Notas de Procesos Estocásticos I
Cadenas de Markov

Prof. Rafael Miranda Cordero

Semestre 2024-1

En general en un proceso estocástico $\mathbf{X} : \Omega \times I \to S$ las variables aleatorias $\{X_i : i \in I\}$ no son independientes,*a priori* podemos pensar que son todas dependientes unas de las otras. Los procesos de Markov lidean con esta posible complejidad asumiendo que la interdependencia de las variables del proceso tiene una forma especifica, a dicha forma se le conoce como *propiedad de Markov*. Particularmente, un proceso a tiempo discreto $(X_n)_{n\in\mathbf{N}}$ y con espacio de estados discreto se denomina *cadena de Markov*, en este caso la propiedad de Markov se expresa como

$$\mathbf{P}\left(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_n = x_n\right) = \mathbf{P}\left(X_{n+1} = x_{n+1} | X_n = x_n\right). \tag{1}$$

La propiedad de Markov se suele interpretar como que los estados futuros $(n+1)$ unicamente dependen del estado presente (n) y no del pasado $(m < n)$.

Definición 1. Se dice que un proceso estocástico $\mathbf{X} : \Omega \times T \to S$ es una cadena de Markov si T y S son espacios discretos (los espacios parametral y de estados son discretos) y satisface la propiedad de Markov (ecuación (1)). En este contexto el proceso se suele escribir como $(X_n)_{n\in\mathbf{N}}$ y el espacio de estados como $S \subseteq \mathbf{N}$ segmento inicial.

Aunque las cadenas de Markov pueden paracer simples,estas guardan la suficiente complejidad como para modelar una gran cantidad de fenómenos aleatorios y ser empleadas en distintos contextos de estadística, finanzas y *machine learning*.

La importancia de la propiedad de Markov es que a pesar de guardar cierta complejidad nos permite calcular con relativa facilidad las probabilidades de transición entre estados, lo cuál a su vez es empleado para obtener todas las distribuciones conjuntas posibles y por lo tanto caracterizar el proceso, es decir, para las cadenas de Markov las probabilidades de transición contienen, de forma muy simplificada, toda la información del proceso.

La probabilidad de transición en un paso del estado i al estado j (x_i a x_j) en el tiempo n se define como

$$p_{ij}(n, n+1) = \mathbf{P}\left(X_{n+1} = j | X_n = i\right). \tag{2}$$

En general las probabilidades de transición en un paso depende del tiempo en el que nos encontramos, *id est*, dependen de n, no obstante en muchas aplicaciones es posible suponer que dicha dependencia no existe.

Definición 2. Decimos que una cadena de Markov $(X_n)_{n\in\mathbf{N}}$ es estacionaria si $p_{ij}(n.n+1) = p(i,j)$ para toda $n \in \mathbf{N}$, con $p(i,j) \in [0,1]$, esto es, las probabilidades de transición no dependen del tiempo.

Ejemplos.

1. **Cadena de Markov de dos estados.** Considere una máquina que al inicio de cualquier día puede estar fuera de servicio (descompuesta) o en condiciones operativas (funcionando). Si la máquina esta descompuesta al inicio del día n, la probabilidad de que sea reparada de forma exitosa y este operando al inicio del siguiente día $(n+1)$ es $p \in [0,1]$. Por otro lado, si la máquina esta funcionando al inicio del n-ésimo día la probabilidad de que este descompuesta al siguiente día es de $q \in [0,1]$. A esta situación le podemos asociar un proceso estocásticos $(X_n)_{n\in\mathbf{N}}$ que nos indica si la máquina esta descompuesta (0) o funcional (1) al inicio de cada día.

Notemos que

$$\begin{aligned}
\mathbf{P}\left(X_{n+1} = 1 | X_n = 0\right) =& p \\
\mathbf{P}\left(X_{n+1} = 0 | X_n = 1\right) =& q \\
\mathbf{P}\left(X_{n+1} = 0 | X_n = 0\right) =& 1 - p \\
\mathbf{P}\left(X_{n+1} = 1 | X_n = 1\right) =& 1 - q
\end{aligned}$$

por lo que la cadena $(X_n)_{n\in\mathbf{N}}$ es estacionaria.

La dinámica del proceso en cada uno de los días esta determinada por la distribución de X_n, la distribución en el momento inicial X_0 es en principio arbitraria pero es necesario conocerla para poder calcular el resto de las distribuciones. Supongamos que π_0 es la función de probabilidad de X_0, esto es, $\pi_0(1) = \mathbf{P}\left(X_0 = 1\right)$ y $\pi_0(0) = \mathbf{P}\left(X_0 = 0\right)$. Por otro lado observemos que

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

De esto obtenemos que

$$\begin{aligned}
\mathbf{P}(X_1 = 0) &= (1-p-q)\,\pi_0(0) + q \\
\mathbf{P}(X_2 = 0) &= (1-p-q)\,\mathbf{P}(X_1 = 0) + q \\
&= (1-p-q)^2\pi_0(0) + q\,(1 + (1-p-q)) \\
\mathbf{P}(X_3 = 0) &= (1-p-q)\mathbf{P}(X_2 = 0) + q \\
&= (1-p-q)^3\pi_0(0) + q\left[1 + (1-p-q) + (1-p-q)^2\right]
\end{aligned}$$

en general

$$\mathbf{P}(X_n = 0) = (1-p-q)^n\,\pi_0(0) + q\sum_{j=0}^{n-1}(1-p-q)^j,$$

además como

$$\sum_{j=0}^{n-1}(1-p-q)^j = \frac{1-(1-p-q)^n}{p+q}$$

tenemos

$$\mathbf{P}(X_n = 0) = \frac{q}{p+q} + (1-p-q)^n\left[\pi_0(0) - \frac{q}{p+q}\right] \tag{3}$$

$$\mathbf{P}(X_n = 1) = \frac{p}{p+q} + (1-p-q)^n\left[\pi_0(0) - \frac{p}{p+q}\right]. \tag{4}$$

En el caso sencillo en el que $p = q = 0$ tenemos que $\mathbf{P}(X_n = 0) = \pi_0(0)$ y $\mathbf{P}(X_n = 1) = \pi_0(1)$, más aun, la dinámica de la cadena solo tiene dos posibilidades: si empiezas en 0 te quedas en 0 para siempre (esto sucede con probabilidad $\pi_0(0)$) y si empiezas en 1 te quedas ahí para siempre (esta posibilidad se da con probabilidad $\pi_0(1)$). Similarmente, si $p = q = 1$ la dinamica de la cadena será simplemente alternar entre 0 y 1: $0, 1, 0, 1, \ldots$ si se empieza en 0 (esto se da con probabilidad $\pi_0(0)$) y $1, 0, 1, 0, \ldots$ si se empieza en 1 (esto se da con probabilidad $\pi_0(1)$).

Si asumimos que $0 < p + q < 2$ (estos es $p, q \in (0, 1)$) entonces tenemos que

$$\lim_{n\to\infty} \mathbf{P}(X_n = 0) = \frac{q}{p+q} \qquad\qquad \lim_{n\to\infty} \mathbf{P}(X_n = 1) = \frac{p}{p+q} \tag{5}$$

esto nos dice que en esta situación la dinámica de la cadena converge a una distribución determinada, dicho de otra forma, a tiempos grandes la distribución de X_n es aproximadamente la distribución límite de la ecuación anterior.

Observación 1. *Note que a partir de las probabilidades de transción y de la distribución inicial fue posible calcular las distribuciones de cada paso en el tiempo.*

2. **Caminatas Aleatorias.** Considere una partícula que se mueve sobre la recta real pero solo sobre los enteros, podemos iniciar en cualquier posición y con probabilidad p saltamos al siguiente entero (a la derecha) y con probabilidad q vamos al entero anterior (a la izquierda), con $p, q \in [0,1]$ y $p+q=1$.

Si denotamos por X_n nuestra posición en el n-ésimo paso, entonces $(X_n)_{n\in\mathbf{N}}$ es un procesos a tiempo discreto son espacio de estados discreto $S=\mathbf{Z}$. Además tiene sentido asumir que se trata de una cadena de Markov (esto lo probaremos rigurosamente) puesto que el estado siguiente sólo depende del estado actual. Es claro que

$$\mathbf{P}\left(X_{n+1}=j|X_n=i\right)=P(i,j)=\begin{cases} p & \text{si } j=i+1 \\ q & \text{si } j=i-1 \\ 0 \text{ en otro caso.} & \end{cases}$$

Un modelo matemático para la caminata aleatoria puede ser dado de la siguiente forma: supongamos que tenemos una sucesión de variables aleatorias independientes idénticamente distribuidas $(\xi_n)_{n\in\mathbf{N}}$, de tal modo que $\mathbf{P}(\xi=1)=p$ y $\mathbf{P}(\xi=-1)=q$. Entonces X_n esta dado como

$$X_n = X_0+\xi_1+\cdots+\xi_n$$

donde X_0 representa la posición inicial de la partícula.

1. Función de Transición y Distribución Inicial

Definición 3. Sea $(X_n)_{n\in\mathbf{N}}$ una cadena de Markov estacionaria. Definicmos la función de transición en un paso $P: S\to[0,1]$ como

$$P(x,y)=\mathbf{P}\left(X_1=y|X_0=x\right) \text{ con } x,y\in S. \tag{6}$$

Definición 4. Sea $(X_n)_{n\in\mathbf{N}}$ una cadena de Markov. La función de probabilidad π_0 de X_0 la llamamos la distribución inicial de la cadena.

Es claro que si la cadena $(X_n)_{n\in\mathbf{N}}$ es estacionaria se tiene que

$$P(x,y)=\mathbf{P}\left(X_1=y|X_0=x\right)=\mathbf{P}\left(X_{n+1}=y|X_n=x\right)$$

y además se sigue de la propiedad de Markov que

$$\mathbf{P}\left(X_{n+1}=y|X_0=x_0,\ldots,X_{n-1}=x_{n-1},X_n=x\right)=\mathbf{P}\left(X_{n+1}=y|X_n=x\right)=P(x,y). \tag{7}$$

Observación 2. *Observe que para la distribución inicial de una cadena de Markov la función π_0 satisface*

1. $\pi_0(x)\geq 0$ *para todo* $x\in S$.

2. $\sum_x \pi_0(x) = 1$.

Similarmente, la función de transición de una cadena de Markov estacionaria se tiene que

1. $P(x,y) \geq 0$ *para todo* $x, y \in S$.

2. $\sum_y P(x,y) = 1$.

A partir de la función de transición y la distribución inicial podemos calcular las distribuciones conjuntas de la cadena de Markov, por ejemplo

$$\mathbf{P}\left(X_0 = x_0, X_1 = x_1\right) = \mathbf{P}\left(X_0 = x_0\right)\mathbf{P}\left(X_1 = x_1 | X_0 = x_0\right) = \pi_0(x_0)P(x_0, x_1)$$

del mismo modo

$$\begin{aligned}\mathbf{P}\left(X_0 = x_0, X_1 = x_1, X_2 = x_2\right) =& \mathbf{P}\left(X_0 = x_0, X_1 = x_1\right)\mathbf{P}\left(X_2 = x_2 | X_0 = x_0, X_1 = x_1\right)\\ =& \pi_0(x_0)P(x_0, x_1)\mathbf{P}\left(X_2 = x_2 | X_1 = x_1\right)\\ =& \pi_0(x_0)P(x_0, x_1)P(x_1, x_2)\end{aligned}$$

en general tenemos

$$\mathbf{P}\left(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\right) = \pi_0(x_0)P(x_0, x_1)\ldots P(x_{n-1}, x_n). \tag{8}$$

La fórmula anterior nos dice que para calcular la probabilidad de tomar una trayectoria particular hasta el tiempo n basta con calcular el producto de las probabilidades de donde empezamos y hacia donde nos movemos en cada paso, esto resulta muy intuitivo, no obstante, note que la identidad (8) es posible debido a la propiedad de Markov y que las probabilidades de transición en un paso son estacionarias.

Ejercicio 1. *Para la cadena de Markov de dos estados vista calcule las probabilidades de transición en dos pasos y las densidades conjuntas de las primeras dos y tres variables, es decir, debe calcular:*

- $\mathbf{P}(X_{n+2} = y | X_n = x)$ *para toda* $n \in \mathbf{N}$ *y cualesquiera* $x, y \in S = \{0, 1\}$.
- $\mathbf{P}(X_0 = x_0, X_1 = x_1)$ *con* $x_0, x_1 \in S$.
- $\mathbf{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2)$ *con* $x_0, x_1, x_2 \in S$.

además calcule $\mathbf{P}(X_1 \neq X_2)$.

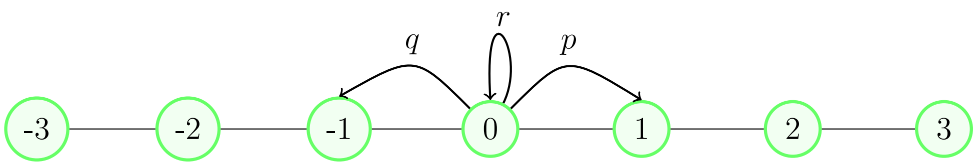
3. **Caminatas Aleatorias (Continuación).** En el caso más general, una caminata aleatoria se puede definir como

$$X_n = X_0 + \xi_1 + \cdots + \xi_n \tag{9}$$

donde $(\xi_n)_{n \in \mathbf{N}}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas con función de probabilidad común f y a su vez todas ellas son independientes de X_0. En el caso de una caminata aleatoria simple se asume que

$$f(x)=\begin{cases} p & \text{si } x=1\\ q & \text{si } x=-1\\ r & \text{si } x=0\\ 0 & \text{en otro caso.}\end{cases} \tag{10}$$

con $p,q,r\in[0,1]$ y $p+q+r=1$. Esto es, en el caso de una caminata aleatoria estamos considerando que el movimiento en cada paso del tiempo de la partícula sobre los enteros sólo puede ser de una unidad a la derecha o una unidad a la izquierda o quedarse quieto, cada una de estas acciones ocurren con probabilidad p, q y r respectivamente.



Observe que la independencia e idéntica distribución de las variables $(\xi_n)_{n\in\mathbf{N}}$ implican la propiedad de Markov, esto pues

$$\begin{aligned}
&\mathbf{P}\left(X_{n+1}=x_{n+1}|X_0=x_0,X_1=x_1,\dots,X_n=x_n\right)\\
&=\mathbf{P}\left(\xi_{n+1}=x_{n+1}-x_n|X_0=x_0,\xi_1=x_1-x_0,\xi_2=x_2-x_1,\dots,\xi_n=x_n-x_{n-1}\right)\\
&=\mathbf{P}\left(\xi_{n+1}=x_{n+1}-x_n\right)=\mathbf{P}\left(\xi_{n+1}=x_{n+1}-x_n|X_n=x_n\right)\\
&=\mathbf{P}\left(X_{n+1}=x_{n+1}|X_n=x_n\right).
\end{aligned}$$

Aunado a lo anterior, es posible demostrar que la función de transición asociada a la cadena esta dada por

$$\begin{aligned}
P(x,y)=&\mathbf{P}\left(X_{n+1}=y|X_n=x\right)\\
=&\mathbf{P}\left(\xi_{n+1}=y-x|X_n=x\right)\\
=&\mathbf{P}\left(\xi_{n+1}=y-x\right)\\
=&f(y-x).
\end{aligned}$$

La forma del proceso $(X_n)_{n\in\mathbf{N}}$ nos permite calcular de manera sencilla la esperanza y la varianza de cada paso en el tiempo cuando asumimos que la caminata empieza en cero

$$\mathbf{E}\left[X_n|X_0=0\right]=\mathbf{E}\left[\sum_{i=1}^n\xi_i\right]=n\mathbf{E}\left[\xi\right]=n(p-q) \tag{11}$$

$$\mathbf{Var}\left[X_n|X_0=0\right]=\mathbf{Var}\left[\sum_{i=1}^n\xi_i\right]=n\mathbf{Var}\left[\xi\right]=n\left[p+q-(p-q)^2\right]=n\left[1-r-(p-q)^2\right]. \tag{12}$$

Las fórmulas anteriores nos dicen que en general si $p > q$ la caminata tenderá a estar a la derecha del 0, pues $\mathbf{E}[X_n|X_0 = 0] > 0$, esto resulta intuitivamente claro, pues al tener la partícula una mayor probabilidad de moverse a la derecha se espera que en promedio la encontremos a la derecha de la posición inicial (en este caso el 0). Sucede lo análogo cuando $p < q$, en esta situación la caminata tenderá a estar a la izquierda del cero, pues $\mathbf{E}[X_n|X_0 = 0] < 0$. Además, es claro que la varianza del procesos aumenta conforme aumenta el número de pasos n, esto también esta acorde con nuestra intuición, pues conforme aumenta el número de pasos aumentan nuestras posibilidades de llegar más lejos de la posición inicial. Para el caso de una **caminata aleatoria simple**, $\mathbf{r = 0}$, tenemos que

$$\mathbf{Var}\left[X_n|X_0 = 0\right] = n\left[1 - (p-q)^2\right] = n\left[(p+q)^2 - (p-q)^2\right] = 4npq, \tag{13}$$

donde podemos notar más claramente que la varianza aumenta conforme aumenta el número de pasos.

Por otro lado, si $p = q$ (**caminata aleatoria simétrica**) entonces $\mathbf{E}\left[X_n|X_0 = 0\right] = 0$ por lo que en promedio nos quedamos donde comenzamos y además $\mathbf{Var}\left[X_n|X_0 = 0\right] = n(1 - r)$, esto quiere decir que aunque la varianza crezca conforme aumenta el número de pasos, la cantidad que aumenta en cada paso queda controlada por la probabilidad de permanecer en el mismo estado (esto es r), si r es muy pequeño la varianza aumentara rápidamente conforme pasa el tiempo, mientras que si r es cercano a 1 la varianza aumentara mas lentamente a través del tiempo.

Es claro que la varianza máxima de una **caminata aleatoria simétrica** se da cuando $r = 0$, es decir, cuando tenemos una ***caminata aleatoria simple y simétrica.***[1] De la ecuación (13), la varianza de una caminata aleatoria simple, también podemos demostrar que la máxima varianza se alcanza cuando $p = q$, de cualquiera de las dos formas esto necesariamente implica que $p = q = \frac{1}{2}$.

Ejercicio 2. *Demuestre a partir de la ecuación* (13) *que la varianza máxima de una caminata aleatoria simple se da cuando* $p = q$.

Al igual que en el caso de la cadena de dos estados, también podemos calcular las distribuciones de X_n a partir de la función de transición de la cadena y de la distribución inicial de la misma. **Abordaremos el caso de una caminata aleatoria simple** $r = 0$, primeramente note que para cada $n \in \mathbf{N}$ el máximo valor que puede tomar X_n es n y el mínimo es $-n$ por lo que $\mathbf{P}\left(X_n = m|X_0 = 0\right) = 0$ si $|m| > n$. Además, observe que si $X_0 = 0$ entonces después de un número par de pasos necesariamente nos encontraremos en una posición par y después de un número impar de pasos necesariamente estaremos en una posición impar, esto quiere decir que $\mathbf{P}\left(X_n = m|X_0 = 0\right) = 0$ si $n + m$ es impar y unicamente sucede que $\mathbf{P}\left(X_n = m|X_0 = 0\right) \geq 0$ si $n + m$ es par.

Para calcular $\mathbf{P}\left(X_n = m|X_0 = 0\right)$ para $n + m$ par con $|m| \leq n$ consideramos lo siguiente

$$R_n = \text{núm. pasos dados a la derecha hasta el tiempo } n$$
$$L_n = \text{núm. pasos dados a la izquierda hasta el tiempo } n.$$

[1] Usualmente nos referimos a una caminata aleatoria simple y simétrica simplemente como una caminata aleatora simétrica.

Observe que $X_n = R_n - L_n$ y que $n = R_n + L_n$, de esta forma

$$R_n = \frac{1}{2}(n + X_n) = \sum_{i=1}^{n} \frac{1}{2}(1 + \xi_i). \tag{14}$$

Puesto que $\mathbf{P}(\xi_n = 1) = p$ entonces $\frac{1}{2}(1 + \xi_n) \sim$ Bernoulli(p), al ser las ξ_n independientes esto implica que $\left(\frac{1}{2}(1 + \xi_n)\right)_{n \in \mathbf{N}}$ son independientes por lo cual $R_n \sim$ Binomial(n,p). Por último note que:

$$\mathbf{P}(X_n = m | X_0 = 0) = \mathbf{P}\left(R_n = \frac{1}{2}(n + m)\right) = \binom{n}{\frac{1}{2}(n + m)} p^{\frac{1}{2}(n+m)} q^{\frac{1}{2}(n-m)}$$

Con base en todo lo anterior podemos concluir el siguiente resultado.

Teorema 1. *Para una caminata aleatoria simple y simétrica ($p = q = 1/2$) la distribución de la posición al tiempo n, es decir X_n, dado que el proceso al tiempo inicial se encuentra en 0, $X_0 = 0$, esta dada por*

$$\mathbf{P}(X_n = m | X_0 = 0) = \begin{cases} 0 & \textit{si } |m| > n \textit{ o } n + m \textit{ es impar} \\ \binom{n}{\frac{1}{2}(n+m)} \frac{1}{2^n} & \textit{en otro caso.} \end{cases} \tag{15}$$

Ejercicio 3. *Usando el resultado anterior demuestre que*

$$\mathbf{P}(X_n = m | X_0 = j) = \binom{n}{\frac{1}{2}(n + m - j)} \frac{1}{2^n} \tag{16}$$

siempre que $n + m - j$ sea par.

Ejercicio 4. *Haciendo un análisis similar al realizado en este ejemplo de muestre que para una caminata aleatoria simple ($r = 0$ pero no necesariamente $p = q$) se tiene que*

$$\mathbf{P}(X_n = m | X_0 = j) = \binom{n}{\frac{1}{2}(n + m - j)} p^{\frac{1}{2}(n+m-j)} q^{\frac{1}{2}(n-m+j)} \tag{17}$$

siempre que $n + m - j$ sea par.

Sugerencia: *Empiece con el caso $j = 0$.*

Ejercicio 5. *¿Como sería $\mathbf{P}(X_n = m | X_0 = j)$ para una caminata aleatoria no necesariamente simétrica ($p \geq q$ o $q \geq p$) y simple ($r \geq 0$)?*

2. Cálculos con la Función de Transición.

En los ejemplos anteriores hemos visto como a partir de la función de transición de un paso y la distribución inicial podemos obtener las distribuciones conjuntas de la cadena y las distribuciones en cada paso del tiempo. En esta sección veremos como generalizar estos cálculos para cualquier cadena de Markov con probabilidades de transición estacionarias.

Considere $\{X_n : n \in \mathbf{N}\}$ una cadena de Markov con espacio de estados S y función de transición P estacionaria. Calcularemos varias probabilidades condicionales en términos de la función de transición.

Primeramente demostraremos que

$$\begin{aligned}\mathbf{P}\left(X_{n+m} = x_{n+m}, \ldots, X_{n+1} = x_{n+1} | X_n = x_n, \ldots, X_0 = x_0\right) = \\ P(x_n, x_{n+1}) P(x_{n+1}, x_{n+2}) \cdots P(x_{n+m-1}, x_{n+m}),\end{aligned} \tag{18}$$

esto es una implicación de la propiedad de Markov y nos dice que la trayectoria que recorrerá el proceso en los siguientes m pasos, dado que ya ha avanzado n pasos, unicamente depende de donde se encuentre en el paso n y no de los $n-1$ pasos anteriores, en otras palabras, el futuro de la trayectoria solo depende del presente y no del pasado. Además de ello la formula nos dice, como antes, que la probabilidad de recorrer dicha trayectoria futura se puede calcular como el producto de las probabilidades de ir del estado actual al siguiente en un paso.

Para demostrar la propiedad de la ecuación (18) observe que por la ecuación (8) se tiene que

$$\begin{aligned}\mathbf{P}\left(X_{n+m} = x_{n+m}, \ldots, X_1 = x_1, X_0 = x_0\right) =& \pi_0(x_0) P(x_0, x_1) \cdots P(x_{n+m-1}, x_{n+m}) \\ \mathbf{P}\left(X_n = x_n, \ldots, X_1 = x_1, X_0 = x_0\right) =& \pi_0(x_0) P(x_0, x_1) \cdots P(x_{n-1}, x_n)\end{aligned}$$

por lo cual

$$\begin{aligned}&\mathbf{P}\left(X_{n+m} = x_{n+m}, \ldots, X_{n+1} = x_{n+1} | X_n = x_n, \ldots, X_0 = x_0\right) = \\ &\qquad \frac{\mathbf{P}\left(X_{n+m} = x_{n+m}, \ldots, X_1 = x_1, X_0 = x_0\right)}{\mathbf{P}\left(X_n = x_n, \ldots, X_1 = x_1, X_0 = x_0\right)} = \frac{\pi_0(x_0) P(x_0, x_1) \cdots P(x_{n+m-1}, x_{n+m})}{\pi_0(x_0) P(x_0, x_1) \cdots P(x_{n-1}, x_n)} \\ &\qquad = P(x_n, x_{n+1}) \cdots P(x_{n+m-1}, x_{n+m}).\end{aligned}$$

A partir de la ecuación (18) se deriva el siguiente resultado más general; considere $A_0, A_1, \ldots, A_{n-1} \subseteq S$ no vacíos, entonces

$$\mathbf{P}\left(X_{n+m} = y_m, \ldots, X_{n+1} = y_1 | X_n = x, X_{n-1} \in A_{n-1} \ldots, X_0 \in A_0\right) = P(x, y_1) \cdots P(y_{m-1}, y_m). \tag{19}$$

Este resultado se deriva del siguiente lema.

Lema 1. *Dado $(\Omega, \mathscr{F}, \mathbf{P})$ un espacio de probabilidad y $\{D_i : i \in \mathbf{N}\} \subseteq \mathscr{F}$ ajenos no vacios. Si $C \in \mathscr{F}$ es tal que $\mathbf{P}\left(C | D_i\right) = p$ es independiente de $i \in \mathbf{N}$ entonces $\mathbf{P}\left(C | \cup_i D_i\right) = p$.*

Ejercicio 6. *Demuestre el lema anterior.*

Ejercicio 7. *Demuestre la identidad en* (19) *usando* (18) *y el lema 1.*

Finalmente observe que de la ecuación (19) y mediante la propiedad aditiva de la probabilidad podemos demostrar que

$$\begin{aligned}\mathbf{P}\left(X_{n+m} \in B_m, \ldots, X_{n+1} \in B_1 | X_n = x, X_{n-1} \in A_{n-1}, \ldots, X_0 \in A_0\right) = \\ \sum_{y_m \in B_m} \cdots \sum_{y_1 \in B_1} P(x, y_1) P(y_1, y_2) \cdots P(y_{m-1}, y_m),\end{aligned} \tag{20}$$

para cualesquiera $B_1, \ldots, B_m, A_1, \ldots, A_n \subseteq S$ no vacíos.

Ejercicio 8. *En la ecuación* (20) *tome* $B_m = \{y\}$ *y* $B_1 = B_2 = \cdots = B_{m-1} = S$. *¿Cómo se interpreta dicha ecuación bajo estas consideraciones?.*

2.1. La ecuación de Chapman-Kolmogorov

Definición 5. Dada una cadena de Markov con probabilidades de transición estacionarias definimos las función de transición en $m \in \mathbf{N}$ pasos como

$$P^m(x, y) = \mathbf{P}\left(X_m = y | X_0 = x\right) \tag{21}$$

para $m \geq 1$ y como

$$P^0(x, y) = \begin{cases} 1 & \text{si } x = y \\ 0 & \text{en otro caso} \end{cases} \tag{22}$$

para $m = 0$.

Note que por la ecuación (20) se tiene que

$$\begin{aligned}\mathbf{P}\left(X_m = y | X_0 = x\right) =& \mathbf{P}\left(X_m = y, X_{m-1} \in S, \ldots, X_1 \in S | X_0 = x\right) \\ =& \sum_{y_{m-1} \in S} \cdots \sum_{y_1 \in S} P(x, y_1) P(y_1, y_2) \cdots P(y_{m-1}, y)\end{aligned}$$

y además

$$\begin{aligned}\mathbf{P}\left(X_{n+m} = y | X_n = x\right) =& \mathbf{P}\left(X_{n+m} = y, X_{n+m-1} \in S, \ldots, X_{n+1} \in S | X_n = x\right) \\ =& \sum_{y_{m-1} \in S} \cdots \sum_{y_1 \in S} P(x, y_1) P(y_1, y_2) \cdots P(y_{m-1}, y)\end{aligned}$$

por lo tanto

$$P^m(x,y) = \mathbf{P}\left(X_m = y | X_0 = x\right) = \mathbf{P}\left(X_{n+m} = y | X_n = x\right), \tag{23}$$

esto nos dice que para una cadena de Markov con probabilidades de transición (en un paso) estacionarias se tiene necesariamente que la probabilidades de transición en cualquier número de pasos también son estacionarias.

Ahora observe que si en (20) elegimos $B_1 = \cdots = B_{m-1} = S$ y $B_m = \{y\}$ entonces

$$\begin{aligned}\mathbf{P}\left(X_{n+m} = y | X_n = x, X_{n-1} \in A_{n-1}, \ldots, X_0 \in A_0\right) =& \sum_{y_{m-1} \in S} \cdots \sum_{y_1 \in S} P(x,y_1)P(y_1,y_2)\cdots P(y_{m-1},y) \\ =& P^m(x,y)\end{aligned}$$

de esta forma

$$\begin{aligned}\mathbf{P}\left(X_{n+m} = y | X_n = x, X_{n-1} \in A_{n-1}, \ldots, X_0 \in A_0\right) =& \mathbf{P}\left(X_{n+m} = y | X_n = x\right) \\ =& \mathbf{P}\left(X_m = y | X_0 = x\right)\end{aligned} \tag{24}$$

lo cuál resumen la propiedad de Markov y la estacionalidad de las probabilidades en una sola expresión.

Por otro lado, usando la fórmula de probabilidad total tenemos lo siguiente

$$\mathbf{P}\left(X_{n+1} = y\right) = \sum_x \mathbf{P}\left(X_{n+1} = y | X_n = x\right)\mathbf{P}\left(X_n = x\right) = \sum_x \mathbf{P}\left(X_n = x\right)P(x,y) \tag{25}$$

esta fórmula nos permite calcular la distribución de las X_n de manera recursiva conociendo las probabilidades de transición en un paso la distribución inicial π_0 de X_0. De forma análoga podemos derivar una fórmula usando las probabilidades de transición en n pasos

$$\mathbf{P}\left(X_{n+1} = y\right) = \sum_x \mathbf{P}\left(X_{n+1} = y | X_0 = x\right)\mathbf{P}\left(X_0 = x\right) = \sum_x \pi_0(x)P^{n+1}(x,y). \tag{26}$$

Le ecuación (25) nos dice que, para una cadena de Markov estacionaria, la probabilidad de que en el tiempo $n+1$ nos encontremos en un estado específico y es igual a la suma de las probabilidades de encontrarme en un estado x cualquiera en el tiempo anterior (n) y de luego llegar en un paso al estado específico y. Mientras que la ecuación (26) nos dice que lo anterior es exactamente lo mismo que la suma de las probabilidades de empezar la trayectoria de la cadena en un estado cualquiera x y después de $n+1$ pasos llegar al estado específico y.

Nuevamente por la fórmula de probabilidad total y usando el hecho de que la probabilidad condicional es una función de probabilidad por propio derecho se tiene que

$$\begin{aligned}P^{n+m}(x,y) =& \mathbf{P}\left(X_{n+m} = y | X_0 = x\right) \\ =& \sum_z \mathbf{P}\left(X_n = z | X_0 = x\right)\mathbf{P}\left(X_{n+m} = y | X_n = z, X_0 = x\right) \\ =& \sum_z \mathbf{P}\left(X_n = z | X_0 = x\right)\mathbf{P}\left(X_{n+m} = y | X_n = z\right) \\ =& \sum_z \mathbf{P}\left(X_n = z | X_0 = x\right)\mathbf{P}\left(X_m = y | X_0 = z\right),\end{aligned}$$

las simplificiones en la condicional y en el paso de tiempo se justifican por la ecuación (24). De esta manera

$$P^{n+m}(x,y) = \sum_z P^n(x,z)P^m(z,y) \tag{27}$$

lo cuál se conoce como *la ecuación de Chapman-Kolmogorov.* Esta nos dice que la probabilidad de ir del estado x al estado y en $n+m$ pasos se puede expresar como la suma de las probabilidades de ir de x a un estado z cualquiera en n pasos y luego ir de estado z al estado y en m pasos.

2.2. Matrices de Transición

Cuando el espacio de estados de la cadena de Markov es finito, es decir $S = \{0, 1, \ldots, d\}$, podemos usar notación matricial para expresar las principales propiedades obtenidas en la sección anterior.

Definición 6. Para una cadena de Markov con espacio de estados finito $S = \{0, 1, \ldots, d\}$ y función de transición P definimos la matriz de transición como

$$\mathbf{P} = \begin{pmatrix} P(0,0) & P(0,1) & \cdots & P(0,d) \\ P(1,0) & P(1,1) & \cdots & P(1,d) \\ \vdots & \vdots & \ddots & \vdots \\ P(d,0) & P(d,1) & \cdots & P(d,d) \end{pmatrix}. \tag{28}$$

Más en general, definimos la matriz de transición de m pasos como

$$\mathbf{P}_m = \begin{pmatrix} P^m(0,0) & P^m(0,1) & \cdots & P^m(0,d) \\ P^m(1,0) & P(1,1) & \cdots & P^m(1,d) \\ \vdots & \vdots & \ddots & \vdots \\ P^m(d,0) & P^m(d,1) & \cdots & P^m(d,d) \end{pmatrix}. \tag{29}$$

De la ecuación de Chapman-Kolmogorov (27) tenemos que

$$P^n(i,j) = \sum_{k=0}^{d} P^{n-1}(i,k)P(k,j)$$

lo anterior implica que

$$\mathbf{P}_n = \mathbf{P}_{n-1}\mathbf{P} \tag{30}$$

de donde

$$\mathbf{P}_n = \mathbf{P}^n, \tag{31}$$

es decir, la matriz de transición en n pasos se calcula simplemente como el producto matricial de la matriz de transición de un paso consigo misma n veces.

De forma análoga, si definimos los vectores renglón $\pi_n = (\mathbf{P}(X_n = 0), \ldots, \mathbf{P}(X_n = d))$ de la ecuación (25) tenemos

$$\pi_n(i) = \sum_{j=0}^{d} \pi_{n-1}(j) P(j,i)$$

por lo cuál

$$\pi_n = \pi_{n-1}\mathbf{P} \tag{32}$$

y por ende

$$\pi_n = \pi_0 \mathbf{P}^n = \pi_0 \mathbf{P}_n. \tag{33}$$

Ejercicio 9. *Para el ejemplo de la cadena de dos estados (ejemplo 1).*

- *Encuentre la matriz de transición en n pasos.*
- *Calcule la distribución π_n usando la distribución inicial y las matrices de transición.*
- *Calcule el $\mathbf{P}(X_n = 0)$ y $\mathbf{P}(X_n = 1)$ cuando n tiende a infinito.*

La matriz de transición de una cadena de Markov con espacio de estados finito satisface

- Todas sus entradas son no negativas.
- La suma de sus entradas por renglón es igual 1.

A una matriz cuadrada que satisface ambas propiedades se le llama *Matriz Estocástica*. Cualquier matriz de transición es estocástica y el reciproco también es cierto, es decir, para cualquier matriz estocástica podemos construir una cadena de Markov con espacio de estados finito cuya matriz de transición sea la dada.

3. Tiempos de Paro

Definición 7. Sea $A \subseteq S$ no vacío. El tiempo de llega T_A de la cadena al subconjunto A se define como

$$T_A = \min\{n > 0 : X_n \in A\} \tag{34}$$

si $\{n > 0 : X_n \in A\} \neq \emptyset$ y $T_A = \infty$ en caso contrario.

La definición anterior nos dice que T_A es el primer momento en el que la cadena de Markov llega al subconjunto de estados A. Si $A = \{a\}$ entonces denotamos $T_{\{a\}}$ simplemente como T_a.

Observe que por la fórmula de probabilidad total se tiene que

$$\mathbf{P}(X_n = y | X_0 = x) = \sum_{m=1}^{\infty} \mathbf{P}(X_n = y | T_y = m, X_0 = x)\, \mathbf{P}(T_y = m | X_0 = x)$$

puesto que $\mathbf{P}\left(X_n = y | T_y = m, X_0 = x\right) = 0$ si $m > n$, entonces

$$\mathbf{P}\left(X_n = y | X_0 = x\right) = \sum_{m=1}^{n} \mathbf{P}\left(X_n = y | T_y = m, X_0 = x\right) \mathbf{P}\left(T_y = m | X_0 = x\right).$$

Aunado a lo anterior considere que

$$\{T_y = m\} = \{X_1 \neq y, X_2 \neq y, \ldots, X_{m-1} \neq y, X_m = y\}$$

por lo cual

$$\begin{aligned}\mathbf{P}\left(X_n = y | T_y = m, X_0 = x\right) =& \mathbf{P}\left(X_n = y | X_0 = x, X_1 \neq y, X_2 \neq y, \ldots, X_{m-1} \neq, X_m = y\right) \\ =& \mathbf{P}\left(X_n = y | X_m = y\right) = \mathbf{P}\left(X_{n-m} = y | X_0 = y\right) = P^{n-m}(y, y)\end{aligned}$$

de esta forma

$$\mathbf{P}\left(X_n = y | X_0 = x\right) = \sum_{m=1}^{n} P^{n-m}(y, y) \mathbf{P}\left(T_y = m | X_0 = x\right).$$

Para simplificar se suele denotar $\mathbf{P}\left(\text{Evento} | X_0 = x\right) = \mathbf{P}_x\left(\text{Evento}\right)$. Por ende, la ecuación anterior toma la forma

$$P^n(x, y) = \sum_{m=1}^{n} \mathbf{P}_x\left(T_y = m\right) P^{n-m}(y, y). \tag{35}$$

La ecuación anterior es la versión de la ecuación de Chapman-Kolmogorov usando tiempos de paro y nos dice que la probabilidad de ir de x a y en n pasos es igual a la suma de las probabilidades de llegar por primera vez a y desde x en m pasos y luego regresas a y en $n - m$ pasos.

Una versión a similar a (35) pero que nos permite calcular la probabilidad del tiempo de paro en lugar de la probabilidad de transición es la siguiente

$$\mathbf{P}_x\left(T_y = n + 1\right) = \sum_{z \neq y} P(x, z) \mathbf{P}_z\left(T_y = n\right). \tag{36}$$

Dicha ecuación nos indica que la probabilidad de que el tiempo de llegada a y desde x sea $n + 1$ es igual a la suma de las probabilidades de que x vaya a un estado cualquiera z que no sea y en un paso y luego que el tiempo de llegada a y desde z sea n.

Ejercicio 10. *Demuestra la ecuación* (36).

Sugerencia: *Usa probabilidad total con los eventos* $\{X_1 = z\}_{z \in S}$ *como partición de* Ω.

Solución.

Note que

$$
\begin{aligned}
\mathbf{P}_x(T_y = n+1) &= \sum_{z\in S} \mathbf{P}_x(T_y = n+1|X_1 = z)\mathbf{P}_x(X_1 = z) \\
&= \sum_{z\neq y} P(x,z)\mathbf{P}_x(T_y = n+1|X_1 = z)
\end{aligned}
$$

lo anterior pues $\mathbf{P}_x(T_y = n+1|X_1 = y) = 0$. Por otro lado

$$
\begin{aligned}
\mathbf{P}_x(T_y = n+1|X_1 = z) =& \mathbf{P}\left(X_{n+1} = y, X_n \neq y, \ldots, X_2 \neq y|X_1 = z, X_0 = x\right) \\
=& \mathbf{P}\left(X_{n+1} = y, X_n \neq y, \ldots, X_2 \neq y|X_1 = z\right) \\
=& \mathbf{P}\left(X_n = y, X_{n-1} \neq y, \ldots, X_1 \neq y|X_0 = z\right) \\
=& \mathbf{P}_z(T_y = n)
\end{aligned}
$$

de donde se concluye la igualdad buscada.

Ejercicios Adicionales.

1. Considere un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbf{P})$, sea I un conjunto de índices finito o a lo más numerable y asume que los conjuntos a continuación mencionados pertenecen a $\mathscr{F}$.

   a) Muestre que si $(D_i)_{i\in I}$ es una sucesión de conjuntos ajenos dos a dos tales que $\mathbf{P}(C|D_i) = p$ independientemente de $i \in I$ entonces $\mathbf{P}(C|\bigcup_i D_i) = p$.

   b) Muestre que si $(E_i)_{i\in I}$ es una partición de Ω, entonces:

   $$\mathbf{P}(C|D) = \sum_i \mathbf{P}(E_i|D)\mathbf{P}(C|E_i \cap D).$$

   c) Demuestre que si $(C_i)_{i\in I}$ es una sucesión de conjuntos ajenos dos a dos y $\mathbf{P}(A|C_i) = \mathbf{P}(B|C_i)$ para toda i, entonces $\mathbf{P}(A|\bigcup_i C_i) = \mathbf{P}(B|\bigcup_i C_i)$.

2. Se dice que un proceso estocástico a tiempo discreto $\{X_n : n \in \mathbf{N}\}$ tiene incrementos independientes si para cualesquiera $0 \leq n_1 < n_2 < \cdots < n_k$ las variables aleatorias $X_{n_1}, X_{n_2} - X_{n_1}, X_{n_3} - X_{n_2}, \ldots, X_{n_k} - X_{n_{k-1}}$ son independientes. Demuestre que cualquier proceso estocástico a tiempo discreto con incrementos independientes satisface la propiedad de Markov. Demuestre ahora que una caminata aleatoria (definida como en el ejemplo 1 de la sección 1.3 del libro de Hoel, Port, Stone) tiene incrementos independientes y concluya que es por ende una cadena de Markov. ¿Una caminata aleatoria tiene probabilidades de transición en un paso estacionarias?.

3. Suponga que tiene dos cajas y $2d$ bolas, de las cuales d son negras y d son rojas. Inicialmente, d de las bolas son puestas en la caja 1 y el resto en la caja 2. En cada ensayo una bola es elegida de cada una de las cajas y cada una de ellas son colocadas en la caja opuesta. Sea X_0 el número de bolas negras que se encuentran inicialmente en la caja 1, y para $n \geq 1$, sea X_n el número de bolas negras en la caja 1 después de n ensayos. Encuentre la función de transición de la cadena de Markov $\{X_n : n \geq 0\}$. ¿Este proceso es una caminata aleatoria?, ¿puede usted determinar la función de densidad de X_n?.

4. Una partícula realiza una caminata aleatoria simétrica (se mueve a la derecha o a la izquierda con la misma probabilidad) sobre $\mathbf{Z}$ empezando en cero. Encuentre la probabilidad de que la partícula se encuentre nuevamente en el origen en el sexto paso.

5. Un algoritmo aleatorio de búsqueda del estado cero opera del siguiente modo: si se encuentra en el estado k en algún momento, entonces el estado del algoritmo al siguiente paso es aleatorio con distribución uniforme en el conjunto $\{0, 1, \ldots, k-1\}$.¿Es este proceso una caminata aleatoria?. Encuentre el número esperado de pasos en los que el algoritmo alcanza el estado cero cuando inicia en el estado k.

6. Demuestre que la probabilidad de una caminata aleatoria simple (que en cada paso se mueve a la derecha con probabilidad p y a la izquierda con probabilidad q, $p + q = 1$) es una función simétrica de x si, y sólo si, la caminata es simétrica. La simetria de la probabilidad se expresa de la forma siguiente
$$\mathbf{P}(X_n = x | X_0 = 0) = \mathbf{P}(X_n = -x | X_0 = 0).$$

7. Decimos que una matriz cuadrada P es doblemente estocástica si tanto ella como su matriz transpuesta son matrices estocásticas.

   a) Si P y Q dos matrices estocásticas (o doblemente estocásticas) de la misma dimensión, entonces PQ también es estocástica (o doblemente estocástica).

   b) Si la matriz de transición de una cadena de Markov es simétrica, ¿Cómo es la función de transición de la cadena? y ¿Cómo es la función de transición de n pasos de la cadena?.

   c) Demuestre que una matriz estocástica simétrica es doblemente estocástica. ¿Es cierto el recíproco?.

   d) Demuestre que toda matriz estocástica tiene siempre al número uno como valor propio.

8. Sea $\{X_n : n \in \mathbf{N}\}$ una cadena de Markov con espacio de estados $S \subseteq \{0, 1, 2 \ldots\}$ y cuya función de transición P es tal que:
$$\sum_y yP(x, y) = Ax + B \text{ con } x \in S$$
para algunas constantes A y B.

   a) Muestre que $\mathbf{E}[X_{n+1}] = A\mathbf{E}[X_n] + B$.

   b) Si $A \neq 1$, entonces
$$\mathbf{E}[X_n] = \frac{B}{1-A} + A^n \left(\mathbf{E}[X_0] - \frac{B}{1-A} \right).$$

9. Sea $\{X_n : n \in \mathbf{N}\}$ una cadena de Markov con espacio de estados $S = \{0, 1, 2\}$ y con matriz de probabilidades de transición
$$P = \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.2 & 0.5 \\ 0.7 & 0 & 0.3 \end{pmatrix}.$$

Calcule

a) $\mathbf{P}(X_2 = 0, X_1 = 0 | X_0 = 1)$.

b) $\mathbf{P}(X_3 = 1, X_2 = 1 | X_1 = 2)$.

10. Considere una cadena de Markov con espacio de estados $\{0, 1, 2\}$ y matriz de transición

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

a) Haga el diagrama de la cadena.

b) Encuentre P^2.

c) Demuestre que $P^4 = P^2$.

d) Encuentre P^n.

11. Para la cadena de la ruina del jugador, con probabilidad p de ganar y espacio de estados $\{0, 1, \ldots, d\}$ calcule:

a) La matriz de transición P.

b) P^2,P^3 y P^4.

c) Puede dar una forma general para P^n.

12. Para la cadena de Markov de dos estados vista en clase haga lo siguiente.

a) Demuestre que las variables aleatorias $\{X_n : n \in \mathbf{N}\}$ son independientes si y solo si $p+q = 1$.

b) Calcule $\mathbf{P}_1(T_0 = n)$ para cada $n \in \mathbf{N}$.

c) ¿Bajo que condiciones $\mathbf{P}_0(T_0 = \infty) = 0$ y $\mathbf{P}_1(T_0 = \infty) = 0$?

d) Calcule $\mathbf{E}[T_0]$. ¿Es posible que $\mathbf{E}[T_0] = \infty$?, si es así, ¿cuándo sucede esto?

Sugerencia: Utilice esperanza condicional para calcular la esperanza de T_0.

13. Demueste que

a) $\mathbf{P}_x(T_y \leq n+1) = P(x, y) + \sum_{z \neq y} P(x, z) \mathbf{P}_z(T_y \leq n)$.

b) $\rho_{xy} = P(x, y) + \sum_{z \neq y} P(x, z) \rho_{zy}$.

Donde $\rho_{xy} = \mathbf{P}_x(T_y < \infty)$.

14. Demueste que $\rho_{xy} > 0$ si y sólo si $P^n(x, y) > 0$ para algún $n > 0$.

4. Clasificación de Estados y Dinámica de la Cadena

Definición 8. Para una cadena de Markov $\{X_n : n \in \mathbf{N}\}$ con espacio de estados S, $y \in S$ se dice ser recurrente si el evento de visitar ese estado es un evento seguro dado que empezamos en dicho estado, y se dice que es transitorio en caso contrario. De modo más formal, $y \in S$ es recurrente si $\rho_{yy} = \mathbf{P}_y (T_y < \infty) = 1$ y es transitorio si $\rho_{yy} < 1$.

Los conceptos de recurrencia y transitoriedad son fundamentales para clasificar los estados y entender la dinámica de la cadena. Estas definiciones no son inmediatamente intuitivas y trabajaremos con varios ejemplos y resultados para comenzar a entenderlas a profundidad.

A partir de la definición lo que podemos decir *a priori* de los estados recurrentes es que cuando la cadena empieza en un estado de este tipo lo visitara una cantidad infinita de veces, esto pues el hecho de que la cadena regrese a dicho estado en un tiempo finito es un evento seguro y una vez que haya regresado la probabilidad de que vuelva a regresar sigue siendo uno, pues por la propiedad de Markov es como si la cadena volviera a empezar. Por intuición lo anterior nos diría que los estados transitorios son lo contrario, es decir, que cuando la cadena empieza en un estado transitorio lo tendrá que visitar a lo más una cantidad finita de veces; esto resultará ser así pero no se puede derivar directamente de la definición (como en el caso de los recurrentes).

Ejemplo 1. ***La Ruina del Jugador.*** *Considere una seria de apuestas, en cada una de ellas se puede ganar o perder* \$1.00 *con probabilidad p y q respectivamente ($p + q = 1$). Comienzas con una cantidad* \$$N$ *de dinero y adoptas la regla de retirarte si tu fortuna llega a* \$0.00 *o a d.*

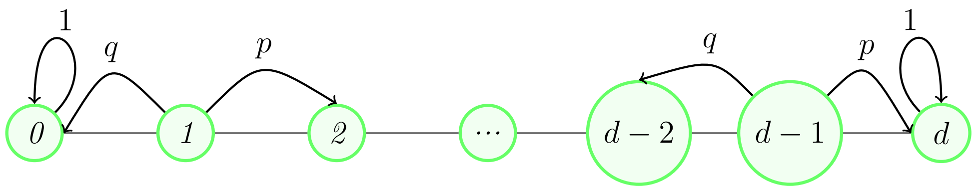
El proceso estocásticos $\{X_n : n \in \mathbf{N}\}$ que denota tu fortuna en cada tiempo n resulta ser una cadena de Markov, note que la función de transición esta dada por

$$P(x,y) = \begin{cases} p & \textit{si } y = x+1 \\ q & \textit{si } y = x-1 \\ 0 & \textit{en otro caso} \end{cases} \tag{37}$$

cuando $0 < x < d$ *y*

$$P(0,y) = \begin{cases} 1 & \textit{si } y = 0 \\ 0 & \textit{en otro caso.} \end{cases} \tag{38}$$

$$P(d,y) = \begin{cases} 1 & \textit{si } y = d \\ 0 & \textit{en otro caso.} \end{cases} \tag{39}$$



*Note que una vez que llegamos a los estados 0 o d ya no es posible salir de ellos pues $P(0,0) = 1 = P(d,d)$, a este tipo de estados se les denomina **estados absorbentes.***

Note que en general si una cadena de Markov tiene un estado absorbente $x \in S$ entonces x es recurrente, pues $1 = P(x,x) = \mathbf{P}_x(T_x = 1)$ y como el evento $\{T_x = 1\} \subseteq \{T_x < \infty\}$ entonces $1 = P(x,x) \leq \rho_{xx} \leq 1$ por lo tanto $\rho_{xx} = 1$. Los estados absorbentes son casos "extremos"de estados recurrentes, evidentemente, una vez que lleguemos a un estado recurrente tendremos que visitarlo una cantidad infinita de veces pues de hecho ya no podremos escapar de el.

Ejercicio 11. *En el ejemplo de la ruina del jugador, considere que el ganar o perder en la serie de apuestas esta dado por la seria de variables aleatorias $(\xi_i)_{i\in\mathbf{N}}$ independientes e idénticamente distribuidas ($\mathbf{P}(\xi = 1) = p$ y $\mathbf{P}(\xi = -1) = q$). Con base en ello definimos*

$$X_{n+1} = \begin{cases} X_n + \xi_{n+1} & si\ X_n \notin \{0,d\} \\ X_n & si\ X_n \in \{0,d\} \end{cases} \tag{40}$$

con $X_0 = N$ ($0 < N < d$). Bajo esta definición demuestre que la cadena de la ruina del jugador es una cadena de Markov con la función de transición descrita en el ejemplo.

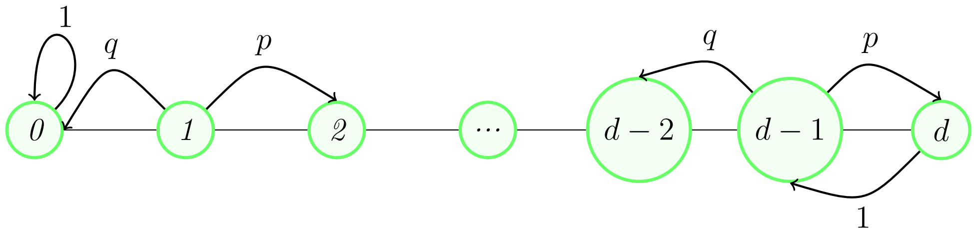
Ejercicio 12. *Calcule la matriz de transición de la ruina del jugador.*

Ejercicio 13. *Demuestre que si $x \in S$ es absorbente, entonces $P^n(x,x) = 1$.*

Ejemplo 2. ***Cadena de Nacimiento y Muerte (simple y con límite poblacional).*** *Considere que en una población cada cierta unidad de tiempo existe una probabilidad p de que nazca un nuevo individuo (el número de individuos aumente en 1) y una probabilidad $q = 1 - p$ de que muera un individuo (el número de individuos disminuya en 1). Además de ello asumiremos que si la población se extingue no hay forma de que nazcan nuevos individuos y si esta llega a una cantidad $d > 0$ necesariamente en el siguiente paso del tiempo disminuye en uno, pues este es el límite máximo de la población. En este contexto, si consideramos $\{X_n : n \in \mathbf{N}\}$ el proceso que nos dice el número de* individuos de la población en cada paso de tiempo, entonces su función de transición esta dada por

$$P(x,y) = \begin{cases} p & \textit{si } y = x+1 \\ q & \textit{si } y = x-1 \\ 0 & \textit{en otro caso.} \end{cases} \tag{41}$$

si $x \notin \{0, d\}$, $P(0,0) = 1$, $P(0,y) = 0$ con $y \neq 0$, $P(d,d) = 0$, $P(d,d-1) = 1$ y $P(d,y) = 0$ si $y < d-1$.



Como en el ejemplo de la ruina del jugador, el estado 0 es absorbente y por lo tanto recurrente, pero en este ejemplo se tiene que $P(d, d-1) = 1$, es decir, una vez que llegamos al estado d regresamos al estado anterior con probabilidad 1, a este tipo de estado se le denomina ***estado reflejante***.

Para esta cadena particular podemos demostrar que su estado reflejante d es transitorio siempre que $p, q > 0$; esto pues la probabilidad de que d llegue en un tiempo finito a 0 es positiva, y una vez que la cadena llegue a 0 no hay forma de que regrese a d, por ende la probabilidad de que empezando en d la cadena nunca regrese a d es positiva, lo que implica que la probabilidad de que empezando en d la cadena regrese a d en un tiempo finito no puede ser 1, es decir, $\rho_{dd} < 1$.

La demostración del párrafo anterior la podemos escribir en términos formales; la probabilidad de que d llegue en un tiempo finito a 0 es simplemente $\mathbf{P}_d\left(T_0 < \infty\right)$, note ahora que

$$\{X_1 = d-1, \ldots, X_{d-1} = 1, X_d = 0 | X_0 = d\} \subseteq \{T_0 = d | X_0 = d\} \subseteq \{T_0 < \infty | X_0 = d\},$$

consecuentemente

$$\mathbf{P}_d\left(X_1 = d-1, \ldots, X_{d-1} = 1, X_d = 0\right) \leq \mathbf{P}_d\left(T_0 < \infty\right),$$

y luego

$$\mathbf{P}_d\left(X_1 = d-1, \ldots, X_{d-1} = 1, X_d = 0\right) = P(d, d-1)P(d-1, d-2)\cdots P(2,1)P(1,0) = q^{d-1} > 0$$

por lo cual $\mathbf{P}_d\left(T_0 < \infty\right) > 0$.

De forma similar, la probabilidad que la cadena nunca regrese a d empezando en d es $\mathbf{P}_d\left(T_d = \infty\right)$. Como antes, observe que

$$\{T_0 = d, T_d = \infty | X_0 = d\} \subseteq \{T_d = \infty | X_0 = d\},$$

de donde $\mathbf{P}_d\left(T_0 = d, T_d = \infty\right) \leq \mathbf{P}_d\left(T_d = \infty\right)$, por otro lado

$$\mathbf{P}_d\left(T_0 = d, T_d = \infty\right) = \mathbf{P}_d\left(T_0 = d\right)\mathbf{P}_0\left(T_d = \infty\right) \geq q^{d-1}\mathbf{P}_0\left(T_d = \infty\right)$$

pero $\mathbf{P}_0(T_d = n) \le \mathbf{P}_0(X_n = d) = 0$ *(pues* $P^n(0,0) = 1$*), de donde* $\mathbf{P}_0(T_d = n) = 0$*, por lo cual* $\mathbf{P}_0(T_d < \infty) = 0$ *y así* $\mathbf{P}_0(T_d = \infty) = 1$*. Esto muestra que* $\mathbf{P}_d(T_0 = d, T_d = \infty) \ge q^{d-1} > 0$*, de este modo* $\mathbf{P}_d(T_d = \infty) > 0$ *por lo que concluimos que* $\rho_{dd} = \mathbf{P}_d(T_d < \infty) < 1$*,* id est*,* d *es transitorio.*

En el ejemplo anterior, el argumento de que el estado reflejante d es transitorio se base fuertemente en que existe una posibilidad de llegar de d a 0 pero no existe forma de llegar de 0 a d. El concepto de que existe posibilidad de llegar de un estado a otro se denomina **accesibilidad**, y en esencia la transitoriedad de d se basa en que d puede acceder al 0 pero el 0 no puede acceder a d. En este mismo sentido, tenemos por ejemplo que d accede a $d-1$ y $d-1$ también accede a d, cuando esto sucede decimos que los estados tienen **comunicación** entre si; de esta forma d y $d-1$ se comunican pero 0 y d no lo hacen.

Aunado a lo anterior, note que el hecho de que d sea reflejante no influye en la forma del argumento, por lo que bajo una argumentación análoga podríamos demostrar que cualquier estado $0 < x < d$ es transitorio, pues puede acceder a 0 pero no se comunica con 0. Esto nos dice que en la cadena de nacimiento y muerte del ejemplo 2 el cero es el único estado recurrente y el resto de estados son transitorios.

Ejercicio 14. *De una construcción formal para la cadena de nacimiento y muerte del ejemplo 2 y con ella demuestre que dicho proceso satisface la propiedad de Markov.*

Sugerencia: *Base la construcción de esta cadena en la hecha para la cadena de la ruina del jugador.*

Ejercicio 15. *Usando una argumentación análoga a la dada en el ejemplo, demuestre que cualquier estado* $0 < x \le d$ *de la cadena de nacimiento y muerte (simple y con límite poblacional) es transitorio.*

Ejercicio 16. *Usando una argumentación similar a la dada para la cadena de nacimiento y muerte, demuestre que, excepto por los estados absorbentes, todos los estados de la cadena de la ruina del jugador son transitorios.*

4.1. Accesibilidad y Comunicación

Los conceptos de accesibilidad y comunicación, aunados a los de recurrencia y transitoriedad son fundamentales para entender la dinámica de la cadena. Ya hemos definido formalmente y discutido brevemente estos dos últimos, ahora definiremos formalmente la accesibilidad y comunicación.

Definición 9. En una cadena de Markov decimos que un estado y es accesible desde x o que x accede a y siempre y cuando $\rho_{xy} > 0$. Decimos que x y y se comunican si x accede a y y y accede a x, es decir, si y solo si $\rho_{xy} > 0$ y $\rho_{yx} > 0$. Esto lo denotamos de la siguiente forma

- $x \mapsto y$ si x accede a y.
- $x \leftrightarrow y$ si x se comunica con y.

Algo que en general resulta intuitivo es que un estado x puede acceder a un estado y si es posible llegar de x a y en una cantidad finita de pasos, es decir, si existe $n \in \mathbf{N}^+$ tal que $P^n(x, y) > 0$. Este hecho queda demostrado en la siguiente proposición.

Teorema 2. *En una cadena de Markov, para $x, y \in S$ se tiene que $x \mapsto y$ si y solo si existe $n \in \mathbf{N}^+$ tal que $P^n(x, y) > 0$.*

Demostración. Observe que

$$\mathbf{P}_x\left(T_y < \infty\right) = \sum_{n=1}^{\infty} \mathbf{P}_x\left(T_y = n\right) \tag{42}$$

por lo que $0 < \rho_{xy} = \mathbf{P}_x\left(T_y < \infty\right)$ si y solo si existe $n_0 \in \mathbf{N}^+$ tal que $\mathbf{P}_x\left(T_y = n_0\right) > 0$. Luego entonces como $\{X_{n_0} = y | X_0 = x\} \supseteq \{T_y = n_0 | X_0 = x\}$ se sigue que $0 < \mathbf{P}_x\left(T_y = n_0\right) \leq P^{n_0}(x, y)$.

Reciprocamente, si existe $n_0 > 0$ tal que $P^{n_0}(x, y) > 0$ entonces por la ecuación (35) se tiene que

$$0 < P^{n_0}(x, y) = \sum_{m=1}^{n_0} \mathbf{P}_x\left(T_y = m\right) P^{n_0 - m}(y, y)$$

por lo que existe $m_0 \in \{1, \ldots, n_0\}$ tal que $\mathbf{P}_x\left(T_y = m_0\right) > 0$ y así $\rho_{xy} > 0$. □

El teorema anterior nos da una forma más general de definir la accesibilidad y comunicación.

Definición 10. En una cadena de Markov, para dos estados $x, y \in S$ se dice que x accede a y ($x \mapsto y$) si existe $n \in \mathbf{N}$ tal que $P^n(x, y,) > 0$. Como antes, x se comunica con y ($x \leftrightarrow y$) siempre y cuando x accede a y y y accede a x.

Cuando $x \neq y$ las definiciones 9 y 10 son equivalentes, pero bajo la definición 10 necesariamente $x \leftrightarrow x$ (pues $P^0(x, x) = 1 > 0$) mientras que bajo la definición 9 esto no siempre es así. Adoptaremos la definición 10 por ser más general y conveniente.

Teorema 3. *La relación de comunicación es una relación de equivalencia sobre el espacio de estados S, esto es*

1. $x \leftrightarrow x$.

2. *$x \leftrightarrow y$ si y solo si $y \leftrightarrow x$.*

3. *Si $x \leftrightarrow y$ y $y \leftrightarrow z$ entonces $x \leftrightarrow z$.*

Demostración. La reflexividad y simetría son fáciles de demostrar. Para la transitividad note que por la ecuación de Chapman-Kolmogorov se tiene que

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

De este modo, como $x \mapsto y$ existe $n \in \mathbf{N}$ tal que $P^n(x,y) > 0$ y como $y \mapsto z$ existe $m \in \mathbf{N}$ tal que $P^m(y,z) > 0$, por ende $P^{n+m}(x,z) > 0$ lo que implica que $x \mapsto z$. $\square$

A las clases de equivalencia obtenidas mediante la relación de comunicación se les denomina **clases de comunicación** y denotamos para $x \in S$ su clase de comunicación como

$$C_x = \{y \in S : x \leftrightarrow y\}. \tag{43}$$

En una clase de comunicación C_x sucede que todos los estados dentro de dicha clase se comunican entre si, mientras que si dos clases de comunicación C_x y C_y son distintas entonces necesariamente sucede alguna de las siguientes 3 cosas

- Los estados de C_x pueden acceder a estados en C_y pero ningún estado de C_y puede acceder a algún estado de C_x.
- Los estados de C_y pueden acceder a estados en C_x pero ningún estado de C_x puede acceder a algún estado de C_y.
- Ningún estado de C_x puede acceder a estados de C_y y ningún estado de C_y puede acceder a algún estado de C_x.

Note que las 3 condiciones dadas son simplemente el resultado de negar la comunicación entre estados de C_x y C_y, esto pues, en caso de que existieran un par de estados $x \in C_x$ y $y \in C_y$ tales que $x \leftrightarrow y$ necesariamente $C_x = C_y$.

Para ejemplificar lo anterior observe que para la cadena de la ruina del jugador (ejemplo 1) existen sólo 3 clases de comunicación, a saber $C_0 = \{0\}$, $C_d = \{d\}$ y $C_1 = \{1, 2, \ldots, d-1\}$. Las clases C_0 y C_d son de estados recurrentes (particularmente absorbentes por lo que son unitarios), mientras que la clase C_1 es de estados transitorios(note que $C_1 = C_x$ con $0 < x < d$); los estados de de C_1 puede acceder a C_0 o C_d pero no sucede que los estados de C_0 o C_d accedan a los de C_1 ni que C_0 acceda a C_d o C_d acceda a C_0. De forma similar, la cadena de nacimiento y muerte del ejemplo 2 unicamente tiene 2 clases de comunicación, estas son $C_0 = \{0\}$ y $C_1 = \{1, 2, \ldots, d\}$; la clase C_1 es de estados transitorios y la clase C_0 es de estados recurrentes, los estados de l clase C_1 pueden acceder a los de la clase C_0 pero la clase C_0 no puede acceder a la clase C_1.

4.2. Número de Visitas a un Estado

Recordemos como se define la función indicadora (o caractrística de un conjunto)

$$\mathbb{I}_y(z) = \begin{cases} 1 & \text{si } z = y \\ 0 & \text{e.o.c.} \end{cases}. \tag{44}$$

Ahora considere la variable aleatoria $N(y)$ definida como el número de veces, $n \geq 1$, que la cadena está o visita el estado y. Es posible definir a $N(y)$ como

$$N(y) = \sum_{n=1}^{\infty} \mathbb{I}_y (X_n), \tag{45}$$

lo anterior pues $\mathbb{I}_y (X_n) = 1$ si y sólo si $X_n = y$.

Note que

$$\mathbf{P}_x (N(y) \geq 1) = \mathbf{P}_x (T_y < \infty). \tag{46}$$

Por otro lado

$$\begin{aligned} \mathbf{P}_x (N(y) \geq 2) =& \mathbf{P}_x \left(N(y) \geq 2, \bigcup_{m=1}^{\infty} \{T_y = m\} \right) \\ =& \sum_{m=1}^{\infty} \mathbf{P}_x (N(y) \geq 2, T_y = m) \end{aligned}$$

pero

$$\begin{aligned} \mathbf{P}_x (N(y) \geq 2, T_y = m) =& \mathbf{P}_x (N(y) \geq 2 | T_y = m) \mathbf{P}_x (T_y = m) \\ =& \mathbf{P} (N(y) \geq 2 | X_m = y, X_{m-1} \neq y, \ldots, X_1 \neq y, X_0 = x) \mathbf{P}_x (T_y = m) \\ =& \mathbf{P} (N(y) \geq 2 | X_m = y) \mathbf{P}_x (T_y = m) \\ =& \mathbf{P}_y (N(y) \geq 1) \mathbf{P}_x (T_y = m) \end{aligned}$$

de esta foma

$$\begin{aligned} \mathbf{P}_x (N(y) \geq 2) =& \sum_{m=1}^{\infty} \mathbf{P}_y (N(y) \geq 1) \mathbf{P}_x (T_y = m) \\ =& \sum_{m=1}^{\infty} \mathbf{P}_y (T_y < \infty) \mathbf{P}_x (T_y = m) \\ =& \mathbf{P}_y (T_y < \infty) \sum_{m=1}^{\infty} \mathbf{P}_x (T_y = m) \end{aligned}$$

en consecuencia

$$\mathbf{P}_x (N(y) \geq 2) = \mathbf{P}_x (T_y < \infty) \mathbf{P}_y (T_y < \infty)$$

y así

$$\mathbf{P}_x\left(N(y)\geq 2\right)=\rho_{xy}\rho_{yy}. \tag{47}$$

En general para $k\geq 1$ se tiene que

$$\begin{aligned}\mathbf{P}_x\left(N(y)\geq k\right)=&\sum_{m=1}^{\infty}\mathbf{P}_y\left(N(y)\geq k-1\right)\mathbf{P}_x\left(T_y=m\right)\\ =&\mathbf{P}_y\left(N(y)\geq k-1\right)\sum_{m=1}^{\infty}\mathbf{P}_x\left(T_y=m\right)\\ =&\rho_{xy}\mathbf{P}_y\left(N(y)\geq k-1\right)\end{aligned}$$

por lo que siguiendo un argumento recursivo se tiene que

$$\mathbf{P}_x\left(N(y)\geq k\right)=\rho_{xy}\rho_{yy}^{k-1}. \tag{48}$$

De este modo, de la ecuación (48) obtenemos que

$$\mathbf{P}_x\left(N(y)=k\right)=\mathbf{P}_x\left(N(y)\geq k\right)-\mathbf{P}_x\left(N(y)\geq k+1\right)$$

y por ende

$$\mathbf{P}_x\left(N(y)=k\right)=\rho_{xy}\rho_{yy}^{k-1}(1-\rho_{yy}). \tag{49}$$

Además para $k=0$ tenemos

$$\mathbf{P}_x\left(N(y)=0\right)=1-\mathbf{P}_x\left(N(y)\geq 1\right)=1-\rho_{xy}. \tag{50}$$

Ahora observe que

$$\mathbf{E}_x\left[\mathbb{I}_y\left(X_n\right)\right]=\mathbf{P}_x\left(X_n=y\right)=P^n(x,y)$$

por lo que

$$\begin{aligned}\mathbf{E}_x\left[N(y)\right]=&\mathbf{E}_x\left[\sum_{n=1}^{\infty}\mathbb{I}_y\left(X_n\right)\right]\\ =&\sum_{n=1}^{\infty}\mathbf{E}_x\left[\mathbb{I}_y\left(X_n\right)\right]\\ =&\sum_{n=1}^{\infty}P^n(x,y),\end{aligned}$$

lo anterior gracias al teorema de convergencia monótona.

Definimos la función G como

$$G(x,y)=\mathbf{E}_x\left[N(y)\right]=\sum_{n=1}^{\infty}P^n(x,y). \tag{51}$$

El siguiente resultado nos ayuda a ligar de manera formal las nociones que hemos discutido entre recurrencia, transitividad y el número de veces que la cadena para por un estado.

Teorema 4. *Sea* $\{X_n|n \in \mathbf{N}\}$ *una cadena de Markov homogenea con espacio de estados* S*. Sean* $x, y \in S$ *dos estados cualesquiera, entonces:*

1. *Si* y *es un estado transitorio se tiene que*

$$\mathbf{P}_x\left(N(y) < \infty\right) = 1 \tag{52}$$

$$G(x,y) = \frac{\rho_{xy}}{1-\rho_{yy}}. \tag{53}$$

2. *Si* y *es un estado recurrente tenemos*

$$\mathbf{P}_y\left(N(y) = \infty\right) = 1 \tag{54}$$

$$G(y,y) = \infty. \tag{55}$$

Además

$$\mathbf{P}_x\left(N(y) = \infty\right) = \rho_{xy} \tag{56}$$

$$G(x,y) = \begin{cases} 0 & si\ \rho_{xy} = 0 \\ \infty & si\ \rho_{xy} > 0. \end{cases} \tag{57}$$

Demostración. De la ecuación (48) se tiene que

$$\mathbf{P}_x\left(N(y) = \infty\right) = \lim_{k\to\infty} \mathbf{P}_x\left(N(y) \geq k\right) = \lim_{k\to\infty} \rho_{xy}\rho_{yy}^{k-1} = \begin{cases} 0 \text{ si } \rho_{yy} < 1 \\ \rho_{xy} \text{ si } \rho_{yy} = 1 \end{cases}$$

de esta forma si y es transitororio se tiene que $\rho_{yy} < 1$ y por lo tanto

$$\mathbf{P}_x\left(N(y) = \infty\right) = 0 \text{ y eso si y sólo si } \mathbf{P}_x\left(N(y) < \infty\right) = 1.$$

Similarmente, si y es recurrente necesariamente $\rho_{yy} = 1$ por lo que

$$\mathbf{P}_x\left(N(y) = \infty\right) = \rho_{xy} \text{ y esto si y sólo si } \mathbf{P}_x\left(N(y) < \infty\right) = 1 - \rho_{xy}$$

en particular

$$\mathbf{P}_y\left(N(y) = \infty\right) = \rho_{yy} = 1.$$

Ahora, si y es transitorio entonces $\rho_{yy} < 1$ de donde

$$\begin{aligned} G(x,y) =& \mathbf{E}_x\left[N(y)\right] \\ =& \sum_{m=1}^{\infty} m\mathbf{P}_x\left(N(y) = m\right) \\ =& \sum_{m=1}^{\infty} m\rho_{xy}\rho_{yy}^{m-1}(1-\rho_{yy}) \\ =& \rho_{xy}(1-\rho_{yy}) \sum_{m=1}^{\infty} m\rho_{yy}^{m-1} \\ =& \frac{\rho_{xy}}{1-\rho_{yy}} < \infty. \end{aligned}$$

Si $\rho_{xy} = 0$ entonces
$$\mathbf{P}_x\left(T_y = m\right) \leq \mathbf{P}_x\left(T_y < \infty\right) = \rho_{xy} = 0$$
por lo que
$$\mathbf{P}_x\left(T_y = m\right) = 0 \tag{58}$$
luego
$$P^n(x,y) = \sum_{m=1}^{n} \mathbf{P}_x\left(T_y = m\right) P^{n-m}(x,y) = 0$$
por lo que de acuerdo a la ecuación (51) se sigue
$$G(x,y) = 0.$$

Si y es recurrente entonces
$$\mathbf{P}_x\left(N(y) = \infty\right) = \rho_{xy}$$
por lo que si $\rho_{xy} > 0$ necesariamente
$$G(x,y) = \mathbf{E}_x\left[N(y)\right] = \infty$$
en particular como $\rho_{yy} = 1 > 0$ se tiene que
$$G(y,y) = \infty.$$

□

Observación 3. *Si y es un estado tal que*
$$\mathbf{P}_y\left(N(y) < \infty\right) = 1$$
entonces
$$\mathbf{P}_y\left(N(y) = \infty\right) = 0$$
pero como
$$\mathbf{P}_y\left(N(y) = \infty\right) = \lim_{n\to\infty} \rho_{yy}^m,$$
por ende
$$\lim_{n\to\infty} \rho_{yy}^m = 0,$$
lo que implica que $\rho_{yy} < 1$ y por lo tanto y es transitorio. Más en general si
$$\rho_{xy} > 0 \text{ y } \mathbf{P}_x\left(N(y) < \infty\right) = 1$$
por el mismo argumento se sigue que $\rho_{yy} < 1$ y por lo tanto y es transitorio.

Lo anterior nos dice que si empezando en algún estado x, si el evento que el número de visitas a y es finito es un evento seguro, entonces necesariamente y es transitorio.

Observación 4. *Si $G(x,y) < \infty$ y $\rho_{xy} > 0$ para algún $x \in S$ entonces necesariamente*

$$G(x,y) = \sum_{n=1}^{\infty} \mathbf{P}_x\left(N(y) \geq n\right) = \sum_{n=1}^{\infty} \rho_{xy}\rho_{yy}^{n-1} < \infty$$

lo cuál sucede siempre y cuando

$$\sum_{n=1}^{\infty} \rho_{yy}^{n-1} < \infty \Leftrightarrow \rho_{yy} < 1$$

por lo tanto y es transitorio.

La observación 4 es similar a la observación 3, solo que esta nos indica que la transitoriedad de un estado también es equivalente a que el promedio de visitas al estado y, cuando la cadena empieza en algún estado x, sea finito.

Observación 5. *Si*

$$\mathbf{P}_y\left(N(y) = \infty\right) = 1$$

entonces

$$\lim_{n\to\infty} \rho_{yy}^{n} = \mathbf{P}_y\left(N(y) = \infty\right) = 1$$

por lo tanto $\rho_{yy} = 1$ y así y es recurrente. Del mismo modo que antes, en general si $\rho_{xy} > 0$ y $\mathbf{P}_x\left(N(y) = \infty\right) = 1$ entonces y es recurrente.

Observación 6. *Si $G(x,y) = \infty$ esto pasa si y sólo si*

$$G(x,y) = \sum_{n=1}^{\infty} \rho_{xy}\rho_{yy}^{n-1} = \infty \Leftrightarrow \sum_{n=1}^{\infty} \rho_{yy}^{n-1} = \infty$$

por lo que $\rho_{yy} = 1$ y por ende y es recurrente.

Corolario 4.1. *Si $\{X_n | n \in \mathbf{N}\}$ es una cadena de Markov homogenea con espacio de estados finito $S = \{0, 1, \ldots, d\}$, entonces existe al menos un estado recurrente.*

Demostración. Si todos los estados de la cadena son transitorios entonces para todo $y \in S$ se tiene que

$$G(x,y) < \infty \ \forall x \in S$$

pero

$$G(x,y) = \sum_{n=1}^{\infty} P^n(x,y) < \infty$$

esto implica

$$\lim_{n\to\infty} P^n(x,y) = 0 \ \forall y \in S$$

luego

$$0 = \sum_{y\in S} \lim_{n\to\infty} P^n(x,y) = \lim_{n\to\infty} \sum_{y\in S} P^n(x,y) = \lim_{n\to\infty} \mathbf{P}_x\left(X_n \in S\right) = \lim_{n\to\infty} 1 = 1$$

lo cuál es una contradicción, por lo tanto debe existir al menos un estado recurrente. □

Lema 2. *Si $x \leftrightarrow y$ y x es recurrente, entonces y es recurrente.*

Demostración. Como $x \leftrightarrow y$ entonces existen $n_1, n_0 \in \mathbf{N}^+$ tales que $P^{n_0}(x,y) > 0$ y $P^{n_1}(y,x) > 0$. Además, como x es recurrente se tiene que

$$G(x,x) = \sum_{n=1}^{\infty} P^n(x,x) = \infty,$$

luego

$$P^{n_0+n+n_1}(y,y) \geq P^{n_1}(y,x)P^n(x,x)P^{n_0}(x,y),$$

de este modo

$$\begin{aligned} G(y,y) = \sum_{n=1}^{\infty} P^n(y,y) \geq \sum_{n=1}^{\infty} P^{n_0+n+n_1}(y,y) \geq \sum_{n=1}^{\infty} P^{n_1}(y,x)P^n(x,x)P^{n_0}(x,y) \\ \geq & P^{n_1}(y,x)P^{n_0}(x,y)\sum_{n=1}^{\infty} P^n(x,x) = P^{n_1}(y,x)P^{n_0}(x,y)G(x,x) = \infty, \end{aligned}$$

por lo tanto y es recurrente. □

Corolario 4.2. *Si $x \leftrightarrow y$ y x es transitorio, entonces y es transitorio.*

Demostración. Si suponemos que y es recurrente entonces se tendría que x es recurrente por el lema anterior, esto es una contradicción, de este modo necesariamente y es transitorio. □

El corolario y lema anterior nos dicen que la propiedad de ser recurrente o transitorio es una propiedad de clases de comunicación, esto es, dentro de una clase de comunicación todos los estados son del mismo tipo, ya sea transitorios o recurrentes pero no pueden existir estados transitorios y recurrentes dentro de la misma clase.

Teorema 5. *Si $x \mapsto y$ y x es recurrente entonces y es recurrente. Además $\rho_{xy} = \rho_{yx} = 1$.*

Demostración. Note que

$$\begin{aligned}\mathbf{P}_x\left(T_x=\infty\right) \geq & \mathbf{P}_x\left(T_y<\infty, T_x=\infty\right)\\ \geq & \mathbf{P}_x\left(T_x=\infty \mid T_y<\infty\right) \mathbf{P}_x\left(T_y<\infty\right)\\ \geq & \mathbf{P}\left(T_x=\infty \mid T_y<\infty, X_0=x\right) \rho_{xy}\\ \geq & \rho_{xy} \mathbf{P}_y\left(T_x=\infty\right)=\rho_{xy}\left(1-\rho_{yx}\right)\end{aligned}$$

por lo que si $\rho_{yx}<1$ se tendría que $\mathbf{P}_x\left(T_x=\infty\right)>0$ por lo cual $\rho_{xx}<1$, esto es una contradicción. Consecuentemente, podemos concluir que $\rho_{yx}=1$, por lo que $y \mapsto x$, y así $x \leftrightarrow y$, esto implica por el lema anterior que y es recurrente. Cambiando los papeles de x y y en la ecuación anterior podemos concluir también que $\rho_{xy}=1$. □

El teorema anterior nos dice que la accesibilidad entre recurrentes es equivalente a la comunicación, esto implica que no es posible que un estado recurrente acceda a un estado transitorio, pues de lo contrario tendrían que ser ambos recurrentes. Más en general, el teorema anterior nos dice que las clases de comunicación de estados recurrentes no pueden acceder a las clases de comunicación de estados transitorio. En nuestros ejemplos ya hemos visto que si es posible que un estado transitorio acceda a un estado recurrente por lo que el teorema anterior no es valido para estados transitorios, es decir, entre los estados transitorios la accesibilidad no es equivalente a la comunicación.

Definición 11. Decimos que un conjunto $C \subseteq S$ es irreducible si para cualesquiera $x, y \in C$ se tiene que $x \leftrightarrow y$, es decir, todos los estados de C se comunican entre si. Del mismo modo, decimos que una cadena de Markov es irreducible si S es irreducible, esto es, todos los estados de la cadena se comunican entre si. En tal caso, diremos que la cadena es recurrente si todos sus estados son recurrentes y que la cadena es transitoria si todos sus estados son transitorios.

Definición 12. Decimos que un conjunto $C \subseteq S$ es cerrado si para cualesquier $x \in C$ y $y \notin C$ se tiene que $P^n(x,y)=0$ para cualquier $n \in \mathbf{N}^+$.

Corolario 5.1. *Sea $\{X_n : n \in \mathbf{N}\}$ una cadena de Markov y $C \subseteq S$ un conjunto cerrado y finito. Entonces C tiene al menos un estado recurrente.*

Demostración. La demostración es completamente análoga al del corolario 4.1 simplemente note que si $x \in C$ y $y \notin C$ se tiene que $G(x,y)=0$. □

Note que si $C \subseteq S$ es una clase de comunicación de recurrentes, entonces no existe $x \in C$ tal que $x \mapsto y$ para cualquier $y \in S \setminus C$, pues de lo contrario se tendtria que $y \leftrightarrow x$ y por ende $y \in C$, lo anterior implica que $P^n(x,y)=0$ para cualquier $x\in C$, $y\in S\setminus C$ y $n\in\mathbf{N}^+$, dicho de otro modo

$$1=\sum_{y\in S}P^n(x,y)=\sum_{y\in C}P^n(x,y). \tag{59}$$

Por lo tanto las clases de comunicación de recurrentes son conjuntos irreducibles y cerrados. El reciproco no es cierto en general.

Ejercicio 17. *Demuestre que un conjunto cerrado e irreducible es una clase de comunicación.*

Ejercicio 18. *Dé ejemplos de:*

1. *Un conjunto irreducible que no sea cerrado.*
2. *Un conjunto cerrado que no sea irreducible.*
3. *Un conjunto irreducible que no sea clase de comunicación.*
4. *Un conjunto cerrado que no sea clase de comunicación.*
5. *Un conjunto cerrado e irreducible que sea transitorio.*

Corolario 5.2. *Sea $S_R\neq\emptyset$ el conjunto de estados recurrentes de una cadena de Markov homogenea. Entonces S_R es la unión de una colección finita o a lo más numerable de clases cerradas irreducibles y ajenas, esto es, clases de comunicación de recurrentes.*

Observación 7. *De acuerdo al teorema anterior podemos entender el comportamiento de la dinámica de comunicación de una cadena de Markov.*

1. *Si la cadena empieza en una clase de comunicación de recurrentes entonces se queda ahí para siempre con probabilidad 1 y visita a cada estado de la clase una cantidad infinita de veces con probabilidad 1.*
2. *Si la cadena de Markov inicia en un estado transitorio entonces puede permanecer entre los estados transitorios para siempre (si estos son una cantidad infinita) o absorberce en algún momento en una clase de comunicación de recurrentes.*

4.3. Probabilidades de Absorción

Sea $C\subseteq S$ un conjunto. Como antes, denotamos

$$\rho_C(x)=\mathbf{P}_x\left(T_C<\infty\right).$$

Si C es una clase de recurrentes tenemos que si $x \in C$ entonces $\rho_C(x) = 1$ y si $x \notin C$ y es recurrente entonces $\rho_C(x) = 0$. El único caso interesante es cuando x es un estado transitorio.

Lema 3. *Sea $C \subseteq S$ una clase de comunicación de recurrentes y $x \in S_T$, entonces*

$$\rho_C(x) = \sum_{y \in C} P(x,y) + \sum_{y \in S_T} P(x,y)\rho_C(y).$$

Demostración.

$$\begin{aligned}
\rho_C(x) =& \mathbf{P}_x\left(T_C < \infty\right) \\
=& \mathbf{P}_x\left(T_C < \infty, X_1 \in C\right) + \mathbf{P}_x\left(T_C < \infty, X_1 \notin C\right) \\
=& \sum_{y \in C} \mathbf{P}_x\left(T_C < \infty, X_1 = y\right) + \sum_{y \in S_T} \mathbf{P}_x\left(T_C < \infty, X_1 = y\right) + \sum_{y \in S_R \setminus C} \mathbf{P}_x\left(T_C < \infty, X_1 = y\right) \\
=& \sum_{y \in C} P(x,y) + \sum_{y \in S_T} \mathbf{P}_x\left(T_C < \infty | X_1 = y\right) \mathbf{P}_x\left(X_1 = y\right) \\
=& \sum_{y \in C} P(x,y) + \sum_{y \in S_T} \mathbf{P}_y\left(T_C < \infty\right) P(x,y) \\
=& \sum_{y \in C} P(x,y) + \sum_{y \in S_T} P(x,y)\rho_C(y).
\end{aligned}$$

□

Teorema 6. *Suponga que el conjunto S_T de estos transitorios es finito y sea $C \subseteq S_R$ clase de comunicación. Entonces el sistema de ecuaciones*

$$f(x) = \sum_{y \in C} P(x,y) + \sum_{y \in S_T} P(x,y) f(y) \ \text{con } x \in S_T \tag{60}$$

tiene una única solución

$$f(x) = \rho_C(x) \ \text{con } x \in S_T.$$

Demostración. Sea $y \in S_T$, entonces

$$f(y) = \sum_{z \in C} P(y,z) + \sum_{z \in S_T} P(y,z) f(z)$$

sustituyendo lo anterior en la ecuación (60) se tiene

$$\begin{aligned}
f(x) =& \sum_{y \in C} P(x,y) + \sum_{y \in S_T} P(x,y) \left[\sum_{z \in C} P(y,z) + \sum_{z \in S_T} P(y,z) f(z)\right] \\
=& \sum_{y \in C} P(x,y) + \sum_{y \in S_T} \sum_{z \in C} P(x,y)P(y,z) + \sum_{y \in S_T} \sum_{z \in S_T} P(x,y)P(y,z) f(z) \\
=& \sum_{y \in C} P(x,y) + \sum_{z \in C} P^2(x,z) + \sum_{z \in S_T} P^2(x,z) f(z) \\
=& \mathbf{P}_x\left(T_C \leq 2\right) + \sum_{z \in S_T} P^2(x,z) f(z)
\end{aligned}$$

si sustituimos nuevamente la ecuación (60) (usando z en lugar de x) en la expresión anterior obtenemos

$$\begin{aligned}
f(x) =& \mathbf{P}_x\left(T_C \leq 2\right) + \sum_{z\in S_T} P^2(x,z) f(z) \\
=& \mathbf{P}_x\left(T_C \leq 2\right) + \sum_{z\in S_T} P^2(x,z) \left[\sum_{u\in C} P(z,u) + \sum_{u\in S_T} P(z,u) f(u)\right] \\
=& \mathbf{P}_x\left(T_C \leq 2\right) + \sum_{z\in S_T} P^2(x,z) \left[\sum_{u\in C} P(z,u) + \sum_{u\in S_T} P(z,u) f(u)\right] \\
=& \mathbf{P}_x\left(T_C \leq 2\right) + \sum_{u\in C} \sum_{z\in S_T} P^2(x,z) P(z,u) + \sum_{u\in S_T} \sum_{z\in S_T} P^2(x,z) P(z,u) f(u) \\
=& \mathbf{P}_x\left(T_C \leq 2\right) + \sum_{u\in C} P^3(x,u) + \sum_{u\in S_T} P^3(x,u) f(u) \\
=& \mathbf{P}_x\left(T_C \leq 3\right) + \sum_{u\in S_T} P^3(x,u) f(u)
\end{aligned}$$

por inducción tenemos que

$$f(x) = \mathbf{P}_x\left(T_C \leq n\right) + \sum_{z\in S_T} P^n(x,z) f(z)$$

tomamos limite de $n \to \infty$

$$f(x) = \mathbf{P}_x\left(T_C < \infty\right) = \rho_C(x)$$

lo anterior pues ya que $z \in S_T$ se sigue que $G(x,z) < \infty$ de donde $\sum_{n=1}^{\infty} P^n(x,z) < \infty$ y así $\lim_{n\to\infty} P^n(x,z) = 0$, además como S_T es finito se concluye que $\lim_{n\to\infty} \sum_{z\in S_T} P^n(x,z) f(z) = \sum_{z\in S_T} \lim_{n\to\infty} P^n(x,z) f(z) = 0$. □

5. Cadenas de Markov Notables

A continuación desarrollaremos algunas cuestiones referentes a ciertas cadenas de Marko que sulen ser relevantes dentro de la teoría o en aplicaciones.

5.1. Cadenas de Nacimiento y Muerte

Suponga que la dinámica de crecimiento de una población pueder ser modelada por una cadena de Markov, de tal modo que en cada paso del tiempo la población sólo puede aumnetar en una unidad, disminuir en una unidad o quedarse igual. La población puede tener o no un límite de crecimiento, esto es, su espacio de estados puede ser $S = \mathbf{N}$ o $S = \{0, 1, \ldots, d\}$. *A priori* podemos suponer que las probabilidades de transición depende del estado en que nos entramos, de esta forma

$$P(x,y) = \begin{cases} q_x & \text{si } y = x-1 \\ r_x & \text{si } y = x \\ p_x & \text{si } y = x+1 \\ 0 & \text{en otro caso} \end{cases} \tag{61}$$

donde $p_x, q_x, r_x \in [0,1]$ y $p_x + q_x + r_x = 1$. La cadena de Ehrenfest y la ruina del jugador son casos particulares de cadenas de nacimiento y muerte.

Una pregunta de interes es si la población llegará a quedar extinta en algún momento, dicho de otro modo estamos interesados en calcular $\mathbf{P}_x\left(T_0 < \infty\right)$. Note que si la cadena inicia en el estado $0 < x$ como en el siguiente paso solo puede estar en $x+1$, x o $x-1$ se tiene que

$$0 < T_x < T_{x+1} < T_{x+2} < \cdots \tag{62}$$

esta ecuación nos dice que sin importar el estado en el que empecemos se tiene que $T_n \to \infty$ como $n \to \infty$. Este hecho implica que la sucesión $\left(\{T_0 < T_{n+x}\}\right)_{n\in\mathbf{N}}$ es creciente por lo cual

$$\lim_{n\to\infty} \mathbf{P}_x\left(T_0 < T_{n+x}\right) = \mathbf{P}_x\left(T_0 < \infty\right) = \lim_{n\to\infty} \mathbf{P}_x\left(T_0 < T_n\right). \tag{63}$$

Lo anterior nos dice que para resolver la cuestión de llegar en un tiempo finito al cero tenemos que calcular las probabilidades de la forma $\mathbf{P}_x\left(T_a < T_b\right)$, con $a, b \in S$. Para este fin definimos

$$u(x) = \mathbf{P}_x(T_a < T_b) \text{ con } a < x < b. \tag{64}$$

Note que $u(a) = 1$ y $u(b) = 0$, por otro lado

$$\begin{aligned}
u(y) =& \mathbf{P}_y\left(T_a < T_b\right) \\
=& \mathbf{P}_y\left(T_a < T_b | X_1 = y+1\right) \mathbf{P}_y\left(X_1 = y+1\right) + \mathbf{P}_y\left(T_a < T_b | X_1 = y\right) \mathbf{P}_y\left(X_1 = y\right) \\
&+ \mathbf{P}_y\left(T_a < T_b | X_1 = y-1\right) \mathbf{P}_y\left(X_1 = y-1\right) \\
=& p_y \mathbf{P}_{y+1}\left(T_a < T_b\right) + r_y \mathbf{P}_y\left(T_a < T_b\right) + q_y \mathbf{P}_{y-1}\left(T_a < T_b\right) \\
=& p_y u(y+1) + r_y u(y) + q_y u(y-1)
\end{aligned}$$

puesto que $r_y = 1 - p_y - q_y$ tenemos que

$$u(y) = q_y u(y-1) + u(y) - q_y u(y) - p_y u(y) + p_y u(y+1)$$

de donde

$$u(y+1) - u(y) = \frac{q_y}{p_y}\left(u(y) - u(y-1)\right). \tag{65}$$

Definimos $\gamma_0 = 1$ y

$$\gamma_y = \frac{q_1 \cdots q_y}{p_1 \cdots p_y} \tag{66}$$

por lo que

$$u(y+1) - u(y) = \frac{\gamma_y}{\gamma_{y-1}}\left(u(y) - u(y-1)\right) \tag{67}$$

usando de manera recursiva la ecuación anterior tenemos

$$u(y+1) - u(y) = \frac{\gamma_{a+1} \cdots \gamma_y}{\gamma_a \cdots \gamma_{y-1}}\left(u(a+1) - u(a)\right) = \frac{\gamma_y}{\gamma_a}\left(u(a+1) - u(a)\right) \tag{68}$$

luego si sumamos las ecuaciones generadas por la identidad anterior tenemos

$$
\begin{aligned}
u(b-1)-u(b) =& \frac{\gamma_{b-1}}{\gamma_a}\left(u(a)-u(a+1)\right)\\
+u(b-2)-u(b-1) =& \frac{\gamma_{b-2}}{\gamma_a}\left(u(a)-u(a+1)\right)\\
+&\vdots\vdots\\
+u(a)-u(a+1) =& \frac{\gamma_{a}}{\gamma_a}\left(u(a)-u(a+1)\right)\\
\hline
u(a)-u(b) =& \sum_{j=a}^{b-1}\frac{\gamma_{j}}{\gamma_a}\left(u(a)-u(a+1)\right)
\end{aligned}
$$

como $1 = u(a) - u(b)$ se sigue que

$$
u(a)-u(a+1) = \frac{\gamma_a}{\sum_{y=a}^{b-1}\gamma_y}, \tag{69}
$$

usando las ecuaciones (68) y (69) concluimos que

$$
u(y)-u(y+1) = \frac{\gamma_y}{\sum_{y=a}^{b-1}\gamma_y} \tag{70}
$$

finalmente sumando sobre $y = x, \ldots, b-1$ tenemos

$$
\mathbf{P}_x\left(T_a < T_b\right) = u(x) = \frac{\sum_{y=x}^{b-1}\gamma_y}{\sum_{y=a}^{b-1}\gamma_y} \text{ con } a < x < b. \tag{71}
$$

Con base en lo anterior podemos calcular la probabilidad de extinción de la población, observe que

$$
\mathbf{P}_x\left(T_0 < \infty\right) = \lim_{n\to\infty}\mathbf{P}_x\left(T_0 < T_n\right) = \lim_{n\to\infty}\frac{\sum_{y=x}^{n-1}\gamma_y}{\sum_{y=0}^{n-1}\gamma_y} = \frac{\sum_{y=x}^{\infty}\gamma_y}{\sum_{y=0}^{\infty}\gamma_y} = 1 - \frac{\sum_{y=0}^{x-1}\gamma_y}{\sum_{y=0}^{\infty}\gamma_y} \tag{72}
$$

de esta forma la extinsión es segura desde cualquier estado $x > 0$, es decir, $\mathbf{P}_x\left(T_0 < \infty\right) = 1$ si y solo si $\sum_{y=0}^{\infty}\gamma_y = \infty$.

Si $p_x, q_x > 0$ para todo $x \in S$ entonces tenemos que todos los estados se comunican entre si, particularmente tenemos

$$
\mathbf{P}_0\left(T_0 < \infty\right) = P(0,0) + P(0,1)\mathbf{P}_1\left(T_0 < \infty\right) \tag{73}
$$

por lo tanto $\rho_{00} = \mathbf{P}_0\left(T_0 < \infty\right) = 1$ si y solo si $\mathbf{P}_1\left(T_0 < \infty\right) = 1$ lo que sucede si y solo si $\sum_{y=0}^{\infty}\gamma_y = \infty$, por tanto esta condición es necesaria y suficiente para que la cadena sea recuerrente.

Un pequeño ejemplo de aplicación de lo antes desarrollado es el siguiente. Consideremos la cadena de la ruina del jugador con $X_0 = 10$, $p_x = 9/19$ y $q_x = 10/19$ (por lo que $r_x = 0$) y $S = \{0, 1, 2, \ldots, 35\}$.

Como siempre 0 y $d = 35$ son absorventes. Con base en las fórmulas anteriores podemos calcular la probabilidad de ganar 25 antes de perder 10, esto es, $\mathbf{P}_x(T_{35} < T_0)$.

Con base en lo anterior tenemos que $\gamma_y = (10/9)^y$ con $0 < y < 35$ por lo que de acuerdo a la ecuación (71) por lo que

$$\mathbf{P}_{10}(T_{35} < T_0) = \frac{\sum_{y=0}^{9}(10/9)^y}{\sum_{y=0}^{34}(10/9)^y} = \frac{(10/9)^{10} - 1}{(10/9)^{35} - 1} \approx 0.47.$$

5.2. Cadena de Fila de Espera

Considere un establecimiento de servicios tal como la caja de un supermercado o un banco (o un proceso como las consultas a un servidor). Las personas llegan a la caja en distintos tiempos y son atendidas en algun momento. Aquellos quienes ya han llegado a la caja pero no han sido atendidos forman una fila o cola de espera. Existe una gran variedad de modelos para describir estos sistemas, nosotros consideraremos uno muy simple.

1. Medimos el tiempo en forma discreta en periodos convenientes, digamos minutos.
2. Si hay clientes esperando al inicio de cualquier periodo entonces exactamente uno será atendido en este periodo.
3. Si no hay clientes esperando al inicio de cualquier periodo entonces no se atiende a ninguno en ese periodo, dicho de otro modo, los clientes que llegan posterior al inicio del periodo y antes del termino del mismo serán atendidos hasta el inicio del siguiente periodo.

Este modelo simple de una fila de espera lo podemos formalizar a traves de una sucesión de variables aleatorias $(\xi_n)_{n\in\mathbf{N}}$, cada ξ_n representa el número de clientes que llega durante el $n-$esimo periodo. Dicha sucesión es de variables independientes (los clientes llegan en cantidades independientes en cada periodo de tiempo) e identicamente distribuidas sobre los naturales con función de densidad común f. El proceso $\{X_n : n \in \mathbf{N}\}$ asociado al modelo de filas de espera es tal que cada X_n representa el número de clientes presentes al final del $n-$esimo periodo. En este contexto definimos $X_{n+1} = \xi_{n+1}$ si $X_n = 0$ y $X_{n+1} = X_n + \xi_{n+1} - 1$ si $X_n > 0$. Es sencillo demostrar que bajo esta definición el proceso es una cadena de Markov.

Ejercicio 19. *Demuestre que la cadena de fila de espera como se ha definido es una cadena de Markov.*

Con base en lo anterior para que la cadena vaya del estado 0 (ninguna persona espera en la fila) al estado y (hay y personas esperando en la fila) tendrían que haber llegado y personas en el periodo de espera en cuestión, lo cuál sucede con probabilidad $f(y)$. Similarmente, para que la cadena pase del estado $x > 0$ al estado y, y puesto que se atiende a una persona en cada periodo de tiempo, tendrían que haber llegado $y - x + 1$ personas en el periodo en cuestión, lo que sucede con probabilidad $f(y - x + 1)$. De este modo, la función de transición en un paso queda de la siguiente forma,

$$P(0,y) = f(y)$$
$$P(x,y) = f(y-x+1).$$

Intuitivamente podemos argumentar que sucede con los estados de la cadena en función de la media μ de las variables ξ_i asumiendo que la cadena es irreducible.

1. Si $\mu > 1$ y ya que a lo más se atiende a una persona en cada periodo de tiempo, tendriamos que conforme aumenta el tiempo también lo hace la fila de espera, pues en promedio nos llegan más personas de las que podemos atender en cada periodo de tiempo. De este modo, en general una vez que pasamos por el estado n (hay n personas esperando en la fila) no volveremos a regresar a el, por ende cada estado es visitado a lo mas una cantidad finita de veces, lo que implica que todos los estados son transitorios.

2. Si $\mu < 1$ y ya que a lo más se atiende a una persona en cada periodo de tiempo, en promedio tendriamos que atendemos más personas de las que llegan a la fila en cada periodo de tiempo, por ende inminentemente la fila llega a cero y en promedio se mantiene en cero, lo que nos dice que el cero es un estado recurrente y por lo tanto la cadena es recurrente.

3. Si $\mu = 1$ entonces en promedio atendemos tantas personas como llegan, por lo tanto en general la cadena no aumenta ni disminuye, esto nos dice que en promedio la fila mantiene el mismo tamaño inicial x_0 lo que nos dice que x_0 es recurrente y por ende la cadena es recurrente.

Las observaciones anteriores son todas correctas, pero para demostrarlas necesitamos acudir a argumentos más técnicos, particularmente el siguiente resultado referente a las funciones generadoras de probabilidad para variables aleatorias sobre naturales.

Teorema 7. *Sea Φ la función generadora de probabilidades de una variable aleatoria ξ que toma valores enteros no negativos y $\mu = \mathbf{E}[\xi]$. Si $\mu \leq 1$ y $\mathbf{P}(\xi = 1) < 1$ la ecuación*

$$\Phi(t) = t \tag{74}$$

no tiene raíces en $[0,1)$. Si $\mu > 1$ entonces (74) tiene una única raíz $\rho_0 \in [0,1)$.

Este resultado nos habla del comportamiento de las raíces de la función generadora de probabilidades. Demostraremos este resultado posteriormente, por ahora lo utilizaremos para probar formalmente las observaciones que realizamos respecto a la cadena de fila de espera.

Note que $P(0,z) = p(1,z)$ para todo $z \in S$, recordemos que en general

$$\rho_{xy} = P(x,y) + \sum_{z \neq y} P(x,z)\rho_{xz}$$

por lo que

$$\rho_{00} = P(0,0) + \sum_{y=1}^{\infty} P(0,y)\rho_{y0} = P(1,0) + \sum_{y=1}^{\infty} P(1,y)\rho_{y0} = \rho_{10}. \tag{75}$$

Sea $\rho = \rho_{00} = \rho_{10}$, demostraremos que

$$\Phi(\rho) = \rho$$

donde Φ es la función generadora de ξ_i. Si 0 es recurrente entonces $\rho = 1$ por lo que se satisface $\Phi(\rho) = \rho$ (pues en general $\Phi(1) = 1$).Para verificar la identidad en cuestión en general observe que

$$\rho = \rho_{00} = P(0,0) + \sum_{y=1}^{\infty} P(0,y)\rho_{y0} = f(0) + \sum_{y=1}^{\infty} f(y)\rho_{y0}. \tag{76}$$

Ahora calcularemos ρ_{y0}, note que si la cadena empieza en y el evento $\{T_{y-1} = n\}$ ocurre si y solo si

$$n = \text{mín}\{m > 0 : y + (\xi_1 - 1) + \cdots + (\xi_m - 1) = y - 1\} = \text{mín}\{m > 0 : \xi_1 + \cdots + \xi_m = m - 1\}$$

esto nos dice que $\mathbf{P}_y(T_{y-1} = n)$ no depende de y, es decir, $\mathbf{P}_z(T_{z-1} = n)$ para cualesquiera $x, z \geq 1$. Esto implica que $\rho_{y,y-1} = \mathbf{P}_y(T_{y-1} < \infty)$, consecuentemente

$$\rho_{y,y-1} = \rho_{y-1,y-2} = \cdots = \rho_{10} = \rho_{00} = \rho.$$

Puesto que la cadena solo puede ir un paso a la izquierda en cada periodo de tiempo para ir de y a 0 necesariamente tiene que pasar por $y-1, y-2, \ldots, 1$. Aplicando la propiedad de Markov se sigue que

$$\rho_{y0} = \rho_{y,y-1}\rho_{y-1,y-2}\cdots\rho_{10} = \rho^y \tag{77}$$

de este modo

$$\rho = f(0) + \sum_{y=1}^{\infty} f(y)\rho_{y0} = f(0) + \sum_{y=1}^{\infty} f(y)\rho^y = \Phi(\rho). \tag{78}$$

Por el teorema enunciado tenemos

1. Si $\mu \leq 1$ entonces $\Phi(t)$ no tiene raices en $[0,1)$ (siempre que $f(1) < 1$), de este modo ya que $\Phi(\rho) = \rho$ esto implica necesariamente que $\rho = 1$ por lo que 0 es recurrente y así la cadena es recurrente.
2. Si $\mu > 1$ entonces $\Phi(t)$ tiene una única raíz $\rho_0 \in [0,1)$. Entonces $\rho = \rho_0$ o $\rho = 1$, de hecho $\rho = \rho_0 < 1$ por lo que el 0 es transitorio y así toda la cadena es transitoria.

Para demostrar que $\rho = \rho_0$ note que

$$\mathbf{P}_y(T_0 \leq n+1) \leq \mathbf{P}_y(T_{y-1} \leq n)\,\mathbf{P}_{y-1}(T_{y-2} \leq n)\cdots\mathbf{P}_1(T_0 \leq n) = (\mathbf{P}_1(T_0 \leq n))^y \tag{79}$$

y además recordemos la identidad

$$\mathbf{P}_x(T_y \leq n+1) = P(x,y) + \sum_{z \neq y} P(x,z)\mathbf{P}_z(T_y \leq n) \tag{80}$$

por lo que

$$\mathbf{P}_1(T_0 \leq n+1) = f(0) + \sum_{y=1}^{\infty} f(y)\mathbf{P}_y(T_0 \leq n) \leq f(0) + \sum_{y=1}^{\infty} f(y)\mathbf{P}_1(T_0 \leq n)^y = \Phi(\mathbf{P}_1(T_0 \leq n)). \tag{81}$$

Por otro lado por inducción podemos demostrar que $\mathbf{P}_1(T_0 \leq n) \leq \rho_0$, para ello basta ver que

1. Base de Inducción. $\mathbf{P}_1(T_0 \le 0) = 0 \le \rho_0$.
2. Hipótesis de Inducción. $\mathbf{P}_1(T_0 \le n) \le \rho_0$.
3. Paso Inductivo. $\mathbf{P}_1(T_0 \le n+1) \le \Phi(\mathbf{P}_1(T_0 \le n)) \le \Phi(\rho_0) = \rho_0$.

Finalmente $\rho = \rho_{10} = \lim_{n\to\infty} \mathbf{P}_1(T_0 \le n) \le \rho_0$, como la raíz de Φ es única en $[0,1)$ entonces $\rho = \rho_0$.

6. Distribución Estacionaria

Definición 13. Sea $\{X_n : n \in \mathbf{N}\}$ una cadena de Markov y S su espacio de estados. Decimos que una función de densidad discreta $\pi : S \to [0,\infty)$ es una distribución estacionaria para la cadena si

$$\sum_x \pi(x)P(x,y) = \pi(y) \text{ para todo } y \in S.$$

Recordemos que una función de densidad discreta $\pi : S \to [0,\infty)$ satisface

- $\pi(x) \ge 0$ para todo $x \in S$.
- $\sum_{x\in S} \pi(x) = 1$.

Teorema 8 (Convergencia Acotada.)**.** *Dados $a(x)$, $x \in S$ (finito o numerable), números no negativos con suma finita, y dado $b_n(x)$, $x \in S$ y $n \ge 1$, tales que $|b_n(x)| \le 1$, $x \in S$ y $n \ge 1$, y*

$$\lim_{n\to\infty} b_n(x) = b(x).$$

Entonces,

$$\lim_{n\to\infty} \sum_x a(x)b_n(x) = \sum_x a(x)b(x).$$

Teorema 9 (Propiedades Elementales)**.** *Considere π la distribución estacionaria de una cadena de Markov, entonces esta satisface*

1. *$\sum_x \pi(x)P^n(x,y) = \pi(y)$ para cada $y \in S$ y $n \in \mathbf{N}^+$.*
2. *Si $X_0 \sim \pi$ entonces $\mathbf{P}(X_n = y) = \pi(y)$ para todo $y \in S$ y $n \in \mathbf{N}^+$.*
3. *Si $\mathbf{P}(X_n = y) = \pi_1(y)$ para todo $y \in S$ y $n \in \mathbf{N}$ entonces π_1 es estacionaria.*

4. *Si* $\lim_{n\to\infty} P^n(x,y) = \pi(y)$ *para todo* $y \in S$ *entonces* $\lim_{n\to\infty} \mathbf{P}(X_n = y) = \pi(y)$.

Demostración. Cada uno de los incisos siguientes contiene la demostración del inciso correspondiente del teorema.

1. Note que

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

   la primera igualdad se debe a la ecuación de Chapman-Kolmogorov y la ultima se debe a que π es distribución estacionaria. Aplicando la identidad anterior podemos deducir de forma recursiva que $\sum_x \pi(x) P^{n+1}(x,y) = \pi(y)$.

2. Observe que

$$\mathbf{P}(X_n = y) = \sum_x \mathbf{P}(X_0 = x) P^n(x,y) = \pi(y)$$

   la ultima igualdad se debe a que π es estacionaria.

3. Al igual que antes tenemos que

$$\mathbf{P}(X_1 = y) = \sum_x \mathbf{P}(X_0 = x) P(x,y),$$

   por hipótesis sabemos que $\pi_1(y) = \mathbf{P}(X_1 = y)$ y $\pi_1(x) = \mathbf{P}(X_0 = x)$, por lo tanto

$$\pi_1(y) = \sum_x \pi_1(x) P(x,y)$$

   de donde π_1 es estacionaria.

4. Nuevamente usamos el hecho que

$$\mathbf{P}(X_n = y) = \sum_x \pi(x) P^n(x,y),$$

   por el teorema de convergencia acotada

$$\lim_{n\to\infty} \mathbf{P}(X_n = y) = \lim_{n\to\infty} \sum_x \pi(x) P^n(x,y) = \sum_x \pi(x) \lim_{n\to\infty} P^n(x,y) = \sum_x \pi(x)\pi(y) = \pi(y),$$

   lo anterior pues por hipótesis $\lim_{n\to\infty} P^n(x,y) = \pi(y)$.

$\square$

6.1. Ejemplos

1. **Cadena de dos Estados.** Recordemos que la matriz de transición en un paso de la cadena de dos estados es la siguiente

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

Demostramos que cuando $0 < p + q < 2$ se tiene que

$$\lim_{n\to\infty} \mathbf{P}(X_n = 0) = \frac{q}{p+q} \qquad \lim_{n\to\infty} \mathbf{P}(X_n = 1) = \frac{p}{p+q}$$

del hecho que

$$\mathbf{P}(X_n = 0) = P^n(0,0)\pi_0(0) + P^n(1,0)\pi_0(1) \qquad \mathbf{P}(X_n = 1) = P^n(0,1)\pi_0(0) + P^n(1,1)\pi_0(1)$$

tenemos que si la cadena empieza en 0, es decir, π_0 esta concentrada en 0[2] entonces $P^n(x,0)$ tiende a $q/(p+q)$ y si la cadenca empiza en 1, es decir, π_0 esta concentrada en 1 tenemos que $P^n(x,1)$ tiende a $p/(p+q)$. En resumen

$$\lim_{n\to\infty} P^n(x,0) = \frac{q}{p+q} \qquad \lim_{n\to\infty} P^n(x,1) = \frac{p}{p+q}.$$

Si tomamos $\pi(0) = q/(p+q)$ y $\pi(1) = p/(p+q)$ por lo anterior es claro que π es la distribución estacionaria de la cadena de dos estados. Además esta distribución es única pues si existiera otra distribución estacionaria π_1 tendriamos que

$$P^n(0,x)\pi_1(0) + P^n(1,x)\pi_1(1) = \pi_1(x)$$

tomando límite de $n \to \infty$ obtenemos que $\pi(x)\pi_1(0) + \pi(x)\pi_1(1) = \pi_1(x)$ por lo tanto $\pi(x) = \pi_1(x)$.

2. **Cadena irreducible de 3 estados.** Considere una cadena de Markov con espacio de estados $S = \{0, 1, 2\}$ y matriz de transición

$$P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \end{bmatrix}. \tag{82}$$

Veamos que esta cadena tiene una única distribución estacionaria. Para este fin basta notar que si π es una distribución estacionaria de la cadena se tiene que

$$\sum_x \pi(x)P(x,y) = \pi(y) \text{ para toda } y \in S \tag{83}$$

esto deriva en un sistema de ecuaciones lineales, a saber

$$\begin{aligned} P(0,0)\pi(0) + P(1,0)\pi(1) + P(2,0)\pi(2) &= \pi(0) \\ P(0,1)\pi(0) + P(1,1)\pi(1) + P(2,1)\pi(2) &= \pi(1) \\ P(0,2)\pi(0) + P(1,2)\pi(1) + P(2,2)\pi(2) &= \pi(2) \end{aligned}$$

[2]Una función de probabilidad esta concentrada en un valor x si $\pi_0(x) = 1$ y $\pi_0(y) = 0$ si $y \neq x$

esto es equivalente a

$$\frac{1}{3}\pi(0) + \frac{1}{4}\pi(1) + \frac{1}{6}\pi(2) = \pi(0) \tag{84}$$
$$\frac{1}{3}\pi(0) + \frac{1}{2}\pi(1) + \frac{1}{3}\pi(2) = \pi(1) \tag{85}$$
$$\frac{1}{3}\pi(0) + \frac{1}{4}\pi(1) + \frac{1}{2}\pi(2) = \pi(2) \tag{86}$$

además no olvidemos que

$$\pi(0) + \pi(1) + \pi(2) = 1. \tag{87}$$

Como es de esperarse en el sistema anterior las ecuaciones no son independientes es posible poner una de las 4 ecuaciones en función de las 3 restantes. Sin embargo, una vez reducido el sistema de 4 a 3 ecuaciones este tiene una única solución, esta es $\pi = (6/25, 10/25, 9/25)$.

3. **Cadena reducible con dos clases recurrentes**. Considere la cadena de Markov con espacio de estados $S = \{0, 1, 2, 3\}$ y matriz de transición

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}. \tag{88}$$

Tenemos 2 clases de comunicación $C_0 = \{0, 1\}$ y $C_2 = \{2, 3\}$. Ambas son clases de recurrentes pues son cerradas ya que $P^n(0, 2) = 0$, $P^n(0, 3) = 0$, $P^n(1, 2) = 0$, $P^n(1, 3) = 0$, $P^n(2, 0) = 0$, $P^n(2, 1) = 0$, $P^n(3, 0) = 0$ y $P^n(3, 1) = 0$ para todo $n \in \mathbf{N}$. Al igual que antes establecemos un sistema de ecuaciones

$$\frac{1}{2}\pi(0) + \frac{1}{2}\pi(1) = \pi(0) \tag{89}$$
$$\frac{1}{2}\pi(0) + \frac{1}{2}\pi(1) = \pi(1) \tag{90}$$
$$\frac{1}{2}\pi(2) + \frac{1}{2}\pi(3) = \pi(2) \tag{91}$$
$$\frac{1}{2}\pi(2) + \frac{1}{2}\pi(3) = \pi(3) \tag{92}$$

considerando que

$$\pi(0) + \pi(1) + \pi(2) + \pi(3) = 1. \tag{93}$$

Note que en este caso $\pi_1 = (1/2, 1/2, 0, 0)$ y $\pi_2 = (0, 0, 1/2, 1/2)$.

4. **Cadena de Nacimiento y Muerte.** Consideremos la cadena de nacimiento y muerte y asumimos que la cadena es irreducible. Para encontrar la distribución estacionaria procedemos como antes, plantemos un sistema de ecuaciones a partir de la identidad que debe satisfacer la distribución estacionaria

$$\sum_x \pi(x)P(x, y) = \pi(y) \tag{94}$$

puesto que

$$P(x,y) = \begin{cases} p_x & \text{si } y = x+1 \\ r_x & \text{si } y = x \\ q_x & \text{si } y = x-1 \\ 0 & \text{en otro caso} \end{cases} \tag{95}$$

tenemos

$$\pi(0)r_0 + \pi(1)q_1 = \pi(0) \tag{96}$$

pues $q_0 = 0$, y

$$\pi(y-1)p_{y-1} + \pi(y)r_y + \pi(y+1)q_{y+1} = \pi(y) \text{ si } y > 0 \tag{97}$$

como $p_x + r_x + q_x = 1$ las ecuaciones anteriores se reducen a

$$q_1\pi(1) - p_0\pi(0) = 0 \tag{98}$$

$$q_{y+1}\pi(y+1) - p_y\pi(y) = q_y\pi(y) - p_{y-1}\pi(y-1) \text{ si } y > 0 \tag{99}$$

usando de forma recursiva la ecuación (99) y dada la ecuación (98) podemos concluir que

$$q_{y+1}\pi(y+1) - p_y\pi(y) = 0 \text{ con } y \geq 0, \tag{100}$$

por lo tanto

$$\pi(x) = \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x}\pi(0). \tag{101}$$

Definimos

$$\pi_x = \begin{cases} 1 & \text{si } x = 0 \\ \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} & \text{si } x \geq 1 \end{cases} \tag{102}$$

de esta forma $\pi(x) = \pi_x\pi(0)$ para toda $x \in S$. Puesto que π es distribución estacionaria tenemos que

$$\sum_x \pi(x) = 1 \text{ lo que sucede si y solo si } \sum_x \pi_x = \frac{1}{\pi(0)}. \tag{103}$$

La ecuación anterior nos dice que la distribución estacionaria existe si y solo si $\sum_x \pi_x < \infty$. De este modo la distribución estacionaria esta dada por

$$\pi(x) = \frac{\pi_x}{\sum_{y=0}^{\infty} \pi_y}. \tag{104}$$

Del mismo modo, la ecuación (103) nos dice que si $\sum_x \pi_x = \infty$ entonces $\pi(0) = 0$ por lo que $\pi(x) = 0$ lo que implica que no existe distribución estacionaria.

6.2. Número promedio de visitas a un estado recurrente

El inciso 4 del teorema 9 nos da una condición para tener la convergencia de la cadena a una distribución estacionaria, para ello es requerido que primero el límite $\lim_{n\to\infty} P^n(x,y)$ exista para cada $y \in S$ independientemente del estado x en el que iniciemos.En el ejemplo de la cadena de dos estados hemos visto que en efecto este límite se tiene, de igual forma en el ejemplo de la cadena irreducible de 3 estados también es posible demostrar sin mucha dificultad que dicho límite existe. Por otro lado, en el caso de la cadena reducible con dos clases recurrentes podemos ver facilmente que este límite no se tiene en general, pues $P^n(2,0) = 0$ y $P^n(3,0) = 0$ para todo $n \in \mathbf{N}$, pero $P^n(0,0) = 1/2$ y $P^n(1,0) = 1/2$ para todo $n \in \mathbf{N}$ por lo que tenemos dos límites distintos dependiendo de donde inicie la cadena; si nos restringimos a cada una de las clases recurrentes el límite queda bien definido independientemente de donde incie la cadena y esto tiene todo el sentido desde el hecho que esta cadena no tiene una única distribución estacionaria.

El caso de la cadena de nacimiento y muerte es más interesante, note que si $\sum_x \pi_x = \infty$ entonces no existe $\lim_{n\to\infty} P^n(x,y)$ para todo $y \in S$ (e independientemente de x), pues de ser asi tendriamos que existe una distribución estacionaria lo cual contradice lo expuesto. Por otro lado si $\sum_x \pi_x < \infty$ y $r_x = 0$ para todo $x \in S$ entonces no existe en general el límite $\lim_{n\to\infty} P^n(x,y)$ independientemente de x a pesar de que existe una única distribución estacionaria; esto pues $P^n(x,y) = 0$ si n es par y $|x-y|$ es impar o viceversa.

Los ejemplos anteriores nos dicen que aunque es necesario conocer la convergencia de $\lim_{n\to\infty} P^n(x,y)$, para determinar la convergencia a la distribución estacionaria, esta no siempre existe a pesar de que si pueda existir la distribución estacionaria. Una forma de lidear con este problema es en lugar de tratar de obtener el límite directamente veamos como se comporta su promedio, esta es una condición más débil pues dada una suceción $(a_n)_{n\in\mathbf{N}}$ si

$$\lim_{n\to\infty} a_n = L \tag{105}$$

entonces

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} a_m = L \tag{106}$$

pero si se satisface (106) no necesariamente se satisface (105). *Exempli gratia*, $a_{2n+1} = 1$ y $a_{2n} = 0$, es claro que $\lim a_n$ no existe pero $\lim \frac{1}{n}\sum_{m=1}^{n} a_m = 1/2$. A la convergencia segun la ecuación (106) se le denomina convergencia de Cesáro o que la sucesión es sumable segun Cesáro.

Lo que haremos ver en esta sección es que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(x,y) \tag{107}$$

siempre existe para toda $y \in S$ independientemente de x cuando hay distribución estacionaria.

Definimos el número de visitas a un estado y hasta el tiempo n como

$$N_n(y) = \sum_{m=1}^{n} \mathbf{1}_y(X_m) \tag{108}$$

y definimos la función $G_n(x, y)$ como la esperanza de $N_n(y)$ dado que la cadena empieza en x, es decir,

$$G_n(x, y) = \mathbf{E}_x \left[N_n(y) \right], \tag{109}$$

puesto que $\mathbf{1}_y (X_n) = 1$ con probabilidad $P^n(x, y)$ tenemos, como antes, que

$$G_n(x, y) = \sum_{m=1}^{n} P^m(x, y). \tag{110}$$

Cuando y es un estado transitorio tenemos que $\mathbf{P}_x (N(y) < \infty) = 1$ por lo cual

$$\lim_{n \to \infty} N_n(y) = N(y) \text{ c.s.} \tag{111}$$

y puesto que $G(x, y) < \infty$ por el teorema de convergencia acotada tenemos

$$\lim_{n \to \infty} G_n(x, y) = G(x, y). \tag{112}$$

Lo anterior implica que

$$\lim_{n \to \infty} \frac{1}{n} N_n(y) = 0 \text{ c.s.} \tag{113}$$

y

$$\lim_{n \to \infty} \frac{1}{n} G_n(x, y) = 0. \tag{114}$$

Note que $N_n(y)/n$ es la proporción del número de visitas a y en los primeros n pasos de la cadena y $G_n(x, y)/n$ es simplemente la esperanza de $N_n(y)/n$. Cuando y es recurrente $N(y) = \infty$ c.s. y $G(x, y)$ es 0 o infinito segun si x puede acceder o no a y, es por ello que cuando y es recurrente no es tan simple derivar que sucede con los limites anteriores.

Teorema 10. *Sea $\{X_n | n \in \mathbf{N}\}$ una cadena de Markov con espacio de estados S y $y \in S$ un estado recurrente de la cadena. Entonces*

$$\lim_{n \to \infty} \frac{N_n(y)}{n} = \frac{\mathbf{1}_{\{T_y < \infty\}}}{m_y} \quad \textit{c.s.} \tag{115}$$

$$\lim_{n \to \infty} \frac{G_n(x, y)}{n} = \frac{\rho_{xy}}{m_y} \quad \textit{para todo } x \in S, \tag{116}$$

donde $m_y = \mathbf{E}_y \left[T_y \right]$.

Estas fórmulas son intuitivamente muy razonables. Una cadena que inicia en y regresa a y en promedio cada m_y unidades de tiempo. Así, si $T_y < \infty$ y n es grande, la proporción de las primeras n unidades de tiempo que la cadena esta en y debe ser alrededor de $1/m_y$ pues en promedio visitaremos a y una cantidad n/m_y de veces por lo que $N_n(y)/n$ debe ser en promedio igual a $(n/m_y)/n = 1/m_y$. La segunda fórmula se obtiene al tomar la esperanza en la primera.

Demostración. Sea $r \geq 1$, denotamos al tiempo de la $r-$ésima visita a y como T_y^r, es decir,

$$T_y^r = \text{mín}\left\{n \geq 1 | N_n(y) = r\right\}.$$

Sea $W_y^1 = T_y^1 = T_y$ y para $r \geq 2$ tomamos $W_y^r = T_y^r - T_y^{r-1}$. Las W_y^r, con $r \in \mathbf{N}$, se denominan tiempos interarrivo y representan el tiempo de espera entre la $(r-1)-$ésima y la $r-$ésima visita. Claramente

$$T_y^r = W_y^1 + W_y^2 + \cdots + W_y^r.$$

Los tiempos interarrivo son variables aleatorias independientes e identicamente distribuidas con media común

$$\mathbf{E}_y\left[W_1^y\right] = \mathbf{E}_y\left[T_y\right] = m_y.$$

De este modo, por la ley de los grandes numeros se tiene que

$$\lim_{k\to\infty} \frac{W_1^y + \cdots + W_y^k}{k} = m_y \text{ c.s.}, \tag{117}$$

esto es

$$\lim_{k\to\infty} \frac{T_y^k}{k} = m_y \text{ c.s.}$$

Sea $r = N_n(y)$, es decir, al tiempo n la cadena ha hecho exactamente r visitas al estado y. La $r-$ésima visita sudecede en o antes del tiempo n y la $(r+1)-$ésima visita sucede depues del tiempo n. Por lo tanto

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

y por ende

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

Puesto que $N_n(y) \to \infty$ como $n \to \infty$ tenemos

$$\lim_{n\to\infty} \frac{T_y^{N_n(y)}}{N_n(y)} = \lim_{n\to\infty} \frac{n}{N_n(y)}$$

pero

$$\lim_{n\to\infty} \frac{T_y^{N_n(y)}}{N_n(y)} = \lim_{n\to\infty} \frac{T_y^n}{n} = m_y \text{ c.s.},$$

consecuentemente

$$\lim_{n\to\infty} \frac{n}{N_n(y)} = m_y \text{ c.s.}$$

Note que el argumento anterior es valido sin importar si m_y es finito o intinito. Si $m_y = \infty$ esto necesariamente implica que

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = 0 \text{ c.s.},$$

por el contrario, si m_y es finito entonces

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{1}{m_y} \text{ c.s.},$$

lo que implica que

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{\mathbf{1}_{\{T_y<\infty\}}}{m_y} \text{ c.s.}$$

Por último, usando el teorema de convergencia dominada de Lebesgue se tiene que

$$\lim_{n\to\infty} G_n(x,y) = \lim_{n\to\infty} \mathbf{E}_x\left[\frac{N_n(y)}{n}\right] = \mathbf{E}_x\left[\lim_{n\to\infty} \frac{N_n(y)}{n}\right] = \mathbf{E}_x\left[\frac{\mathbf{1}_{\{T_y<\infty\}}}{m_y}\right] = \frac{\mathbf{P}_x\left(T_y<\infty\right)}{m_y} = \frac{\rho_{xy}}{m_y}$$

$\square$

6.3. Recurrencia Nula y Recurrencia Positiva

Definición 14. Sea $\{X_n | n \in \mathbf{N}\}$ una cadena de Markov con espacio de estados S. Si $x \in S$ es recurrente decimos que es **recurrente nulo** si y sólo si

$$m_x = \mathbf{E}_x\left[T_x\right] = \infty$$

si $m_x < \infty$ decimos ques x es **recurrente positivo**.

Observación 8. *Decimos que un estado recurrente nulo es un estado que se visita una infinidad de veces pero que el tiempo entre cada visita es infinito*

Si $y \in S$ es un estado recurrente nulo, $m_y = \infty$, del teorema 10 se tiene que

$$\lim_{n\to\infty} \frac{G_n(x,y)}{n} = \frac{1}{m_y} = 0$$

por ende

$$\lim_{n\to\infty} \frac{\sum_{m=1}^{n} P^m(x,y)}{n} = 0.$$

Análogamente si y es un estado recurrente positivo del teorema 10 se tiene que

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

consecuentemente

$$\lim_{n\to\infty} \frac{\sum_{m=1}^{n} P^m(x,y)}{n} = \frac{\rho_{xy}}{m_y}.$$

Teorema 11. *Sea x un estado recurrente positivo y suponga que $x \mapsto y$. Entonces y es recurrente positivo.*

Demostración. Particularmente, como x es recurrente se sigue que $y \mapsto x$. Entonces existe $n_1, n_2 \in \mathbf{N}^+$ tales que

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

De la ecuación de Kolmogorov se sigue que

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

sumando sobre $m = 1, \ldots, n$ obtenemos

$$\sum_{m=1}^{n} P^{n_1+m+n_2}(y,y) \geq P^{n_1}(y,x)P^{n_2}(x,y) \sum_{m=1}^{n} P^m(x,x)$$

pero

$$\sum_{m=1}^{n} P^{n_1+m+n_2}(y,y) = \sum_{m=1}^{n_1+n+n_2} P^m(y,y) - \sum_{m=1}^{n_1+n_2} P^m(y,y) = G_{n_1+n+n_2}(y,y) - G_{n_1+n_2}(y,y)$$

consecuentemente

$$\frac{G_{n_1+n+n_2}(y,y)}{n} - \frac{G_{n_1+n_2}(y,y)}{n} \geq P^{n_1}(y,x)P^{n_2}(x,y)\frac{G_n(x,x)}{n}$$

cuando $n \to \infty$ se tiene que

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

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

$\square$

Corolario 11.1. *Sea x un estado recurrente nulo. Si $x \mapsto y$ entonces y es recurrente nulo.*

Corolario 11.2. *Sea $C \subseteq S$ una clase de comunicación. Entonces todos los estados de C son transitorios, recurrentes positivos o recurrentes nulos.*

Teorema 12. *Sea $C \subseteq S$ una clase de comunicación cerrada (conjunto cerrado e irreducible) y finita. Entonces todo estado de C es recurrente positivo.*

Demostración. Como $C \subseteq S$ es cerrado se sigue que

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

sumando $m = 1, \ldots, n$ y como C es finito se sigue que

$$\begin{aligned}
\sum_{y \in C} \frac{G_n(x,y)}{n} =& \frac{1}{n} \sum_{y \in C} \sum_{m=1}^{n} P^m(x,y) \\
=& \frac{1}{n} \sum_{m=1}^{n} \sum_{y \in C} P^m(x,y) \\
=& \frac{1}{n} \sum_{m=1}^{n} 1 = 1
\end{aligned}$$

si todo estado de C fuese transitorio o recurrente nulo de la finitud de C obtenemos

$$1 = \lim_{n \to \infty} \sum_{y \in C} \frac{G_n(x,y)}{n} = \sum_{y \in C} \lim_{n \to \infty} \frac{G_n(x,y)}{n} = 0$$

lo cual es una contradicción. Por lo tanto existe al menos un estado recurrente positivo en C, por ende todo estado de C es recurrente positivo. $\square$

Corolario 12.1. *Una cadena de Markov irreducible con espacio de estados finito es recurrente positiva.*

Corolario 12.2. *Una cadena de Markov con espacio de estados finito no tiene estados recurrentes nulos.*

6.4. Existencia y Unicidad de Distribuciones Estacionarias

Sea π una distribución estacionaria y $m \in \mathbf{N}^+$. Entonces:

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

sumando desde $m = 1, ..., n$, y dividiendo entre n se tiene:

$$\sum_z \pi(z)\frac{G_n(z,x)}{n} = \pi(x). \tag{118}$$

Teorema 13. *Sea π una distribución estacionaria. Si $x \in S$ es un estado recurrente nulo o transitorio entonces $\pi(x) = 0$.*

Demostración. Como x es un estado recurrente nulo o transitorio se tiene por el teorema 10 que:

$$\lim_{n\to\infty} \frac{G_n(z,x)}{n} = 0 \qquad \forall z \in S,$$

por la ecuación en (118) y el teorema de convergencia acotada se sigue que:

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

□

Corolario 13.1. *Si $\{X_n | n \in \mathbf{N}\}$ es una cadena que no tiene estados recurrentes positivos entonces no existe distribución estacionaria.*

Teorema 14. *Una cadena irreducible de estados recurrentes positivos tiene una única distribución estacionaria π, dada por:*

$$\pi(x) = \frac{1}{m_x} \qquad \forall x \in S.$$

Demostración. Como todo estado es recurrente positivo se tiene del teorema 10 que:

$$\lim_{n\to\infty} \frac{G_n(z,x)}{n} = \frac{1}{m_x} \qquad \forall z \in S.$$

Primero, veamos la unicidad. Si π es una distribución estacionaria de la cadena se tiene que:

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

Luego,

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

por

es distribución estacionaria, es decir, cumple con lo siguiente:

a) $\sum_z \pi(z) = 1$, es decir, $\sum_z \dfrac{1}{m_z} = 1$.

b) $\pi(x) = \sum_z \pi(z)P(z,x) \iff \dfrac{1}{m_x} = \sum_z P(z,x)\dfrac{1}{m_z}$.

Es claro que,

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

sumando sobre $m = 1, ..., n$ y dividiendo entre n, concluimos:

$$\sum_x \frac{G_n(z,x)}{n} = 1 \qquad \forall x \in S. \tag{119}$$

Por otro lado, también sabemos que:

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

sumando sobre $m = 1, ..., n$ y dividiendo entre n se sigue:

$$\sum_x \frac{G_n(z,x)}{n}P(x,y) = \frac{G_{n+1}(z,y)}{n} - \frac{P(z,y)}{n}. \tag{120}$$

Primero veamos el caso si S es finito, entonces de (120) se tiene que:

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

y luego

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

por lo tanto π es estacionaria.

Ahora veamos el caso si S es no finito. El argumento es más complejo (de hecho no puede aplicar el teorema de convergencia acotada). Sea $S = \{x_1, x_2, x_3, ...\} = \{x_n : n \in \mathbf{N}\}$, tenemos que:

$$\sum_{i=1}^{k} \frac{G_n(z,x_i)}{n} = \frac{1}{n}\sum_{i=1}^{k}\sum_{m=1}^{n} P^m(z,x_i) = \frac{1}{n}\sum_{m=1}^{n}\sum_{i=1}^{k} P^m(z,x_i) \le \frac{1}{n}\sum_{m=1}^{n} 1 \le 1 \qquad \forall z \in S,$$

tomando $\lim_{n\to\infty}$:

$$\sum_{i=1}^{k} \frac{1}{m_{x_i}} \leq 1 \qquad \forall z \in S,$$

esto implica que:

$$\sum_{i=1}^{\infty} \frac{1}{m_{x_i}} \leq 1. \tag{121}$$

Similarmente de (120) se tiene que:

$$\sum_{i=1}^{k} \frac{G_n(z,x_i)}{n} P(x_i,y) \leq \frac{G_{n+1}(z,y)}{n} - \frac{P(z,y)}{n},$$

tomando $\lim_{n\to\infty}$, concluimos que:

$$\sum_{i=1}^{k} \frac{1}{m_{x_i}} P(x_i,y) \leq \frac{1}{m_y};$$

y así,

$$\sum_{i=1}^{\infty} \frac{1}{m_{x_i}} P(x_i,y) \leq \frac{1}{m_y}. \tag{122}$$

Ahora veamos que en (122) se da la igualdad, supongamos que existe $y_0 \in S$ tal que,

$$\frac{1}{m_{y_0}} > \sum_{i=1}^{\infty} \frac{1}{m_{x_i}} P(x_i,y_0),$$

sumando sobre todos los valores de y del lado derecho de la igualdad (122) se sigue que:

$$\sum_y \frac{1}{m_y} > \sum_y \left(\sum_{i=1}^{\infty} \frac{1}{m_{x_i}} P(x_i,y) \right) = \sum_{i=1}^{\infty} \frac{1}{m_{x_i}} \left(\sum_y P(x_i,y) \right) = \sum_{i=1}^{\infty} \frac{1}{m_{x_i}} = \sum_y \frac{1}{m_y} \quad !$$

Por lo tanto en (122) debe darse la igualdad,

$$\sum_x \frac{1}{m_x} P(x,y) = \frac{1}{m_y} \qquad \forall y \in S. \tag{123}$$

De la ecuación (121) se tiene que $\sum_x \dfrac{1}{m_x} = c^{-1} < +\infty$, así,

$$\pi_s(x) = \frac{c}{m_x} \qquad \forall x \in S,$$

define una distribución estacionaria debido a (122), pues:

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

Por la unicidad de la distribución estacionaria se sigue que:

$$\pi_s = \pi, \quad i.e. \quad \frac{c}{m_x} = \frac{1}{m_x} \quad \forall x \in S \quad \Rightarrow \quad c = 1,$$

de donde se concluye que

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

y que π, como lo define el teorema, es una distribución estacionaria.

$\square$

Corolario 14.1. *Una cadena de Markov irreducible es recurrente positiva si y sólo si tiene una distribución estacionaria.*

Demostración. $\Rightarrow$) Teorema anterior.

$\Leftarrow$) Como la cadena es irreducible sólo tiene estados de un sólo tipo, es decir, todos son transitorios, todos recurrentes nulos o recurrentes positivos. Sabemos que si la cadena no tiene estados recurrentes positivos entonces no tiene distribución estacionaria y por ende todos los estados de la cadena son recurrentes positivos. $\square$

Corolario 14.2. *Si una cadena de Markov es irreducible y con espacio de estados finito entonces tiene una única distribución estacionaria.*

Corolario 14.3. *Sea $\{X_n | n \in \mathbf{N}\}$ una cadena de Markov irreducible recurrente positiva con distribución estacionaria π. Entonces*

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \lim_{n\to\infty} \frac{G_n(x,y)}{n} = \pi(y) \text{ para todo } y \in S. \tag{124}$$

6.5. Cadenas Reducibles

Definición 15. Sea π una distribución estacionaria sobre S y $C \subseteq S$. Decimosq ue π esta concentrada en C si

$$\pi(x) = 0 \text{ para toda } x \notin C. \tag{125}$$

Teorema 15. *Sea $C \subseteq S$ una clase de comunicación de recurrentes positivos. Entonces la cadena de Markov tiene una única distribución estacionaria concentrada en C la cuál esta dada por*

$$\pi(x) = \begin{cases} \frac{1}{m_x} & si\ x \in C \\ \\ 0 & e.o.c. \end{cases} \tag{126}$$

Demostración. Primero note que la distribución definida es estacionaria sobre la cadena. Si $y \in S$ entonces

$$\sum_x \pi(x)P(x,y) = \sum_{x \in C} \pi(x)P(x,y) = \pi(y)$$

pues por el teorema 14 π es distribución estacionaria sobre C. Si $y \notin C$ entonces

$$\sum_x \pi(x)P(x,y) = \sum_{x \in C} \pi(x)P(x,y) = 0 = \pi(y)$$

pues $P(x,y) = 0$ siempre que $x \in C$ y por definición $\pi(y) = 0$.

Ahora veamos la unicidad. Si existe otra distribución estacionaria π_s concentrada en C se tiene que si $y \notin C$ entonces $0 = \pi_s(y) = \pi(y)$ y si $y \in C$ entonces $\pi_s(y) = \pi(y)$ pues sobre C existe una única distribución estacionaria de acuerdo al teorema 14. ☐

Corolario 15.1. *Suponga que $C_0, C_1 \subseteq S$ son dos clases de comunicación de recurrentes positivos para una cadena de Markov dada. Entonces la cadena tiene una infinidad de distribuciones estacionarias.*

Demostración. Por el teorema anterior existen únicas distribuciones π_0 y π_1 concentradas en C_0 y C_1 respectivamente. Luego para cada $\alpha \in [0,1]$ podemos definir la función de probabilidad

$$\pi_\alpha = \alpha\pi_0 + (1-\alpha)\pi_1$$

y π_α es distribución estacionaria sobre la cadena,

$$\sum_x \pi_\alpha(x)P(x,y) = \alpha\sum_x \pi_0(x)P(x,y) + (1-\alpha)\sum_x \pi_1(x)P(x,y) = \alpha\pi_0(x) + (1-\alpha)\pi_1(x) = \pi_\alpha(x).$$

De este modo existe una infinidad de distribuciones estacionarias para la cadena. ☐

Corolario 15.2. *Sea $S_P \subseteq S$ el conjunto de estados recurrentes positivos de una cadena de Markov.*

1. *Si $S_P = \emptyset$ entonces la cadena no tiene distribución estacionaria.*
2. *Si $S_P \neq \emptyset$ y es irreducible entonces la cadena tiene una única distribución estacionaria.*
3. *Si $S_P \neq \emptyset$ y no es irreducible entonces la cadena tiene una infinidad de distribuciones estacionarias.*

6.6. Convergencia a la Distribución Estacionaria

En las secciones anteriores ya hemos demostrado que

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} P^m(x,y) = \lim_{n\to\infty} \frac{G_n(x,y)}{n} = \pi(y)$$

cuando la cadena es irreducible, es decir, sabemos que la sucesión de probabilidades de trancisión en n pasos converge segun Cesáro. En esta sección final discutiremos cuando la convergencia de dicha sucesión se da en forma clasica o fuerte, esto es, cuando en efecto

$$\lim_{n\to\infty} P^n(x,y) = \pi(y).$$

Para este fin es necesario abordar el concepto de periodicidad y ver como esta afecta la convergencia clasica de las probabilidades de transición.

Definición 16. Considere una cadena de Markov y un estado $x \in S$ tal que $\rho_{xx} > 0$. Definimos el periodo de x como

$$d_x = g.c.d\left\{n \in \mathbf{N}^+ : P^n(x,x) > 0\right\}$$

esto es, el periodo es el máximo común divisor de los naturales para los cuales la funciones de transición son positivas. Si $d_x = 1$ decimos que el estado no tiene periodo o es aperiodico.

Lema 4. *Si $x \leftrightarrow y$ entonces $d_x = d_y$.*

Demostración. Como $x \leftrightarrow y$ existen $n_1, n_2 \in \mathbf{N}$ tales que

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

en consecuencia

$$P^{n_1+n_2}(x,x) \geq P^{n_1}(x,y)P^{n_2}(y,x) > 0$$

de donde $d_x | n_1 + n_2$. Por otro lado, si $P^n(y,y) > 0$ se tiene que

$$P^{n_1+n+n_2}(x,x) \geq P^{n_1}(x,y)P^n(y,y)P^{n_2}(y,x) > 0$$

por lo cual $d_x | n_1 + n + n_2$ de donde $d_x | n$. Así, d_x divida a cada $n \in \mathbf{N}$ tal que $P^n(y,y) > 0$ lo que implica que $d_x \leq d_y$. *Mutatis mutandis* se puede demostrar que $d_y \leq d_x$ por lo que $d_x = d_y$. ☐

Observación 9. *Debido a que el periodo es una propiedad de clase de comunicación tiene sentido definir el periodo de una cadena de Markov irreducible como el periodo de cualquiera de sus estados, pues este será igual para todos ellos.*

Ejemplo 3. *Calculemos la periodicidad de la cadena de nacimiento y muerte cuando es irreducible. Si $r_x > 0$ para algun $x \in S$ entonces $P(x,x) = r_x > 0$ por lo que $d_x = 1$, de donde la cadena es aperiodica. Si $r_x = 0$ para todo $x \in S$ entonces $P^{2n+1}(0,0) = 0$ y $P^2(0,0) = p_0q_1 > 0$ por lo que $d_0 = 2$ de donde la cadena tiene periodo 2.*

Teorema 16. *Sea* $\{X_n | n \in \ N\}$ *una cadena de Markov irreducible recurrente positiva con distribución estacionaria* π*.*

1. *Si la cadena es aperiodica* ($d = 1$) *entonces*

$$\lim_{n \to \infty} P^n(x, y) = \pi(y) \text{ para todo } y \in S. \tag{127}$$

2. *Si la cadena tiene periodo* $d > 1$ *entonces para cada par de estados* $x, y \in S$ *existe un entero* r*,* $0 \leq r < d$*, tal que* $P^n(x, y) = 0$ *a menos que* $n = md + r$ *para algun* $m \geq 0$*, y entonces*

$$\lim_{m \to \infty} P^{md+r}(x, y) = d\pi(y). \tag{128}$$