Markov Chains

1873 Sir Francis Galton (Darwin's cousin)

"Educational times" - Will a particular English peerage survive?

Reverend Watson — Mathematical notion and provided a solution. [Branching Processes]

1906 - Markov proposed a model of dependent random variables. Motivation:- Is IID crucial for weak law of large numbers? [Borel - 1909 SLLN, Bernoulli($\frac{1}{2}$)]

Example
- distrib

~1907

- Markov ch

Today

Example Markov Considered:- Pushkin's 'Eugene.. "
- distribution of alternation of consonants & vowels

[Two state Markov chain]

~1907 Poincare - studied Markov chains on finite groups
- [Gambling Probability - French Mathematics]

- Paul Ehrenfest - proposed "Markov" model to show
Tatiana
Thermodynamics was irreversible.

- Markov chains :- challenged Doeblin, Kolmogorov, Frechet

Today :- clean & very mathematically deep & widely applied in sciences/application ain)

List of Applications

Modelling tool – Ordinary differential equations
Markov chains

- Social Sciences
- Biology (Population Genetics) are built on Markov chains

Physics – Resolves several complex phenomena.

Electrical Engineering.

Computer Science

Markov chains ..

- Simple model
- Used to Describe Complex behaviour
- most successfully applied idea in mathematics.

- Easily accessible
- not too technical

Background:

- Basic Probability
  - Conditional probability / Expectation
  - Conditional independence

tion

1. Discrete time, Discrete state space Markov chains

Intuition. - Markov chain $\{X_n : n \geq 0\}$ $\quad X_n \in S$

states

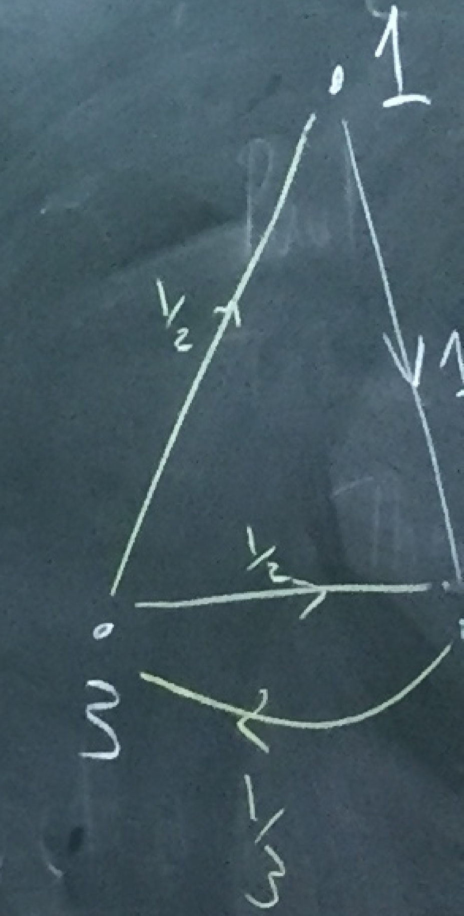
$n \in \{0\} \cup \mathbb{N}$

Key characteristics

- X_n - describing the state of the system at time n.
- It has the property of being memoryless

- best understood

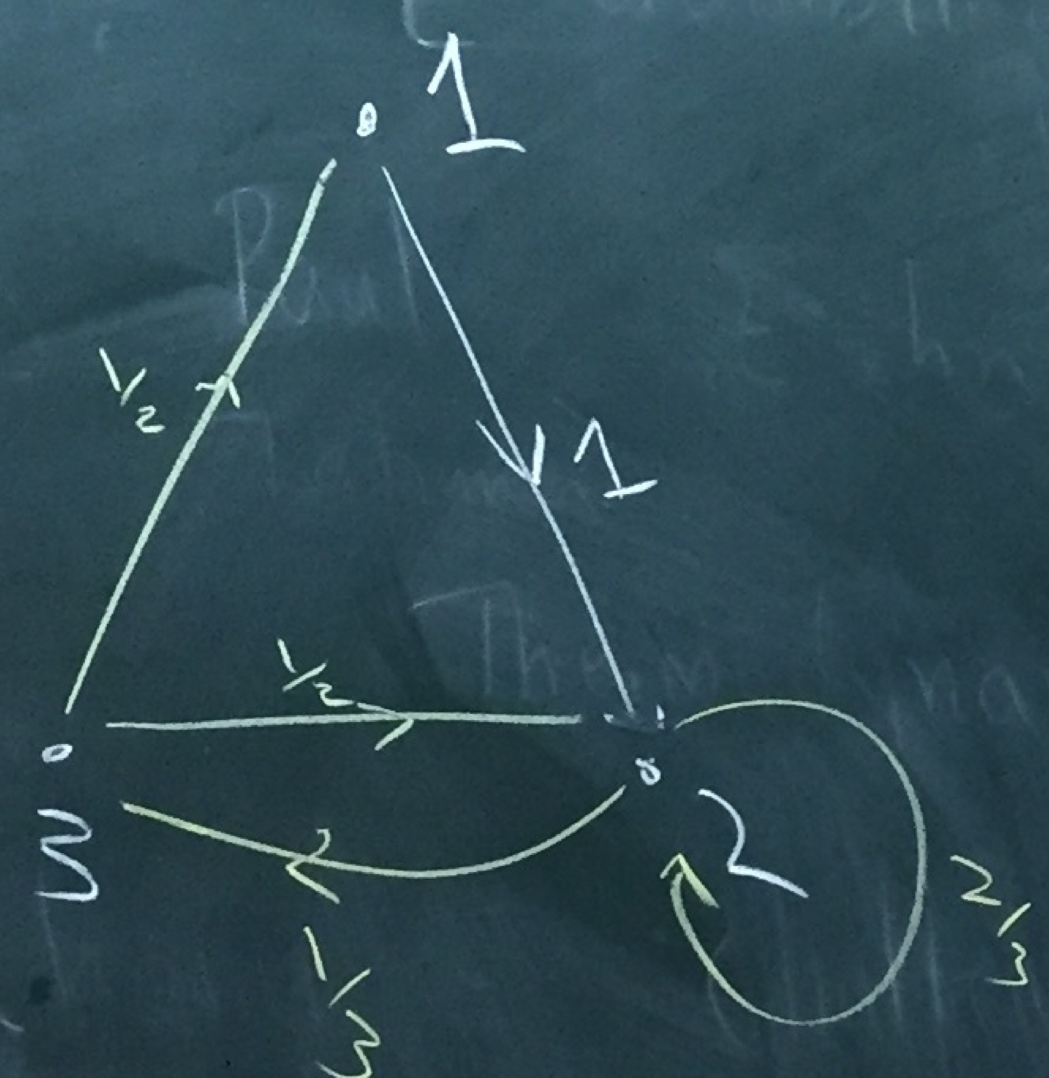
Example 1.1.1



– best understood via graphs

Example 1.1.1

$S = \{1, 2, 3\}$



Dynamics

$1 \xrightarrow{} 2$ w.p 1

$2 \xrightarrow{2/3} 2$

$2 \xrightarrow{1/3} 3$

$3 \xrightarrow{1/2} 1$

$3 \xrightarrow{1/2} 2$

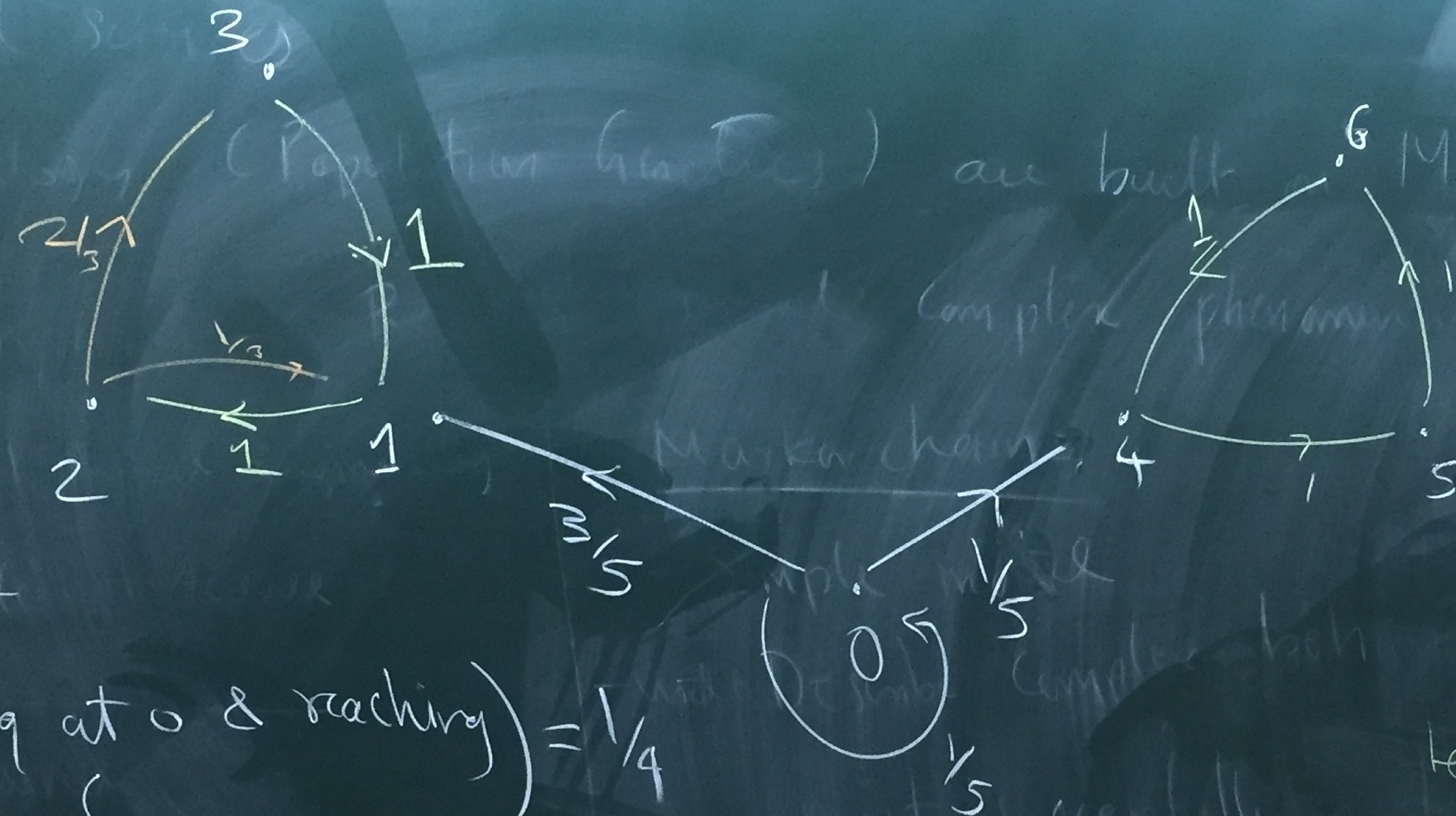
Say $X_0 = 1$

$X_n \equiv$ position of particle time n

State of system at time n

Example 1.1.2

$S = \{0,1,2,3,4,5,6\}$



- Starting at 1, on an average it takes 3 steps to reach 3.
- $p^n_{ij} = \mathbb{P}($ start at i & at time n in $j)$
- $\lim_{n\to\infty} p^n_{01} = 9/32$
- p^n_{04} does not converge, $\lim_{n\to\infty} p^{3n}_{04} = \frac{1}{24}$

$S = \{0\} \cup \{4,5,6\} \cup \{1,2,3\}$

leave '0' after finite time — periodic — aperiodic

Facts :

- $\mathbb{P}($starting at 0 & reaching $6) = 1/4$
- $\mathbb{P}($starting at 1 & reaching $3) = 1$

1.2 Discrete time Markov chain

S (set of all possible states) — a countable set ($|S| < \infty$; S – countably infinite)

$P = [p_{ij}]_{i,j \in S}$ (stochastic Matrix)

$$\begin{cases} \forall\, i,j \in S \quad p_{ij} \in [0,1] \\ \sum_{j \in S} p_{ij} = 1 \quad \text{for all } i \in S \end{cases}$$

let μ be a probability on S. That is, $\mathcal{P}(S)$ is power set of S and

Definition

let

for all $k \in \mathbb{N}$

$\{X_n\}_{n \geq}$

Ex 1. $\mu : \mathcal{P}(S) \longrightarrow [0,1]$ & (1) $\mu(S) = 1$. (2) $\{A_n\}_{n\geq 1} \in \mathcal{P}(S)$ $A_m \cap A_n = \phi$

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Definition 1.2.1 :-

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be another Probability space & for $n \in W = \{0\} \cup \mathbb{N}$

$X_n : \Omega \to S$ be random variables such that

for all $k \in W$ $\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_k = i_k) = \mu(\{i_0\}) \, p_{i_0 i_1} \, p_{i_1 i_2} \cdots p_{i_{k-1} i_k}$ (1.2.2)

where $i_0, i_1, \ldots i_k \in S$.

$\{X_n\}_{n\geq 0}$ is called a Markov chain on state space S, with initial distribution

μ, and transition matrix $P \doteq [p_{ij}]_{(i,j) \in S}$.

Remarks on Definition 1.2.1

- Assume that one can construct $\{X_n\}_{n \geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying (1.2.2) given $S, \mu(\cdot), P$
  [See chapter 4 & 6, 5]

- $\mathbb{P}(X_0 = i_0) = \mu(\{i_0\})$ by (1.2.2), $\mu(\cdot)$ specifies distribution of X_0!

(b) Distribution of X_n for some $n \in W$ is determined by $\mu \in P$

$$\mathbb{P}(X_n = j) = \mathbb{P}\Big(\bigcup_{\substack{i_k \in S \\ k=0,1,\dots n-1}} \{X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = j\} \Big)$$

($\mathbb{P}$ is Probability)

$$= \sum_{\substack{i_k \in S \\ k \in \{0,1,\dots n-1\}}} \mathbb{P}(X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = j)$$

$$\overset{(1.2.2)}{=} \sum_{\substack{i_k \in S \\ k \in \{0,1,\dots, n-1\}}} \mu(\{i_0\})\, p_{i_0 i_1}\, p_{i_1 i_2} \dots p_{i_{n-1} j}$$

(c)

LHS $\mathbb{P}$

(c) Suppose $n \in \mathbb{N}$ and $i_0, i_1 \ldots i_{n-1} \in S$ are s.t. $\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) > 0$

$$\mathbb{P}(X_n = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = \mathbb{P}(X_n = j \mid X_{n-1} = i_{n-1}) = p_{i_{n-1} j}$$

ie $X_n \mid X_0, \ldots X_{n-2}, X_{n-1} \stackrel{d}{=} X_n \mid X_{n-1}$

LHS $$\mathbb{P}(X_n = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = \frac{\mathbb{P}(X_n = j, X_0 = i_0, \ldots, X_{n-1} = i_{n-1})}{\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1})}$$

$$\overset{(12.2)}{=} \frac{\cancel{\mu(\{i_0\})}\, \cancel{p_{i_0 i_1}} \cdots \cancel{p_{i_{n-2} i_{n-1}}}\, p_{i_{n-1} j}}{\cancel{\mu(\{i_0\})}\, \cancel{p_{i_0 i_1}} \cdots \cancel{p_{i_{n-2} i_{n-1}}}} = p_{i_{n-1} j}$$

RHS

Given i_{n-1}

independent of n

$$\mathbb{P}(X_{n-1}=i_{n-1}) \geq \mathbb{P}(X_0=i_0, \dots, X_{n-2}=i_{n-2}, X_{n-1}=i_{n-1}) > 0$$

RHS $$\mathbb{P}(X_n=j \mid X_{n-1}=i_{n-1}) = \frac{\mathbb{P}(X_n=j, X_{n-1}=i_{n-1})}{\mathbb{P}(X_{n-1}=i_{n-1})}$$

$$\overset{(b)}{\equiv} \frac{\left(\sum_{\substack{i_k \in S \\ k \in \{0,1,\dots,n-2\}}} \cancel{\mu(\{i_0\}) p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}}}\right) \cdot p_{i_{n-1} j}}{\sum_{i_k \in S, k \in \{0,\dots,n-2\}} \cancel{\mu(\{i_0\}) p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}}}} = p_{i_{n-1} j}$$

or (1.2.2)

i	i_2	Past
Given	i_{n-1}	present
independent of	i_n	future