^{i} q_j} < \infty.$$

En la primera parte de cadenas de Markov, se vio que

$$\sum_{i=0}^{\infty} \frac{\prod_{j=1}^{i} q_j}{\prod_{j=0}^{i-1} p_j} < \infty$$

es condición necesaria y suficiente para que una cadena de nacimiento y muerte irreducible sea transitoria. Así pues, la cadena de nacimiento y muerte será recurrente nula si y sólo si

$$\sum_{i=0}^{\infty} \frac{\prod_{j=0}^{i-1} p_j}{\prod_{j=1}^{i} q_j} = \infty \quad \text{y} \quad \sum_{i=0}^{\infty} \frac{\prod_{j=1}^{i} q_j}{\prod_{j=0}^{i-1} p_j} = \infty.$$

Así las cosas, hemos logrado clasificar todas las cadenas de nacimiento y muerte irreducibles en transitorias, recurrentes positivas y recurrentes nulas.

**Ejemplo 1.6.** Supongamos que $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ es una cadena de nacimiento y muerte irreducible en $\mathcal{E} = \{0, 1, \ldots, d\}$, con $d \geq 1$. Siguiendo un procedimiento similar al del ejemplo 1.5, la cadena tendrá una única distribución estacionaria $\pi$ dada por

$$\pi(i) = \frac{\eta_i}{\sum_{j=0}^{d} \eta_j},$$

donde

$$\eta_i = \frac{\prod_{j=0}^{i-1} p_j}{\prod_{j=1}^{i} q_j}.$$

**Ejemplo 1.7.** Consideremos el modelo de difusión de Bernoulli−Laplace con 6 partículas, que es una cadena de Markov $\boldsymbol{X}$ con espacio de estados $\mathcal{E} = \{0, 1, 2, 3\}$ y matriz de transición

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/9 & 4/9 & 4/9 & 0 \\ 0 & 4/9 & 4/9 & 1/9 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Esta cadena resulta ser una cadena de nacimiento y muerte. Por el ejemplo 1.6, $\eta_0 = 1$ y

$$\eta_1 = \frac{1}{\frac{1}{9}} = 9, \quad \eta_2 = \frac{1 \times \frac{4}{9}}{\frac{1}{9} \times \frac{4}{9}} = 9 \quad \text{y} \quad \eta_3 = \frac{1 \times \frac{4}{9} \times \frac{1}{9}}{\frac{1}{9} \times \frac{4}{9} \times 1} = 1,$$

de donde la única distribución única está dada por

$$\pi = \left(\frac{1}{20}, \frac{9}{20}, \frac{9}{20}, \frac{1}{20}\right).$$

**Ejemplo 1.8.** Consideremos ahora una modificación de la cadena de Ehrenfest con 4 bolas etiquetadas del 1 al 4. Supongamos que inicialmente algunas bolas están en la urna 1 y el resto se encuentra en la urna 2. Se selecciona una entero de $\{1, 2, 3, 4\}$ y se saca la bola etiquetada con tal entero de la urna en la que se encuente, para después seleccionar de forma equiprobable a qué urna se meterá la bola. Si denotamos por $X_n$ al número de bolas en la urna 1 después del $n$-ésimo experimento, $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ es una cadena de Markov con espacio de estados $\mathcal{E} = \{0, 1, 2, 3, 4\}$ y matriz de transición

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

Notemos que nuevamente tendremos una cadena de nacimiento y muerte irreducible. Siguiendo el ejemplo 1.6, su única distribución estacionaria estará dada por

$$\pi = \left( \frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16} \right).$$

**Ejemplo 1.9.** Consideremos una cadena de rachas con probabilidad $0 < p < 1$, es decir $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ es una cadena de Markov con espacio de estados $\mathbb{N}$ y probabilidades de transición $P_{i,i+1} = 1 - P_{i0} = p$. Es claro que esta cadena es irreducible y además

$$m_0 = \mathbb{E}[T_0 \mid X_0 = 0] = \frac{p}{1-p} < \infty,$$

por lo que la cadena es recurrente positiva. Así, debe tener una única distribución estacionaria $\pi$. De $\pi = \pi P$ obtenemos el sistema de ecuaciones,

$$\begin{aligned} \pi(0) &= \sum_{i=0}^{\infty} (1-p)\pi(i), \\ \pi(i) &= p\pi(i-1) \quad \text{para } i \geq 1. \end{aligned}$$

Como $\sum_{i=0}^{\infty} \pi(i) = 1$, se deduce que $\pi(i) = p^i(1-p)$ para toda $i \in \mathbb{N}$.

**Ejemplo 1.10.** Supongamos que una partícula se mueve sobre los vértices de un cubo de la siguiente manera: en cada paso la partícula se mueve a cualquier vértice adyacente al que se encuentre, que se escoge de forma equiprobable, independientemente de su trayectoria pasada. Sea $I$ el vértice inicial y $O$ el vértice opuesto a $I$. Vamos a calcular las siguientes cantidades:

1) El número esperado de pasos para regresar a $I$.

2) El número esperado de visitas a $O$ antes de regresar a $I$.

3) El número esperado de pasos para visitar $O$ por primera vez.

Denotemos por $A$ al conjunto de vértices adyacentes a $I$ y por $B$ al conjunto de vértices adyacentes a $O$. Si $X_n$ es la posición de la partícula después del $n$-ésimo movimiento, $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ es una cadena de Markov con espacio de estados $\mathcal{E} = \{I, A, B, O\}$ y matriz de transición $P$ dada por

$$\begin{array}{c} \\ I \\ A \\ B \\ O \end{array} \begin{array}{c} \begin{array}{cccc} I & A & B & O \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}.$$

Por lo tanto tenemos una cadena de nacimiento y muerte irreducible con distribución estacionaria

$$\pi = \left(\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}\right).$$

Recordando que $m_j = 1/\pi(j)$, la solución de (1) está dada por $m_I = 8$. Para (2), estamos buscando $\gamma^I_O$. Recordemos que $\gamma^I$ es una medida invariante con $\gamma^I_I = 1$ tal que $\gamma^I \leq \mu$ para cualquier otra medida invariante $\mu$ con $\mu(I) = 1$. De hecho, por ser $\boldsymbol{X}$ irreducible, la desigualdad es igualdad. Notemos que $\mu = 8\pi$ es una medida invariante tal que $\mu(I) = 1$ y, por lo tanto, $\gamma^I_O = 8\pi(O) = 1$. Finalmente buscamos $\mathbb{E}[T_O \mid X_0 = I]$, para lo que basta resolver el sistema de ecuaciones

$$\begin{aligned}
\mathbb{E}[T_O \mid X_0 = I] &= \mathbb{E}[T_O \mid X_0 = A] + 1, \\
\mathbb{E}[T_O \mid X_0 = A] &= \frac{1}{3}\mathbb{E}[T_O \mid X_0 = I] + \frac{2}{3}\mathbb{E}[T_O \mid X_0 = B] + 1 \\
\mathbb{E}[T_O \mid X_0 = B] &= \frac{2}{3}\mathbb{E}[T_O \mid X_0 = I] + 1,
\end{aligned}$$

de donde $\mathbb{E}[T_O \mid X_O = I] = 10$.

Para poder entender los siguientes ejemplos, necesitamos un par de definiciones nuevas, además de un resultado auxiliar.

**Definición 1.23.** Sea $\boldsymbol{X}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Diremos que $\boldsymbol{X}$, o bien la matriz $P$, es *doblemente estocástica* si

$$\sum_{i \in \mathcal{E}} P_{ij} = 1$$

se satisface para toda $j \in \mathcal{E}$.

**Definición 1.24.** Sean $\boldsymbol{X}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$, además de $\pi$ una medida, *i.e.* $\pi : \mathcal{E} \to \mathbb{R}$ con $\pi(i) \geq 0$ para todo estado $i \in \mathcal{E}$. Diremos que $\pi$ y $\boldsymbol{X}$ están en *balance detallado* si para cualesquiera $i, j \in \mathcal{E}$ se cumple que

$$\pi(i)P_{ij} = \pi(j)P_{ji}.$$

**Proposición 1.25.** *Si $\boldsymbol{X}$ y $\pi$ están en balance detallado entonces $\pi$ es una medida invariante para $\boldsymbol{X}$.*

*Demostración.* Para cada $j \in \mathcal{E}$,

$$\sum_{i \in \mathcal{E}} \pi(i)P_{ij} = \sum_{i \in \mathcal{E}} \pi(j)P_{ji} = \pi(j). \qquad ⁜$$

**Ejemplo 1.11.** Si $\boldsymbol{X}$ es una cadena de Markov irreducible, con espacio de estados $\mathcal{E}$ finito y doblemente estocástica, entonces tiene una única distribución estacionaria $\pi$, donde

$$\pi(j) = \frac{1}{K},$$

con $K$ la cardinalidad de $\mathcal{E}$.

**Ejemplo 1.12.** Un crupier se encuentra aburrido porque no llegan jugadores a su mesa de blackjack, por lo que se pone a barajar una baraja de 52 cartas repetidamente. De tan aburrido que está se hace la pregunta de cuántas veces deberá barajar los naipes para regresar al orden original bajo dos esquemas:

1) Baraja todo el mazo de cartas y cualquier orden de las cartas es equiprobable como resultado.

2) Saca una carta al azar y la coloca en una posición distinta a la original, escogida de forma equiprobable.

Consideremos que $X_n$ es el orden de la baraja después de la $n$-ésima barajada y denotemos por $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ la cadena de Markov que se obtiene, con espacio de estados $\mathcal{E}$ dado por los posibles ordenamientos de los naipes. Observemos que si existe un barajeo que lleve del estado $i$ al estado $j$, ese mismo barajeo llevará de un estado $k$ al estado $j$, de donde el $j$-ésimo renglón de la matriz $P$ asociada a $\boldsymbol{X}$ es la $j$-ésima columna en un orden distinto, por lo que $\boldsymbol{X}$ es una cadena doblemente estocástica. Así pues, en promedio se requerirán 52! barajeos, bajo cualquiera de los esquemas, para regresar al orden en el que se encontraba la baraja de cartas inicialmente.

**Ejemplo 1.13** (Caminata aleatoria sobre gráficas). Una gráfica es un par $G = (V, E)$ de conjuntos tal que $E \subset [V]^2$, donde $[V]^k = \{A \subset V : |A| = k\}$; esto es, los elementos de $E$, llamados aristas, son subconjuntos de dos elementos de $V$, llamados vértices. Diremos que $i \in V$ y $j \in V$ son vecinos si $\{i, j\} \in E$. Para $i, j \in V$ diremos que podemos ir de $i$ a $j$ si existen elementos $i_0, i_1, \ldots, i_n \in V$, con $i_0 = i$ y $j = i_n$ tales que $\{i_{k-1}, i_k\} \in E$ para $k \in \{1, \ldots, n\}$. La gráfica será conexa si para cualesquiera $i, j \in V$, se puede ir de $i$ a $j$. Además se define la valencia $v_i$ de $i \in V$ como la cantidad de vecinos que tiene $i$, *i.e.* $v_i = |\{e \in E : i \in e\}|$.

Definimos una caminata aleatoria sobre una gráfica $G$ conexa, con $V$ finito de cardinalidad al menos 2, de la siguiente manera: si nos encontramos en $i \in V$, iremos a cualquiera de sus vecinos de forma equiprobable, independientemente de la trayectoria pasada. Así, si $X_n$ es el vértice después del $n$-ésimo movimiento, tendremos que $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ es una cadena de Markov irreducible con espacio de estados $V$ y probabilidades de transición

$$P_{ij} = \begin{cases} \frac{1}{v_i} & \text{si } \{i, j\} \in E, \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces $\mu = \{v_i : i \in V\}$ es una medida que está en balance detallado con $\boldsymbol{X}$. En efecto $0 < v_i \leq |V| - 1 < \infty$ y para cualesquiera $i, j \in V$,

$$\mu(i)P_{ij} = v_i \frac{1}{v_i} = 1 = v_j \frac{1}{v_j} = \mu(j)P_{ji}$$

si $\{i, j\} \in E$ y

$$\mu(i)P_{ij} = \mu(j)P_{ji} = 0$$

en otro caso. Ahora observemos que $0 < \sum_{i\in V} v_i \leq |E| \leq \binom{|V|}{2} < \infty$ y, por lo tanto, si definimos $\pi$ por

$$\pi(i) = \frac{v_i}{\sum_{j\in V} v_j},$$

obtendremos la única distribución estacionaria de la cadena.

0.207879576350761938641894445615708196827575653285222461198725984262...

**Ejemplo 1.14.** Consideremos un caballo en un tablero de ajedrez que comienza en una esquina y se mueve aleatoriamente de la siguiente manera: puede ir, de forma equiprobable, a cualquier casilla que le sea permitido bajo su movimiento en L, ver la figura 1, y nos preguntamos cuántos movimientos tardará en promedio el caballo en regresar a la esquina de la que comenzó. Así, los movimientos aleatorios del caballo forman una caminata aleatoria sobre una gráfica conexa y finita.

De la figura 1, vemos que hay 4 vértices de valencia 2, 8 vértices de valencia 3, 20 vértices de valencia 4, 16 vértices de valencia 6 y 16 vértices de valencia 8. Luego, por el ejemplo 1.13, si denotamos por $c$ la esquina en la que comenzó el caballo y por $X_n$ su posición tras el $n$-ésimo movimiento,

$$\mathbb{E}[T_c \,|\, X_0 = c] = \frac{1}{\pi(c)} = \frac{\sum_{i\in V} v_i}{v_c} = \frac{8+24+80+96+128}{2} = 168.$$

Es decir que el caballo tardará en promedio 168 movimientos en regresar a la esquina en la que comenzó.

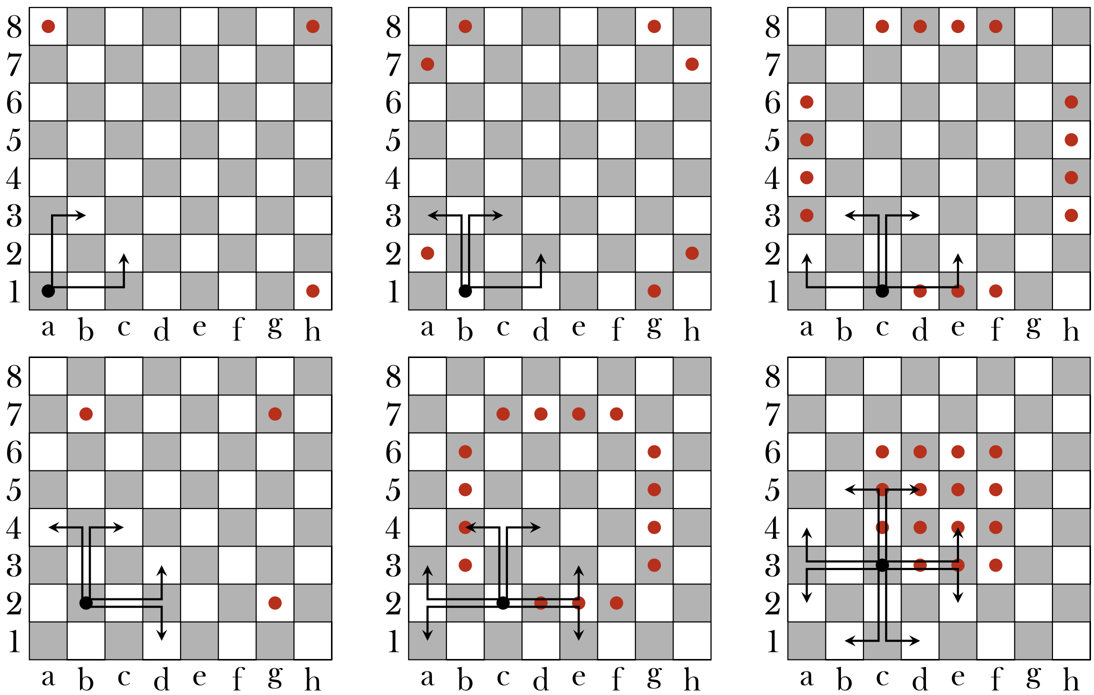


Figura 1: Las flechas muestran los posibles movimientos del caballo comenzando en el círculo negro, y los círculos rojos son posiciones con *la misma cantidad de movimientos* —similares— a los del círculo negro.

Hasta el momento hemos obtenido resultados para la existencia y unicidad de las distribuciones estacionarias así como para el comportamiento asintótico de $n^{-1}\sum_{m=1}^{n} P_{ij}^{m}$. Para finalizar las notas, en la siguiente sección estudiaremos el comportamiento de $P_{ij}^{n}$ para cadenas de Markov irreducibles.

0.20787957635076193864189444561570819682757565328522246119872598426 2...

## §1.4. Convergencia a la distribución estacionaria

En la sección 1.2, vimos que si $\boldsymbol{X}$ es una cadena de Markov irreducible y positiva recurrente, con distribución estacionaria $\pi$, entonces para cualesquiera $i, j \in \mathcal{E}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij} = \pi(j) = \frac{1}{m_j}.$$

Ahora estudiaremos el comportamiento de $P_{ij}^n$ cuando $n \to \infty$. En la proposición vimos que si el espacio de estados es finito y el límite existe para alguna $i \in \mathcal{E}$ y toda $j \in \mathcal{E}$ entonces debe ser una distribución estacionaria. Aquí veremos una refinación de ese resultado, cuando es que

$$\lim_{n\to\infty} P_{ij}^n = \pi(j)$$

se satisface para cualesquiera $i, j \in \mathcal{E}$ y qué pasa cuando este límite no existe, restringiéndonos a cadenas irreducibles y recurrentes positivas. Veamos que, en efecto, el límite puede no existir.

**Ejemplo 1.15.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E} = \{0, 1\}$ y matriz de transición

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

Es claro que

$$P^n = \begin{cases} P & \text{si } n \text{ es impar,} \\ \mathrm{Id} & \text{si } n \text{ es par,} \end{cases}$$

donde Id es la matriz identidad. Así, $\lim_{n\to\infty} P_{ij}^n$ no existe. Sin embargo, observemos que tanto $\lim_{n\to\infty} P_{ij}^{2n}$ como $\lim_{n\to\infty} P_{ij}^{2n+1}$ sí existen para cualesquiera $i, j \in \mathcal{E}$.

El ejemplo anterior da pauta para la siguiente definicion.

**Definición 1.26.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $i \in \mathcal{E}$ es tal que exista $n \geq 1$ tal que $P_{ii}^n > 0$, se define su *perído* como

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

Si $d_i = 1$ se dice que $i$ es *aperiódico*.

Regresando al ejemplo 1.15, vemos que $P_{ii}^{2n} = 1$ y que $P_{ii}^{2n+1} = 0$, por lo que el período de cada uno de los estados es 2. Por otra parte, no es difícil convencerse de que todo estado absorbente es aperiódico. Ahora veremos que la periodicidad es una propiedad de clase.

**Proposición 1.27.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Sean $i \in Ee$ y $j \in \mathcal{E}$ tales que $P_{ii}^n, P_{jj}^m > 0$ para algunas $n, m \geq 1$. Si $\rho_{ij}, \rho_{ji} > 0$, entonces $d_i = d_j$.*

*Demostración.* Como $\rho_{ij}, \rho_{ji} > 0$ existen $n_1, n_2 \geq 1$ tales que $P_{ij}^{n_1}, P_{ji}^{n_2} > 0$. Entonces, por Chapman–Kolmogorov,

$$P_{ii}^{n_1+n_2} = \sum_{k\in\mathcal{E}} P_{ik}^{n_1} P_{ki}^{n_2} \geq P_{ij}^{n_1} P_{ji}^{n_2} > 0,$$

0.207879576350761938641894445615708196827575653285222461198725984262…

de donde $d_i$ es divisor de $n_1 + n_2$. Ahora supongamos que $m$ es tal que $P_{jj}^m > 0$, entonces, nuevamente por Chapman–Kolmogorov,

$$P_{ii}^{n_1+m+n_2} \geq P_{ij}^{n_1} P_{jj}^m P_{ji}^{n_2} > 0,$$

por lo que $d_i$ divide a $(n_1 + n_2) + m$ y, como $d_i$ también divide a $n_1 + n_2$, concluimos que $d_i$ divide a $m$. Por lo tanto, $d_i$ es divisor de todos los números en el conjunto

$$\{m \geq 1 : P_{jj}^m > 0\},$$

y por definición $d_j$ es el máximo de los divisores de tal conjunto. Luego, $d_i \leq d_j$. De forma similar se deduce que $d_j \leq d_i$, por lo que $d_i = d_j$. ⁂

En particular hemos obtenido que si $\mathcal{C} \subset \mathcal{E}$ es una clase irreducible de estados, entonces todos los estados en la clase $\mathcal{C}$ tienen el mismo período $d$. Cuando $d = 1$, se dice que la clase es *aperiódica*. Más aún, si $\mathcal{C} = \mathcal{E}$ entoces $\boldsymbol{X}$ será irreducible y, por lo tanto, todos sus estados tendrán el mismo período $d$. En este caso diremos que $\boldsymbol{X}$ es una cadena de período $d$ si $d > 1$ y diremos que es una cadena *aperiódica* si $d = 1$.

**Ejemplo 1.16.** Consideremos una cadena de rachas con $p \in (0, 1)$, la cual es irreducible. Por hipótesis, $P_{00} = 1 - p > 0$, por lo que $d_0 = 1$. Así pues, la cadena de rachas resulta ser aperiódica.

**Ejemplo 1.17.** Consideremos la cadena de Ehrenfest. Es fácil convencerse de que $P_{ii}^{2n+1} = 0$ para cualquier $n \in \mathbb{N}$ mientras que $P_{ii}^{2n} > 0$ para toda $n \geq 1$, de donde el período de la cadena de Ehrenfest es 2.

Ahora veremos que, si tenemos una cadena irreducible, aperiódica y recurrente positiva, entonces su comportamiento asintótico será "ideal" y veremos lo que sucede para cadenas irreducibles y recurrentes positivas pero con período $d > 1$. La técnica con la que se demostrará el teorema es conocida como *acoplamiento*, en la que la idea general es comparar el comportamiento de dos procesos estocásticos $\boldsymbol{X}$ y $\boldsymbol{Y}$ construyendo copias $\tilde{X}$ y $\tilde{Y}$ de estos procesos en un espacio de probabilidad común. Las pruebas que usan la técnica de acoplamiento suelen ser transparentes y simples.

**Teorema 1.28.** *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov irreducible, recurrente positiva, con distribución estacionaria $\pi$ y distribución inicial $\mu$. Si la cadena es aperiódica, entonces*

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \pi(j) \tag{9}$$

*para toda $j \in \mathcal{E}$. En particular*

$$\lim_{n\to\infty} P_{ij}^n = \pi(j) \tag{10}$$

*para cualesquiera $i, j \in \mathcal{E}$.*

*Si la cadena tiene período $d > 1$, entonces para cada $i \in \mathcal{E}$ y cada $j \in \mathcal{E}$ existe un entero $r \in \{0, 1, \ldots, d-1\}$ tal que*

$$\lim_{n\to\infty} P_{ij}^{nd+r} = d\pi(j). \tag{11}$$

*Demostración.* **Caso aperiódico.** Sea $\boldsymbol{Y} = \{Y_n\}_{n\in\mathbb{N}}$ una cadena de Markov, irreducible, recurrente positiva, con distribución inicial $\pi$ e independiente[3] de $\boldsymbol{X}$ tal que $\mathbb{P}(Y_1 = j \,|\, Y_0 = i) = \mathbb{P}(X_1 = j \,|\, X_0 = i) = P_{ij}$ para cualesquiera $i, j \in \mathcal{E}$. Fijemos $k \in \mathcal{E}$ y definamos

$$\tau = \inf\{n \geq 1 : X_n = Y_n = k\},$$

[3] Esto nos dice que $\boldsymbol{X}$ y $\boldsymbol{Y}$ están definidas en el mismo espacio de probabilidad.

la primera vez que ambas cadenas son iguales a $k$.

*Paso 1.* Mostraremos que $\mathbb{P}(\tau < \infty) = 1$. Con este fin en mente definamos $Z_n = (X_n, Y_n)$ para toda $n \in \mathbb{N}$. Entonces $\boldsymbol{Z} = \{Z_n\}_{n\in\mathbb{N}}$ es una cadena de Markov con espacio de estados $\mathcal{E} \times \mathcal{E}$, probabilidades de transición

$$P_{(i,i')(j,j')} = P_{ij}P_{i'j'}$$

y distribución inicial $\lambda$ dada por

$$\lambda(i, i') = \mu(i)\pi(i).$$

Como $\boldsymbol{X}$ es aperiódica, $\boldsymbol{Y}$ también lo es y consecuentemente $\boldsymbol{Z}$ es aperiódica. Más aún, $\boldsymbol{Z}$ es irreducible, pues $P^n_{(i,i')(j,j')} = P^n_{ij}P^n_{i'j'}$ para toda $n \in \mathbb{N}$. Finalmente observamos que $\boldsymbol{Z}$ tiene distribución estacionaria $\tilde{\pi}$ dada por

$$\tilde{\pi}(i, i') = \pi(i)\pi(i'),$$

entonces, por el teorema 1.13, $\boldsymbol{Z}$ es recurrente positiva. Entonces $\tau$ es la primera vez que $\boldsymbol{Z}$ visita el estado $(k, k)$ y, por lo tanto

$$\mathbb{P}(\tau < \infty) = \sum_{(i,i')\in\mathcal{E}\times\mathcal{E}} \mathbb{P}(\tau < \infty \mid Z_0 = (i, i'))\lambda(i, i') = \sum_{(i,i')\in\mathcal{E}\times\mathcal{E}} \lambda(i, i') = 1,$$

donde la segunda igualdad se da por ser $\boldsymbol{Z}$ irreducible y recurrente.

*Paso 2.* Probaremos que para $n \geq 1$ y para $j \in \mathcal{E}$ se cumple

$$\mathbb{P}(X_n = j, \tau \leq n) = \mathbb{P}(Y_n = j, \tau \leq n). \tag{12}$$

En efecto, para $m \in \{1, \dots, n\}$,

$$\mathbb{P}(X_n = j \mid \tau = m, X_m = Y_m = k) = P^{n-m}_{kj} = \mathbb{P}(Y_n = j \mid \tau = m, X_m = Y_m = k),$$

y como $\{\tau \leq n\} = \bigcup_{m=1}^n \{\tau_m, X_m = Y_m = k\}$, por el ejercicio 4 del capítulo 1 de [Hoe72],

$$\mathbb{P}(X_n = j \mid \tau \leq n) = \mathbb{P}(Y_n = j \mid \tau \leq n).$$

Consecuentemente (12) se satisface.

*Paso 3.* Demostraremos el teorema para el caso aperiódico. Por (12) tendremos que

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

De manera análoga

$$\mathbb{P}(Y_n = j) \leq \mathbb{P}(X_n = j) + \mathbb{P}(\tau > n).$$

Así pues, para $n \geq 1$ se cumplirá que

$$|\mathbb{P}(X_n = j) - \mathbb{P}(Y_n = j)| \leq \mathbb{P}(\tau > n)$$

para cualquier $j \in \mathcal{E}$. Por hipótesis $\pi$ es la distribución inicial y estacionaria de $\boldsymbol{Y}$, de donde $\mathbb{P}(Y_n = j) = \pi(j)$ y además $\tau < \infty$ casi seguramente. Así las cosas,

$$\limsup_{n\to\infty} |\mathbb{P}(X_n = j) - \pi(j)| \leq \lim_{n\to\infty} \mathbb{P}(\tau > n) = \mathbb{P}(\tau = \infty) = 0,$$

de donde se deduce que

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \pi(j)$$

se satisface para cualquier $j \in \mathcal{E}$, demostrando (9). En particular (10) se satisface cuando consideramos $\mu$ dada por

$$\mu(k) = \begin{cases} 1 & \text{si } k = i, \\ 0 & \text{en otro caso.} \end{cases}$$

Esto prueba el resultado en el caso aperiódico.

**Caso periódico.** *Paso 1.* Extenderemos el resultado del caso aperiódico como sigue: Consideremos $\mathcal{C} \subset \mathcal{E}$ cerrada, irreducible, aperiódica y recurrente positiva y sea $\pi^{\mathcal{C}}$ la única distribución estacionaria concentrada en $\mathcal{C}$. Si $\mathbb{P}(X_0 \in \mathcal{C}) = 1$ podemos pensar en que $\boldsymbol{X}$ es una cadena con espacio de estados $\mathcal{C}$ y por el caso aperiódico,

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \pi^{\mathcal{C}}(j) = \frac{1}{m_j}.$$

para toda $j \in \mathcal{C}$, en particular si $j$ es un estado recurrente aperiódico,

$$\lim_{n\to\infty} P_{jj}^n = \frac{1}{m_j}. \tag{13}$$

*Paso 2.* Probaremos que para $i = j \in \mathcal{E}$, (11) se satisface con $r = 0$. Sea entonces $\boldsymbol{Y} = \{Y_n\}_{n\in\mathbb{N}}$, donde $Y_n = X_{nd}$, una cadena de Markov con matriz de transición $Q = P^d$. Para $j \in \mathcal{E}$,

$$\operatorname{mcd}\{m \geq 1 : Q_{jj}^m > 0\} = \operatorname{mcd}\{m \geq 1 : P_{jj}^{md} > 0\} = d^{-1}\operatorname{mcd}\{n \geq 1 : P_{jj}^n > 0\} = 1,$$

es decir $j \in \mathcal{E}$ es aperiódico. Por otra parte tenemos que $\mathbb{P}(T_j = m \mid X_0 = j) = 0$ para toda $m \neq nd$, con $n \geq 1$, donde $T_j = \inf\{n \geq 1 : X_n = j\}$; además, condicional en $X_0 = j$, $T_j = nd$ si y sólo si $T_j' = n$, donde $T_j' = \inf\{n \geq 1 : Y_n = j\}$, por lo que

$$\begin{aligned} m_j = \mathbb{E}[T_j \mid X_0 = j] &= \sum_{n=1}^{\infty} nd\mathbb{P}(T_j = nd \mid X_0 = j) \\ &= d\sum_{n=1}^{\infty} n\mathbb{P}(T_j' = n \mid Y_0 = j) = d\mathbb{E}[T_j' \mid Y_0 = j], \end{aligned}$$

y de esta forma, por (13),

$$\lim_{n\to\infty} P_{jj}^{nd} = \lim_{n\to\infty} Q_{jj}^n = \frac{d}{m_j} = d\pi(j).$$

*Paso 3.* Para finalizar demostraremos que (11) se satisface en general. Así, sean $i, j \in \mathcal{E}$ cualquier par de estados y sea

$$r_1 = \inf\{n \geq 1 : P_{ij}^n > 0\}.$$

Sea $n_1$ tal que $P_{ji}^{n_1} > 0$, por Chapman–Kolmogorov,

$$P_{jj}^{n_1+r_1} \geq P_{ji}^{n_1} P_{ij}^{r_1} > 0,$$

por lo que $d$ divide a $n_1 + r_1$. De forma similar, si $n \geq 1$ es tal que $P_{ij}^n > 0$, $d$ dividirá a $n_1 + n$ y, por lo tanto $n - r_1 = m_1 d$ para alguna $m_1 \in \mathbb{N}$. Además existen $m_2 \in \mathbb{N}$ y

$r \in \{0, 1, \ldots, d-1\}$ únicas tales que $r_1 = m_2 d + r$, por lo que si $P_{ij}^n > 0$ entonces $n = md + r$ para alguna $m \in \mathbb{N}$. Es ahora claro que

$$P_{ij}^{md+r} = \sum_{k=0}^{m} \mathbb{P}(T_j = kd + r \mid X_0 = i) P_{jj}^{(m-k)d} = \mathbb{E}[\xi_m],$$

donde $\xi_m$ toma el valor $P_{jj}^{(m-k)d}$, con $k \in \{0, \ldots, m\}$, con probabilidad

$$\mathbb{P}(T_j = kd + r \mid X_0 = i)$$

y $\xi_m = 0$ con probabilidad $\mathbb{P}(T_j > md + r \mid X_0 = i)$. Fijemos $k \in \mathbb{N}$ y notemos que

$$\lim_{m\to\infty} P_{jj}^{(m-k)d} = d\pi(j),$$

por lo que

$$\begin{aligned}
\mathbb{P}\left(\lim_{m\to\infty} \xi_m = d\pi(j)\right) &= 1 - \mathbb{P}\left(\lim_{m\to\infty} \xi_m = 0\right) \\
&= 1 - \lim_{m\to\infty} \mathbb{P}(T_j > md + r \mid X_0 = i) \\
&= 1 - (1 - \rho_{ij}) \\
&= 1,
\end{aligned}$$

pues $\boldsymbol{X}$ es irreducible y recurrente positiva. Además $|\xi_m| \le 1$ para toda $m \in \mathbb{N}$. Así pues, por el teorema A.3,

$$\lim_{m\to\infty} P_{ij}^{md+r} = \lim_{m\to\infty} \mathbb{E}[\xi_m] = d\pi(j),$$

finalizando la demostración del teorema. ⋇

Una versión aún más general del teorema 1.28 es el teorema 1.8.5 de [Nor97], pues se relajan los supuestos de aperiodicidad y recurrencia positiva. Para los propósitos de estas notas, tanta generalidad no es necesaria. Concluimos las notas con ejemplos en los que se aplica el teorema anterior.

**Ejemplo 1.18.** Consideremos la cadena de rachas con $p \in (0, 1)$. Hemos visto que esta cadena es recurrente positiva y aperiódica, con distribución estacionaria $\pi$ dada por $\pi(i) = p^i(1-p)$ para toda $i \in \mathbb{N}$. Por el teorema 1.28,

$$\lim_{n\to\infty} P_{ij}^n = p^j(1-p).$$

**Ejemplo 1.19.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Ehrenfest con espacio de estados $\mathcal{E} = \{0, 1, 2, 3\}$. La matriz de transición es

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

por lo que tiene período 2 y distribución estacionaria $\pi = (1/8, 3/8, 3/8, 1/8)$. Por lo tanto,

$$\lim_{n\to\infty} P^{2n} = \begin{pmatrix} 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \end{pmatrix} \quad \text{y} \quad \lim_{n\to\infty} P^{2n+1} = \begin{pmatrix} 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \end{pmatrix}.$$

0.207879576350761938641894445615708196827575653285222461198725984262…

**Ejemplo 1.20.** Un dado justo se lanza repetidamente, los lanzamientos siendo independientes entre sí. Sea $X_n$ la suma de los primeros $n$ lanzamientos. Calcularemos

$$\lim_{n\to\infty} \mathbb{P}(X_n \text{ es un múltiplo de } 13).$$

Denotemos por $R_n$ el residuo que queda al dividir $X_n$ entre 13. Entonces $\boldsymbol{R} = \{R_n\}_{n\in\mathbb{N}}$ es una cadena de Markov con espacio de estados $\mathcal{E} = \{0, 1, \ldots, 12\}$ y probabilidades de transición

$$P_{ij} = \begin{cases} \frac{1}{6} & \text{si } 1 \le j - i \le 6 \text{ o } i - j \ge 7, \\ 0 & \text{en otro caso.} \end{cases}$$

Como $P_{i,i+1} > 0$ para $0 \le i \le 11$ y $P_{12,0} > 0$ se sigue que la cadena es irreducible. De las probabilidades de transición tenemos que

$$\sum_{i=0}^{12} P_{ij} = 1$$

para toda $j \in \mathcal{E}$, por lo que $\boldsymbol{R}$ es doblemente estocástica y, por lo tanto, tiene una única distribución estacionaria $\pi$ con $\pi(i) = 1/13$ para toda $i \in \mathcal{E}$. Por otra parte, $P_{00}^3 \ge P_{06}P_{6,12}P_{12,0} > 0$ y $P_{00}^4 \ge P_{06}P_{67}P_{7,12}P_{12,1}$, y así vemos que $\boldsymbol{X}$ es aperiódica. Finalmente notemos que $X_n$ es múltiplo de 13 si y sólo si $R_n = 0$, por lo que

$$\lim_{n\to\infty} \mathbb{P}(X_n \text{ es un múltiplo de } 13) = \lim_{n\to\infty} \mathbb{P}(R_n = 0) = \frac{1}{13}.$$

**Ejemplo 1.21.** Consiremos dos vectores $r^0 = (r_0^0, r_1^0, \ldots, r_K^0)$ y $p = (p_0, p_1, \ldots, p_k)$, con $K \ge 1$, con entradas no negativas tales que $\sum_{i=0}^K r_i^0 = \sum_{i=0}^K p_i = 1$ y $p_0, p_1 > 0$. Para $i, j \in \{0, 1, \ldots, K\}$, denotemos por $\kappa^i(j)$ como el único entero en $\{0, 1, \ldots, K\}$ tal que $i + \kappa^i(j) - j$ es divisible entre $K + 1$. Definamos para cada $n \ge 1$ e $i \in \{0, 1, \ldots, K\}$,

$$r_i^n = \sum_{j=0}^K r_j^{n-1} p_{\kappa^i(j)}.$$

Probaremos que $r_i^n \to 1/(K+1)$ cuando $n \to \infty$ para cualquier $i \in \{0, 1, \ldots, K\}$.

Para esto consideremos una cadena de Markov $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ con espacio de estados $\mathcal{E} = \{0, 1, \ldots, K\}$, con probabilidades de transición dadas por $P_{ij} = p_{\kappa^i(j)}$, lo que da la siguiente matriz de transición:

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & \cdots & p_{K-1} & p_K \\ p_K & p_0 & p_1 & \cdots & p_{K-2} & p_{K-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p_2 & p_3 & p_4 & \cdots & p_0 & p_1 \\ p_1 & p_2 & p_3 & \cdots & p_K & p_0 \end{pmatrix}.$$

Es claro que $\boldsymbol{X}$ es doblemente estocástica, mientras que las condiciones $p_0 > 0$ y $p_1 > 0$ son suficientes para que la cadena sea irreducible y aperiódica. Como $\mathcal{E}$ es finito, $\boldsymbol{X}$ es recurrente positiva. Por ejemplos anteriores, sabemos que la única distribución estacionaria está dada por $\pi$, con $\pi(i) = 1/(K+1)$ para cada $i \in \mathcal{E}$. Por el teorema 1.28,

$$\lim_{n\to\infty} \mathbb{P}(X_n = i) = \frac{1}{K+1}$$

para cualquier distribución inicial $\pi_0$. En particular, si consideramos $\pi_0 = r^0$ tendremos que $\pi_n = (r_0^n, r_1^n, \ldots, r_K^n)$ para cada $n \in \mathbb{N}$. Por lo tanto podemos concluir que para cualquier $i \in \{0, 1, \ldots, K\}$ se satisfará

$$\lim_{n\to\infty} r_i^n = \frac{1}{K+1}.$$

## § A. Apéndice de resultados

En esta sección enunciaremos algunos resultados que se usarán en las notas. Cabe aclarar que varios de estos resultados no se enuncian en su generalidad puesto que se necesitan herramientas de teoría de la medida para enunciarlos y entenderlos. Asimismo, por su naturaleza, varios de los resultados aquí expuestos no se demuestran en las notas, aunque se da la referencia pertinente para el lector entusiasta.

Los primeros temas que se enunciarán tienen que ver con convergencia de esperanzas, sus demostraciones se pueden consultar en el capítulo 9 de [JP02].

**Teorema A.1** (Teorema de convergencia monótona). *Sea $\{Y_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias no negativas tales que $Y_n \leq Y_{n+1}$ para toda $n \in \mathbb{N}$. Si además $Y$ es una variable aleatoria tal que $Y_n \to Y$ con probabilidad uno entonces*

$$\lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}[Y].$$

**Teorema A.2** (Teorema de convergencia dominada). *Sea $\{Y_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tales que $Y_n \to Y$ con probabilidad uno. Si existe una variable aleatoria no negativa $\xi$ con $\mathbb{E}[\xi] < \infty$ tal que $|Y_n| \leq \xi$ para toda $n \in \mathbb{N}$, entonces $\mathbb{E}[|Y_n|] < \infty$ para toda $n \in \mathbb{N}$, $\mathbb{E}[|Y|] < \infty$ y*

$$\lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}[Y].$$

Como caso particular del teorema A.2 es cuando $\xi$ es una constante $M > 0$. Dentro de probabilidad este caso particular resulta ser tan útil que merece su propio enunciado.

**Corolario A.3** (Teorema de convergencia acotada). *Sea $\{Y_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tales que $Y_n \to Y$ con probabilidad uno. Si existe $M > 0$ tal que $|Y_n| \leq M$ para toda $n \in \mathbb{N}$, entonces $\mathbb{E}[|Y_n|] < \infty$ para toda $n \in \mathbb{N}$, $\mathbb{E}[|Y|] < \infty$ y*

$$\lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}[Y].$$

El siguiente teorema, la ley fuerte de grandes números, es uno de los resultados principales de la teoría de probabilidad. La primera parte del enunciado requiere de técnicas más avanzadas de las que se pueden cubrir en estas notas y se puede consultar la demostración en la sección 12.10 de [Wil91] o en el artículo [Ete81]. La segunda parte se probará suponiendo que la primera parte es cierta.

**Teorema A.4** (Ley fuerte de grandes números). *Sea $\{Y_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas tales que $\mathbb{E}[|Y_1|] < \infty$. Entonces*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_i = \mu$$

*con probabilidad uno, donde $\mu = \mathbb{E}[Y_1]$. Si $\{Y_n\}_{n\in\mathbb{N}}$ son no negativas y $\mu = \infty$, el resultado sigue siendo válido.*

*Demostración.* Consideremos una sucesión de variables aleatorias no negativas $\{Y_n\}_{n\in\mathbb{N}}$ independientes e idénticamente distribuidas tales que $\mathbb{E}[Y_1] = \infty$. Para $N < \infty$ definamos $Y_n^N = Y_n \wedge n := \text{mín}\{Y_n, N\}$. Entonces por la primera parte del teorema

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \geq \frac{1}{n}\sum_{i=1}^{n} Y_i^N \to \mathbb{E}[Y_1^N]$$

cuando $n \to \infty$ con probabilidad uno. Notemos que si $N_1 < N_2$ entonces $Y_1^{N_1} \leq Y_1^{N_2}$ y que $Y_1^N \to Y_1$ cuando $N \to \infty$, por lo que del teorema A.1,

$$\lim_{N\to\infty} \mathbb{E}[Y_1^N] = \mathbb{E}[Y_1] = \infty.$$

Luego entonces,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_i = \infty$$

con probabilidad uno. ⋇

Situándonos en un contexto de procesos estocásticos, consideremos una cadena de Markov $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ con espacio de estados $\mathcal{E}$ y consideremos un conjunto $\mathcal{C} \subset \mathcal{E}$ finito. Definamos $T_{\mathcal{C}} = \text{ínf}\{n \geq 0 : X_n \in \mathcal{C}\}$, es decir el tiempo en el que la cadena está por primera vez en $\mathcal{C}$. Notemos que $T_{\mathcal{C}}$ toma valores en $\mathbb{N} \cup \{\infty\}$ y que para $n \in \mathbb{N}$,

$$\{T_{\mathcal{C}} = n\} = \{X_0 \notin \mathcal{C}, X_1 \notin \mathcal{C}, \ldots, X_{n-1} \notin \mathcal{C}, X_n \in \mathcal{C}\},$$

de donde el evento $\{T_{\mathcal{C}} = n\}$ depende únicamente de $X_0, X_1, \ldots, X_n$. Este es un ejemplo de un *tiempo de de paro*, concepto que definimos ahora.

**Definición A.5.** Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ definida en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Diremos que $T : \Omega \to \mathbb{N} \cup \{\infty\}$ es un *tiempo de paro* si el evento $\{T = n\}$ depende únicamente de $X_0, X_1, \ldots, X_n$ para toda $n \in \mathbb{N}$.

Recordemos que la propiedad de Markov nos dice que para toda $n \in \mathbb{N}$,

$$\mathbb{P}(X_{n+1} = i_{n+1} \,|\, X_n = i_n, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+1} = i_{n+1} \,|\, X_n = i_n).$$

Es natural preguntarnos si en lugar índices deterministas consideramos índices aleatorios, ¿la propiedad de Markov se seguirá cumpliendo? La respuesta es sí cuando el índice aleatorio es un tiempo de paro. A este resultado se le conoce como *propiedad fuerte de Markov*, nombre que se le da por ser una extensión de la propiedad de Markov.

**Teorema A.6** (Propiedad fuerte de Markov). *Sea $\boldsymbol{X} = \{X_n\}_{n\in\mathbb{N}}$ una cadena de Markov con espacio de estados $\mathcal{E}$ y matriz de transición $P$. Si $T$ es un tiempo de paro, entonces condicional el $T < \infty$ y $X_T = i$, al definir $Y_n = X_{T+n}$ con $n \in \mathbb{N}$, $\boldsymbol{Y} = \{Y_n\}_{n\in\mathbb{N}}$ es una cadena de Markov, independiente de $X_0, \ldots, X_T$, con espacio de estados $\mathcal{E}$ y matriz de transición $P$.*

*Demostración.* Sea $E$ un evento determinado por $X_0, X_1, \ldots, X_T$. Como $E \cap \{T = n\} \subset \{T = n\}$, entonces $E \cap \{T = n\}$ dependerá únicamente de $X_0, X_1, \ldots, X_n$ por ser $T$ tiempo de paro. Entonces,

$$\begin{aligned}
&\mathbb{P}(\{Y_0 = j_0, Y_1 = j_1, \ldots, Y_m = j_m\} \cap E \cap \{T = n\} \cap \{X_T = i\}) \\
&\quad = \mathbb{P}(X_n = j_0, X_{n+1} = j_1, \ldots, X_{n+m} = j_m \,|\, X_n = i)\mathbb{P}(E \cap \{T = n\} \cap \{X_T = i\}) \\
&\quad = \mathbb{P}(X_0 = j_0, X_1 = j_1, \ldots, X_m = j_m \,|\, X_0 = i)\mathbb{P}(E \cap \{T = n\} \cap \{X_T = i\}),
\end{aligned}$$

donde se ha usado la propiedad de Markov en la primera igualdad y la homogeneidad en la segunda. Sumando sobre $n \in \mathbb{N}$ y dividiendo entre $\mathbb{P}(T < \infty, X_T = i)$ obtenemos

$$\begin{aligned}&\mathbb{P}(\{Y_0 = j_0, Y_1 = j_1, \dots, Y_m = j_m\} \cap E \,|\, T < \infty, X_T = i) \\ &\quad = \mathbb{P}(X_0 = j_0, X_1 = j_1, \dots, X_m = j_m \,|\, X_0 = i)\mathbb{P}(E \,|\, T < \infty, X_T = i).\end{aligned}$$ ⁂

Intuitivamente podemos pensar en esta propiedad como en un reinicio de la cadena después de que suceda un evento aleatorio.

## Referencias


[Ete81] N. Etemadi. «An elementary proof of the strong law of large numbers». En: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 55.1 (1981), págs. 119-122.

[Hoe72] P. G. Hoel. *Introduction to Stochastic Processes*. Houghton Mifflin Harcourt, 1972.

[JP02] J. Jacod y P. Protter. *Probability Essentials*. 2.ª ed. Springer, 2002.

[Nor97] J. R. Norris. *Markov Chains*. 1.ª ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[Wil91] D. Williams. *Probability with martingales*. 17.ª ed. Cambridge Mathematical Textbooks. Cambridge University Press, 1991.