Capítulo 1

Cadenas de Markov

1.1. Introducción

Sea $\{X_n\}_{n\geq 0}$ una secuencia de variables aleatorias, donde X_n representa el estado de un sistema en el instante n y el conjunto de estados $\mathbf{E}$ esta dado por $\mathbf{E} = \{E_j : j = 0, 1, 2, \ldots\}$.

Notación

$X_n = j$. valor que toma la variable aleatoria X en el instante n.

Definición 1.1.1 *Se dice que las variables aleatorias $X_i, i = 0, 1, 2, \ldots, n$ son **mutuamente independientes** si y sólo si*

$$P\left(\{X_n = j/X_0 = i_0, X_1 = i_1, X_2 = i_2, \ldots, X_{n-1} = i_{n-1}\}\right) = P\left(\{X_n = j\}\right)$$

Definición 1.1.2 (Cadena de Markov) *Es una familia de variables aleatoria, $\{X_n : n \in \mathbb{N}\}$, tales que para todo n y para todos los valores posibles de los X_n se cumple*

$$P\left(\{X_n = j/X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\}\right) = P\left(\{X_n = j/X_{n-1} = i_{n-1}\}\right)$$

1.2. Ejemplos de algunas Cadenas de Markov

1.2.1. Cadenas de Ramificación

Considere una partícula que es capaz de producir nuevos hijos de cada tipo. Para cada miembro sea $p_k, k = 0, 1, 2, ...$ la probabilidad de crear k nuevos miembros. Los descendientes directos de la generación n forman la generación $(n + 1)$. Los miembros de cada generación son independientes de otros. Supongamos que X_n es el tamaño de la n-esima generación, es claro que X_n solo depende d X_{n-1} ya que $X_n = \sum_{i=1}^{n-1} Y_i$ donde Y_i es el número de hijos del i-esimo miembro de la generación $n - 1$. Por lo tanto, X_n representa una cadena de Markov. Las reacciones nucleares en cadena, la sobrevida de los apellidos familiares, la mutación de gens y las líneas de espera en sistemas de colas son ejemplos de procesos de ramificación.

Ejemplo 1.2.1 (Reacción Nuclear en Cadena) *En una reacción nuclear en cadena una partícula tal como un neutron se activa con probabilidad p, creando m partículas nuevas, y $q = 1 - p$ representa la probabilidad de que esta permanezca inactiva sin descendientes. En este caso, el único número posible de descendientes es cero o m con probabilidad q y p respectivamente. Si p es cercano a uno, es muy probable que el número de partículas se incrementan indefinidamente, llevando a una explosión, mientras que si p es cercano a cero el proceso puede nunca iniciarse.*

Ejemplo 1.2.2 (Sobrevida del Apellido Familiar) *En la sociedad victoriana producía preocupación la posibilidad de que los apellidos aristocráticos se estuviesen extinguiendo.*
El francés I.J. Bienaymé fue el primer estadístico del que se tiene constancia que consideró este problema, en 1845. Paralelamente a Bienaymé, los investigadores británicos Francis Galton y Henry William Watson estudiaron de modo propio el problema de la propagación de los apellidos, publicando sus resultados en Galton y Watson (1874), entre otros documentos. Para una historia detallada, ver (Kendall, 1966) y (Kendall, 1975).
Para la realización del estudio se considera una población con M varones adultos en la que, en cada generación, el a_0 por ciento de los varones adultos

no tiene hijos varones que alcancen la vida adulta, el a_1 *por ciento tiene un hijo varón que llega a adulto, el* a_2 *tiene dos y así sucesivamente hasta el* a_5 *por ciento que tiene 5. El problema plantea, en esas circunstancias, dos cuestiones:*

- *¿Qué proporción de los apellidos se ha extinguido tras r generaciones?*
- *¿En cuántos casos el mismo apellido será compartido por m personas?*

Si bien Galton y Watson realizaron grandes avances en la investigación del Problema, la solución definitiva al mismo llegó bastante más tarde, en 1930, de la mano de J.F. Steffensen. Este matemático demostró, entre otras cosas, que la probabilidad de extinción de una familia (del apellido que comparte dicha familia) depende del número medio de hijos de cada individuo (con la notación anterior, de la cantidad $r = a_1 + 2a_2 + 3a_3 + ...$ *Si* $r \leq 1$ *la probabilidad de que un apellido familiar se extinga es 1 y si* $r > 1$*, dicha probabilidad de extinción tomará un valor entre 0 y 1.*
Estos estudios crearon el llamado proceso de Galton-Watosn que en un principio se aplicó únicamente al problema de la extinción de los apellidos. Sin embargo, pronto se empezó a emplear en el campo de la biología, para modelizar la extensión de los seres vivos. Hoy en día, se aplica en multitud de disciplinas, desde la teoría de colas hasta la propagación de virus informáticos.

Ejemplo 1.2.3 (Genética) *En el problema de la mutación genética, todos los genes tienen un chance de reaparecer en sus k descendientes directos con probabilidad* $p_k, k = 1, 2, ...$ *Una mutación espontánea produce un gen simple que juega el rol de una partícula en la generación cero. La dispersión del gen a través de las siguientes generaciones sigue una cadena de Markov. Una cuestión de interés puede ser que tan probable es que el gen este presente en k nacimientos.*

Ejemplo 1.2.4 (Líneas de Espera) *En cualquier tipo de colas o líneas de espera, los clientes llegan aleatoriamente y esperan a ser servidos. Un cliente* *que llega cuando el servidor está libre recibe inmediatamente el servicio, de lo contrario el cliente se une a la cola y espera por el servicio. El servidos atiende a los clientes de acuerdo a alguna disciplina de la cola, como por ejemplo primero en llegar primero en salir. La duración total de servicio ininterrumpido desde que inicia en $t = 0$ es conocido como el período de ocupación. Bajo el supuesto de que el tiempo de varias llegadas y el tiempo de servicio son variables aleatorias independdientes, entonces podría ser de interés conocer el número total de clientes durante el período de ocupación, la duración del período de ocupación, la probabilidad del tiempo de terminación, entre otras. Como se verá, las colas están universalmente caracterizadas dependiendo de la distribución del tiempo de servicio y el número total de servidores. Por ejemplo, los tiempos entre llegadas y servicio pueden ser distribuidos exponencial, mientras que el número de servidores puede ser uno o varios en paralelo.*

Sea X_n el número de clientes que esperan en la cola en el instante t_n cuando el n-esimo cliente se va del sistema después de completar el servicio. Si consideramos que el primer cliente que llega y recibe de inmediato el servicio como la generación cero, entonces los descendientes directos son los X_1 clints que llegan durante el tiempo de servicio del primer cliente y forman la cola. Y de esta manera el proceso continua. Es claro que el número de clientes, X_n que esperan servicio en l tiempo t_n forman una cadena de Markov.

1.2.2. Caminatas Aleatorias

Una caminata aleatoria es una cadena de Markov que en su versión más sencilla considera que una partícula, en el tiempo 0, está en el origen. Luego en cada unidad de tiempo, aleatoriamente la partícula se mueve a la derecha o a la izquierda. Por lo tanto, la variable aleatoria X_n denota la posición de la partícula después de un tiempo n, entonces el espacio de estado es $\{0, 1, 2, 3, ...\}$. Es claro que la posición siguiente de la partícula solo depende de la posición actual, entonces se cumple la propiedad de Markov. Si estando la partícula en el estado i esta sólo se puede mover a los estados $i+1$ y $i-1$, con probabilidades p y q respectivamente, entonces se tiene

$$P\left(\{X_n = j/X_{n-1} = i\}\right) = \begin{cases} p, & \text{j=i+1;} \\ q, & \text{j=i-1.} \end{cases}$$

Una generalización de la caminata aleatoria es considerar que la partícula puede permanecer en el mismo estado en una transición, si denotamos por r la probabilidad de que eso ocurra se tiene que

$$P\left(\{X_n = j/X_{n-1} = i\}\right) = \begin{cases} p, & \text{j=i+1;} \\ q, & \text{j=i-1;} \\ r, & \text{j=i.} \end{cases}$$

Otra manera de generalizar la caminata aleatoria es asumir que las probabilidades p y q (y r en el caso anterior) pueden depender del estado i donde la partícula esta. Así, tenemos

$$P\left(\{X_n = j/X_{n-1} = i\}\right) = \begin{cases} p_i, & \text{j=i+1;} \\ q_i, & \text{j=i-1.} \\ r_i, & \text{j=i.} \end{cases}$$

Ejemplo 1.2.5 (Ruina del Jugador) *Supongamos que 2 jugadores Antonio y Bonifacio, participan en un juego solamente ellos dos. Antonio dispone de a bolívares y Bonifacio de b bolívares, por lo tanto se dispone de* $c = a + b$ *Bs. Sea* X_n *la fortuna de Antonio en el instante* n. *Suponga que Antonio gana con probabilidad p y pierde con probabilidad q. Como la fortuna de Bonifacio solo depende de su fortuna actual entonces* $\{X_n\}$ *es una cadena de Markov, cuyo espacio de estado es* $\{0, 1, 2, 3, ..., c = a + b\}$.

Ejemplo 1.2.6 (Borrachito) *Un borracho se mueve aleatoriamente en el eje de las abscisas con paso unitario. Sea*

r_i : *probabilidad de quedarse en* E_i.

p_i : *probabilidad de pasar de* E_i *a* E_{i+1}.

q_i : *probabilidad de pasar de* E_i *a* E_{i-1}.

Supongamos que

$p_i > 0; q_i > 0; r_i \geq 0; \forall i \geq 1$

$p_i + q_i + r_i = 1; \forall i \geq 1$

$r_0 + p_0 = 1; r_0 \geq 0; p_0 \geq 0.$

Sea X_n la posición del borracho en el instante n. Como la posición siguiente del borrachito solamente depende de la posición actual, $\{X_n\}$ una cadena de Markov con espacio de estados $E = \{E_0, E_1, E_2, ...\}$

Ejemplo 1.2.7 (Vendedor) *Un vendedor vive en el pueblo $'a'$ y es responsable de los pueblos $'a','b'$ y $'c'$. Cada semana se requiere visitar un pueblo diferente. Cuando él esta en casa, es indiferente el pueblo que el visitará a la semana siguiente, por lo tanto lanza una moneda y si cae cara el visita al pueblo $'b'$ de lo contrario visita al pueblo $'c'$. Sin embargo, cuando el esta por fuera tiene cierta preferencia de ir a casa, por lo tanto, él lanza dos monedas, si las dos monedas caen cara el visita el otro pueblo que no está en su casa. Los sucesivos pueblos que el visita forman una cadena de Markov con espacio de estados $E = \{a, b, c\}$ donde la variable aleatoria X_n es igual a $'a','b'$ o $'c'$ de acuerdo a su ubicación en la semana n.*

1.2.3. Otros Ejemplos

Ejemplo 1.2.8 (Acciones) *Considere el siguiente modelo para el valor de una acción. Al final de un día dado se registra el precio. Si la acción subió, la probabilidad de que suba mañana es p_1. Si la acción bajó, la probabilidad de que suba mañana es p_2. Evidentemente el valor de la acción al día siguiente depende el valor actual, por lo tanto si X_n representa el valor de la acción el día n, entonces $\{X_n\}$ es una cadena de markov, cuyo espacio de estado es $\{s, b\}$ donde s representa que la acción subió y b que la acción bajo.*

Ejemplo 1.2.9 (Inventario) *Este es un ejemplo tomado del Hillier- Lieberman (Investigación de operaciones). Una tienda de cámaras tiene en almacén un modelo especial de cámara que se puede ordenar cada semana.*

Sean $D_1, D_2, ...$ las demandas respectivas de esta cámara duran te la primera, segunda,..., semana. Se supone que las D_i son variables aleatorias independientes e idénticamente distribuidas que tienen distribución Poisson con media 1. Sea X_0 el número de cámaras que se tienen en el momento de iniciar el proceso, X_1 el número de cámaras que se tienen al final de la semana 1, X_2 el número de cámaras que se tienen al final de la semana 2, etc. Suponga que $X_0 = 3$. El Sábado en la noche la tienda hace un pedido que le entregan en el momento de abrir la tienda el lunes. La tienda usa la siguiente política para ordenar: si no hay cámaras en el inventario, ordena 3 cámaras, de otra manera, si se cuenta con cámaras en el almacén, no se hace el pedido. Se supone que las ventas se pierden cuando la demanda excede el inventario. Entonces $\{X_n; n = 0, 1, 2...\}$ es una cadena de Markov de la forma que se acaba de describir. Los estados posibles de la cadena son los enteros $0, 1, 2, 3, ...$ *que representan el número posible de cámaras en inventario al final de la semana.*

Ejemplo 1.2.10 (Mantenimiento de una maquinaria) *Considere un granjero que usa un viejo tractor. El tractor esta a menudo en reparación pero solo dura un día para entrar nuevamente en funcionamiento. El primer día fuera de reparación el siempre trabaja pero un día cualquiera, independientemente de su historia previa, hay cierto de chance de que se dañe y por lo tanto sea enviada a reparación. Sean $X_0, X_1, ...$ variables aleatorias que denotan la condición diaria del tractor, donde un uno denota la condición trabajando y cero la condición fallando. Como el estado siguiente del tractor depende del estado actual, entonces $\{X_n; n = 0, 1, 2...\}$ es una cadena de Markov.*

1.3. Definiciones

Definición 1.3.1 (Probabilidades Iniciales) *Sea*

$$p_j^{(0)} = P(X_0 = j) \quad \forall j = 0, 1, 2, ...$$

Los $p_j^{(0)}$ *reciben el nombre de* ***probabilidades iniciales****. En forma matricial se tiene*

$$p^{(0)} = [p_0^{(0)}, p_1^{(0)}, p_2^{(0)}, ...]$$

Definición 1.3.2 (Probabilidades de Estado) *Sea*

$$p_j^{(n)} = P(X_n = j) \quad \forall j = 0,1,2,...\forall n = 0,1,2,...$$

$p_j^{(n)}$ *es la probabilidad de que en el momento* n *el sistema está en el estado* j*. En forma matricial*

$$\underline{\mathbf{P}}^{(n)} = [p_0^{(n)}, p_1^{(n)}, p_2^{(n)}, ...]$$

Definición 1.3.3 (Probabilidades de Transición) *Sea*

$$p_{ij} = P(X_n = j/X_{n-1} = i) \quad \forall i,j = 0,1,2,...\forall n = 1,2,...$$

p_{ij} *es la probabilidad de que en el momento* n *el sistema está en el estado* j *dado que en el instante anterior estaba en el estado* i*.*

Nota: Si los p_{ij} no dependen de n, se dice que la cadena es homogénea.

Definición 1.3.4 (Matriz de Probabilidad de Transición)

$$\mathbf{P} = \begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \\ \vdots \\ E_r \end{array} \begin{array}{c} \begin{array}{ccccc} E_0 & E_1 & E_2 & \cdots & E_r \end{array} \\ \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0r} \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1r} \\ p_{20} & p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ p_{r0} & p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \end{array}$$

Características de P:

1. $0 \leq p_{ij} \leq 1 \; \forall i = j$
2. $\sum_{j=0}^{r} p_{ij} = 1 \; \forall i = 0,1,2,...,r$

1.3.1. Ejemplos

Ejemplo 1.3.1 *Considere un granjero que usa un viejo tractor. El tractor esta a menudo en reparación pero solo dura un día para entrar nuevamente en funcionamiento. El primer día fuera de reparación el siempre trabaja pero un dia cualquiera, independientemente de su historia previa, hay un 10 % de chance de que se dañe y por lo tanto sea enviada a reparación. Sean* $X_0, X_1, ...$ *variables aleatorias que denotan la condición diaria del tractor, donde un uno denota la condición trabajando y cero la condición fallando. En otras palabras,*

$$X_n = \begin{cases} 1, & \textit{el tractor esta trabajando;} \\ 0, & \textit{el tracto esta en reparación.} \end{cases}$$

Por lo tanto, $X_0, X_1, ...$ *es una cadena de Markov cuyo espacio de estado* $E = \{0, 1\}$*. Las probabilidades de transición son*

$$P(X_{n+1} = 0/X_n = 0) = 0$$

$$P(X_{n+1} = 1/X_n = 0) = 1$$

$$P(X_{n+1} = 0/X_n = 1) = 0{,}1$$

$$P(X_{n+1} = 1/X_n = 1) = 0{,}9$$

Por lo tanto, la matriz de transición está dada por

$$\begin{array}{cc} & \begin{array}{cc} E_0 & E_1 \end{array} \\ \begin{array}{c} E_0 \\ E_1 \end{array} & \begin{pmatrix} 0 & 1 \\ 0{,}1 & 0{,}9 \end{pmatrix} \end{array}$$

Ejemplo 1.3.2 (Ruina del Jugador) *Supongamos que 2 jugadores Antonio y Bonifacio, participan en un juego solamente ellos dos. Antonio dispone de a bolívares y Bonifacio de b bolívares, por lo tanto se dispone de* $c = a + b$ *Bs. Sea* X_n *la fortuna de Antonio en el instante* n*. Suponga que Antonio gana con probabilidad* p *y pierde con probabilidad* q*. Como la fortuna de Bonifacio solo depende de su fortuna actual entonces* $\{X_n\}$ *es una cadena de*

Markov, cuyo espacio de estado es $\{0, 1, 2, 3, ..., c = a + b\}$. *Por lo tanto, las probabilidades de transición de un paso son:*

$$p_{00} = P(X_n = 0/X_{n-1} = 0) = 1$$

$$p_{0j} = P(X_n = j/X_{n-1} = 0) = 0 \quad \forall j > 0$$

$$p_{10} = P(X_n = 0/X_{n-1} = 1) = q$$

$$p_{11} = P(X_n = 1/X_{n-1} = 2) = 0$$

$$p_{12} = P(X_n = 2/X_{n-1} = 1) = p$$

$$p_{1j} = P(X_n = j/X_{n-1} = 2) = 0 \quad \forall j > 2$$

$$p_{20} = P(X_n = 0/X_{n-1} = 2) = 0$$

$$p_{21} = P(X_n = 1/X_{n-1} = 2) = q$$

$$p_{22} = P(X_n = 2/X_{n-1} = 1) = 0$$

$$p_{23} = P(X_n = 3/X_{n-1} = 2) = p$$

$$p_{2j} = P(X_n = j/X_{n-1} = 2) = p \quad \forall j > 3$$

y así sucesivamente. La matriz de transición de estados es

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & ... & 0 \\ q & 0 & p & 0 & ... & 0 \\ 0 & q & 0 & p & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & ... & 1 \end{pmatrix}$$

Ejemplo 1.3.3 (Borrachito) *Un borracho se mueve aleatoriamente en el eje de las abscisas con paso unitario. Sea*

r_i : *probabilidad de quedarse en* E_i.

p_i : *probabilidad de pasar de* E_i *a* E_{i+1}.

q_i : *probabilidad de pasar de* E_i *a* E_{i-1}.

Supongamos que

$p_i > 0; q_i > 0; r_i \geq 0; \forall i \geq 1$

$p_i + q_i + r_i = 1; \forall i \geq 1$

$r_0 + p_0 = 1; r_0 \geq 0; p_0 \geq 0.$

Sea

$$X_n : \{\textit{Es la posición del borracho en el instante } t\}$$

Como la posición siguiente del borrachito solamente depende de la posición actual, $\{X_n\}$ *una cadena de Markov con espacio de estados* $E = \{E_0, E_1, E_2, ...\}$. *Las probabilidades de transición de estados son*

$p_{00} = P(X_n = 0/X_{n-1} = 0) = r_0$

$p_{01} = P(X_n = 1/X_{n-1} = 0) = p_0$

$p_{0j} = P(X_n = j/X_{n-1} = 0) = 0 \quad \forall j > 1$

$p_{10} = P(X_n = 0/X_{n-1} = 1) = q_1$

$p_{11} = P(X_n = 1/X_{n-1} = 2) = r_1$

$p_{12} = P(X_n = 2/X_{n-1} = 1) = p_1$

$p_{1j} = P(X_n = j/X_{n-1} = 2) = 0 \quad \forall j > 3$

$p_{20} = P(X_n = 0/X_{n-1} = 2) = 0$

$p_{21} = P(X_n = 1/X_{n-1} = 2) = q_2$

$p_{22} = P(X_n = 2/X_{n-1} = 1) = r_2$

$p_{23} = P(X_n = 3/X_{n-1} = 2) = p_2$

$$p_{2j} = P(X_n = j/X_{n-1} = 2) = p \quad \forall j > 4$$

y así sucesivamente. La matriz de transición de estados es

$$P = \begin{pmatrix} r_0 & p_0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

Ejemplo 1.3.4 (Vendedor) *Un vendedor vive en el pueblo $'a'$ y es responsable de los pueblos $'a','b'$ y $'c'$. Cada semana se requiere visitar un pueblo diferente. Cuando él esta en casa, es indiferente el pueblo que el visitará a la semana siguiente, por lo tanto lanza una moneda y si cae cara el visita al pueblo $'b'$ de lo contrario visita al pueblo $'c'$. Sin embargo, cuando el esta por fuera tiene cierta preferencia de ir a casa, por lo tanto, él lanza dos monedas, si las dos monedas caen cara el visita el otro pueblo que no está en su casa. Los sucesivos pueblos que el visita forman una cadena de Markov con espacio de estados $E = \{a, b, c\}$ donde la variable aleatoria X_n es igual a $'a','b'$ o $'c'$ de acuerdo a su ubicación en la semana n. Las probabilidades de transición de estados están dadas por*

$$p_{aa} = P(X_n = a/X_{n-1} = a) = 0$$

$$p_{ab} = P(X_n = b/X_{n-1} = a) = 0{,}5$$

$$p_{ac} = P(X_n = c/X_{n-1} = a) = 0{,}5$$

$$p_{ba} = P(X_n = a/X_{n-1} = b) = 0{,}75$$

$$p_{bb} = P(X_n = b/X_{n-1} = b) = 0$$

$$p_{bc} = P(X_n = c/X_{n-1} = b) = 0{,}25$$

$$p_{ca} = P(X_n = a/X_{n-1} = c) = 0{,}75$$

$$p_{cb} = P(X_n = b/X_{n-1} = c) = 0{,}25$$

$$p_{cc} = P(X_n = c/X_{n-1} = c) = 0$$

La matriz de transición está dada por

$$\begin{array}{c} \\ a \\ b \\ b \end{array} \begin{array}{c} \begin{array}{ccc} \quad a \quad & \quad b \quad & \quad c \quad \end{array} \\ \begin{pmatrix} 0 & 0{,}5 & 0{,}5 \\ 0{,}75 & 0 & 0{,}25 \\ 0{,}75 & 0{,}25 & 0 \end{pmatrix} \end{array}$$

Ejemplo 1.3.5 (Modelo de Colas en Tiempo Discreto.) *Supongamos que observamos el número de clientes que están en una cola delante de un cajero automático. Sea X_n el número de clientes en el momento $\in \{0, 1, ...\}$. Es decir, observamos el sistema en instantes de tiempo separados por una unidad de tiempo, por ejemplo cada hora o cada 10 minutos (una unidad de tiempo puede ser un número arbitrario de minutos), etc. Supongamos que si un cliente está usando el cajero automático en el momento n, entonces la probabilidad que este finalice antes del momento $n+1$ es $q \in (0, 1)$. Finalmente, suponga que la probabilidad de que k clientes lleguen durante una unidad de tiempo está dado por*

$$P(Y_n = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

para $k = 0, 1, 2, ...$ y que para cualquier $n \in \{0, 1, 2, ...\}$. Es decir el número Y_n de llegadas al sistema en el intervalo $[n, n+1)$ se distribuye Poisson con parámetro λ para todo n. Por lo tanto, tenemos que

$$X_n = \begin{cases} X_n + Y_n, & \textit{con probabilidad } p; \\ (X_n - 1) + Y_n, & \textit{con probabilidad } q. \end{cases}$$

Hemos encontrado que $X_n, n = 0, 1, 2, ...$ es una cadena de Markov cuyas probabilidades de transición son:

$$P[X_{n+1} = j/X_n = i] = \begin{cases} p\frac{e^{-\lambda}\lambda^k}{k!} + q\frac{e^{-\lambda}\lambda^{k+1}}{(k+1)!}, & \textit{si } j = i + k,\ i \in N; \\ qe^{-\lambda}, & \textit{si } j = i - 1,\ i \in N; \\ \frac{e^{-\lambda}\lambda^k}{k!}, & \textit{si } j = k,\ i = 0. \end{cases}$$

$j = i+k$ ocurre cuando no sale nadie y llegan k personas o cuando alguien sale y llegan $k+1$.

$$P[X_{n+1} = j/X_n = i] = pP(Y_n = k) + qP(Y_n = k+1) = p\frac{e^{-\lambda}\lambda^k}{k!} + q\frac{e^{-\lambda}\lambda^{k+1}}{(k+1)!}$$

$j = i-1$ no llega nadie y alguien se fue

$$P[X_{n+1} = j/X_n = i] = qP(Y_n = 0) = qe^{-\lambda}$$

$j = k$, $i = 0$ no había nadie en el sistema y llegaron k

$$P[X_{n+1} = j/X_n = i] = P(Y_n = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Ejemplo 1.3.6 *Este ejemplo es tomado de Parzen, ilustra que el parámetro n no se refiere necesariamente al tiempo. Considere una página de texto y represente las vocales por ceros y las consonantes por unos. Por lo tanto la página es una cadena de ceros y unos. Esto indica que la secuencia de vocales y consonantes en el lenguaje Samoan forman una cadena de Markov, donde una vocal siempre sigue una consonante y una vocal sigue otra vocal con probabilidad de 0.51. Por lo tanto, la secuencia de ceros y unos de una página de texto Samoan se comporta de acuerdo a la siguiente matriz de transición*

$$\begin{array}{c} \\ E_0 \\ E_1 \end{array} \begin{array}{c} \begin{array}{cc} E_0 & E_1 \end{array} \\ \begin{pmatrix} 0{,}51 & 0{,}49 \\ 1 & 0 \end{pmatrix} \end{array}$$

Teorema 1.3.1 *Una cadena de Markov esta completamente definida una vez que la matriz de probabilidad de transición y las probabilidades iniciales están especificadas. Es decir,*

- $P(X_0 = i_0, X_1 = i_1, \cdots X_{n-1} = i_{n-1}, X_n = i_n) = p_{i_0}p_{i_0i_1}\cdots p_{i_{n-2}i_{n-1}}p_{i_{n-1}i_n}$
- $\underline{\mathbf{P}}^{(n)} = \underline{\mathbf{P}}^{(0)}\mathbf{P}^n$

Para la primera parte se tiene que

$$
\begin{aligned}
P &= P(X_n = i_n/X_0 = i_0, X_1 = i_1, \cdots X_{n-1} = i_{n-1}) \\
&* \ P(X_0 = i_0, X_1 = i_1, \cdots X_{n-1} = i_{n-1}) \\
&= P(X_n = i_n/X_{n-1} = i_{n-1})P(X_0 = i_0, X_1 = i_1, \cdots X_{n-1} = i_{n-1}) \\
&= p_{i_{n-1},i_n} P(X_{n-1} = i_{n-1}/X_0 = i_0, X_1 = i_1, \cdots, X_{n-2} = i_{n-2}) \\
&* \ P(X_0 = i_0, X_1 = i_1, \cdots, X_{n-2} = i_{n-2}) \\
&= p_{i_{n-1},i_n} P(X_{n-1} = i_{n-1}/X_{n-2} = i_{n-2})P(X_0 = i_0, X_1 = i_1, \cdots, X_{n-2} = i_{n-2}) \\
&= p_{i_{n-1},i_n} p_{i_{n-2}i_{n-1}} P(X_0 = i_0, X_1 = i_1, \cdots, X_{n-2} = i_{n-2}) \\
&\vdots \\
&= p_{i_{n-1}i_n} p_{i_{n-2}i_{n-1}} \cdots p_{i_2i_1} p_{i_1i_0} p_{i_0}
\end{aligned}
$$

Para la segunda parte, recordando el teorema de probabilidad total $p_j^{(n)}$, probabilidad de estado, puede escribirse como

$$
\begin{aligned}
p_j^{(n)} &= P(X_n = j) = \sum_{k=0}^{\infty} P(X_n = j/X_{n-1} = k)P(X_{n-1} = k) \\
&= \sum_{k=0}^{\infty} P(X_{n-1} = k)P(X_n = j/X_{n-1} = k) \\
&= \sum_{k=0}^{\infty} p_k^{(n-1)} p_{kj} \\
&= \begin{pmatrix} p_0^{(n-1)} & p_0^{(n-1)} & \cdots \end{pmatrix} \begin{pmatrix} p_{0j} \\ p_{1j} \\ \vdots \end{pmatrix}
\end{aligned}
$$

Ahora, $p_j^{(n)}$ es la componente j del vector $p^{(n)}$, por lo tanto tomando en cuenta todas las componentes se tiene que el vector $p^{(n)}$ (vector de las probabilidades de estado) está dado por

$$\begin{aligned}
\underline{\mathbf{P}}^{(n)} &= \begin{pmatrix} p_0^{(n-1)} & p_0^{(n-1)} & \cdots \end{pmatrix} \begin{pmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \\
&= \underline{\mathbf{P}}^{(n-1)} * \mathbf{P} \\
&= \underline{\mathbf{P}}^{(n-2)} * \mathbf{P} * \mathbf{P} = \underline{\mathbf{P}}^{(n-2)} * \mathbf{P}^2 \\
&\vdots \\
&= \underline{\mathbf{P}}^{(0)} * \mathbf{P}^n
\end{aligned}$$

Ejemplo 1.3.7 *Suponga que se inspecciona una máquina la cual se puede encontrar en los siguientes estados*

$E_0 =$ *Funcionando*

$E_1 =$ *En reparación de inmediato*

Supongamos que

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} E_0 & E_1 \end{array} \\ \begin{array}{c} E_0 \\ E_1 \end{array} & \begin{pmatrix} 0{,}6 & 0{,}4 \\ 0{,}8 & 0{,}2 \end{pmatrix} \end{array}$$

Si hoy la máquina está funcionando ¿Cuál es la probabilidad de que pasado mañana este en reparación?
Es evidente que se esta interesado en obtener a $p_1^{(2)}$. *Para ello usamos el resultado obtenido anteriormente, es decir,* $p^{(n)} = p^{(0)} * P^n$ *para* $n = 2$. *El vector de probabilidades iniciales está dado por*

$$\underline{\mathbf{P}}^{(0)} = \begin{pmatrix} p_0^{(0)} & p_0^{(1)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

y $\mathbf{P}^2$ *sería*

$$\mathbf{P}^2 = \mathbf{P} * \mathbf{P} = \begin{pmatrix} 0{,}6 & 0{,}4 \\ 0{,}8 & 0{,}2 \end{pmatrix} \begin{pmatrix} 0{,}6 & 0{,}4 \\ 0{,}8 & 0{,}2 \end{pmatrix} = \begin{pmatrix} 0{,}68 & 0{,}32 \\ 0{,}64 & 0{,}36 \end{pmatrix}$$

Por lo tanto,

$$\underline{\mathbf{P}}^{(2)} = \underline{\mathbf{P}}^{(0)} * \mathbf{P}^2 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0{,}68 & 0{,}32 \\ 0{,}64 & 0{,}36 \end{pmatrix} = \begin{pmatrix} 0{,}68 & 0{,}32 \end{pmatrix}$$

Entonces $p_1^{(2)} = 0{,}32$

Luego de que una cadena de Markov ha sido formulada, hay muchas preguntas que podrían ser de interés. Por ejemplo, ¿En qué estado estará la cadena de Markov después de 5 pasos? ¿Qué porcentaje de tiempo se mantendrá en un estado dado? ¿A partir de un estado, será que la cadena no llegará a otro estado?

1.4. Probabilidad de Transición de n pasos

Las cadenas de Markov proveen información directa acerca de las probabilidades de transición de un paso y estas a su vez pueden ser usadas para calcular las probabilidades para transiciones de mas de un paso. Por ejemplo suponga que en el caso del vendedor se quiere calcular la probabilidad de que estando en el pueblo ”b” este en su casa en la segunda semana, esta no es una probabilidad de un paso pues los cambios de localidad del vendedor son semanalmente, por lo tanto, estamos en presencia del cálculo de una probabilidad de dos pasos. Para resolverlo veamos la siguiente definición.

Definición 1.4.1 (Probabilidad de Transición de n pasos) *Denotaremos con* $p_{ij}^{(n)}$ *la probabilidad de ir del estado* i *al estado* j *en exactamente* n *pasos. Es decir,* $p_{ij}^{(n)}$ *es la probabilidad condicional de entrar al estado* E_j *en el* $n-$*ésimo paso dado que estaba inicialmente en el estado* E_i*. Simbólicamente*

$$p_{ij}^{(n)} = P(X_{n+m} = j / X_m = i) \quad n, m \geq 0$$

donde

$$p_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

Ejemplo 1.4.1 *Volviendo al ejemplo del vendedor se quiere calcular* $P(X_2 = a/X_0 = b)$*. Al tomar todas las posibles combinaciones se obtiene la siguiente expresión*

$$
\begin{aligned}
P(X_2 = a/X_0 = b) &= P(X_1 = a/X_0 = b)P(X_2 = a/X_1 = a) \\
&+ P(X_1 = b/X_0 = b)P(X_2 = a/X_1 = b) \\
&+ P(X_1 = c/X_0 = b)P(X_2 = a/X_1 = c) \\
&= P(b,a)P(a,a) + P(b,b)P(b,a) + P(b,c)P(c,a) \\
&= (0,75)(0) + (0)(0,75) + (0,25)(0,75) = 0,1875
\end{aligned}
$$

Definición 1.4.2 (Matriz de Transición de n pasos (etapas)) *Sea* $p_{ij}^{(n)}$ *la probabilidad de transición de* n *pasos, entonces*

$$\mathbf{P}^{(n)} = \left[p_{ij}^{(n)}\right]$$

Teorema 1.4.1 (Ecuaciones de Chapman Kolmogorov) *Sea* $P^{(m+n)}$ *la matriz de transición de* $m+n$ *pasos, entonces la componente* (i,j) *de la matriz* $P^{(m+n)}$ *dada por*

$$p_{ij}^{(m+n)} = \sum_{k=0}^{\infty} p_{ik}^{(m)} p_{kj}^{(n)} \tag{1.4.1}$$

Prueba:

$$
\begin{aligned}
p_{ij}^{(m+n)} &= P(X_{m+n}=j/X_0=i)=\frac{P(X_{m+n}=j,X_0=i)}{P(X_0=i)}\\
&= \sum_{k=0}^{\infty}\frac{P(X_{m+n}=j,X_0=i,X_m=k)}{P(X_0=i)}\\
&= \sum_{k=0}^{\infty}\frac{P(X_{m+n}=j/X_0=i,X_m=k)P(X_m=k,X_0=i,)}{P(X_0=i)}\\
&= \sum_{k=0}^{\infty}\frac{P(X_{m+n}=j/X_m=k)P(X_m=k/X_0=i)P(X_0=i)}{P(X_0=i)}\\
&= \sum_{k=0}^{\infty}P(X_{m+n}=j/X_m=k)P(X_m=k/X_0=i)\\
&= \sum_{k=0}^{\infty}p_{ik}^{(m)}p_{kj}^{(n)}
\end{aligned}
$$

A partir de este teorema obtenemos el siguiente resultado matricial

Teorema 1.4.2 *Sea $P^{(h)}$ la matriz de transición de h pasos, entonces*

$$
\mathbf{P}^{(m+n)}=\mathbf{P}^{(m)}\mathbf{P}^{(n)} \tag{1.4.2}
$$

Prueba:

La prueba se basa en el resultado de las ecuaciones de Chapman Kolmogorov.

Teorema 1.4.3 *Sea $\mathbf{P}^{(n)}$ la matriz de probabilidad de transición de n pasos y $\mathbf{P}^n$ la n-esima potencia de la matriz de probabilidad de transición, entonces*

$$
\mathbf{P}^{(n)}=\mathbf{P}^n
$$

Prueba:*Por el teorema anterior*

$$
\begin{aligned}
\mathbf{P}^{(n)} &= \mathbf{P}^{(n-1)}\mathbf{P}^{(1)}\\
&= \mathbf{P}^{(n-2)}\mathbf{P}^{(1)}\mathbf{P}^{(1)}\\
&\vdots\\
&= \mathbf{P}^{(1)}\mathbf{P}^{(1)}\cdots\mathbf{P}^{(1)} \qquad \textit{n veces}\\
&= \mathbf{P}\mathbf{P}\cdots\mathbf{P}=\mathbf{P}^n
\end{aligned}
$$

Ejemplo 1.4.2 *Volviendo al ejemplo 1.3.7 en el cual se tiene los siguientes estados*

$$E_0 = \{Funcionando\}$$

$$E_1 = \{En\ reparación\ de\ inmediato\}$$

con la siguiente matriz de transición de estados

$$\mathbf{P} = \begin{array}{c} \\ E_0 \\ E_1 \end{array} \begin{array}{c} \begin{array}{cc} E_0 & E_1 \end{array} \\ \begin{pmatrix} 0{,}6 & 0{,}4 \\ 0{,}8 & 0{,}2 \end{pmatrix} \end{array}$$

Si quisieramos hallar todas las probabilidades del estado de la máquina en dos dias, tendríamos que calcular $\mathbf{P}^{(2)}$*, para ello sólo basta con calcular* $\mathbf{P}^2$*, el cual está dado por*

Si hoy la máquina está funcionando ¿Cuál es la probabilidad de que pasado mañana este en reparación?
Es evidente que se esta interesado en obtener a $p_1^{(2)}$*. Para ello usamos el resultado obtenido anteriormente, es decir,* $p^{(n)} = p^{(0)} * P^n$ *para* $n = 2$*. El vector de probabilidades iniciales está dado por*

$$\mathbf{P}^2 = \mathbf{P} * \mathbf{P} = \begin{pmatrix} 0{,}6 & 0{,}4 \\ 0{,}8 & 0{,}2 \end{pmatrix} \begin{pmatrix} 0{,}6 & 0{,}4 \\ 0{,}8 & 0{,}2 \end{pmatrix} = \begin{pmatrix} 0{,}68 & 0{,}32 \\ 0{,}64 & 0{,}36 \end{pmatrix}$$

Por ejemplo si quisieramos conocer la probabilidad de que la máquina este funcionado pasado mañana dado que hoy esta funcionando, es decir, $p_{00}^{(2)}$*, tomamos la celda correspondiente a la intersección de la fila 1 y la columna 1 de la matriz* $\mathbf{P}^2$*, el cual esta dado por* 0,68*. De igual manera, para conocer la probabilidad de que la máquina este funcionado pasado mañana dado que hoy esta reparación, es decir,* $p_{10}^{(2)}$*, tomamos la celda correspondiente a la intersección de la fila 2 y la columna 1 de la matriz* $\mathbf{P}^2$*, el cual esta dado por* 0,64.

Por lo general no es necesario calcular completamente $\mathbf{P}^n$ para obtener la probabilidad que queremos calcular, solo basta con multiplicar la fila de la primera matriz y la columna de la segunda matriz que se requieren para obtener el valor deseado. Para una ilustración, consideremos que se quiere calcular $p_{ij}^{(n)}$, entonces se tiene varias alternativas de cálculo, veamos algunas de ellas

- Se tiene que $\mathbf{P}^n = \mathbf{P}^{n-1}\mathbf{P}$ entonces se toma la i-esima fila de $\mathbf{P}^{n-1}$ y se multiplica por la j-esima columna de $\mathbf{P}$
- Se tiene que $\mathbf{P}^n = \mathbf{P}\mathbf{P}^{n-1}$ entonces se toma la i-esima fila de $\mathbf{P}$ y se multiplica por la j-esima columna de $\mathbf{P}^{n-1}$
- Se tiene que $\mathbf{P}^n = \mathbf{P}^3\mathbf{P}^{n-3}$ entonces se toma la i-esima fila de $\mathbf{P}^3$ y se multiplica por la j-esima columna de $\mathbf{P}^{n-3}$

Teorema 1.4.4 *Sea* $\underline{\mathbf{P}}^{(n+m)}$ *el vector de probabilidades de estado en el momento* $(n+m)$ *y* $\mathbf{P}$ *la matriz de transición de un paso, entonces*

$$\underline{\mathbf{P}}^{(n+m)} = \underline{\mathbf{P}}^{(n)}\mathbf{P}^m$$

Prueba:

$$\underline{\mathbf{P}}^{(n+m)} = \underline{\mathbf{P}}^{(0)}\mathbf{P}^{n+m} = \underline{\mathbf{P}}^{(0)}\mathbf{P}^{n}\mathbf{P}^{m} = \underline{\mathbf{P}}^{(n)}\mathbf{P}^m$$

Ejemplo 1.4.3 *En el ejemplo de la máquina considere que se tiene* $\underline{\mathbf{P}}^{(20)} = \begin{pmatrix} 0{,}5 & 0{,}5 \end{pmatrix}$ *y se quiere calcular* $\underline{\mathbf{P}}^{(22)}$*, entonces*

$$\underline{\mathbf{P}}^{(22)} = \underline{\mathbf{P}}^{(20)}\mathbf{P}^2 = \begin{pmatrix} 0{,}5 & 0{,}5 \end{pmatrix}\begin{pmatrix} 0{,}68 & 0{,}32 \\ 0{,}64 & 0{,}36 \end{pmatrix} = \begin{pmatrix} 0{,}66 & 0{,}34 \end{pmatrix}$$

Cómo obtener P^n

Supongamos que se puede encontrar k vectores característicos (vectores propios) linealmente independientes para la matriz P. sean $\mathbf{X}_i$ $(i = 1, 2, ..., k)$ vectores linealmente independientes de la matriz P. Llamemos H la matriz formada por los k vectores

$$H = \begin{pmatrix} X_1 & X_2 & ... & X_k \end{pmatrix}$$

Puesto que H es no singular entonces

$$H^{-1}PH = D = diag(\lambda_i * \delta ij)$$

además

$$HH^{-1}PH = HD \Rightarrow PH = HD \Rightarrow PHH^{-1} = HDH^{-1} \Rightarrow P = HDH^{-1}$$

En consecuencia

$$P^n = HD^nH^{-1}$$

ya que

$$P^2 = P * P = HDH^{-1}HDH^{-1} = HD^2H^{-1}$$

1.5. Costos o Beneficios Esperados

Las cadenas de Markov son utilizadas para analizar el costo o beneficio esperado de una operación y por lo tanto necesitamos considerar una función de costo o beneficio impuesta al proceso.

Por ejemplo, supóngase que toda la semana en el pueblo "a" resulta en un beneficio de 1000 unidades monetarias (UM), toda la semana en el pueblo "b" resulta en un beneficio de 1200 UM y toda la semana en el pueblo "c" resulta en un beneficio de 1250 UM. Entonces podríamos preguntarnos ¿Cuánto debería ser el beneficio esperado si inicialmente estamos en el pueblo "a"?. O de manera más general, ¿Qué beneficio debería esperarse en la n-esima semana si inicialmente estaba en el pueblo a?.

No debe ser difícil convencerse que los siguientes cálculos deben ser apropiados

$$E[\text{Beneficio para la semana} n] = 1000P_{aa}^{(n)} + 1200P_{ab}^{(n)} + 1250P_{ac}^{(n)}$$

Por lo tanto se tiene el siguiente resultado

Teorema 1.5.1 *Sea $\{X_n, n = 0, 1, 2, ...\}$ una cadena de Markov con espacio de estado E, matriz de transición P, vector de probabilidades iniciales $\underline{P}^{(0)}$ y función de beneficio B (es decir cada vez que se visite al estado i, un beneficio $B(i)$ es obtenido). El beneficio esperado en el n-esimo paso está dado por:*

$$E[B(X_n)/X_0 = i] = \underline{P}^{(0)}P^nB(i)$$

Por lo tanto en el ejemplo del vendedor, el beneficio esperado durante la segunda semana dado que estaba inicialmente en el pueblo "a" es

$$
\begin{aligned}
E[B(X_2)/X_0 = a] &= \underline{P}^{(0)} P^2 B(a) \\
&= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0{,}75 & 0{,}125 & 0{,}125 \\ 0{,}1875 & 0{,}4375 & 0{,}375 \\ 0{,}1875 & 0{,}375 & 0{,}4375 \end{pmatrix} \begin{pmatrix} 1000 \\ 1200 \\ 1250 \end{pmatrix} \\
&= \begin{pmatrix} 0{,}75 & 0{,}125 & 0{,}125 \end{pmatrix} \begin{pmatrix} 1000 \\ 1200 \\ 1250 \end{pmatrix} \\
&= 1000(0{,}75) + 1200(0{,}125) + 1250(0{,}125) = 1056{,}25
\end{aligned}
$$

Hasta ahora hemos visto que el estado inicial era conocido. Sin embargo no siempre es así. El gerente del vendedor puede no saber la ubicación del vendedor, por el contrario, sabe que hay una función de masa de probabilidad que describe su ubicación inicial.
Nuevamente usando el ejemplo del vendedor, supongamos que no se esta seguro de la ubicación inicial del vendedor pero se sabe que tiene un chance del 50 % de que este en el pueblo "a", un 30 % en el pueblo "b" y 20 % en el pueblo "c". Ahora preguntamos ¿Cuál es el beneficio del vendedor en la segunda semana?

$$
\begin{aligned}
E[B(X_2)] &= \underline{P}^{(0)} P^2 B(a) \\
&= \begin{pmatrix} 0{,}5 & 0{,}3 & 0{,}2 \end{pmatrix} \begin{pmatrix} 0{,}75 & 0{,}125 & 0{,}125 \\ 0{,}1875 & 0{,}4375 & 0{,}375 \\ 0{,}1875 & 0{,}375 & 0{,}4375 \end{pmatrix} \begin{pmatrix} 1000 \\ 1200 \\ 1250 \end{pmatrix} \\
&= 1119{,}375
\end{aligned}
$$

Ejemplo 1.5.1 (Tomado de Feldman) *Un análisis de mercado sobre el comportamiento de clientes de autos ha sido realizado. El cambio del estilo de los carros ha sido registrado con los siguientes resultados:*

Se asume que estos datos son representativos del comportamiento promedio de los clientes y se asume que los supuestos de Markov se cumplen.

Desarrollamos una Cadena de Markov para describir el cambio del tipo de carro de un cliente (se asume que la edad no influye). Definimos la Cadena

Tabla 1.1:

N° de clientes	Cambio de Tipo
275	Sedan por Sedan
180	Sedan por Camioneta
45	Sedan por Convertible
80	Camioneta por Sedan
120	Camioneta por Camioneta
150	Convertible por Sedan
50	Convertible por Convertible

de Markov como el tipo de carro que el cliente tiene inmediatamente después de cambiarlo, por lo tanto el espacio de estado es $E = \{s, w, c\}$ *donde* s *es para identificar el sedan,* w *para la camioneta y* c *para el convertible.*

- ***Determinación de la matriz de transición de estados.*** *Se observa que de 500 clientes 275 se quedarán con un sedan, por lo tanto el elemento* $s-s$ *de la matriz de probabilidades de transición será* 275/500. *Por lo tanto* P *será* $\mathbf{P}=$

$$\begin{array}{c c} & \begin{array}{ccc} s & w & c \end{array} \\ \begin{array}{c} s \\ w \\ c \end{array} & \begin{pmatrix} 275/500 & 180/500 & 45/500 \\ 80/200 & 120/200 & 0 \\ 150/200 & 0 & 50/200 \end{pmatrix} \end{array}$$

Asumamos que tenemos un cliente cuyo comportamiento es descrito por este modelo. Además, el cliente siempre compra un carro nuevo en enero de todos los años. Suponiendo que estamos comenzando enero de 1997, ¿cuál sería la probabilidad de que compre un sedan si actualmente tiene un sedan?. De acuerdo con la matriz anterior se tiene que

$$P(X_{1997} = s | X_{1996} = s) = 275/500 = 0{,}55$$

- *¿Cuál es la probabilidad de que un cliente pase esta año (1997) a un convertible y el próximo año a un sedan?. Es evidente que esta probabilidad esta dada por*

$$P_{sc}P_{cs} = 0{,}09 * 0{,}75 = 0{,}0675$$

- *Ahora queremos predecir el tipo de carro que tendrá comenzando enero del 2000. Note que la pregunta envuelve 3 transiciones de la cadena de Marrkov, por lo tanto se debe calcular $P^{(3)} = P^3$, haciendo los cálculos se tiene que* $\mathbf{P}^{(3)} =$

$$\begin{array}{c c} & \begin{array}{ccc} s & w & c \end{array} \\ \begin{array}{c} s \\ w \\ c \end{array} & \begin{pmatrix} 0{,}5023 & 0{,}4334 & 0{,}0643 \\ 0{,}4816 & 0{,}4680 & 0{,}0504 \\ 0{,}5355 & 0{,}3780 & 0{,}0865 \end{pmatrix} \end{array}$$

  Por lo tanto, hay aproximadamente un 50 % de chance de que el cliente tenga un sedan, 43 % una camioneta y casi 7 % un convertible

- *¿Cuál es la probabilidad de que un cliente quien comenzó 1997 con un seda, se mantenga con un sedan y en el año 200 tenga un sedan?. En este caso queremos calcular $PP(X_{1997} = s, X_{2000} = s|X_{1996} = s)$, para ello se tiene que*

$$\begin{aligned} P(X_{1997} = s, X_{2000} = s|X_{1996} = s) &= \frac{P(X_{1996} = s, X_{1997} = s, X_{2000} = s)}{P(X_{1996} = s)} \\ &= P(X_{2000} = s|X_{1997} = s)P(X_{1997} = s|X_{1996} = s) \\ &= p_{ss}^3 * p_{ss} = 0{,}28 \end{aligned}$$

  Note que no se hace mención del tipo de carro en los años intermedios, de este modo, el cliente pudo o no cambiar el carro en los años 1998 y 1999.

- *Suponiendo que un sedan deja un beneficio de 1200 UM, una camioneta 1500 UM y un convertible 2500 UM. ¿Cuál sería el beneficio esperado* *en el año 1999 quien comenzó con un sedan en 1997? En este caso calculamos*

$$
\begin{aligned}
E[B(X_{1999})|X_{1996}=s] &= \underline{\mathbf{P}}^{(0)}\mathbf{P}^{(3)}B(s) \\
&= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0{,}5023 & 0{,}4334 & 0{,}0643 \\ 0{,}4816 & 0{,}4680 & 0{,}0504 \\ 0{,}5355 & 0{,}3780 & 0{,}0865 \end{pmatrix} \begin{pmatrix} 1200 \\ 1500 \\ 2500 \end{pmatrix} \\
&= \begin{pmatrix} 0{,}5023 & 0{,}4334 & 0{,}0643 \end{pmatrix} \begin{pmatrix} 1200 \\ 1500 \\ 2500 \end{pmatrix} \\
&= 1416{,}61
\end{aligned}
$$

1.6. Comunicación

Para dar una tipología de las cadenas de Markov y de los estados presentes en dichas cadenas es necesario definir algunas terminologías, a continuación se define el concepto de comunicación que es necesario para definir cadenas cerradas y este último caracteriza las cadenas de Markov.

Definición 1.6.1 (Estado Accesible) *Un estado E_j se dice accesible desde un estado E_i si existe algún entero $n \geq 1$ tal que $p_{ij}^{(n)} > 0$*

Gráficamente esto se representa

$$E_i \longrightarrow E_j$$

Ejemplo 1.6.1 *Considere la siguiente cadena de Markov $\{X_n\}$ con matriz de probabilidad de transición* $\mathbf{P}=$

$$\begin{array}{c c} & \begin{array}{ccc} 0 & \quad 1 & \quad 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} 0{,}7 & 0{,}2 & 0{,}1 \\ 0 & 0{,}6 & 0{,}4 \\ 0{,}5 & 0 & 0{,}5 \end{pmatrix} \end{array}$$

Es evidente que el estado E_2 es accesible desde todos los estados pues $p_{i2}^{(n)} > 0$ para $n = 1$ y para todo $i = 0, 1, 2$. El estado E_0 es accesible desde el estado E_2 en $n = 1$ y accesible desde el estado E_1 para $n = 2$, ya que, $p_{10}^{(2)} = p_{12}P_{20} > 0$. El estado E_1 es accesible desde el Estado E_0 en $n = 1$ y accesible desde el estado E_2 para $n = 2$, ya que, $p_{21}^{(2)} = p_{20}P_{01} > 0$.

Definición 1.6.2 (Intercomunicación) *Dos estados E_i y E_j cada uno accesible al otro, se dice que ellos se intercomunican, si y sólo si*

$$E_i \longleftrightarrow E_j$$

Ejemplo 1.6.2 *Consideremos el ejemplo anterior, es evidente que todos los estados se intercomunican.*

Una vez teniendo claras estas dos definiciones es propicio representar gráficamente una cadena de Markov, ya que dicha representación se basa en esos conceptos.

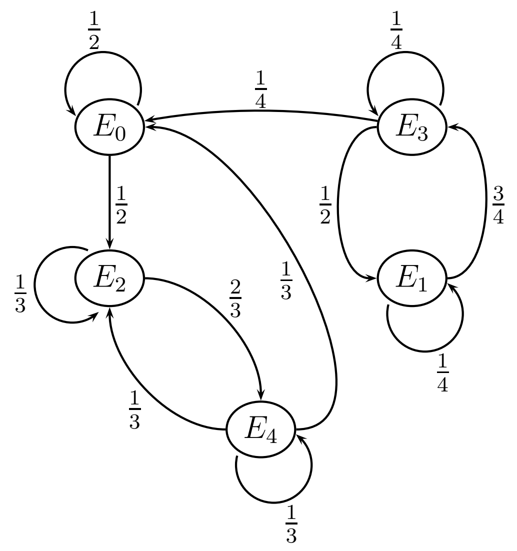
Representación Gráfica de una Cadena de Markov
La representación gráfica de una cadena de Markov es conocida como grafo, donde

- Los estados están representados por nodos.
- La probabilidad de comunicación entre los estados por flechas.

Ejemplo 1.6.3 *Sea* $\{X_t\}_{t\in T}$ *una cadena de Markov con matriz de probabilidad de transición dada por*

$$\begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array}\begin{array}{c} \begin{array}{ccccc} E_0 & E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 0 & 3/4 & 0 \\ 0 & 0 & 1/3 & 0 & 2/3 \\ 1/4 & 1/2 & 0 & 1/4 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix} \end{array}$$

La siguiente figura es la representación gráfica de una cadena de Markov. Los nodos representan los estados de la cadena y las flechas las probabilidades de transición.



Teorema 1.6.1 *El concepto de intercomunicación constituye una relación de equivalencia en el conjunto de estados, es decir,*

1. $E_i \leftrightarrow E_j$. *(Reflexiva)*

2. $E_i \leftrightarrow E_j$ *y* $E_j \leftrightarrow E_i$. *(Simétrica)*

3. $E_i \leftrightarrow E_j$ *y* $E_j \leftrightarrow E_k \Rightarrow E_i \leftrightarrow E_k$. *(Transitiva)*

Prueba:

1. *Es reflexiva, ya que* $E_i \leftrightarrow E_i$, $\forall i$, *al ser* $p_{ii}^{(0)} = 1 > 0$

2. *Es simétrica, ya que* $E_i \leftrightarrow E_j \Leftrightarrow \exists n, m / p_{ij}^{(n)} > 0$ *y* $p_{ji}^{(m)} > 0 \Leftrightarrow p_{ji}^{(m)} > 0$ *y* $p_{ji}^{(m)} > 0 \Leftrightarrow E_j \leftrightarrow E_i$.

3. *Es transitiva, ya que*

$$E_i \leftrightarrow E_j \Leftrightarrow \exists n_1, m_1 / p_{ij}^{(n_1)} > 0 \text{ y } p_{ji}^{(m_1)} > 0 \text{ y}$$
$$E_j \leftrightarrow E_k \Leftrightarrow \exists n_2, m_2 / p_{jk}^{(n_2)} > 0 \text{ y } p_{kj}^{(m_2)} > 0$$

Por lo tanto, $\exists\ n = n1 + n2$ *y* $m = m_1 + m_2$ *tal que*

$$p_{ik}^{(n)} = \sum_{r=0}^{\infty} p_{ir}^{(n_1)} p_{rk}^{(n_2)} \geq p_{ij}^{(n_1)} p_{jk}^{(n_2)} > 0 \text{ y}$$
$$p_{ki}^{(n)} = \sum_{r=0}^{\infty} p_{kr}^{(m_1)} p_{ri}^{(m_2)} \geq p_{kj}^{(m_1)} p_{ji}^{(m_2)} > 0$$

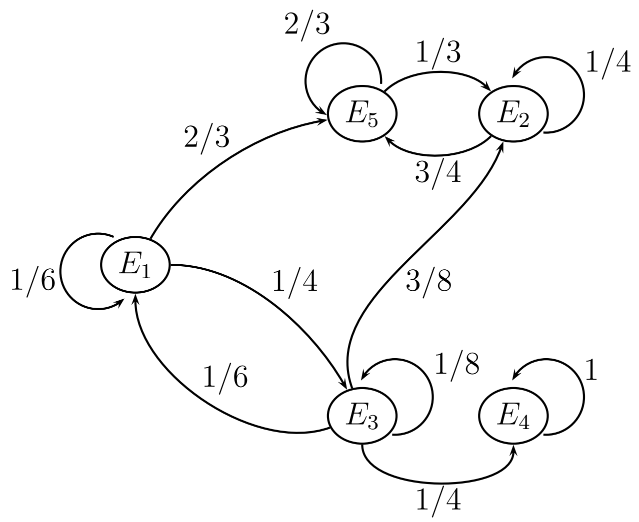
Entonces $E_i \leftrightarrow E_k$

Por lo tanto, como la relación de comunicación es de equivalencia, entonces se puede hacer una partición de los estados en clases de equivalencia. Los estados en una clase de equivalencia son aquellos que se comunican unos con otros. Puede ser posible, comenzar en una clase, entrar a alguna otra clase con probabilidad positiva; de ser así, es evidente que no es posible regresar a la clase inicial, o de lo contrario las dos clases deben juntas formar una sola clase.
Es decir, estando en una clase de equivalencia, se deja para ir a otra clase con una probabilidad positiva pero sin regresar a ella.

Ejemplo 1.6.4 *Sea $\{X_t\}_{t\in T}$ una cadena de Markov con la siguiente matriz de transición*

$$\begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array}
\begin{array}{c} \begin{array}{ccccc} E_0 & E_1 & E_2 & E_3 & E_4 \end{array} \\
\begin{pmatrix} 1/6 & 0 & 1/6 & 0 & 2/3 \\ 0 & 1/4 & 0 & 0 & 3/4 \\ 1/4 & 3/8 & 1/8 & 1/4 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1/3 & 0 & 0 & 2/3 \end{pmatrix} \end{array}$$

El grafo de la cadena se muestra en la siguiente figura



Note que se tienen las siguientes clases:

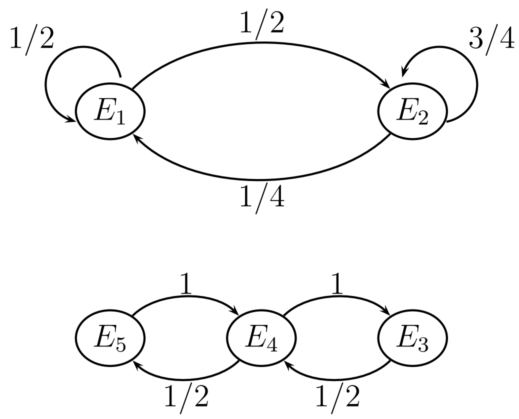
1. $C_1 = \{E_1, E_3\}$
2. $C_2 = \{E_2, E_5\}$
3. $C_3 = \{E_4\}$

Puesto que de la clase 1 se puede ir a la clase 2 y a la clase 3, a la clase 1 se le da el nombre de transición. La clase 2 es una clase cerrada. Las clases 2 y 3 son conocidas como ergógicas

Ejemplo 1.6.5 *Sea $\{X_t\}_{t\in T}$ una cadena de Markov con la siguiente matriz de transición*

$$\begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \\ E_5 \end{array} \begin{array}{c} \begin{array}{ccccc} E_1 & E_2 & E_3 & E_4 & E_5 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/4 & 3/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

El grafo de la cadena se muestra en la siguiente figura



La cadena tiene 2 clases

1. $C_1 = \{E_1, E_2\}$
2. $C_2 = \{E_3, E_4, E_5\}$

Definición 1.6.3 (Conjunto de Estados Cerrados) *Se dice que el conjunto $\mathbb{C}$ es un conjunto de estados cerrado si no es posible pasar de un estado del conjunto $\mathbb{C}$ a otro estado fuera de $\mathbb{C}$, es decir,*

$$p_{ij} = 0 \quad \forall \quad E_i \in \mathbb{C} \quad y \quad E_j \notin \mathbb{C}$$

Los conjuntos cerrados tienen la siguiente propiedad:

Para todo $n > 0$ y para todo $i \in \mathbb{C}$, se tiene $\sum_{j\in\mathbb{C}}^{n} p_{ij}^{(n)} = 1$

Es decir, después de n transiciones, el sistema se encuentra siempre en $\mathbb{C}$. Esta propiedad permite constatar que un conjunto cerrado constituye una **cadena independiente**. A partir de la definición de conjuntos cerrados se puede establecer una relación que estos tienen con los **estados absorbentes**y dar la definición de **cadenas irreducibles**.

Definición 1.6.4 (Relación entre conjuntos cerrados y estados absorbentes) *Si un estado E_i constituye un conjunto cerrado entonces se dice que el **estado es absorbente**, y en este caso ocurre que $p_{ii}^{(n)} = 1$ para todo n.*

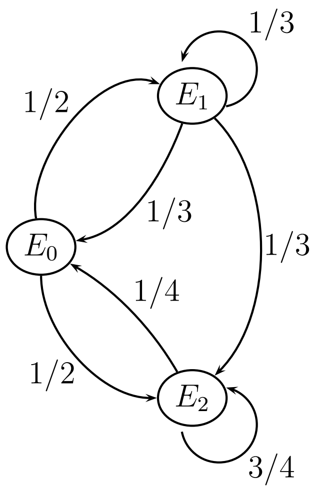
Es decir si un conjunto cerrado contiene un solo estado, dicho estado es un **estado absorbente**

Definición 1.6.5 (Cadena Irreducible) *Es una cadena en la que todos los estados se comunica entre sí, es decir, la cadena no contiene ningún subconjunto cerrado excepto el conjunto de todos sus estados.*

Ejemplo 1.6.6 *Sea $\{X_t\}_{t\in T}$ una cadena de Markov con la siguiente matriz de transición*

$$\begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \end{array} \begin{array}{c} \begin{array}{ccc} E_0 & E_1 & E_2 \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 1/3 \\ 1/4 & 0 & 3/4 \end{pmatrix} \end{array}$$

El grafo de la cadena se muestra en la siguiente figura



Contiene una sola clase $C = \{E_0, E_1, E_2\}$, por lo tanto, la cadena es irreducible.

Ejercicios

1. Demostrar la propiedad de los conjuntos cerrados, es decir

   Para todo $n > 0$ y para todo $i \in \mathbb{C}$, se tiene $\sum_{j\in\mathbb{C}}^{n} p_{ij}^{(n)} = 1$

2. Se consideran cuatro cadenas de Markov, cada una con espacio de estado $E = \{E_1, E_2, E_3, E_4\}$ y con matrices de transición respectivas P_1, P_2, P_3, P_4 dadas por

$$P_1 = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \end{array} \qquad P_2 = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 \end{pmatrix} \end{array}$$

$$P_3 = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} \end{array} \qquad P_4 = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \end{array}$$

Determinar las clases de estas cadenas.

3. Determinar las clases de la cadena de Markov que admite la siguiente matriz de transición

$$P_3 = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \\ E_5 \end{array} \begin{array}{c} \begin{array}{ccccc} E_1 & E_2 & E_3 & E_4 & E_5 \end{array} \\ \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 0 & 3/4 & 0 \\ 0 & 0 & 1/3 & 0 & 2/3 \\ 1/4 & 1/2 & 0 & 1/4 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix} \end{array}$$

1.7. Periodicidad de una Cadena de Markov

Definición 1.7.1 (Período de un Estado (Periodicidad)) *El período de retorno "i" al estado E_i, representado por $d(i)$, es el máximo común divisor del conjunto de enteros $n \geq 1$ para el cual se cumple que $p_{ii}^{(n)} > 0$. Simbólicamente*

$$d(i) = mcd\{n : p_{ii}^{(n)} > 0\}$$

De la definición se puede extraer que

1. Si n es diferente de un múltiplo de $d(i)$, entonces $p_{ii}^{(n)} = 0$.

2. El estado i se reproduce únicamente después de un número de transiciones múltiples de $d(i)$.

3. Si $p_{ii} > 0$, entonces $d(i) = 1$ y se dice que E_i es aperiódica.

Recordemos que el mcd de por ejemplo 48 y 60 es

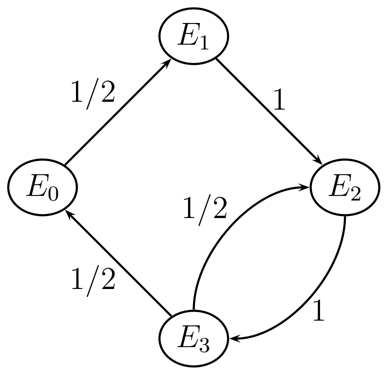
$$mcd\{48, 60\} = 2^2 * 3 = 12$$

De acuerdo a lo anterior, el retorno al estado inicial se produce en las fechas $r_i + sd(i)$, $s = 0, 1, 2, ...$, donde r_i en el mínimo número de etapas para un primer retorno.

Ejemplo 1.7.1 *Considere la siguiente cadena de Markov*

$$\begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{cccc} E_0 & E_1 & E_2 & E_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \end{array}$$

El grafo de la cadena se muestra en la siguiente figura



Se quiere determinar el período de E_0. *Para ello veamos de que manera partiendo de* E_0 *puedo volver a* E_0

$E_0 \to E_1 \to E_2 \to E_3 \to E_0 \Rightarrow n = 4$

$E_0 \to E_1 \to E_2 \to E_3 \to E_2 \to E_3 \to E_0 \Rightarrow n = 6$

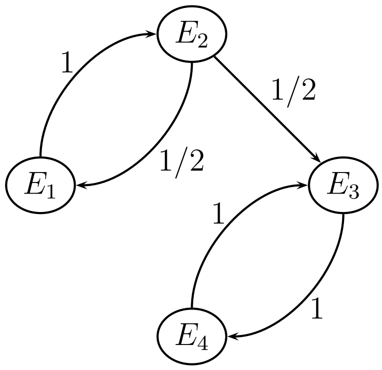
Luego $n = 8, 10, ...$

Por lo tanto, $d(0) = mcd\{4, 6, 8, 10, ...\} = 2$

Ejemplo 1.7.2 *Considere la siguiente cadena de Markov*

$$\begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

El grafo de la cadena se muestra en la siguiente figura



Se quiere determinar el período de E_1. *Para ello veamos de que manera partiendo de* E_1 *puedo volver a* E_1

$E_1 \rightarrow E_2 \rightarrow E_1 \Rightarrow n = 2$

$E_1 \rightarrow E_2 \rightarrow E_1 \rightarrow E_2 \rightarrow E_1 \Rightarrow n = 4$

Luego $n = 6, 8, 10, ...$

Por lo tanto, $d(0) = mcd\{2, 4, 6, 8, 10, ...\} = 2$

A continuación se enuncian sin demostración tres propiedades básicas de las cadenas de Markov.

Teorema 1.7.1 *Todos los estados de una clase de equivalencia tienen el mismo período, es decir si* $i \leftrightarrow j$ *entonces* $d(i) = d(j)$

Este teorema muestra que el período es una constante en cada clase de estados que se comunican.

Teorema 1.7.2 *Si el estado* i *tiene período* $d(i)$ *entonces existe un entero* N *que depende de* i *tal que para todos los enteros* $n \geq N$

$$P_{ii}^{nd(i)} > 0$$

Este teorema asegura que un retorno al estado i puede ocurrir en numeros suficientes grandes que son múltiplos del periodo $d(i)$

Teorema 1.7.3 *Si* $p_{ji}^{m} > 0$*, entonces* $p_{ji}^{m+nd(i)}$ *para todo* n *(entero positivo) suficientemente grande.*

Ejercicios de Período

1. Demostrar que todos los estados de una clase de equivalencia tienen el mismo período.

2. Determinar el período de los estados para las siguientes cadenas de markov

   a)

$$\begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \end{array} \begin{array}{c} \begin{array}{ccc} E_0 & E_1 & E_2 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

   b)

$$\begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \end{array}$$

3. Determinar el período de los estados de la siguiente cadena de Markov

$$\begin{array}{c} \\ E_0 \\ E_1 \\ \vdots \\ E_{n-1} \\ E_n \end{array} \begin{array}{c} \begin{array}{ccccc} E_0 & E_1 & E_2 & \cdots & E_n \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix} \end{array}$$

1.8. Primera Visita

Para dar una clasificación de los estados de Markov es necesario definir algunos términos relacionados con las probabilidades y el número de visitas a los estados de una cadena de Markov. Veamos a continuación tales términos.

Definición 1.8.1 (Tiempo de primera visita) *Sea A un subconjunto del espacio de estados de una cadena de Markov $\{X_n : n \geq 0\}$. El tiempo de primera visita al conjunto A es la variable aleatoria*

$$\tau_A = \begin{cases} \text{mín}\{n \geq 1 : X_n \in A\}, & \textit{si } X_n \in A \textit{ para algún } n \geq 1; \\ \infty, & \textit{otro caso.} \end{cases}$$

Es decir, τ_A es el primer momento positivo en el cual la cadena toma un valor dentro de la colección de estados A, si ello eventualmente sucede. Estaremos interesados principalmente en el caso cuando el conjunto A consta de un solo estado j, y si suponemos que la cadena inicia en i, entonces el tiempo de primera visita al estado j se escribe τ_{ij} . Cuando los estados i y j coinciden se escribe simplemente τ_i. En general no es fácil encontrar la distribución de probabilidad de la variable aleatoria τ_{ij}. Definiremos a continuación como $f_{ij}^{(n)}$ la probabilidad de que τ_{ij} tome el valor n.

Definición 1.8.2 (Probabilidad de la primera visita) *Sea $f_{ij}^{(n)}$ la probabilidad de que estando inicialmente el sistema en el estado i, visite por primera vez el estado j después de n transiciones, es decir*

$$f_{ij}^{(n)} = P(\tau_{ij} = n) = P\left(X_n = j, X_k \neq j \forall k = 1, 2, ..., n-1 / X_0 = i\right)$$

Si $i = j$, entonces

$$f_{jj}^{(n)} = P\left(X_n = j, X_k \neq j \forall k = 1, 2, ..., n-1 / X_0 = j\right)$$

$f_{jj}^{(n)}$ se define como la **probabilidad de retorno por primera vez** y se lee como la probabilidad de que, comenzando en el estado j, retorne por primera vez al estado j en la n-ésima transición.
Note que

1. $f_{ij}^{(0)} = 0$

2. $f_{jj}^{(0)} = 1$

3. $f_{jj}^{(1)} = p_{jj} = P(X_n = j/X_{n-1} = j)$

El uso de la letra f para esta probabilidad proviene del término en inglés *first* para indicar que es la probabilidad de primera visita; saber esto ayuda a recordar su significado.
El siguiente resultado establece que la probabilidad de visitar el estado j, a partir de i, en n pasos, puede descomponerse en las probabilidades de los eventos disjuntos en los que se presenta la primera visita, la cual puede efectuarse en el primer paso, o en el segundo paso, y así sucesivamente, hasta el último momento posible n.

Teorema 1.8.1 (Teorema de la primera visita) *Sea $i \neq j$ entonces para cada $n \geq 1$*

$$p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}$$

Prueba:

$$\begin{aligned}
p_{ij}^{(n)} &= P(X_n = j/X_0 = i) \\
&= P(X_n = j, X_1 \neq j/X_0 = i) + P(X_n = j, X_2 = j, X_1 \neq j/X_0 = i) \\
&+ P(X_n = j, X_3 = j, X_2 \neq j, X_1 \neq j/X_0 = i) + \cdots \\
&= \sum_{k=1}^{n} P(X_n = j/X_k = j) P(X_k = j, X_{k-1} \neq j, \cdots, X_1 \neq j/X_0 = i) \\
&= \sum_{k=1}^{n} P_{jj}^{(n-k)} f_{ij}^{(k)}
\end{aligned}$$

Si $i = j$ entonces

$$p_{jj}^{(n)} = \sum_{k=0}^{n} f_{jj}^{(k)} p_{jj}^{(n-k)}, \quad n \geq 1, f_{jj}^{(0)} = 0$$

El cual será definido como el **teorema del primer retorno.** Recordemos que

1. $p_{jj}^{(n)} = P(X_{m+n} = j/X_m = j)$

2. $p_{jj}^{(n)} = \sum_{k=0}^{n} P_{jj}^{(n-k)} f_{jj}^{(k)} = \sum_{k=1}^{n} P_{jj}^{(n-k)} f_{jj}^{(k)}$

Teorema 1.8.2

$$f_{jj}^{(n)} = p_{jj}^{(n)} - \sum_{k=1}^{n-1} f_{jj}^{(k)} p_{jj}^{(n-k)}$$

Definición 1.8.3 *Sea*

$$f_j = \sum_{n=1}^{\infty} f_{jj}^{(n)} = f_{jj}^{(1)} + f_{jj}^{(2)} + \cdots$$

f_j *se define como* ***la probabilidad de que alguna vez el estado*** j ***sea visitado.***
Si $f_j = 1$ *entonces* $\{f_{jj}^{(n)}\}$ *constituye una distribución de probabilidad*

1.9. Recurrencia y Transitoriedad

Veremos a continuación que los estados de una cadena de Markov pueden ser clasificados, en una primera instancia, en dos tipos, dependiendo si la cadena es capaz de regresar con certeza al estado de partida.

Definición 1.9.1 (Estado recurrente o persistente) *El estado* E_j *se dice recurrente o persistente si la probabilidad de eventualmente regresar a* E_j*, partiendo de* E_j *es 1, es decir si* $f_j = 1$*. El sistema, saliendo de* E_j*, pasará otra vez por el estado* E_j*, durante su evolución.*

Definición 1.9.2 (Estado transitorio) *El estado* E_j *se dice transitorio si y sólo si* $f_j < 1$*. El sistema, saliendo de* E_j*, puede (tiene "chance") de no retornar al estado* E_j*.*

Es decir que un estado que no es recurrente se llama transitorio, y en ese caso la probabilidad es menor a uno.
De manera intuitiva, un estado es recurrente si con probabilidad uno la cadena es capaz de regresar eventualmente a ese estado, y cuando ello ocurre en algún momento finito, por la propiedad de Markov, se puede regresar a él una y otra vez con probabilidad uno. Debido a este comportamiento es que

al estado en cuestión se le llama recurrente. En cambio, el estado se llama transitorio si existe una probabilidad positiva de que la cadena, iniciando en él, ya no regrese nunca a ese estado.

Ejemplo 1.9.1 *Considere una cadena de Markov con espacio de estados* $\{0, 1\}$ *y con matriz de transición*

$$\begin{array}{c c} & \begin{array}{cc} E_0 & E_1 \end{array} \\ \begin{array}{c} E_0 \\ E_1 \end{array} & \begin{pmatrix} 1-a & 1 \\ b & 1-b \end{pmatrix} \end{array}$$

donde $0 \leq a, b \leq 1$*. En esta cadena ambos estados son recurrentes pues:*

- *Para el estado* 0*:*

$$\begin{aligned} f_{00}^{(1)} &= 1-a \\ f_{00}^{(2)} &= ab \\ f_{00}^{(3)} &= a(1-b)b = ab(1-b) \\ f_{00}^{(4)} &= a(1-b)(1-b)b = ab(1-b)^2 \\ \vdots & \quad \vdots \quad \vdots \\ f_{00}^{(n)} &= ab(1-b)^{n-2} \end{aligned}$$

Por lo tanto,

$$f_0 = \sum_{n=1}^{\infty} f_{00}^{(n)} = (1-a) + ab\sum_{n=2}^{\infty} ab(1-b)^{n-2} = 1 - a + \frac{ab}{b} = 1$$

- *Se deja al estudiante como tarea probar la recurrencia del estado 1*

Ejemplo 1.9.2 *Considere la matrix de transición*

$$\begin{array}{c c} & \begin{array}{ccccc} E_0 & E_1 & E_2 & E_3 & \cdots \end{array} \\ \begin{array}{c} E_0 \\ E_1 \\ E_2 \\ \vdots \end{array} & \begin{pmatrix} q & p & 0 & 0 & \cdots \\ q & 0 & p & 0 & \cdots \\ q & 0 & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \ddots & \end{pmatrix} \end{array}$$

la cual corresponde a una cadena de rachas. En esta cadena todos los estados son recurrentes. Solo analizaremos el estado 0

- *Para el estado* 0*:*

$$
\begin{array}{rcl}
f_{00}^{(1)} & = & q \\
f_{00}^{(2)} & = & pq \\
f_{00}^{(3)} & = & ppq = qp^2 \\
f_{00}^{(4)} & = & pppq = qp^3 \\
\vdots & \vdots & \vdots \\
f_{00}^{(n)} & = & qp^{n-1}
\end{array}
$$

  Por lo tanto,

$$f_0 = \sum_{n=1}^{\infty} f_{00}^{(n)} = qp^{n-1} = \frac{q}{1-p} = 1$$

- *Se deja al estudiante como tarea probar la recurrencia del estado 1*

Además de la definición, tenemos el siguiente criterio para determinar si un estado es recurrente o transitorio.

Proposición 1.9.1 *El estado* E_j *es*

a) recurrente si y sólo si $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$

b) transitorio si y sólo si $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$

La recurrencia y la transitoriedad son propiedades de clase, es decir, si dos estados están en una misma clase de comunicación, entonces ambos estados son recurrentes o ambos son transitorios.

Corolario 1.9.0.1 *La recurrencia es una propiedad de clase, es decir*

a) Si $i \leftrightarrow j$ *y si* i *es recurrente entonces* j *es recurrente.*

b) Si $i \leftrightarrow j$ y si i es transitorio entonces j es transitorio.

En consecuencia, cuando una cadena es irreducible y algún estado es recurrente, todos los estados lo son, y se dice que la cadena es recurrente. También puede presentarse la situación en donde el espacio de estados conste de varias clases de comunicación recurrentes, en tal caso la cadena también se llama recurrente. En contraparte, una cadena es transitoria si todos los estados lo son, ya sea conformando una sola clase de comunicación de estados transitorios o varias de ellas. Sin embargo, demostraremos más adelante que cuando el espacio de estados es finito, siempre existe por lo menos un estado recurrente, y por lo tanto no existen cadenas finitas transitorias. De este modo el espacio de estados de una cadena de Markov puede descomponerse en dos subconjuntos ajenos de estados, aquellos que son transitorios y aquellos que son recurrentes. Cada uno de estos subconjuntos puede constar de ninguna, una o varias clases de comunicación.

Proposición 1.9.2 *Toda cadena de Markov finita tiene por lo menos un estado recurrente*

Prueba: *Por contradicción, suponga que todos los estados son transitorios. Entonces para cualesquiera estados i y j, se cumple que la suma $\sum_{n=1}^{\infty} p_{ij}^{(n)}$ es finita. Sumando sobre el conjunto finito de todos los posibles estados j se obtiene*

$$\sum_{j} \sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$$

Por otro lado, intercambiando el orden de las sumas se llega a la afirmación contraria,

$$\sum_{n=1}^{\infty} \sum_{j} p_{ij}^{(n)} = \sum_{n=1}^{\infty} 1 = \infty$$

Por lo tanto es erróneo suponer que todos los estados son transitorios, debe existir por lo menos uno que es recurrente.

En consecuencia, toda cadena finita e irreducible es recurrente. Más adelante demostraremos que en tal caso, con probabilidad uno la cadena visita cada uno de sus estados una infinidad de veces.

Definición 1.9.3 (Función generatriz de la primera visita) *La función generatriz de la sucesión* $\{f_{ij}^{(n)}\}$ *está dada por*

$$F_{ij}(s) = \sum_{n=0}^{\infty} f_{ij}^{(n)} s^n$$

Teorema 1.9.1 *Sea* $P_{ij}(s)$ *la función generatriz de las probabilidades de transición de* E_i *a* E_j, $p_{ij}^{(n)}$, *y sea* $F_{ij}(s)$ *la función generatriz de la probabilidad de la primera visita,* $f_{ij}^{(n)}$, *entonces*

$$F_{ij}(s) = \frac{P_{ij}(s)}{1 + P_{ij}(s)}$$

Prueba: *Del teorema 1.8.2 se tiene que*

$$f_{ij}^{(n)} = p_{ij}^{(n)} - \sum_{k=1}^{n-1} f_{ij}^{(k)} p_{ij}^{(n-k)}$$

Multiplicando por s^n *y aplicando sumatoria*

$$\sum_{n=0}^{\infty} f_{ij}^{(n)} s^n = \sum_{n=0}^{\infty} p_{ij}^{(n)} s^n - \sum_{n=0}^{\infty} \sum_{k=1}^{n-1} f_{ij}^{(k)} p_{ij}^{(n-k)} s^n$$

Como

$\sum_{n=0}^{\infty} f_{ij}^{(n)} s^n = F_{ij}(s)$

$\sum_{n=0}^{\infty} p_{ij}^{(n)} s^n = P_{ij}(s)$

$\sum_{n=0}^{\infty} \sum_{k=1}^{n-1} f_{ij}^{(k)} p_{ij}^{(n-k)} s^n = F_{ij}(s) P_{ij}(s)$

se tiene que

$$F_{ij}(s) = P_{ij}(s) - F_{ij}(s) P_{ij}(s)$$

Agrupando $F_{jj}(s)$

$$F_{ij}(s) + F_{ij}(s) P_{ij}(s) = P_{ij}(s)$$

Sacando factor común y despejando se obtiene

$$F_{ij}(s) = \frac{P_{ij}(s)}{1 + P_{ij}(s)}$$

Cuando $i = j$ se tiene que

$$F_{jj}(s) = \frac{P_{jj}(s)}{1 + P_{jj}(s)}$$

la cual es la **función generatriz de la probabilidad del primer retorno**.

Teorema 1.9.2 *Un estado es recurrente o transiente según la probabilidad de regreso a él en un número infinito de veces es uno o cero.*

Prueba: *Sean*

$$Q_{jj}^{(n)} = P\{(Estando\ el\ sistema\ en\ E_j\ regrese\ a\ E_j\ n\ veces\ )\}$$

$$\begin{aligned} Q_{jj} &= \lim_{n\to\infty} Q_{jj}^{(n)} \\ &= P\{(Estando\ el\ sistema\ en\ E_j\ regrese\ a\ E_j\ un\ número\ infinito\ de\ veces\ )\} \end{aligned}$$

Ahora $Q_{jj}^{(n)}$ *se puede escribir como*

$$\begin{aligned} Q_{jj}^{(n)} &= \sum_{n=1}^{\infty} f_{jj}^{(n)} Q_{jj}^{(n-1)} = Q_{jj}^{(n-1)} \sum_{n=1}^{\infty} f_{jj}^{(n)} \\ &= Q_{jj}^{(n-1)} f_j \end{aligned}$$

Por recurrencia se tiene que

$$Q_{jj}^{(n)} = Q_{jj}^{(n-2)} f_j f_j = Q_{jj}^{(n-2)} (f_j)^2$$

y así sucesivamente, con lo cual se obtiene que

$$Q_{jj}^{(n)} = Q_{jj}^{(1)} (f_j)^{n-1} = (f_j)^n$$

ya que $Q_{jj}^{(1)} = f_j$ *por definición. Por lo tanto*

$$Q_{jj} = \lim_{n\to\infty} Q_{jj}^{(n)} = \lim_{n\to\infty} (f_j)^n = \begin{cases} 1, & si\ E_j\ es\ recurrente; \\ 0, & si\ E_j\ es\ transiente. \end{cases}$$

1.9.1. Tiempo medio de Recurrencia

Hemos visto que si una cadena de Markov inicia en un estado recurrente,entonces regresa a él una infinidad de veces con probabilidad uno. Y hemos definido el tiempo de primera visita a un estado j, a partir de cualquier estado i, como la variable aleatoria discreta $\tau_{ij} = \text{mín}\{n \geq 1 : X_n = j | X_0 = i\}$, con la posibilidad de que tome el valor infinito. Vamos a definir el tiempo medio de recurrencia como la esperanza de esta variable aleatoria en el caso cuando el estado a visitar es recurrente.

Definición 1.9.4 (Tiempo medio de recurrencia) *El tiempo medio de recurrencia de un estado recurrente j, a partir del estado i, se define como la esperaza de τ_{ij}, y se denota por μ_{ij}, es decir*

$$\mu_{ij} = E(\tau_{ij}) = \sum_{n=1}^{\infty} n f_{ij}^{(n)}$$

Recordemos que cuando el tiempo de primera visita se refiere al mismo estado de inicio y de llegada j, se escribe τ_j en lugar de τ_{ij}. En este caso el tiempo medio de recurrencia se escribe simplemente como μ_j . Esta esperanza puede ser finita o infinita, y representa el número de pasos promedio que a la cadena le toma regresar al estado recurrente j.

Ejemplo 1.9.3 *La cadena de Markov de dos estados es irreducible y recurrente cuando $a, b \in (0, 1)$. Vamos a calcular los tiempos medios de recurrencia de estos dos estados. Tenemos que*

$$f_{00}^{(1)} = 1 - a$$

$$f_{00}^{(n)} = a(1 - b)^{n-2} b \text{ para } n \geq 2$$

Por lo tanto,

$$\begin{aligned}
\mu_0 = \sum_{n=1}^{\infty} n f_{00}^{(n)} &= (1 - a) + ab \sum_{n=2}^{\infty} (1 - b)^{n-2} \\
&= (1 - a) + ab \left(\frac{b + 1}{b^2} \right) \\
&= \frac{a + b}{b}
\end{aligned}$$

De manera análoga, se encuentra que $\mu_1 = (a+b)/a$. Observe que estos dos tiempos medios de recurrencia son, en general distintos. Esto ejemplifica ell hecho de que los tiempos medios de recurrencia no son necesariamente idénticos para cada elemento de una misma clase de comunicación recurrente.

1.9.2. Recurrencia Positiva y Nula

Hemos visto que si una cadena de Markov inicia en un estado recurrente, entonces regresa a él una infinidad de veces con probabilidad uno. Sin embargo, esta recurrencia puede presentarse de dos formas: cuando el tiempo promedio de retorno es finito o cuando es infinito. Por lo tanto el tiempo medio de retorno μ_j lleva a la siguiente clasificación de estados recurrentes.

Definición 1.9.5 *Se dice que un estado recurrente j es:*

- *recurrente positivo si $\mu_j < \infty$*
- *recurrente nulo si $\mu_j = \infty$*

Demostraremos a continuación que la recurrencia positiva y la recurrencia nula son propiedades de las clases de comunicación. Es decir, dos estados en una misma clase de comunicación recurrente, son ambos recurrentes positivos o recurrente nulos.

Proposición 1.9.3 *Sea j un estado recurrente. Entonces,*

a) si j es recurrente positivo e $j \leftrightarrow i$ entonces i es recurrente positivo.

1. si j es recurrente nulo e $j \leftrightarrow i$ entonces i es recurrente nulo.

De esta forma, el espacio de estados de toda cadena de Markov puede descomponerse en tres grandes subconjuntos ajenos de estados: transitorios, recurrentes positivos y recurrentes nulos.

Ejemplo 1.9.4 *Demostraremos ahora que la cadena de Markov de rachas de éxitos es recurrente positiva. Recordemos que dicha cadena es irreducible* *y recurrente. Comprobaremos que el tiempo medio de recurrencia del estado 0 es finito. En efecto,*

$$\mu_0 = \sum_{n=1}^{\infty} n f_{00}^{(n)} = \sum_{n=1}^{\infty} n(1-p)p^{n-1} = (1-p)\sum_{n=1}^{\infty} np^{n-1} = \frac{1}{1-p} < \infty$$

Esto demuestra que el estado 0 es recurrente positivo y siendo la cadena irreducible, es recurrente positiva. Por lo tanto, el tiempo medio de recurrencia de cada estado es finito. Hemos aprovechado la facilidad del cálculo de las probabilidades de primer regreso al estado 0.

Definición 1.9.6 (Estado Ergódico) *Se dice que el estado E_j es ergódico si es recurrente-nulo y aperiódico.*

1.10. Evolución de distribuciones

Una matriz estocástica establece una dinámica en el conjunto de las distribuciones de probabilidad definidas sobre el espacio de estados de la correspondiente cadena de Markov. Para explicar la situación de manera simple consideraremos un espacio de estados finito $0, 1, ..., n$ y una distribución de probabilidad inicial $\pi^0 = (\pi_0^0, \pi_1^0, ..., \pi_n^0)$. Después de transcurrida la primera unidad de tiempo, la cadena se encuentre en cualquiera de sus posibles estados de acuerdo a la distribución $\pi^1 = (\pi_0^1, \pi_1^1, ..., \pi_n^1)$, en donde la j-ésima entrada de este vector es

$$\begin{aligned}
\pi_j^1 &= P(X_j = 1) \\
&= \sum_{i=0}^{n} P(X_1 = j | X_0 = i) P(X_0 = i) \\
&= \sum_{i=0}^{n} \pi_i^0 p_{ij}
\end{aligned}$$

Es decir, el vector π^1 se obtiene a partir del vector π^0 y de la matriz de probabilidades de transición $\mathbf{P}$ a través de la multiplicación $\pi^1 = \pi^0 \mathbf{P}$, esto es,

$$\begin{pmatrix} \pi_0^1 & \cdots & \pi_n^1 \end{pmatrix} = \begin{pmatrix} \pi_0^0 & \cdots & \pi_n^0 \end{pmatrix} \begin{pmatrix} p_{00} & \cdots & p_{0n} \\ \vdots & \ddots & \vdots \\ p_{n0} & \cdots & p_{nn} \end{pmatrix}$$

A su vez el vector π^1 se transforma en el vector π^2 a través de la ecuación $\pi^2 = \pi^1\mathbf{P} = \pi^0\mathbf{P}^2$, y así sucesivamente. En general, para $m \geq 1$

$$\pi^m = \pi^{m-1}\mathbf{P} = \pi^0\mathbf{P}^m \qquad (1.10.1)$$

De esta forma se obtiene una sucesión infinita de distribuciones de probabilidad $\pi^0, \pi^1, \pi^2, ...$ en donde cada una de ellas, excepto la primera, es obtenida de la anterior multiplicada a la derecha por la matriz de probabilidades de transición en un paso. Es natural preguntarse si existe algún límite para esta sucesión de distribuciones. En las siguientes secciones estudiaremos tal problema y encontraremos condiciones bajo las cuales existe un único límite para esta sucesión.

Ejemplo 1.10.1 *Considere la matriz de probabilidades de transición*

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

con distribución inicial el vector $\pi^0 = (1/10, 0, 9/10)$. *Los subsecuentes vectores de probabilidad* $pi_1, \pi_2, ...$ *se calculan a continuación*

$$\begin{aligned} \pi^1 &= \pi^0\mathbf{P} = (0{,}45, 0{,}55, 0) \\ \pi^2 &= \pi^1\mathbf{P} = (0, 0{,}45, 0{,}55) \\ \pi^3 &= \pi^2\mathbf{P} = (0{,}275, 0{,}275, 0{,}45) \\ \pi^4 &= \pi^3\mathbf{P} = (0{,}225, 0{,}5, 0{,}275) \\ &\vdots \end{aligned}$$

¿Existirá el límite para esta sucesión de vectores de probabilidad?

Ejemplo 1.10.2 *Considere la matriz de probabilidades de transición*

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

con distribución inicial el vector $\pi^0 = (\alpha, 1-\alpha)$, *con* $0 \leq \alpha \leq 1$. *Los subsecuentes vectores de probabilidad* $pi_1, \pi_2, ...$ *son*

$$\begin{aligned}
\pi^0 &= (\alpha, 1-\alpha) \\
\pi^1 &= (1-\alpha, \alpha) \\
\pi^2 &= (\alpha, 1-\alpha) \\
\pi^3 &= (1-\alpha, \alpha) \\
&\vdots
\end{aligned}$$

que claramente no es convergente, pues tiene un comportamiento oscilatorio, a menos que $\alpha = 1/2$.

Antes de encontrar condiciones bajo las cuales la sucesión de distribuciones de probabilidad definidas por (3.5) es convergente, estudiaremos a continuación el caso particular cuando la distribución inicial no cambia al ser multiplicada por la derecha por $\mathbf{P}$. A tales distribuciones se les llama estacionarias o invariantes en el tiempo

1.11. Distribuciones Estacionarias

Definición 1.11.1 *Una distribución de probabilidad* $\pi = (\pi_0, \pi_1, ...)$ *es estacionaria o invariante para una cadena de Markov con matriz de probabilidades de transición* $\mathbf{P}$ *si*

$$\pi_j = \sum_i \pi_i p_{ij}$$

En términos matriciales, la distribución de probabilidad π es estacionaria si $\pi = \pi\mathbf{P}$. Esta identidad tiene como consecuencia el hecho de que para cualquier número natural n se cumpla que $\pi = \pi\mathbf{P}^n$, es decir, π es también una distribución estacionaria para la matriz $\mathbf{P}^n$. Esto significa que si la variable aleatoria inicial X_0 tiene esa distribución π, entonces la distribución de X_n también es π pues $P(X_n = j) \sum_i \pi_i p_{ij}^{(n)} = \pi_j$, es decir, esta distribución no cambia con el paso del tiempo y por ello es que se le llama estacionaria o invariante. Observe que el vector de ceros cumple la condición $\pi = \pi\mathbf{P}$, sin embargo no corresponde a una distribución de probabilidad. Los siguientes ejemplos muestran que las distribuciones estacionarias pueden no ser unicas y pueden incluso no existir.

Ejemplo 1.11.1 (Existencia Múltiple) *Considere una cadena de Markpv sobre el conjunto de estados* $\{0, 1, 2\}$ *con probabilidades de transición dada por la siguiente matriz*

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

Es fácil probar que el vector $\pi = (1-\alpha, 0, \alpha)$ *satisface el sistema de ecuaciones* $\pi = \pi\mathbf{P}$ *para cada* $\alpha \in [0, 1]$. *Existen entonces una infinidad de distribuciones estacionarias para esta cadena. Observe que el vector* $(1-\alpha, 0, \alpha)$ *se puede escribir como la combinación lineal* $(1-\alpha)(1, 0, 0) + \alpha(0, 0, 1)$.

Ejemplo 1.11.2 (No existencia) *Para la caminata aleatoria simétrica simple sobre* $\mathbb{Z}$ *no existe ninguna distribución estacionaria pues la condición* $\pi = \pi\mathbf{P}$ *se traduce en el sistema de ecuaciones*

$$\pi_j = \frac{1}{2}\pi_{j-1} + \frac{1}{2}\pi_{j+1}, \qquad j \in \mathbb{Z}$$

o bien $\pi_{j+1} - \pi_j = \pi_j - \pi_{j-1}$. *Más explícitamente, iniciando con la identidad* $\pi_1 - \pi_0 = \pi_1 - \pi_0$, *y escribiendo algunas de estas diferencias en términos de la diferencia* $\pi_1 - \pi_0$ *se encuentra que*

$$\begin{aligned} \pi_1 - \pi_0 &= \pi_1 - \pi_0 \\ \pi_2 - \pi_1 &= \pi_1 - \pi_0 \\ \pi_3 - \pi_2 &= \pi_1 - \pi_0 \\ &\vdots \\ \pi_n - \pi_{n-1} &= \pi_1 - \pi_0 \end{aligned}$$

Sumando estas n ecuaciones se llega a que para todo entero $n \geq 1$,

$$\pi_n - \pi_0 = n(\pi_1 - \pi_0)$$

El lado izquierdo es acotado pues es la diferencia de dos probabilidades mientras el lado derecho crece sin límite cuando n es grande, a menos que $\pi_1 - \pi_0 = 0$. Esto demuestra que todas las diferencias $\pi_j - \pi_{j-1}$ con $j \in \mathbb{Z}$ son cero, y por lo tanto π_j es constante para cualquier valor de j. Es decir, el vector constante es la solución del sistema de ecuaciones en las diferencias planteadas, por ello es incompatible con la condición $\sum_j \pi_j = 1$. Por lo tanto no existe ninguna distribución de probabilidad π que cumpla la igualdad $\pi = \pi\mathbf{P}$ para esta cadena.

Ejemplo 1.11.3 (Existencia Única) *La cadena de Markov de dos estados dada por la matriz*

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

tiene una única distribución estacionaria dada por

$$\pi = (\pi_0, \pi_1) = \left(\frac{b}{a+b}, \frac{a}{a+b} \right)$$

cuando $a + b > 0$.

Con base en los ejemplos anteriores haremos algunas observaciones sobre las distribuciones estacionarias. Primeramente observe que para encontrar una posible distribución estacionaria de una cadena con matriz $\mathbf{P}$, un primer método consiste en resolver el sistema de ecuaciones $\pi = \pi\mathbf{P}$, sujeto a la condición $\sum_j \pi_j = 1$. Por otro lado, suponga que π y π' 1 son dos distribuciones estacionarias distintas para una matriz $\mathbf{P}$. Entonces la combinación lineal convexa $\alpha\pi + (1-\alpha)\pi'$, para $\alpha \in [0,1]$, también es una distribución estacionaria pues

$$(\alpha\pi + (1-\alpha)\pi')\mathbf{P} = \alpha\pi\mathbf{P} + (1-\alpha)\pi'\mathbf{P} = \alpha\pi + (1-\alpha)\pi'$$

Por lo tanto, si existen dos distribuciones estacionarias distintas para una cadena, entonces existe una infinidad de ellas. Esto demuestra que el conjunto de distribuciones estacionarias es un conjunto convexo. Tomando en cuenta las observaciones anteriores y de acuerdo a los ejemplos mostrados, sólo hay tres situaciones sobre la existencia de distribuciones estacionarias para una cadena de Markov cualquiera: no existe ninguna distribución estacionaria, existe una distribución estacionaria y es unica, o existe una infinidad de distribuciones estacionarias. Dadas estas consideraciones, es natural plantearse el problema de encontrar condiciones necesarias y suficientes para que una cadena tenga alguna distribución estacionaria. Primeramente demostraremos que cuando existe una distribución estacionaria, ésta tiene como soporte el conjunto de estados recurrentes positivos.

Proposición 1.11.1 (Soporte de una distribución estacionaria) *Sea π una distribución estacionaria para una cadena de Markov. Si j es un estado transitorio o recurrente nulo, entonces $\pi_j = 0$.*

En particular, si $\pi_j > 0$, entonces j es un estado recurrente positivo. Esto es una consecuencia inmediata del resultado anterior y para ello puede usarse un argumento por contradicción. Por otro lado, la proposición también nos ayuda a corroborar nuevamente, por ejemplo, que una caminata aleatoria simétrica simple no tiene distribución estacionaria, pues se trata de una cadena cuyos estados son todos recurrentes nulos y por lo tanto $\pi_j = 0$ para cualquier valor entero de j. Se presenta a continuación una solución al problema de encontrar condiciones suficientes que garanticen la existencia y unicidad de la distribución estacionaria.

Proposición 1.11.2 (Existencia y unicidad de la distribución estacionaria) *Toda cadena de Markov que es irreducible y recurrente positiva tiene una única distribución estacionaria dada por*

$$\pi_j = \frac{1}{\mu_j} > 0$$

en donde μ_j es el tiempo medio de recurrencia del estado j. En particular, toda cadena finita e irreducible tiene una única distribución estacionaria.

Ejemplo 1.11.4 *La cadena de Markov de dos estados es finita e irreducible cuando $a + b > 0$, y por lo tanto es recurrente positiva. Por la Proposición anterior existe una única distribución estacionaria para esta cadena. Resolviendo el sistema de ecuaciones $\pi = \pi\mathbf{P}$ para $\pi = (\pi_0, \pi_1)$ con $\pi_0 + \pi_1 = 1$, se encuentra que*

$$(\pi_0, \pi_1) = \left(\frac{b}{a+b}, \frac{a}{a+b}\right)$$

Como otra aplicación de la proposición anterior encontramos nuevamente que, sin hacer cálculos mayores, los tiempos medios de recurrencia son

$$(\mu_0, \mu_1) = \left(\frac{1}{\pi_0}, \frac{1}{\pi_1}\right) = \left(\frac{a+b}{b}, \frac{a+b}{a}\right)$$

Ejemplo 1.11.5 *La cadena de racha de éxitos, aunque no es finita, es irreducible y recurrente positiva. Por lo tanto tiene una única distribución estacionaria dada por la distribución geometrica con parámetro $1 - p$, es decir el sistema de ecuaciones $\pi = \pi\mathbf{P}$ juntos con $\sum_j \pi_j = 1$ tiene como única solución la distribución*

$$\pi_j = (1-p)p^j, \qquad j = 0, 1, 2, ...$$

Como consecuencia de la proposición anterior, se resuelve un problema dificil de manera inmediata: los tiempos medios de recurrencia para cada uno de los estados de esta cadena son

$$\mu_j = \frac{1}{\pi_j} = \frac{1}{(1-p)p^j}, \qquad j = 0, 1, 2, ...$$

En particular, $\mu_0 = 1/(1-p)$. esto confirma los cálculos realizados para μ_0 a partir de la fórmula $\mu_0 = \sum_{n=1}^{\infty} n f_{00}^{(n)}$.

1.12. Distribución Límite

Como hemos mencionado antes, toda matriz de probabilidades de transición $\mathbf{P}$ determina una sucesión de distribuciones de probabilidad $\pi^0, \pi^1, ...$ sobre el espacio de estados, dada por

$$\pi^n = \pi^{n-1}\mathbf{P} = \pi^0\mathbf{P}^n, \qquad n \geq 1 \tag{1.12.1}$$

Bajo ciertas condiciones tal sucesión es convergente a una distribución de probabilidad límite π. Imaginemos por ahora que tal es el caso y supongamos entonces que

$$\pi = \lim_{n\to\infty} \pi^n$$

Examinaremos algunas propiedades de esta distribución límite. Tomando el límite cuando n Ñ8en las dos igualdades de (3.10) se tiene que

$$\pi = \pi\mathbf{P} \tag{1.12.2}$$

y

$$\pi = \pi^0 \left(\lim_{n\to\infty} \mathbf{P}^n \right) \tag{1.12.3}$$

Estas igualdades revelan varias cosas. Primero, la supuesta distribución límite es una distribución estacionaria, (1.12.2). Segundo, la distribución límite no depende de la distribución inicial, pues nuevamente la igualdad (1.12.2) indica que π se determina a través de la ecuación (1.12.2). Tercero, la distribución límite está dada por el límite de las potencias de $\mathbf{P}$, (1.12.3). Cuarto, a partir de (1.12.3), el límite de las potencias de $\mathbf{P}$ es una matriz con todos sus renglones idénticos, siendo este renglón la distribución límite. En esta sección se establecen condiciones para obtener rigurosamente estos resultados.

Definición 1.12.1 *Considere una cadena de Markov con matriz de probabilidades de transición* $\mathbf{P}$ *y distribución inicial* π^0*. Se llama distribución límite de esta cadena al vector*

$$\pi = \lim_{n\to\infty} \pi^0\mathbf{P}^n = \lim_{n\to\infty} \sum_{n=1}^{\infty} \pi_i^0 p_{ij}^{(n)}$$

Esta convergencia significa que, en corridas grandes $(n \to \infty)$*, la probabilidad de que la cadena de Markov se encuentre en el estado* j *es aproximadamente* π_j *y no importa del estado del que venga.*

Ejemplo 1.12.1 *Consideremos la siguiente matriz de transición*

$$P = \begin{pmatrix} 0{,}33 & 0{,}67 \\ 0{,}75 & 0{,}25 \end{pmatrix}$$

Veamos lo que ocurre con la matriz P a medida que aumenta el valor de n.

$$\mathbf{P}^2 = \begin{pmatrix} 0{,}6114 & 0{,}3886 \\ 0{,}4350 & 0{,}5650 \end{pmatrix} \quad \mathbf{P}^3 = \begin{pmatrix} 0{,}4932 & 0{,}5068 \\ 0{,}5673 & 0{,}4327 \end{pmatrix}$$
$$\mathbf{P}^4 = \begin{pmatrix} 0{,}5428 & 0{,}4572 \\ 0{,}5117 & 0{,}4883 \end{pmatrix} \quad \mathbf{P}^5 = \begin{pmatrix} 0{,}5220 & 0{,}4780 \\ 0{,}5350 & 0{,}4560 \end{pmatrix}$$
$$\mathbf{P}^6 = \begin{pmatrix} 0{,}5307 & 0{,}4693 \\ 0{,}5253 & 0{,}4747 \end{pmatrix} \quad \mathbf{P}^7 = \begin{pmatrix} 0{,}5271 & 0{,}4729 \\ 0{,}5294 & 0{,}4706 \end{pmatrix}$$

Se puede notar que para $n = 7$ los valores de las probabilidades de transición en cada columna están muy cercanos, lo cual indica que, por ejemplo la probabilidad de estar en el estado E_0 es aproximadamente 0.52 (se vera más adelante que es 0.5282) independientemente de si estaba en el estado E_0 o E_1. Lo mismo se aprecia con E_1.

Observe que el vector límite π en la definición anterior podría no ser una distribución de probabilidad verdadera, a pesar de esto mantendremos dicho término en la definición. Como se ha mencionado, toda posible distribución límite es estacionaria pero el recíproco es en general falso, como se ilustra en el siguiente ejemplo.

Ejemplo 1.12.2 *Es inmediato comprobar que la distribución* $(1/2, 1/2)$ *es estacionaria para la cadena con matriz*

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

Sin embargo las potencias de P no convergen, pues para cualquier $n \geq 0$

$$P^{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

El siguiente resultado es válido para espacios de estados finitos o infinitos, y establece que si el límite de las probabilidades $p_{ij}^{(n)}$, cuando $n \to \infty$, existen y no dependen de i, entonces la distribución límite podría ser una distribución estacionaria. Esto es solamente una posibilidad, pues los límites podrían ser todos cero. En el caso finito, sin embargo, demostraremos que tales límites conforman una distribución de probabilidad verdadera.

Proposición 1.12.1 *Considere una cadena de Markov con probabilidades de transición p_{ij} tales que los límites $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$ existen para cada j, y no dependen del estado i. Entonces*

1. $\sum_j \pi_j \leq 1$
2. $\pi_j = \sum_i \pi_i p_{ij}$

Ahora se establecen condiciones suficientes para que exista el límite de las probabilidades de transición cuando el número de pasos n crece a infinito. Este resultado es una especie de recíproco del resultado anterior, pues supone la existencia de una distribución estacionaria para concluir que los límites de las probabilidades existen.

Teorema 1.12.1 (Convergencia a la distribución estacionaria) *Considere una cadena de Markov que es:*

a) irreducible,

b) aperiódica, y

c) con distribución estacionaria π

Entonces para cualesquiera estados i y j, $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$

Ejemplo 1.12.3 *Encontrar la distribución limite de una cadena de Markov cuya matriz de transición es*

$$\begin{array}{c} \\ E_0 \\ E_1 \\ E_2 \end{array} \begin{array}{c} \begin{array}{ccc} E_0 & E_1 & E_2 \end{array} \\ \begin{pmatrix} 0{,}40 & 0{,}50 & 0{,}10 \\ 0{,}05 & 0{,}70 & 0{,}25 \\ 0{,}05 & 0{,}50 & 0{,}45 \end{pmatrix} \end{array}$$

Solución: *Las ecuaciones que determinan la distribución límite son:*

$$0{,}40\pi_0 + 0{,}05\pi_1 + 0{,}05\pi_2 = \pi_0 \tag{1.12.4}$$
$$0{,}50\pi_0 + 0{,}70\pi_1 + 0{,}50\pi_2 = \pi_1 \tag{1.12.5}$$
$$0{,}10\pi_0 + 0{,}25\pi_1 + 0{,}45\pi_2 = \pi_2 \tag{1.12.6}$$
$$\pi_0 + \pi_1 + \pi_2 = 1 \tag{1.12.7}$$

Recordemos que una de las ecuaciones anteriores es redundante. Por lo tanto dejando la tercera ecuación por fuera y resolviendo el sistema de ecuaciones se obtiene

- $\pi_0 = \frac{1}{13}$
- $\pi_1 = \frac{5}{8}$
- $\pi_2 = \frac{31}{104}$

Teorema 1.12.2 (Convergencia para cadenas de Markov) *Considere una cadena de Markov que es:*

a) irreducible,

b) aperiódica, y

c) recurrente positiva

Entonces las probabilidades límite $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} =$ *existe, y están dadas por* $\pi_j = 1/\mu_j$*, y constituyen la única solución al sistema de ecuaciones*

$$\pi_j = \sum_i \pi_i p_{ij}$$

sujeto a las condiciones $\pi_j \geq 0$ *y* $\sum_j \pi_j = 1$