

# Markov Chain Analysis

# What is Markov Model?

- In [probability theory](#), a **Markov model** is a [stochastic model](#) used to [model](#) randomly changing systems where it is assumed that future states depend only on the present state and not on the sequence of events that preceded it (that is, it assumes the [Markov property](#)). Generally, this assumption enables reasoning and computation with the model that would otherwise be [intractable](#).
- Some Examples are:
  - Snake & ladder game
  - Weather system
  - .

# Assumptions for Markov model

- a fixed set of states,
- fixed transition probabilities, and the possibility of getting from any state to another through a series of transitions.
- a Markov process converges to a unique distribution over states. This means that what happens in the long run won't depend on where the process started or on what happened along the way.
- What happens in the long run will be completely determined by the transition probabilities – the likelihoods of moving between the various states.

# Types of Markov models & when to use which model



	System state is fully observable	System state is partially observable
System is autonomous	Markov Chain	Hidden Markov Model
System is controlled	Markov Decision Process	Partially observable Markov decision process

# Markov chain

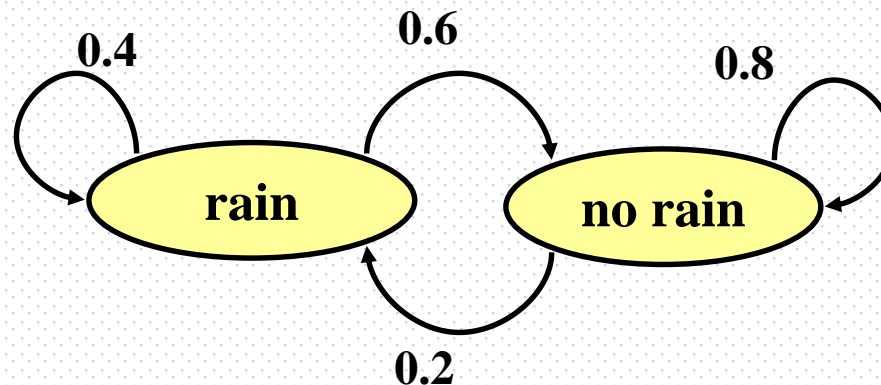
- Here system states are observable and fully autonomous.
- Simplest of all Markov models.
- Markov chain is a random process that undergoes transitions from one state to another on a state space.
- It is required to possess a property that is usually characterized as "memoryless": the probability distribution of the next state depends only on the current state and not on the sequence of events that preceded it.
- Also remember we are considering that time is moving in discrete steps.

# Lets try to understand Markov chain from very simple example

- **Weather:**

- raining today  40% rain tomorrow
- 60% no rain tomorrow
- not raining today  20% rain tomorrow
- 80% no rain tomorrow

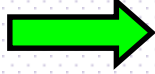
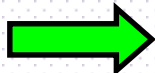

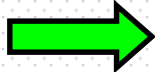
Stochastic Finite State Machine:



# Markov Process

## Simple Example

### Weather:

- raining today  40% rain tomorrow  
 60% no rain tomorrow
- not raining today  20% rain tomorrow  
 80% no rain tomorrow

### The transition matrix:

	Rain	No rain
Rain	$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$	
No rain		

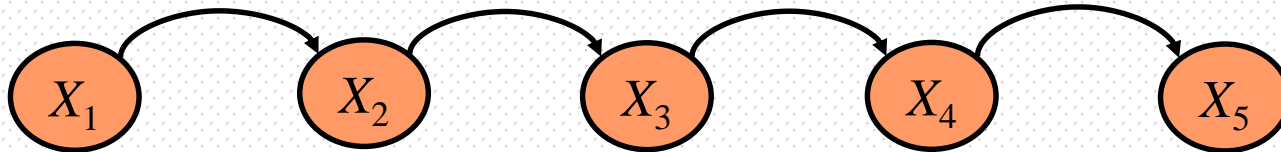
- Stochastic matrix:  
Rows sum up to 1
- Double stochastic matrix:  
Rows and columns sum up to 1

# Markov Process

Let  $X_i$  be the weather of day  $i$ ,  $1 \leq i \leq t$ . We may decide the probability of  $X_{t+1}$  from  $X_i$ ,  $1 \leq i \leq t$ .

- **Markov Property:**  $X_{t+1}$ , the state of the system at time  $t+1$  depends only on the state of the system at time  $t$

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



- **Stationary Assumption:** Transition probabilities are independent of time ( $t$ )

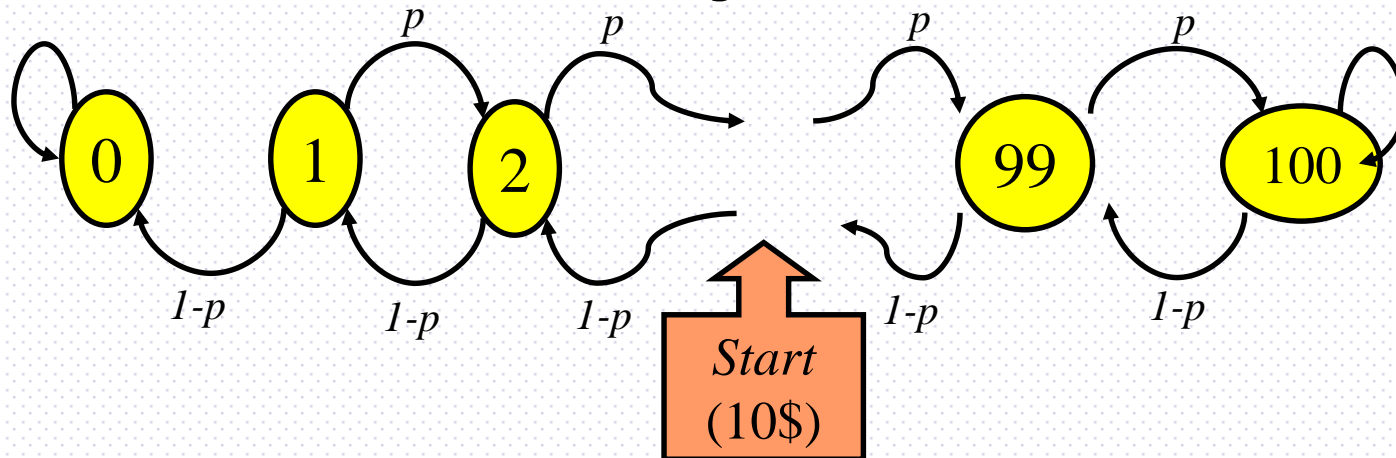
$$\Pr[X_{t+1} = b / X_t = a] = p_{ab}$$



# Markov Process

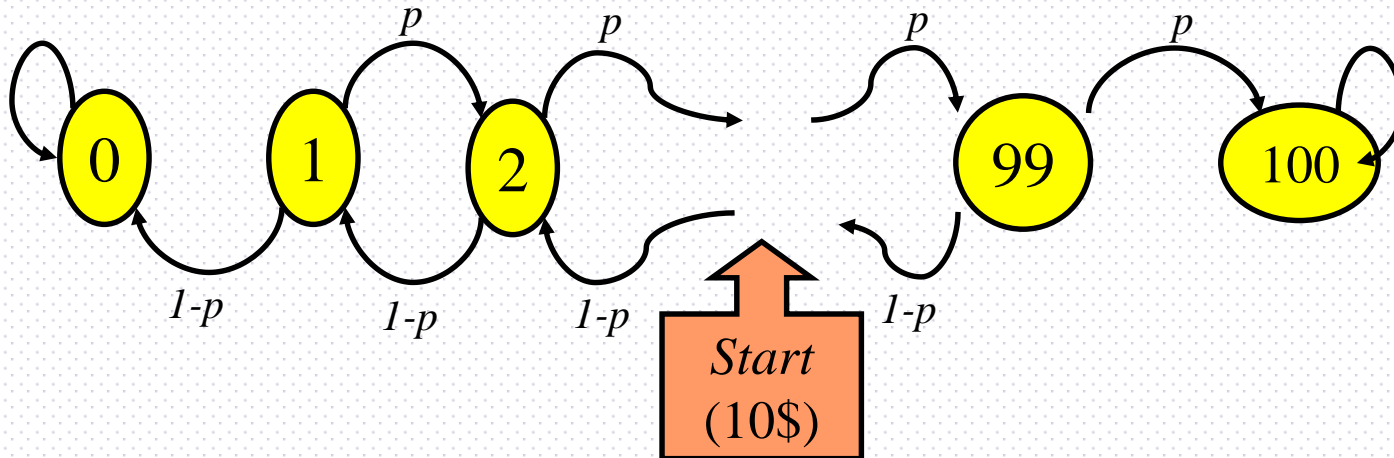
## Gambler's Example

- Gambler starts with \$10 (the initial state)
- At each play we have one of the following:
  - Gambler wins \$1 with probability  $p$
  - Gambler loses \$1 with probability  $1-p$
- Game ends when gambler goes broke, or gains a fortune of \$100  
(Both 0 and 100 are absorbing states)



# Markov Process

- **Markov process** - described by a stochastic FSM
- **Markov chain** - a random walk on this graph  
(distribution over paths)
- Edge-weights give us  $\Pr[X_{t+1}=b / X_t=a] = p_{ab}$
- We can ask more complex questions, like  $\Pr[X_{t+2} = a / X_t = b] = ?$



# Markov Process

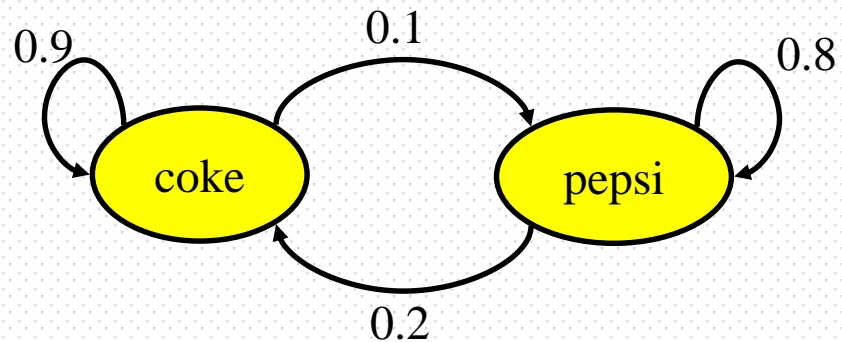
## Coke vs. Pepsi Example

- Given that a person's last cola purchase was **Coke**, there is a **90%** chance that his next cola purchase will also be **Coke**.
- If a person's last cola purchase was **Pepsi**, there is an **80%** chance that his next cola purchase will also be **Pepsi**.

transition matrix:

	coke	pepsi
coke	0.9	0.1
pepsi	0.2	0.8

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$



# Markov Process

## Coke vs. Pepsi Example (cont)

Given that a person is currently a **Pepsi** purchaser, what is the probability that he will purchase **Coke** two purchases from now?

$$\Pr[ \text{Pepsi} \rightarrow ? \rightarrow \text{Coke} ] =$$

$$\Pr[ \text{Pepsi} \rightarrow \text{Coke} \rightarrow \text{Coke} ] + \Pr[ \text{Pepsi} \rightarrow \text{Pepsi} \rightarrow \text{Coke} ] =$$
$$0.2 * 0.9 + 0.8 * 0.2 = 0.34$$

$$P^2 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

$\text{Pepsi} \rightarrow ? \quad ? \rightarrow \text{Coke}$

# Markov Process

## Coke vs. Pepsi Example (cont)

Given that a person is currently a **Coke** purchaser, what is the probability that he will buy **Pepsi** at the third purchase from now?

$$P^3 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

# Markov Process

## Coke vs. Pepsi Example (cont)

- Assume each person makes one cola purchase per week
- Suppose 60% of all people now drink Coke, and 40% drink Pepsi
- What fraction of people will be drinking Coke three weeks from now?

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

$$\Pr[X_3 = \text{Coke}] = 0.6 * 0.781 + 0.4 * 0.438 = 0.6438$$

$Q_i$  - the distribution in week  $i$

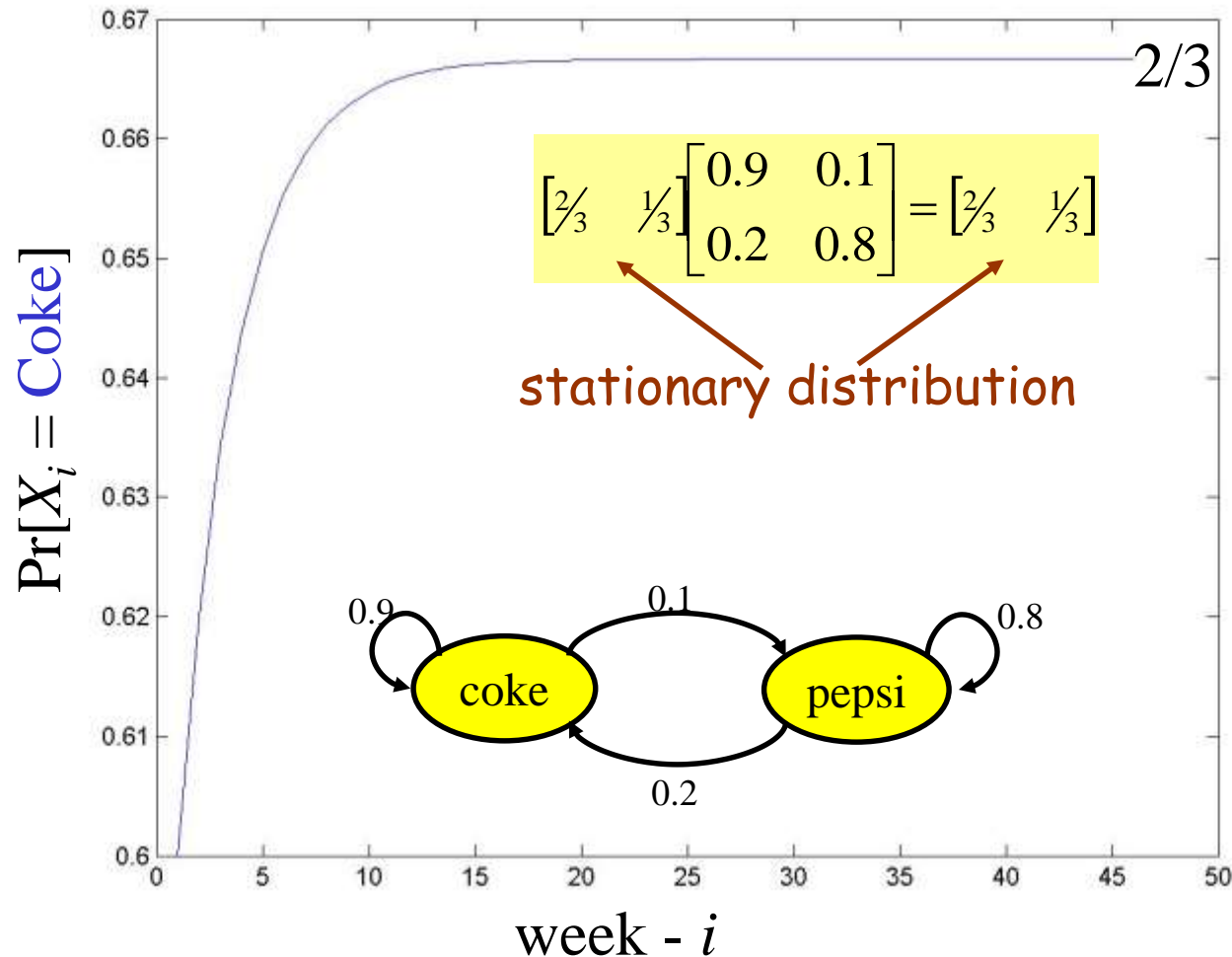
$Q_0 = (0.6, 0.4)$  - initial distribution

$$Q_3 = Q_0 * P^3 = (0.6438, 0.3562)$$

# Markov Process

## Coke vs. Pepsi Example (cont)

Simulation:



# Supervised vs Unsupervised

- ◆ Decision tree learning is “supervised learning” as we know the correct output of each example.
- ◆ Learning based on Markov chains is “unsupervised learning” as we don’t know which is the correct output of “next letter”.



# Implementation Using R

- ◆ **msm** (Jackson 2011) :handles Multi-State Models for panel data;
- ◆ **mcmcR** (Geyer and Johnson 2013) implements Monte Carlo Markov Chain approach;
- ◆ **hmm** (Himmelman and [www.linhi.com](http://www.linhi.com) 2010) fits hidden Markov models with covariates;
- ◆ **mstate** fits Multi-State Models based on Markov chains for survival analysis (de Wreede, Fiocco, and Putter 2011).
- ◆ **markovchain**

# Implementaion using R

## ◆ Example1:Weather Prediction:

The Land of Oz is acknowledged not to have ideal weather conditions at all: the weather is snowy or rainy very often and, once more, there are never two nice days in a row. Consider three weather states: rainy, nice and snowy, Given that today it is a nice day, the corresponding stochastic row vector is  $w_0 = (0, 1, 0)$  and the forecast after 1, 2 and 3 days.

Solution: please refer solution.R attached.

# Source

- ◆ Slideshare
- ◆ Wikipedia
- ◆ Google