# Stochastic Processes

MATH 447

Nicholas Hayek

*Taught by Prof. Louigi Addario-Berry*

CONTENTS

We assume working knowledge of probability and no knowledge of measure theory (though a grasp of analysis is essential). See these MATH 356 notes, also taught by Louigi!

# I Markov Chains

Before providing definitions, we give some examples of stochastic processes:

**Eg. 1.1** A simple random walk: $S_{i+1} = S_i + X_i$, where $X_i \sim \text{Ber}(p)$ and $S_0 = 0$. We might ask: does $S_i$ ever return to 0, i.e.

$$\mathbb{P}(\exists i > 0 : S_i = 0)$$

**Eg. 1.2** A branching process: as in asexual reproduction, we have an initial node. Each node $n$ has a number of children $X_n$, where $\frac{X_n}{2} \sim \text{Ber}(p)$. We denote $Z_i$ to be the number of inididuals in the $i$-th generation. We might ask: does $Z_i$ ever have no children, i.e.

$$\mathbb{P}(\exists i > 0 : Z_i = 0)$$

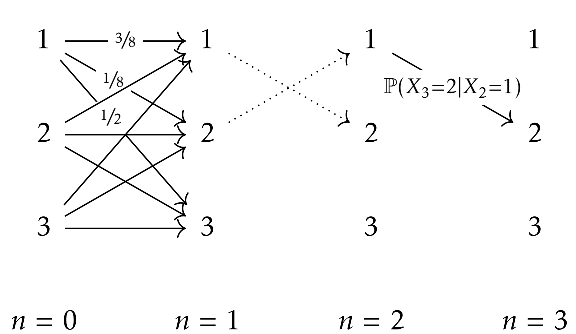
**Eg. 1.3** Choose $k$ independent random points in the square $[0, \sqrt{k}]^2$. On average, then, there is 1 point within any unit square $U \subseteq [0, \sqrt{k}]^2$.

DEF 1.1 DEF 1.2 Given a finite or countable set $V$, a *Markov chain* with *state space* $V$ is a sequence $X_n : n \geq 0$ of random variables, with $X_n \in V$, such that:

$$\mathbb{P}(\underbrace{X_{n+1} = v_{n+1}}_{\text{future}} \mid \underbrace{X_0 = v_0, \ldots, X_{n-1} = v_{n-1}}_{\text{past}}, \underbrace{X_n = v_n}_{\text{present}}) = \mathbb{P}(X_{n+1} = v_{n+1} | X_n = v_n)$$

In other words, the future only depends on the past via the present. This is called the
DEF 1.3 *Markov property*.

Sometimes we allow Markov chains to be only finitely large (i.e. $0 \leq n \leq m$). For instance, we limit ourselves to one weekend of gambling in Las Vegas. A graphical example would look something like:



By repeated Bayes' Law, we observe **PROP 1.1**

$$\begin{aligned}&\mathbb{P}(X_1 = v_1, ..., X_n = v_n | X_0 = v_0)\\ =&\mathbb{P}(X_1 = v_1|X_0 = v_0) \cdot \mathbb{P}(X_2 = v_2|X_0 = v_0, X_1 = v_1) \cdots \mathbb{P}(X_n = v_n|X_0 = v_0, ..., X_{n-1} = v_{n-1})\\ =&\mathbb{P}(X_1 = v_1|X_0 = v_0) \cdot \mathbb{P}(X_2 = v_2|X_1 = v_1) \cdots \mathbb{P}(X_n = v_n|X_{n-1} = v_{n-1}) \quad \text{by Markov property}\end{aligned}$$

### *Time-Homogeneous Markov Chains*

We often write THMC

We say that a Markov chain is *time-homogeneous* if, for all $u, v \in V$ and $n \geq 0$ DEF 1.4

$$\mathbb{P}(X_{n+1} = v|X_n = u) = \mathbb{P}(X_1 = v|X_0 = u)$$

In other words, the chain's behavior is described entirely by $\mathbb{P}(X_1 = v|X_0 = u)$ for each $(v, u) \in V \times V$. In this case, we can describe such probabilities in a *transition matrix* $P$: DEF 1.5

$$P = \left(p_{u,v}\right)_{(u,v)\in V^2} = \left(\mathbb{P}(X_1 = v|X_0 = u)\right)_{(u,v)\in V^2}$$

**Eg. 1.4** Recall the game Snakes and Ladders. A $6 \times 6$ grid is indexed $1, ..., 36$. Players start at the 1 cell. The game ends when a player reaches the 36 cell. A die roll dictates how many spots one advances. The are some directed edges between cells (increasing: "ladders", decreasing: "snakes"). One must follow these edges when one lands at its tail. Suppose a ladder exists from 11 to 27. Then

$$\mathbb{P}(X_{11} = 27|X_{10} = 6, X_9 = 3) = \frac{1}{6} = \mathbb{P}(X_{11} = 27|X_{10} = 6) = \mathbb{P}(X_2 = 27|X_1 = 6)$$

We see that Snakes and Ladders is naturally modeled as a time-homogeneous Markov chain.

**Eg. 1.5** Sampling without replacement is *not* a Markov chain. If we sample from $|X| = 10$, we have

$$\begin{aligned}&\mathbb{P}(X_3 = a\,|X_2 = b) = 1/9\\ &\mathbb{P}(X_3 = a\,|X_2 = b, X_1 = c) = 1/8\\ &\mathbb{P}(X_3 = a\,|X_2 = b, X_1 = a) = 0\end{aligned}$$

so we do not satisfy the Markov property.

**Eg. 1.6** Returning to the Snakes and Ladders example, consider $S \subseteq V$. Let $T_S = \inf\{n \geq 0 : X_n \in S\}$. We may ask...

- What is the average number of rounds to finite? We can write this as $\mathbb{E}[T_{\{36\}}|X_0 = 1]$.
- What is the probability of landing on 18 or 19 before the game ends? We can write this as $\mathbb{P}(T_{\{18,19\}} < T_{\{36\}}|X_0 = 1)$.
- What is the average number of visits to 6 before the game ends? We can write this as
$$\mathbb{E}[\#\{n \in [T_{\{36\}}] : X_n = 6\}|X_0 = 1]$$

- What is the expected proportion of time spent on state 5 before the game ends?
- If we allow two players, what is the probability that player 1 wins? Is this still a Markov chain?

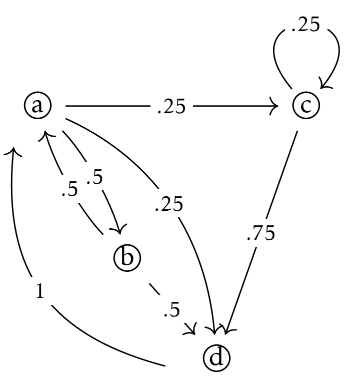
DEF 1.6 A matrix $P = \left(p_{u,v}\right)_{(u,v)\in V^2}$ is called a *stochastic matrix* if every row sums to 1, i.e.

$$\forall\ u \in V, \sum_{v\in V} p_{u,v} = 1$$

Note that any stochastic matrix is the transition matrix of some time-homogeneous Markov chain with state space $V$ and transition probabilities

$$\mathbb{P}(X_{n+1} = v | X_n = u) = \mathbb{P}(X_1 = v | X_0 = u) = p_{u,v}$$

A directed graph, together with its stochastic matrix, can visualize any THMC:




$$\longleftrightarrow \begin{bmatrix} 0 & .5 & .25 & .25 \\ .5 & 0 & 0 & .5 \\ 0 & 0 & .25 & .75 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

**Eg. 1.7** Random walks on an undirected weighted graph, where edge weights dictate the proportional probability of transitioning between two states, are a special class of THMCs. In particular, given a graph $G = (V, E)$ with weights $w(e) > 0 : e \in E$, we set

$$p_{u,v} = \frac{w(\{u,v\})}{\sum_{z\in N(u)} w(\{u,z\})}$$

If there are no edges $u \leftrightarrow v$, we write $p_{u,v} = 0$.

Not every THMC can be represented by a random walk on an undirected weighted graph. In particular, see the directed graph listed above, or any transition matrix which is not symmetric.

As a concrete example, we can consider a random walk on the number line $\mathbb{Z}$, where, if $w(k, k+1) = \alpha$, $w(k-1, k) = \frac{\alpha}{2}$.

$$\cdots \overset{\frac{1}{16}}{\text{——}} -3 \overset{\frac{1}{8}}{\text{——}} -2 \overset{\frac{1}{4}}{\text{——}} -1 \overset{\frac{1}{2}}{\text{——}} 0 \overset{1}{\text{——}} 1 \overset{2}{\text{——}} 2 \overset{4}{\text{——}} 3 \overset{8}{\text{——}} \cdots$$

### *Multi-Step Transition Probabilities*

Given a THMC $X = X_n : n \geq 0$ with a transition matrix $P$, we write

$$\begin{aligned}\mathbb{P}(X_2 = v|X_0 = u) &= \sum_{w\in V} \mathbb{P}(X_2 = v, X_1 = w|X_0 = u) \\ &= \sum_{w\in V} \mathbb{P}(X_1 = w|X_0 = u)\mathbb{P}(X_2 = v|X_1 = w, \cancel{X_0 = u}) \quad \text{by Markov property} \\ &= \sum_{w\in V} p_{u,w}p_{w,v} = (P^2)_{u,v} \quad \text{or write } P^2_{u,v}\end{aligned}$$

Hence, to determine a two-step transition probability, and by extension an $n$-step transition probability from $u$ to $w$, we consider $P^n_{u,v}$. **PROP 1.2**

PROOF. See Prop 1.1 to expand probabilities, using Bayes', as needed. We get that

$$\begin{aligned}\mathbb{P}(X_n = v|X_0 = u) &= \sum_{v_1,\dots,v_{n-1}\in V} \mathbb{P}(X_1 = v_1, \dots, X_{n-1} = v_{n-1}, X_n = v|X_0 = u) \\ &= \sum_{v_1,\dots,v_{n-1}\in V} p_{u,v_1} p(v_1, v_2)\cdots p(v_{n-1}, v) = (P^n)_{u,v} \qquad \square\end{aligned}$$

Thus, if $P$ is a stochastic matrix, then so is $P^n$. **PROP 1.3**

PROOF. $\sum_{v\in V} P^n_{u,v} = \sum_{v\in V} \mathbb{P}(X_n = v|X_0 = u) = 1.$ $\square$

If $X_n : n \geq 0$ is a THMC with state space $V$, then for all $u_0, \dots, u_{n-1}, u, v \in V$, **PROP 1.4**

$$\mathbb{P}(X_{n+m} = v|X_0 = u_0, \dots, X_{n+1} = u_{n-1}, X_n = u) = \mathbb{P}(X_{n+m} = v|X_n = u) = P^m_{u,v}$$

PROOF. One shows this by combining the Markov property with Prop 1.2 via induction. $\square$

## INDEX OF DEFINITIONS