Cadenas de Markov

http://humberto-r-alvarez-a.webs.com

Definición

- Procesos estocásticos: procesos que evolucionan de forma no determinista a lo largo del tiempo en torno a un conjunto de estados. Estos estados está formado por una familia de variables aleatorias $\{X_i\}$
- Cadenas de Markov: conjunto de herramientas para analizar procesos estocásticos.
- Una Cadena de Markov representa un sistema que varía su estado a lo largo del tiempo, siendo cada cambio una transición del sistema.
- Aunque los cambios no están predeterminados, las probabilidades del próximo estado en función del estado anterior son conocidas.
- Las probabilidades son constantes a lo largo del tiempo, el sistema es homogéneo.

Características:

- Número finito y conocido de estados.
- Probabilidades de transición conocidas y constantes.
- Se puede, eventualmente, pasar de un estado a cualquier otro.
- No son ciclos simples.

Elementos de una Cadena de Markov

- Para definirla es necesario determinar:
  - El conjunto de estados del sistema
  - La definición de transición
  - Una ley de probabilidad condicional, que defina la probabilidad como del nuevo estado en función de los anteriores.
- Los estados son una caracterización de un sistema en un instante dado.
- El estado de un sistema en un instante t es una variable cuyos valores sólo pueden pertenecer al conjunto de estados del sistema.
- El estado del sistema en el tiempo t+1 dependerá solo del estado del sistema en el tiempo t

Definición

- Es una colección indexada de varias x_t, donde t denota intervalos de tiempo temporales significativos para el sistema bajo estudio. Los valores de x se toman de un conjunto de categorías o estados, mutuamente excluyentes de las que se conoce la probabilidad asociada y no la certeza del estado, tal que:

$$\Pr\left[X_{t+1} = x_{t+1} \mid X_1 \cdots X_t = x_1 \cdots x_t\right] = \Pr\left[X_{t+1} = x_{t+1} \mid X_t = x_t\right]$$

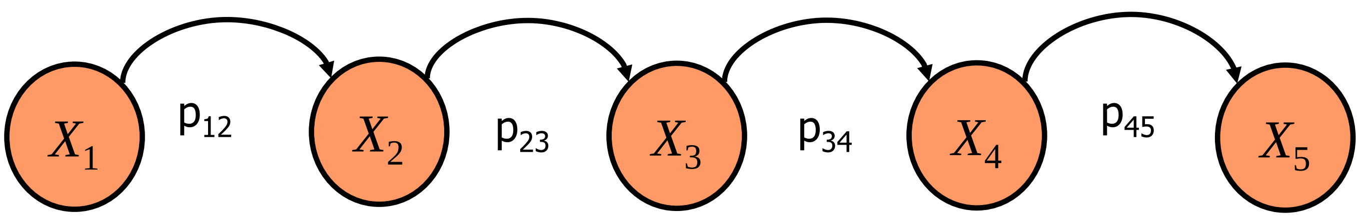
En otras palabras: la probabilidad de existencia de un estado, solamente depende del estado anterior.

Probabilidad de transición

- Sea $\Pr[X_{t+1}=b|X_t=a]$; la probabilidad de que el proceso, estando en el estado a en el tiempo t, pase al estado b en el instante siguiente t+1
- Las probabilidades de transición son independientes del tiempo t, tal que:

$$\boxed{\Pr\left[X_{t+1}=b\,|\,X_t=a\right]=p_{ab}}$$

- Estas son conocidas como cadenas homogéneas de orden 1



Matrices de transición P

- Es una matriz cuadrada de orden n, donde n es el número de estados de la cadena.
- Debido a que el sistema evoluciona de cualquier estado i a j, en cualquier tiempo t, entonces:

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

- Para todo $p_{ij} \geq 0$ por definición, tal que:

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \cdots p_{1n} \\ p_{21} & p_{22} & p_{23} \cdots p_{2n} \\ p_{31} & p_{32} & p_{33} \cdots p_{3n} \\ & \vdots & \\ p_{n1} & p_{n2} & p_{n3} \cdots p_{nn} \end{bmatrix}$$

Ejemplo: (conjunto de problemas 17.7A Taha, p. 1)

- Un profesor de ingeniería adquiere una computadora nueva cada dos años. Si el modelo actual es M_1, la siguiente computadora puede ser M_2 con probabilidad de 0.2 o M_3 con probabilidad de 0.15. Si el modelo actual es M_2, las probabilidades de cambiar a M_1 o M_3 son 0.6 y 0.25, respectivamente. Si el modelo actual es M_3, las probabilidades de comprar M_1 o M_2 son 0.5 y 0.1, respectivamente. Representar la situación como una cadena de Markov.

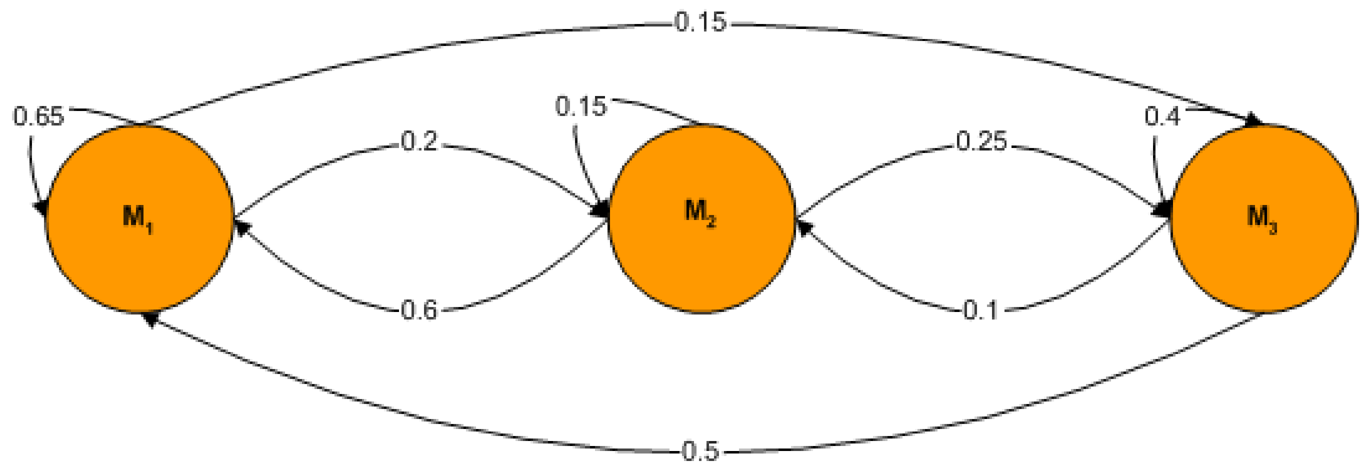
Matriz de transición

t

		M_1	M_2	M_3
	M_1	0.65	0.2	0.15
t+1	M_2	0.6	0.15	0.25
	M_3	0.5	0.1	0.4

- Proceso de Markov: definido por el gráfico (FSM)
  (FSM: Finite Stochastic Machine)
- Cadena de Markov: cualquier recorrido aleatorio en el FSM

Proceso de Markov



Probabilidades de transición para k pasos Ecuaciones de Chapman-Kolmogorov

- Al ser las probabilidades de transición estables en el tiempo, se puede conocer las propiedades después de k pasos:

$$p\{E_{t+k}=j \mid E_t = i\} = p\{E_k = j \mid E_0 = i\} = p_{ij}^{(k)}$$

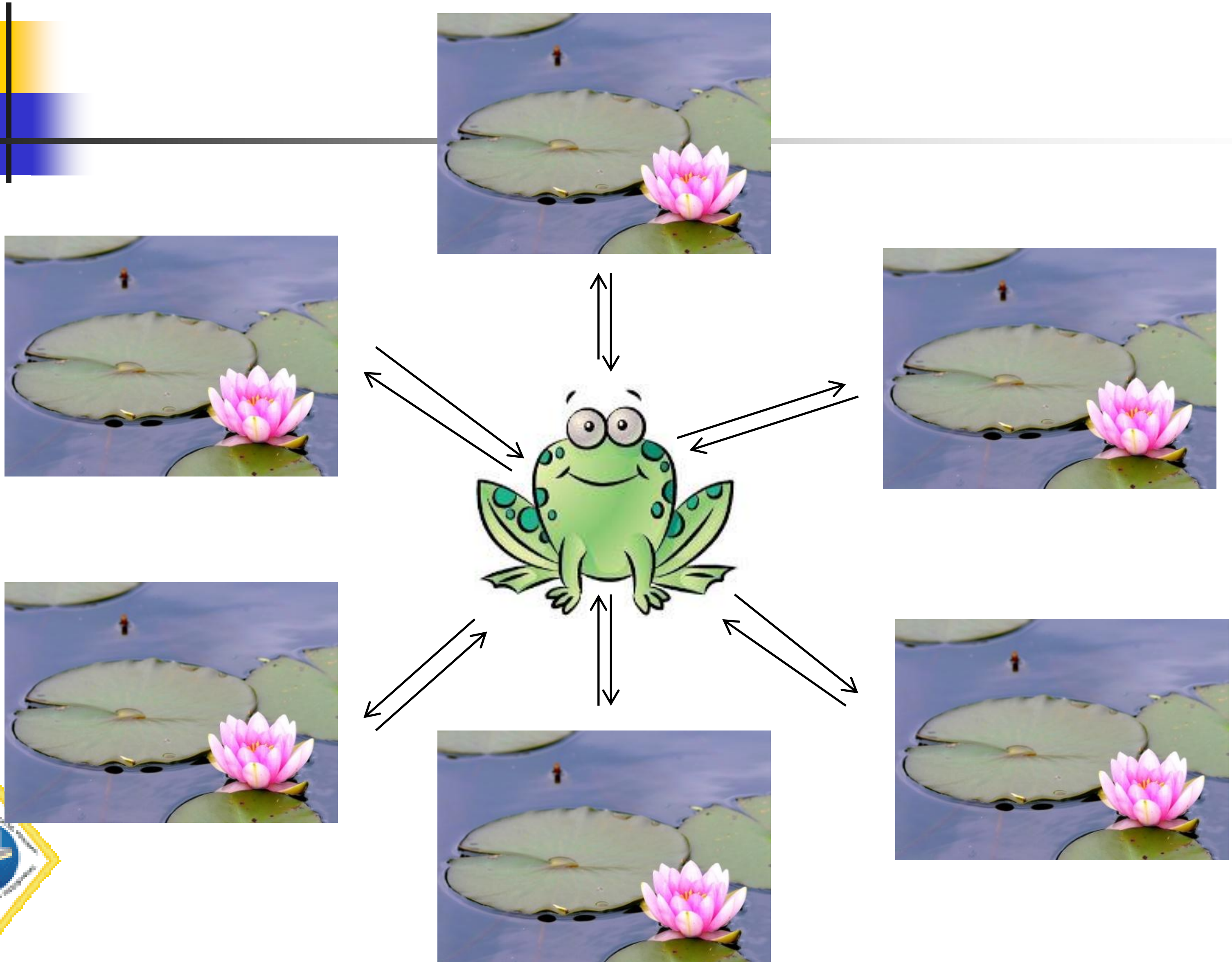
- Esto es, la probabilidad de que el proceso se encuentre en estado j si k etapas antes se encontraba en el estado i.
- Si se conoce p_{ij}, es posible calcular las $p_{ij}^{(k)}$ tal que:

$$p_{ij}^{(k)} = \sum_{e=1}^{n} p_{ie}^{(k-1)} \cdot p_{ej}$$

- Al ser P el conjunto de probabilidades de transición, entonces:

$$P^{(k)} = P.P^{k-1} = P^k$$

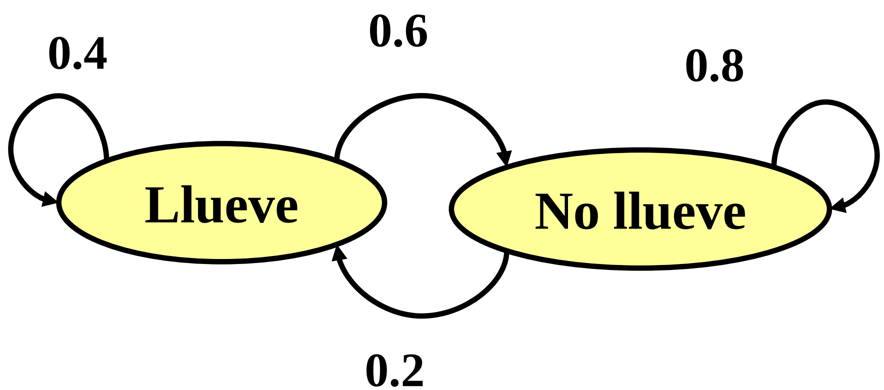
- Siendo P^1 la probabilidad de una transición y $P^0 = I$



Ejemplo:

P =

	Llueva	No llueva
Llueva	40%	60%
No llueva	20%	80%



Ejemplo …

- Si hoy llovió, ¿cuál es la probabilidad de que no llueva en dos días?

P(Llueva → → No llueva) = P(Llueva →Llueva→ No llueva) + P(Llueva →No Llueva→ No llueva);

P(No llueva/Llueva dos días antes) = P(No llueva/(Llueva∪Llueva))+P(No llueva(Llueva∪No llueva)

= 0.4*0.6 + 0.6*0.8 = 0.72 o 72%

	Llueva	**No llueva**
Llueva	40%	60%
No llueva	20%	80%

+

	Llueva	**No llueva**
Llueva	40%	60%
No llueva	20%	80%

Ejemplo….

P =

	Llueva	No llueva
Llueva	40%	60%
No llueva	20%	80%

Primer día	Segundo día	Hoy	
Llueve	Llueve	Llueve	0.4*0.4
Llueve	No llueve	Llueve	0.6*0.2
Llueve	Llueve	No llueve	0.4*0.6
Llueve	No llueve	No llueve	0.6*0.8
No llueve	Llueve	Llueve	0.2*0.4
No llueve	No llueve	Llueve	0.8*0.2
No llueve	Llueve	No llueve	0.2*0.6
No llueve	No llueve	No llueve	0.8*0.8

0.28

0.72

0.24

0.76

Ejemplo….

- Para los posibles escenarios en dos días (k=2):

	t + 2	
t	Llueva	No llueva
Llueva	28%	72%
No llueva	24%	76%

- Puede demostrarse que $P^{(2)} = P^2$

Propiedades a largo plazo de las Cadenas de Markov

- En general, cuando k $\rightarrow \infty$, la probabilidad de estar en un estado j es independiente del estado inicial y se conoce como probabilidad estacionaria o de estado estable π_j
- Estas probabilidades se mantienen constantes después de dichas transiciones.
- Las cadenas de Markov con esta característica se conocen como cadenas ergódicas y su matriz de transición será

$$\mathbf{P}^* = \begin{vmatrix} \pi_1 & \pi_2 & & \pi_n \\ \pi_1 & \pi_2 & & \pi_n \\ & & & \\ & & & \\ \pi_1 & \pi_2 & & \pi_n \end{vmatrix}$$

Volviendo al ejemplo de la lluvia

$P^1 =$

	Llueva	No llueva
Llueva	40%	60%
No llueva	20%	80%

$P^4 =$

	Llueva	No llueva
Llueva	25.1%	74.9%
No llueva	25.0%	75.0%

$P^8 =$

	Llueva	No llueva
Llueva	25.0%	75.0%
No llueva	25.0%	75.0%

$P^{16} =$

	Llueva	No llueva
Llueva	25.0%	75.0%
No llueva	25.0%	75.0%

Propiedades a largo plazo

- Una cadena es ergódica cuando tiene el mismo comportamiento promedio en el tiempo como promedio durante el espacio de todos los estados del sistema.
- En general, para una cadena ergódica existe un límite:

$$\lim_{k \to \infty} P_{ij}^{(k)}$$

tal que el mismo existe es independiente de i.

Además

$$\lim_{k \to \infty} P_{ij}^{(k)} = \pi_j \geq 0$$

donde π_j satisface las siguientes ecuaciones de estado estable:

$$\pi_j = \sum_{i=1}^{M} \pi_i p_{ij,} \forall\, j = 1, \dots, M$$

$$\sum_{j=1}^{M} \pi_j = 1$$

Algunos aspectos importantes

- π_j define la probabilidad de encontrar el sistema en el estado j, después de un gran número de transiciones y tiende a ser constante e independiente del estado inicial.
- Aunque la probabilidad π_j define la probabilidad de estar en el estado j, la probabilidad de la transición entre i y j, en cualquier etapa k, será p_{ij}
- Las ecuaciones de estado estable estará definido por n+1 ecuaciones y n incógnitas, por lo que siempre habrá una ecuación redundante que puede eliminarse, con excepción de la expresión

$$\sum_{j=1}^{M} \pi_j = 1$$

Ejemplo de la lluvia…

$P =$

	Llueva	No llueva
Llueva	40%	60%
No llueva	20%	80%

$P^{*}=$

	Llueva	No llueva
Llueva	π_1	π_2
No llueva	π_1	π_2

$P^{*}=P^{*}P$, por las ecuaciones de Chapman-Kolmogorov, donde se tiene que $P^{*-1}=P^{*}$ en el estado estable y además la ecuación

$$\sum_{j=1}^{M} \pi_j = 1$$

$$[P^*] = [P^*] * [P]$$

$$\begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{bmatrix} * \begin{bmatrix} 40\% & 60\% \\ 20\% & 80\% \end{bmatrix}$$

$$\pi_1 + \pi_2 = 1$$

$$\begin{aligned} 0.4*\pi_1 + 0.2*\pi_2 &= \pi_1 \\ 0.6*\pi_1 + 0.8*\pi_2 &= \pi_2 \\ \pi_1 + \pi_2 &= 1 \end{aligned}$$

Resolviendo:

$\pi_1 = 0.25$

$\pi_2 = 0.75$

y $P^* =$

	Llueva	No llueva
Llueva	25.0%	75.0%
No llueva	25.0%	75.0%

Esta matriz estacionaria coincide con $P^{(k)}$, para $k \to \infty$

El caso cuando hay probabilidades iniciales

- Sea $a^t = (a_0, a_1,\ldots,a_m)$ la distribución en t_0 de la Cadena de Markov (con m estados)
- La probabilidad de que el sistema se encuentre en el estado j después de n etapas estará dada por:
- $P\{X_n = j\}$ = elemento j del vector $a^t P^n$
- En el largo plazo: $P\{X_n = j\} = a^t P^*$

El ejemplo de la lluvia, nuevamente

P =

	Llueva	No llueva
Llueva	40%	60%
No llueva	20%	80%

- Suponga que la probabilidad de que hoy llueva es 65%, ¿cuál será la probabilidad de que no llueva dentro de dos días?
- En este caso, a^t = (65%, 35%)
- Resolviendo para a^tP^2 =(26.6%, 73.4%)
- P(No llueva dentro de dos días): 73.4%

Cuando hay costos asociados

- Sea $c^t = (c_0, c_1, \ldots, c_m)$ la distribución de costos asociados a la Cadena de Markov (con m estados)
- El costo de que el sistema se encuentre en el estado j después de n etapas estará dada por:
- $E(c) = a^t P^{(n)} c^t$
- Esto resultará en el costo esperado del sistema en la etapa n.
- En caso de no tener definida la probabilidad inicial a, se supone igual para cada estado j.

Regresando al ejemplo de la lluvia

P =

	Llueva	No llueva
Llueva	40%	60%
No llueva	20%	80%

- Suponga que el valor de que llueva es de -20 y que no llueva es de 10.
- La probabilidad inicial de lluvia se mantiene en 65%
- ¿Cuánto será el valor esperado después de dos días para los diferentes estados?
- Resolviendo para $E(c)=a^tP^{(n)}c^t$ = 2.02 unidades monetarias.

Clasificación de los estados de una Cadena de Markov

- Existen ciertas propiedades de los estados que son importantes en el comportamiento de las Cadenas de Markov.
- El estado j se dice que es accesible desde el estado i si $p_{ij}^{(n)} > 0$ en cualquier etapa n.
  - Esto quiere decir que el sistema puede entrar al estado j eventualmente, si está en el estado i.
- Si el estado j es accesible desde el estado i, y el estado i, es accesible desde el estado j, se dice entonces que i y j pueden comunicarse.

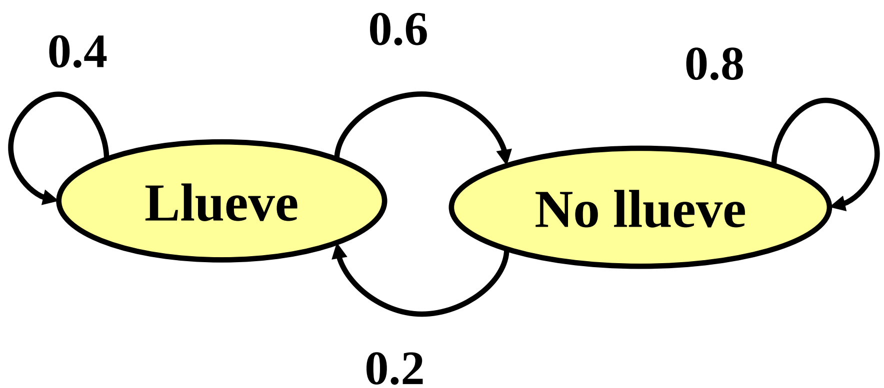
En general

- 1. Cada estado se comunica con si mismo ya que $p_{ii}^{(0)}=P(x_0 = i|x_0 = i) = 1$
- Si el estado i se comunica con el estado j, entonces j se couminca con i.
- 3. Si el estado i se comunica con j, y j se comunica con k, entonces el estado i se comunica con k.
- Como resultado de lo anterior, los estados pueden ser divididos en dos o más clases de tal manera que los estados que se comunican entre ellos puedan estar en una sola clase.
- Si el estado tiene una sola clase se dice que es irreducible.

Clasificación de los estados de una Cadena de Markov

- Se dice que un estado es transitorio si una vez en ese estado, el sistema no retorna a dicho estado otra vez. En otras palabras, el estado i es transitorio si existe un estado j que es accesible desde i pero no lo contrario desde el estado j.
- Un estado es recurrente si, una vez en ese estado, el sistema volverá a ese estado otra vez en algún momento.
- Un estado es absorvente si una vez en él, el sistema nunca lo dejará. Así, en un estado absorvente $p^{ii} = 1$
- Un estado es periódico, con período $t > 1$ si es posible su retorno en períodos definidos. Un período de un estado i se define como un entero t ($t>1$), tal que $p_{ii}^{(n)} = 0$ para cualquier valor de n diferente a t, 2t, 3t, …, siendo t el entero mayor de dicha propiedad.

En el caso de la lluvia



- El sistema es accesible en cada estado y se pueden comunicar.

Estados absorventes

- Como se dijo, un estado k se llama absorvente si´p_{kk} = 1.
- Así, el estado k permanence así por siempre.
- Si el estado k es un estado absorvente y el sistema inició en estado i, la probabilidad de ir en algún momento al estado k se conoce como probabilidad de absorción al estado k, siendo que el sistema inició en el estado I y se denota por f_{ik}.
- Cuando hay dos o más estados absorventes en una cadena de Markov, el sistema será absorvido por uno de esos estados.
- Es necesario conocer las probabilidades de absorción resolviendo el sistema de ecuaciones dado por:

$$f_{ik} = \sum_{M} p_{ij} f_{jk} \text{ para todo } j = 1, 2, \ldots, M$$

- Sujeto a $f_{kk} = 1$, y
- $f_{ik} = 0$, para $i \neq k$, y el estado i sea recurrente

Otro ejemplo

Considere una tienda por departamento que clasifica el balance de las cuentas de sus clientes en totalmente pagado (estado 0), 1 a 30 días de atraso (estado 1), 31 a 60 días (estado 2) o cuenta mala (estado 3).

Los estados de cuenta se verifican cada mes para ver la situación de cada cliente. En general el crédito no se extiende y se espera que los clientes paguen dentro de los primeros 30 días. Muchas veces, los clientes pagan solamente parte de su cuenta. Si esto ocurre cuando el balance está dentro de los primeros treinta días, entonces la tienda lo mantiene dentro del estado uno. Si el pago ocurre cuando el cliente estaba en el estado dos, entonces cambia del estado 1 al estado 2. Los clientes con más de 60 días se consideran dentro de las cuentas malas y se envían a una agencia de cobros.

Otro ejemplo

Aunque cada cliente termina ya sea en estado 0 o 3, la tienda está interesada en determinar la probabilidad de que un cliente caerá como cuenta mala ya sea que esté en el estado 1 o 2.

P =

Estado \ Estado		Pagado 0	1 a 30 días 1	31 a 60 días 2	Cuenta mala 3
Pagado	0	1	0	0	0
1 a 30 días	1	0.7	0.2	0.1	0
31 a 60 días	2	0.5	0.1	0.2	0.2
Cuenta mala	3	0	0	0	1

Solución

- Se requiere calcular f_{13} y f_{23}
- Resolviendo $f_{ik} = \Sigma_M p_{ij} f_{jk}$
- $f_{13} = p_{10}f_{03} + p_{11}f_{13} + p_{12}f_{23} + p_{13}f_{33}$
- $f_{23} = p_{20}f_{03} + p_{21}f_{13} + p_{22}f_{23} + p_{23}f_{33}$
- Se puede suponer que $f_{03} = 0$ (no se puede ir de pagar todo a cuenta mala), $f_{33} = 1$ (si eres cuenta mala, te quedas en cuenta mala)
- Sustituyendo estos valores y las probabilidades definidas en la matriz de transición
- $f_{13} = 0.032$
- $F_{23} = 0.025$