

# Introducción a las cadenas de Markov

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:

# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,

# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,

# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,

# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,

# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,
  - la velocidad de una determinada conexión a internet,

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,
  - la velocidad de una determinada conexión a internet,
  - etc.



# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,
  - la velocidad de una determinada conexión a internet,
  - etc.
- De cara a modelar una determinada situación, necesitamos que las **variables aleatorias** cambien con el **tiempo**.

# Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,
  - la velocidad de una determinada conexión a internet,
  - etc.
- De cara a modelar una determinada situación, necesitamos que las **variables aleatorias** cambien con el **tiempo**.
- Por dicho motivo, vamos a introducir los **procesos estocásticos**, que son **variables aleatorias** que además de depender de los elementos del **espacio muestral**, dependen del **tiempo**.

## Procesos estocásticos

Un **proceso estocástico** es una **variable aleatoria**  $X(w, t)$  que depende de dos argumentos:

- un elemento del **espacio muestral**  $w \in \Omega$ ,
  - el **tiempo**  $t \in \mathcal{T}$ , donde el conjunto  $\mathcal{T}$  puede ser
    - discreto:  $\mathcal{T} = \{0, 1, 2, \dots\}$ , o  $\mathcal{T} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ ,
    - continuo:  $\mathcal{T} = [0, \infty)$  o  $\mathcal{T} = [-\infty, \infty)$ .
- 
- Si fijamos el **tiempo**  $t$ , la **variable aleatoria**  $X(w, t) = X_t(w)$  sería una **variable aleatoria** que modelizaría lo que pasa en el sistema en el **instante**  $t$ .

## Procesos estocásticos

Un **proceso estocástico** es una **variable aleatoria**  $X(w, t)$  que depende de dos argumentos:

- un elemento del **espacio muestral**  $w \in \Omega$ ,
  - el **tiempo**  $t \in \mathcal{T}$ , donde el conjunto  $\mathcal{T}$  puede ser
    - discreto:  $\mathcal{T} = \{0, 1, 2, \dots\}$ , o  $\mathcal{T} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ ,
    - continuo:  $\mathcal{T} = [0, \infty)$  o  $\mathcal{T} = [-\infty, \infty)$ .
- 
- Si fijamos el **tiempo**  $t$ , la **variable aleatoria**  $X(w, t) = X_t(w)$  sería una **variable aleatoria** que modelizaría lo que pasa en el sistema en el **instante**  $t$ .
  - Si fijamos el elemento  $w$  del espacio muestral  $\Omega$ , tenemos la función dependiendo del **tiempo**  $X(w, t) = X_w(t)$ . Dicha función se denomina trayectoria del **proceso estocástico**.

## Ejemplo: lanzamiento de una moneda

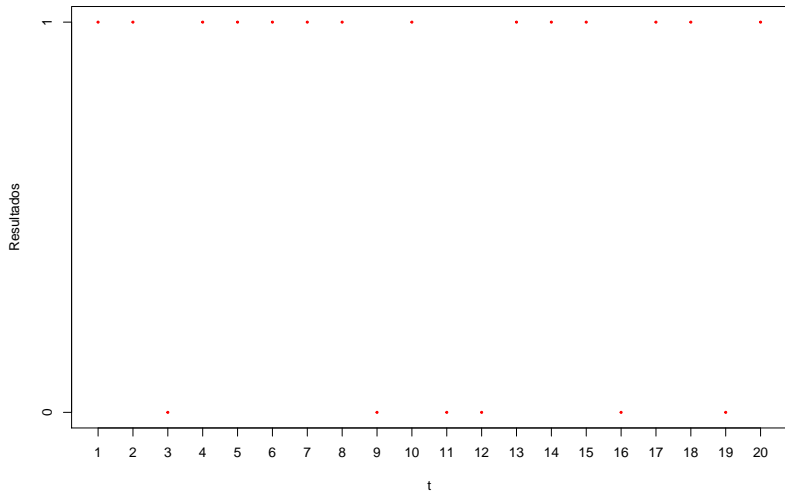
Un **proceso estocástico** sencillo es considerar lanzar una moneda cada cierto espacio de tiempo, por ejemplo cada minuto, y observar el resultado.

En este caso  $\Omega = \{c, +\}$  y  $\mathcal{T} = \{0, 1, 2, \dots\}$ .

$X(w, t)$  sería la variable aleatoria que nos dice el comportamiento de la moneda en el lanzamiento  $t$ -ésimo.

La distribución de dicha variable será de Bernoulli de parámetro  $p$  para cualquier instante  $t$  donde  $p$  es la probabilidad de sacar cara.

# Procesos estocásticos



## Tipos de procesos estocásticos

Diremos que el **proceso estocástico** es **de estado discreto** si la variable aleatoria  $X_t(w)$  es discreta y es **de estado continuo** si la variable aleatoria  $X_t(w)$  es continua para todo valor del tiempo  $t$ . Diremos que el **proceso estocástico** es **de tiempo discreto** si el conjunto  $\mathcal{T}$  de valores del tiempo es discreto y es **de tiempo continuo** si  $\mathcal{T}$  es continuo.

# Procesos estocásticos. Ejemplos

- Ejemplo anterior del lanzamiento de una moneda de parámetro  $p$ : estado discreto y tiempo discreto.



# Procesos estocásticos. Ejemplos

- Ejemplo anterior del lanzamiento de una moneda de parámetro  $p$ : estado discreto y tiempo discreto.
- Número de infectados diarios por una pandemia: estado discreto y tiempo discreto.

# Procesos estocásticos. Ejemplos

- Ejemplo anterior del lanzamiento de una moneda de parámetro  $p$ : estado discreto y tiempo discreto.
- Número de infectados diarios por una pandemia: estado discreto y tiempo discreto.
- Temperatura en un lugar determinado: estado continuo y tiempo continuo.

# Procesos estocásticos. Ejemplos

- Ejemplo anterior del lanzamiento de una moneda de parámetro  $p$ : estado discreto y tiempo discreto.
- Número de infectados diarios por una pandemia: estado discreto y tiempo discreto.
- Temperatura en un lugar determinado: estado continuo y tiempo continuo.
- Tiempo de espera del  $n$ -ésimo cliente de una cola en el supermercado modelado como  $X(n, w)$ . Este ejemplo es “lioso” ya que el tiempo sería el número del cliente y el espacio muestral sería precisamente el tiempo. Pensar que la variable aleatoria “tiempo” dependerá del número de cliente  $n$ : estado continuo y tiempo discreto.

## Cadena de Markov

Un proceso estocástico  $X(t)$  es un **proceso de Markov** si para cualquier secuencia de valores  $t_1 < \dots < t_n < t$  y para cualquier secuencia de sucesos  $A, A_1, \dots, A_n$ ,

$$\begin{aligned} P(X(t) \in A | X(t_1) \in A_1, \dots, X(t_n) \in A_n) \\ = P(X(t) \in A | X(t_n) \in A_n). \end{aligned}$$

- Es decir, que la **distribución condicionada** de la variable  $X(t)$  condicionada a los valores del proceso estocástico en  $n$  **instantes cualesquiera del pasado** sólo depende de la **distribución condicionada** de la variable  $X(t)$  condicionada al proceso correspondiente al **último instante de la secuencia**.

- Esquemáticamente, podemos escribir:

$$\begin{aligned} P(\text{FUTURO}|\text{PASADO}, \text{PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}). \end{aligned}$$

- Esquemáticamente, podemos escribir:

$$\begin{aligned} P(\text{FUTURO}|\text{PASADO}, \text{PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}). \end{aligned}$$

- Ejemplos:

- Esquemáticamente, podemos escribir:

$$\begin{aligned} P(\text{FUTURO}|\text{PASADO}, \text{PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}). \end{aligned}$$

- Ejemplos:
  - Temperatura en un día determinado: no es un **proceso de Markov**.

- Esquemáticamente, podemos escribir:

$$\begin{aligned} P(\text{FUTURO}|\text{PASADO}, \text{PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}). \end{aligned}$$

- Ejemplos:
  - Temperatura en un día determinado: no es un **proceso de Markov**.
  - Número de conexiones registradas en un router de internet en un instante determinado: sí es un **proceso de Markov** ya que la gente se conecta aleatoriamente.



- Esquemáticamente, podemos escribir:

$$\begin{aligned} P(\text{FUTURO}|\text{PASADO}, \text{PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}). \end{aligned}$$

- Ejemplos:
  - Temperatura en un día determinado: no es un **proceso de Markov**.
  - Número de conexiones registradas en un router de internet en un instante determinado: sí es un **proceso de Markov** ya que la gente se conecta aleatoriamente.
  - Precio de un stock en la bolsa en un día determinado: no es un **proceso de Markov**.

- Esquemáticamente, podemos escribir:

$$\begin{aligned} P(\text{FUTURO}|\text{PASADO}, \text{PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}). \end{aligned}$$

- Ejemplos:
  - Temperatura en un día determinado: no es un **proceso de Markov**.
  - Número de conexiones registradas en un router de internet en un instante determinado: sí es un **proceso de Markov** ya que la gente se conecta aleatoriamente.
  - Precio de un stock en la bolsa en un día determinado: no es un **proceso de Markov**.
  - Ejemplo del lanzamiento de la moneda: sí es un **proceso de Markov**.

# Cadenas de Markov. Introducción

- Dentro de los **procesos de Markov** están las denominadas **cadenas de Markov**:

## Cadenas de Markov

Una **cadena de Markov** es un **proceso de Markov** de **estado discreto** y de **tiempo discreto**.

# Cadenas de Markov. Introducción

- Dentro de los **procesos de Markov** están las denominadas **cadenas de Markov**:

## Cadenas de Markov

Una **cadena de Markov** es un **proceso de Markov** de **estado discreto** y de **tiempo discreto**.

- Las cadenas de Markov se aplican a:

# Cadenas de Markov. Introducción

- Dentro de los **procesos de Markov** están las denominadas **cadenas de Markov**:

## Cadenas de Markov

Una **cadena de Markov** es un **proceso de Markov** de **estado discreto** y de **tiempo discreto**.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos meteorológicos básicos.

# Cadenas de Markov. Introducción

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

## Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos meteorológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.

# Cadenas de Markov. Introducción

- Dentro de los **procesos de Markov** están las denominadas **cadenas de Markov**:

## Cadenas de Markov

Una **cadena de Markov** es un **proceso de Markov** de **estado discreto** y de **tiempo discreto**.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos meteorológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.

# Cadenas de Markov. Introducción

- Dentro de los **procesos de Markov** están las denominadas **cadenas de Markov**:

## Cadenas de Markov

Una **cadena de Markov** es un **proceso de Markov** de **estado discreto** y de **tiempo discreto**.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos meteorológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.
  - Economía y finanzas: modelización del colapso de una bolsa de valores.



# Cadenas de Markov. Introducción

- Dentro de los **procesos de Markov** están las denominadas **cadenas de Markov**:

## Cadenas de Markov

Una **cadena de Markov** es un **proceso de Markov** de **estado discreto** y de **tiempo discreto**.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos meteorológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.
  - Economía y finanzas: modelización del colapso de una bolsa de valores.
  - Genética: teoría genética de poblaciones.

# Cadenas de Markov. Introducción

- Dentro de los **procesos de Markov** están las denominadas **cadenas de Markov**:

## Cadenas de Markov

Una **cadena de Markov** es un **proceso de Markov** de **estado discreto** y de **tiempo discreto**.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos meteorológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.
  - Economía y finanzas: modelización del colapso de una bolsa de valores.
  - Genética: teoría genética de poblaciones.
  - Redes neuronales: se utilizan en las máquinas de Boltzmann.

- Vamos a introducir algunas simplificaciones:

- Vamos a introducir algunas simplificaciones:
  - Definiremos  $\mathcal{T} = \{0, 1, 2, \dots\}$ . Entonces la **cadena de Markov** sería una **secuencia aleatoria** de variables aleatorias  $X(0), X(1), X(2), \dots$

- Vamos a introducir algunas simplificaciones:
  - Definiremos  $\mathcal{T} = \{0, 1, 2, \dots\}$ . Entonces la **cadena de Markov** sería una **secuencia aleatoria** de variables aleatorias  $X(0), X(1), X(2), \dots$
  - Llamaremos al conjunto  $\Omega$  **conjunto de estados**. Como es discreto, lo enumeraremos de la forma siguiente  $\Omega = \{1, 2, \dots, n\}$ .

- Por ser una **cadena de Markov**, un **proceso de Markov**, tenemos la denominada **propiedad de Markov** que nos dice que la variable aleatoria en un instante  $t + 1$  sólo depende de los valores que toma la variable aleatoria  $X(t)$  o el proceso en el instante  $t$ .

# Cadenas de Markov. Introducción

- Por ser una **cadena de Markov**, un **proceso de Markov**, tenemos la denominada **propiedad de Markov** que nos dice que la variable aleatoria en un instante  $t + 1$  sólo depende de los valores que toma la variable aleatoria  $X(t)$  o el proceso en el instante  $t$ .
- Es decir,  $p_{ij}(t) = P(X(t + 1) = j | X(t) = i)$  vale:

$$\begin{aligned} p_{ij}(t) &= P(X(t + 1) = j | X(t) = i) \\ &= P(X(t + 1) = j | X(t) = i, X(t - 1) = h, X(t - 2) = g, \dots) \end{aligned}$$

## Probabilidades de transición

Las probabilidades  $p_{ij}(t)$  se llaman **probabilidades de transición**.  
La probabilidad

$$p_{ij}^{(h)}(t) = P(X(t+h) = j | X(t) = i),$$

es la probabilidad de ir desde el **estado**  $i$  hasta el **estado**  $j$  usando  $h$  **transiciones**. Dicha probabilidad se llama **probabilidad de transición de  $h$  pasos**.



## Cadenas de Markov homogéneas

Una **cadena de Markov** es **homogénea** cuando las **probabilidades de transición** no dependen del tiempo  $t$ .

Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:

## Cadenas de Markov homogéneas

Una **cadena de Markov** es **homogénea** cuando las **probabilidades de transición** no dependen del tiempo  $t$ .

Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:
  - $p_{ij} = P(X(t+1) = j | X(t) = i)$ : probabilidad de transición.

## Cadenas de Markov homogéneas

Una **cadena de Markov** es **homogénea** cuando las **probabilidades de transición** no dependen del tiempo  $t$ .

Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:

- $p_{ij} = P(X(t+1) = j | X(t) = i)$ : probabilidad de transición.
- $p_{ij}^{(h)} = P(X(t+h) = j | X(t) = i)$ : probabilidad de transición en  $h$  pasos.

## Cadenas de Markov homogéneas

Una **cadena de Markov** es **homogénea** cuando las **probabilidades de transición** no dependen del tiempo  $t$ .

Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:

- $p_{ij} = P(X(t+1) = j | X(t) = i)$ : probabilidad de transición.
- $p_{ij}^{(h)} = P(X(t+h) = j | X(t) = i)$ : probabilidad de transición en  $h$  pasos.
- $P_t(x) = P(X(t) = x)$ : distribución de la variable aleatoria  $X(t)$ .

## Cadenas de Markov homogéneas

Una **cadena de Markov** es **homogénea** cuando las **probabilidades de transición** no dependen del tiempo  $t$ .

Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:

- $p_{ij} = P(X(t+1) = j | X(t) = i)$ : probabilidad de transición.
- $p_{ij}^{(h)} = P(X(t+h) = j | X(t) = i)$ : probabilidad de transición en  $h$  pasos.
- $P_t(x) = P(X(t) = x)$ : distribución de la variable aleatoria  $X(t)$ .
- $P_0(t) = P(X(0) = x)$ : distribución inicial o en el tiempo 0.

## Ejemplo

Un ordenador es compartido por dos usuarios. Estudiamos **cuantos usuarios hay conectados cada minuto**. Nos dicen que un usuario puede **desconectarse con probabilidad 0.3** y se puede **conectar con probabilidad 0.4**.

Modelizamos la situación anterior con una **cadena de Markov  $X(t)$**  que nos da el **número de usuarios conectados al cabo de  $t$  minutos**. Los valores de  $X(t)$  pueden ser 0, 1 y 2. Por tanto, el conjunto de estados será:  $\Omega = \{0, 1, 2\}$ .

A continuación, calculemos las **probabilidad de transición**:

## Ejemplo

Si  $X(0) = 0$ , o no hay ningún usuario conectado, el número de usuarios que habrá en el siguiente minuto es una variable aleatoria **binomial** de parámetros  $n = 2$  y  $p = 0.4$ :

$$p_{00} = \binom{2}{0} \cdot 0.4^0 \cdot 0.6^2 = 0.36,$$

$$p_{01} = \binom{2}{1} \cdot 0.4^1 \cdot 0.6^1 = 0.48,$$

$$p_{02} = \binom{2}{2} \cdot 0.4^2 \cdot 0.6^0 = 0.16.$$

## Ejemplo

Si  $X(0) = 1$ , o hay un usuario conectado, el número de usuarios nuevos que habrá en el siguiente minuto es una variable aleatoria **binomial** de parámetros  $n = 1$  y  $p = 0.4$  y el número de desconexiones en el minuto siguiente será una variable aleatoria **binomial** de parámetros  $n = 1$  y  $p = 0.3$ :

$$p_{10} = 0.6 \cdot 0.3 = 0.18,$$

$$p_{11} = 0.3 \cdot 0.4 + 0.7 \cdot 0.6 = 0.54,$$

$$p_{12} = 0.7 \cdot 0.4 = 0.28.$$



## Ejemplo

Si  $X(0) = 2$ , o hay los dos usuarios conectados, el número de usuarios desconectados que habrá en el siguiente minuto es una variable aleatoria **binomial** de parámetros  $n = 2$  y  $p = 0.3$ .

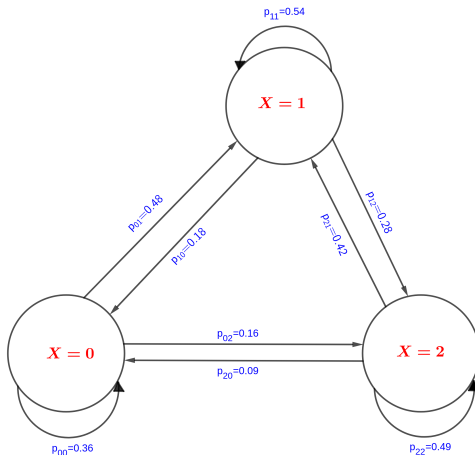
$$p_{20} = \binom{2}{0} \cdot 0.7^0 \cdot 0.3^2 = 0.09,$$

$$p_{21} = \binom{2}{1} \cdot 0.7^1 \cdot 0.3^1 = 0.42,$$

$$p_{22} = \binom{2}{2} \cdot 0.7^2 \cdot 0.3^0 = 0.49.$$

# Cadenas de Markov. Diagrama de transición

Diagrama de transición del ejemplo.



# Cadenas de Markov. Diagrama de transición

## Ejemplo en R

Las probabilidades anteriores se almacenarían en R de la forma siguiente.

Vemos que la suma de las filas vale 1:

```
P=matrix(c(0.36,0.48,0.16,0.18,0.54,0.28,0.09,0.42,0.49),  
          3,3,byrow=T)
```

P

```
##      [,1] [,2] [,3]  
## [1,] 0.36 0.48 0.16  
## [2,] 0.18 0.54 0.28  
## [3,] 0.09 0.42 0.49
```

```
apply(P,1,sum)
```

```
## [1] 1 1 1
```

# Cadenas de Markov. Diagrama de transición

## Ejemplo en python

```
import np
P=np.matrix([[0.36,0.48,0.16],[0.18,0.54,0.28],
[0.09,0.42,0.49]])
P
```

```
## matrix([[0.36, 0.48, 0.16],
##          [0.18, 0.54, 0.28],
##          [0.09, 0.42, 0.49]])
```

```
np.sum(P, axis = 1)
```

```
## matrix([[1.],
##          [1.],
##          [1.]])
```

- Las **probabilidades de transición** se pueden escribir en forma de matriz de la forma siguiente:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix},$$

donde  $\mathbf{P}$  se llama **matriz de probabilidades de transición**.

- Las **probabilidades de transición** se pueden escribir en forma de matriz de la forma siguiente:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix},$$

donde  $\mathbf{P}$  se llama **matriz de probabilidades de transición**.

- El valor  $p_{ij}$  es la **probabilidad de transición** desde el **estado  $i$**  al **estado  $j$** .

- La **matriz de probabilidades de transición** es una matriz **estocástica**, es decir la suma de las filas vale 1:

$$p_{i1} + p_{i2} + \cdots + p_{in} = 1.$$

- La **matriz de probabilidades de transición** es una matriz **estocástica**, es decir la suma de las filas vale 1:

$$p_{i1} + p_{i2} + \cdots + p_{in} = 1.$$

- Recordemos que  $p_{ij} = P(X(1) = j | X(0) = i)$ . Por tanto el vector  $(p_{i1}, \dots, p_{in})$  representaría la **función de masa de probabilidad** de la variable aleatoria discreta  $X(1) | X(0) = i$ .



## Ejemplo anterior

La **matriz de probabilidades de transición** sería la siguiente:

$$\mathbf{P} = \begin{pmatrix} 0.36 & 0.48 & 0.16 \\ 0.18 & 0.54 & 0.28 \\ 0.09 & 0.42 & 0.49 \end{pmatrix}.$$

- La **matriz de probabilidades de transición de  $h$  pasos** para la transición de  $h$  pasos sería la siguiente:

$$\mathbf{P}^{(h)} = \begin{pmatrix} p_{11}^{(h)} & p_{12}^{(h)} & \cdots & p_{1n}^{(h)} \\ p_{21}^{(h)} & p_{22}^{(h)} & \cdots & p_{2n}^{(h)} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1}^{(h)} & p_{n2}^{(h)} & \cdots & p_{nn}^{(h)} \end{pmatrix},$$

donde recordemos que  $p_{ij}^{(h)} = P(X(h) = j | X(0) = i)$ .

- La **matriz de probabilidades de transición de  $h$  pasos** es también una matriz **estocástica**:

$$p_{i1}^{(h)} + p_{i2}^{(h)} + \cdots + p_{in}^{(h)} = 1.$$

- La **matriz de probabilidades de transición de  $h$  pasos** es también una matriz **estocástica**:

$$p_{i1}^{(h)} + p_{i2}^{(h)} + \cdots + p_{in}^{(h)} = 1.$$

- Por tanto el vector  $(p_{i1}^{(h)}, \dots, p_{in}^{(h)})$  representaría la **función de masa de probabilidad** de la variable aleatoria discreta  $X(h)|X(0) = i$ .

- Empezemos con el cálculo de  $\mathbf{P}^{(2)}$ , es decir  $p_{ij}^{(2)}$ , las probabilidades de transición de 2 pasos:

$$\begin{aligned} p_{ij}^{(2)} &= P(X(2) = j | X(0) = i) \\ &= \sum_{k=1}^n P(X(1) = k | X(0) = i) \cdot P(X(2) = j | X(1) = k) \\ &= \sum_{k=1}^n p_{ik} \cdot p_{kj} = (p_{i1}, \dots, p_{in}) \cdot \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix}. \end{aligned}$$

- La probabilidad  $p_{ij}^{(2)}$  se calcula como la suma de probabilidades de ir del estado  $i$  a  $j$  pasando por el estado  $k$ , para  $k$  desde 1 hasta  $n$ :

$$i \longrightarrow k \longrightarrow j.$$

- La probabilidad  $p_{ij}^{(2)}$  se calcula como la **suma de probabilidades** de ir del **estado**  $i$  a  $j$  pasando por el **estado**  $k$ , para  $k$  desde 1 hasta  $n$ :

$$i \longrightarrow k \longrightarrow j.$$

- En resumen  $\mathbf{P}^{(2)} = \mathbf{P} \cdot \mathbf{P} = \mathbf{P}^2$ .

- En general, las **probabilidades de transición de  $h$  pasos** se pueden calcular en función de las **probabilidades de transición de  $h - 1$  pasos** razonando de forma similar:

$$\begin{aligned} p_{ij}^{(h)} &= P(X(h) = j | X(0) = i) \\ &= \sum_{k=1}^n P(X(h-1) = k | X(0) = i) \cdot P(X(h) = j | X(1) = k) \\ &= \sum_{k=1}^n p_{ik}^{(h-1)} \cdot p_{kj} = (p_{i1}^{(h-1)}, \dots, p_{in}^{(h-1)}) \cdot \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix}. \end{aligned}$$



- En general, las **probabilidades de transición de  $h$  pasos** se pueden calcular en función de las **probabilidades de transición de  $h - 1$  pasos** razonando de forma similar:

$$\begin{aligned} p_{ij}^{(h)} &= P(X(h) = j | X(0) = i) \\ &= \sum_{k=1}^n P(X(h-1) = k | X(0) = i) \cdot P(X(h) = j | X(1) = k) \\ &= \sum_{k=1}^n p_{ik}^{(h-1)} \cdot p_{kj} = (p_{i1}^{(h-1)}, \dots, p_{in}^{(h-1)}) \cdot \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix}. \end{aligned}$$

- En conclusión,

$$\mathbf{P}^{(h)} = \mathbf{P}^{(h-1)} \cdot \mathbf{P} = \mathbf{P}^{h-1} \cdot \mathbf{P} = \mathbf{P}^h.$$

## Ejemplo

Calculemos las **probabilidades de transición de 2 y 3 pasos**:

$$\begin{aligned} \mathbf{P}^{(2)} &= \mathbf{P} \cdot \mathbf{P} = \begin{pmatrix} 0.36 & 0.48 & 0.16 \\ 0.18 & 0.54 & 0.28 \\ 0.09 & 0.42 & 0.49 \end{pmatrix} \cdot \begin{pmatrix} 0.36 & 0.48 & 0.16 \\ 0.18 & 0.54 & 0.28 \\ 0.09 & 0.42 & 0.49 \end{pmatrix} \\ &= \begin{pmatrix} 0.2304 & 0.4992 & 0.2704 \\ 0.1872 & 0.4956 & 0.3172 \\ 0.1521 & 0.4758 & 0.3721 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} \mathbf{P}^{(3)} &= \mathbf{P}^{(2)} \cdot \mathbf{P} = \begin{pmatrix} 0.2304 & 0.4992 & 0.2704 \\ 0.1872 & 0.4956 & 0.3172 \\ 0.1521 & 0.4758 & 0.3721 \end{pmatrix} \cdot \begin{pmatrix} 0.36 & 0.48 & 0.16 \\ 0.18 & 0.54 & 0.28 \\ 0.09 & 0.42 & 0.49 \end{pmatrix} \\ &= \begin{pmatrix} 0.197136 & 0.493728 & 0.309136 \\ 0.185148 & 0.490704 & 0.324148 \\ 0.173889 & 0.486222 & 0.339889 \end{pmatrix} \end{aligned}$$

# Cadenas de Markov. Ejemplo en R

```
P2=P%*%P
```

```
P2
```

```
##           [,1]      [,2]      [,3]
## [1,] 0.2304 0.4992 0.2704
## [2,] 0.1872 0.4956 0.3172
## [3,] 0.1521 0.4758 0.3721
```

```
P3=P2%*%P
```

```
P3
```

```
##           [,1]      [,2]      [,3]
## [1,] 0.197136 0.493728 0.309136
## [2,] 0.185148 0.490704 0.324148
## [3,] 0.173889 0.486222 0.339889
```

# Cadenas de Markov. Ejemplo en python

```
P2 = np.dot(P,P)
```

```
P2
```

```
## matrix([[0.2304, 0.4992, 0.2704],  
##         [0.1872, 0.4956, 0.3172],  
##         [0.1521, 0.4758, 0.3721]])
```

```
P3 = np.dot(P2,P)
```

```
P3
```

```
## matrix([[0.197136, 0.493728, 0.309136],  
##         [0.185148, 0.490704, 0.324148],  
##         [0.173889, 0.486222, 0.339889]])
```

# Distribución de $X(h)$

- Dar la **distribución de la variable aleatoria**  $X(h)$ , es decir, la distribución de la transición al cabo de  $h$  **pasos** sería equivalente a dar la **función de probabilidad** que sería una tabla del tipo:

$X(h)$	1	2	$\dots$	$n$
	$P_h(1)$	$P_h(2)$	$\dots$	$P_h(n)$

# Distribución de $X(h)$

- Dar la **distribución de la variable aleatoria**  $X(h)$ , es decir, la distribución de la transición al cabo de  $h$  **pasos** sería equivalente a dar la **función de probabilidad** que sería una tabla del tipo:

$X(h)$	1	2	$\dots$	$n$
	$P_h(1)$	$P_h(2)$	$\dots$	$P_h(n)$



$$P_h(j) = P(X(h) = j) = \sum_{k=1}^n P(X(0) = k) \cdot P(X(j) = j | X(0) = k)$$

$$= \sum_{k=1}^n P_0(k) \cdot p_{kj}^{(h)} = (P_0(1), \dots, P_0(n)) \cdot \begin{pmatrix} p_{1j}^{(h)} \\ \vdots \\ p_{nj}^{(h)} \end{pmatrix}$$

# Distribución de $X(h)$

- La expresión nos da la **función de probabilidad** de  $X(h)$  en función de la **función de probabilidad** de  $X(0)$  y de las **probabilidades de transición**  $\mathbf{P}^{(h)}$ .

# Distribución de $X(h)$

- La expresión nos da la **función de probabilidad** de  $X(h)$  en función de la **función de probabilidad** de  $X(0)$  y de las **probabilidades de transición**  $\mathbf{P}^{(h)}$ .
- Matricialmente  $\mathbf{P}_h = \mathbf{P}_0 \cdot \mathbf{P}^h$ .



## Ejemplo anterior

Vamos a calcular la distribución de los usuarios conectados al cabo de 2 y 3 minutos en dos casos:

- Suponiendo que inicialmente hay un usuario conectado.
- Suponiendo que el número de usuarios conectados inicialmente es equiprobable.

# Distribución de $X(h)$

## Ejemplo anterior

En el primer caso, la **función de probabilidad** de  $X(0)$  será  $X(0) = (0, 1, 0)$ .

La **función de probabilidad** de los usuarios conectados al cabo de 2 y 3 minutos será:

$$\begin{aligned} \mathbf{P}_2 &= (0, 1, 0) \cdot \mathbf{P}^{(2)} = (0, 1, 0) \cdot \begin{pmatrix} 0.2304 & 0.4992 & 0.2704 \\ 0.1872 & 0.4956 & 0.3172 \\ 0.1521 & 0.4758 & 0.3721 \end{pmatrix} \\ &= \begin{pmatrix} 0.1872 \\ 0.4956 \\ 0.3172 \end{pmatrix}, \end{aligned}$$

## Ejemplo anterior

$$\begin{aligned}\mathbf{P}_3 &= (0, 1, 0) \cdot \mathbf{P}^{(3)} = (0, 1, 0) \cdot \begin{pmatrix} 0.197136 & 0.493728 & 0.309136 \\ 0.185148 & 0.490704 & 0.324148 \\ 0.173889 & 0.486222 & 0.339889 \end{pmatrix} \\ &= \begin{pmatrix} 0.185148 \\ 0.490704 \\ 0.324148 \end{pmatrix}.\end{aligned}$$

## Ejemplo anterior

En el segundo caso, la **función de probabilidad** de  $X(0)$  será

$$X(0) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

La **función de probabilidad** de los usuarios conectados al cabo de 2 y 3 minutos será:

$$\begin{aligned} \mathbf{P}_2 &= \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \cdot \mathbf{P}^{(2)} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \cdot \begin{pmatrix} 0.2304 & 0.4992 & 0.2704 \\ 0.1872 & 0.4956 & 0.3172 \\ 0.1521 & 0.4758 & 0.3721 \end{pmatrix} \\ &= \begin{pmatrix} 0.1899 \\ 0.4902 \\ 0.3199 \end{pmatrix}, \end{aligned}$$

## Ejemplo anterior

$$\begin{aligned}\mathbf{P}_3 &= \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \cdot \mathbf{P}^{(3)} \\ &= \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \cdot \begin{pmatrix} 0.197136 & 0.493728 & 0.309136 \\ 0.185148 & 0.490704 & 0.324148 \\ 0.173889 & 0.486222 & 0.339889 \end{pmatrix} \\ &= \begin{pmatrix} 0.185391 \\ 0.490218 \\ 0.324391 \end{pmatrix}.\end{aligned}$$

# Cadenas de Markov. Ejemplo en R

```
X01=c(0,1,0)
(dist.X21 = t(X01)%*%P2)
```

```
##           [,1]    [,2]    [,3]
## [1,] 0.1872 0.4956 0.3172
```

```
(dist.X31 = t(X01)%*%P3)
```

```
##           [,1]    [,2]    [,3]
## [1,] 0.185148 0.490704 0.324148
```

# Cadenas de Markov. Ejemplo en python

```
X01= np.matrix([0,1,0])  
dist_X21 = X01*P2  
dist_X21
```

```
## matrix([[0.1872, 0.4956, 0.3172]])
```

```
dist_X31 = X01*P3  
dist_X31
```

```
## matrix([[0.185148, 0.490704, 0.324148]])
```

# Cadenas de Markov. Ejemplo en R

```
X02=c(1/3,1/3,1/3)
(dist.X22 = t(X02)%*%P2)
```

```
##           [,1]    [,2]    [,3]
## [1,] 0.1899 0.4902 0.3199
```

```
(dist.X32 = t(X02)%*%P3)
```

```
##           [,1]    [,2]    [,3]
## [1,] 0.185391 0.490218 0.324391
```



# Cadenas de Markov. Ejemplo en python

```
X02= np.matrix([1/3,1/3,1/3])  
dist_X22 = X02*P2  
dist_X22
```

```
## matrix([[0.1899, 0.4902, 0.3199]])
```

```
dist_X32 = X02*P3  
dist_X32
```

```
## matrix([[0.185391, 0.490218, 0.324391]])
```

# Distribución en el equilibrio

- La **distribución de la cadena de Markov en el equilibrio** nos dice cuál es la distribución de  $X(h)$ ,  $\mathbf{P}_h$ , al cabo de un número muy grande de **transiciones**  $h$ .

# Distribución en el equilibrio

- La **distribución de la cadena de Markov en el equilibrio** nos dice cuál es la distribución de  $X(h)$ ,  $\mathbf{P}_h$ , al cabo de un número muy grande de **transiciones**  $h$ .
- Es decir, definimos la **distribución de la cadena de Markov en el equilibrio**  $\pi(x)$  de la forma siguiente:

$$\pi(x) = \lim_{h \rightarrow \infty} \mathbf{P}_h(x), \quad x = 1, \dots, n.$$

# Distribución en el equilibrio

- La **distribución de la cadena de Markov en el equilibrio** nos dice cuál es la distribución de  $X(h)$ ,  $\mathbf{P}_h$ , al cabo de un número muy grande de **transiciones**  $h$ .
- Es decir, definimos la **distribución de la cadena de Markov en el equilibrio**  $\pi(x)$  de la forma siguiente:

$$\pi(x) = \lim_{h \rightarrow \infty} \mathbf{P}_h(x), \quad x = 1, \dots, n.$$

- La **distribución de la cadena de Markov en el equilibrio** puede interpretarse como el porcentaje de “tiempo” que una persona pasa en cada estado  $x$  suponiendo que realiza un camino aleatorio por la cadena según la **matriz de transición de probabilidades**.

# Distribución en el equilibrio

- Vamos a dar una forma de calcular  $\pi(x)$  la **distribución de la cadena de Markov en el equilibrio**.

# Distribución en el equilibrio

- Vamos a dar una forma de calcular  $\pi(x)$  la **distribución de la cadena de Markov en el equilibrio**.
- Podemos escribir que **la función de probabilidad** de  $X(h+1)$ ,  $\mathbf{P}_{h+1}$  puede escribirse como:

$$\mathbf{P}_h = \mathbf{P}_h \cdot \mathbf{P}.$$

# Distribución en el equilibrio

- Vamos a dar una forma de calcular  $\pi(x)$  la **distribución de la cadena de Markov en el equilibrio**.
- Podemos escribir que **la función de probabilidad** de  $X(h+1)$ ,  $\mathbf{P}_{h+1}$  puede escribirse como:

$$\mathbf{P}_h = \mathbf{P}_h \cdot \mathbf{P}.$$

- Si hacemos  $h \rightarrow \infty$  en la expresión anterior, obtenemos:

$$\pi = \pi \cdot \mathbf{P}.$$

# Distribución en el equilibrio

- Vamos a dar una forma de calcular  $\pi(x)$  la **distribución de la cadena de Markov en el equilibrio**.
- Podemos escribir que **la función de probabilidad** de  $X(h+1)$ ,  $\mathbf{P}_{h+1}$  puede escribirse como:

$$\mathbf{P}_h = \mathbf{P}_h \cdot \mathbf{P}.$$

- Si hacemos  $h \rightarrow \infty$  en la expresión anterior, obtenemos:

$$\pi = \pi \cdot \mathbf{P}.$$

- Entonces se dice que la **distribución de la cadena de Markov en el equilibrio**  $\pi(x)$  es un **vector propio** de **valor propio 1** **por la izquierda** de la **matriz de transición de probabilidades**  $\mathbf{P}$ .



- En resumen, para calcular la **distribución de la cadena de Markov en el equilibrio**  $\pi(x)$ , hay que resolver:

$$\pi \mathbf{P} = \pi,$$

$$\sum_{k=1}^n \pi = 1.$$

# Distribución en el equilibrio

- En resumen, para calcular la **distribución de la cadena de Markov en el equilibrio**  $\pi(x)$ , hay que resolver:

$$\begin{aligned}\pi \mathbf{P} &= \pi, \\ \sum_{k=1}^n \pi &= 1.\end{aligned}$$

- Hallar un **vector propio por la izquierda** de la matriz  $\mathbf{P}$  es equivalente a hallar un **vector propio por la derecha** de la matriz  $\mathbf{P}^\top$  ya que si:

$$\pi \mathbf{P} = \pi, \Rightarrow \mathbf{P}^\top \cdot \pi^\top = \pi^\top.$$

## Ejemplo

El vector propio de valor propio 1 de la matriz  $\mathbf{P}^T$  cuyas componentes suman 1 y son positivas vale:

$$(0.1836735, 0.4897959, 0.3265306).$$

El vector anterior sería la **distribución de probabilidad en el equilibrio** de la cadena de Markov considerada.

# Distribución en el equilibrio en R

## Ejemplo

Para hallar la **distribución de probabilidad en el equilibrio** en R haríamos lo siguiente:

```
(vectores.propios = eigen(t(P))$vectors)
```

```
##           [,1]      [,2]      [,3]
## [1,] -0.2978565 -0.5883484  0.4082483
## [2,] -0.7942841 -0.1961161 -0.8164966
## [3,] -0.5295227  0.7844645  0.4082483
```

# Distribución en el equilibrio en R

```
(vector.propio.vap.1=vectores.propios[,1])
```

```
## [1] -0.2978565 -0.7942841 -0.5295227
```

```
(equilibrio =  
  vector.propio.vap.1/sum(vector.propio.vap.1))
```

```
## [1] 0.1836735 0.4897959 0.3265306
```

# Distribución en el equilibrio en python

```
from numpy import linalg as LA
vap, vep = LA.eig(P.T)
vep
```

```
## matrix([[ -0.29785653, -0.58834841,  0.40824829],
##         [ -0.79428407, -0.19611614, -0.81649658],
##         [ -0.52952271,  0.78446454,  0.40824829]])
```

```
equilibrio = vep[:,0]/vep[:,0].sum()
equilibrio
```

```
## matrix([[0.18367347],
##         [0.48979592],
##         [0.32653061]])
```