Math 632

2020 Fall

Yuren Sun

9/2

Chap 1

Basic Review of Prob.

- probability space: $\Omega, \mathcal{F}, P$
  - Ω: sample space, set of all possible events.
  - F: subset of Ω (events)
  - P: probability measure
    - P: function $F \to \mathbb{R}$
    - $A \in F$: compute $P(A)$

- Kolmogorov's Axioms
  - $\mathbb{P}(\Omega) = 1$
  - $A_1, A_2 \ldots$ disjoint ($A_i \cap A_j = \phi$ for $i \neq j$) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

- Random Variable X
  - $X: \Omega \to \mathbb{R}$.

Example: roll 3 different dice, let X be the medium.

$\Omega = \{w = (w_1, w_2, w_3) \mid w_i \in \{1, 2, \cdots 6\}\}$

$w = (1, 2, 3)$

$X(w) = \text{medium}(w_1, w_2, w_3)$

$X(1, 2, 3) = 2$

Q: $P(X=1)$

$\underset{A}{\{x=1\}} = \{\text{all three dices are } 1\} \leftarrow A_1$

$\cup \{\text{exactly 2 dice are } 1\} \leftarrow A_2$

$P(A) = P(A_1) + P(A_2)$ $\quad * A_1 \cap A_2 = \phi$

$P(A_1) = \frac{1}{6^3}$

$P(A_2) = \frac{3 \times 5}{6^3}$

$P(X=1) = \frac{16}{6^3}$

Q: $P(X=2)$

$A_i = \{\text{exactly } i \text{ 2' rolled}\}$

$$P(X=2) = P(A_1) + P(A_2) + P(A_3)$$
$$= \frac{3\times4\times2}{6^3} + \frac{15}{6^3} + \frac{1}{6^3}$$
$$= \frac{40}{6^3}$$

- Prob distribution of a R.V. X
  $P(X \in B)$
  $\{w : X(w) \in B\}$ $\quad B \in \mathbb{R}$.

- Stochastic Process - collection of R.V.
  - index set T
  - stochastic process $(X_t)\ t \in T$
  - X_t is a R.V. for each $t \in T$
  - common example $T = \mathbb{N}$
    $(X_n)_{n=1}^{\infty}$ random sequence
    $(X_n(w))_{n=1}^{\infty}$
    vary
  - example: $T = [0, \infty)$
    $(X_t)_{t=0}$

9/4

Finite dimensional distributions of a statistic process

- $(\Omega, \mathcal{F}, P)$ probability space

R.V. $X : \Omega \to \mathbb{R}$ $\quad B \in \mathbb{R}$ $\quad$ ex. $B = [0,1)$

probability distribution : collection of probability

$P(X \in B)$

$N \in \mathbb{N}$

N random variables $X_1 \dots X_n$

$(X_1 \dots X_n) = (X_i)_{i=1}^{N} : \Omega \to \mathbb{R}^N$

↑ joint distribution

$P(X \in B)$, $B \in \mathbb{R}^N$ $\quad$ ex. $B = [0,1] \times [0,1] \times \cdots \times [0,1] \times [-1,1]$

- stochastic process : $(X_t)_{t \in T}$ collection of random variable.

$T = \mathbb{Z}_{\geq 0}$

$(X_n)_{n=0}^{\infty}$, $\omega \in \Omega$

$(X_n(\omega))_{n=0}^{\infty}$ is a $\mathbb{R}$-valued sequence.

$P((X_n)_{n=0}^{\infty} \in B)$ where $B \subset \{$all possible $\mathbb{R}$-valued sequence$\} = \mathbb{R}^{\infty}$

ex. $B = \{(X_n)_{x=0}^{\infty} : \lim_{n \to \infty} X_n = 0\}$ stochastic process converges to 0.

$P(B)$ complicated.

- finite-demensional distribution.

$K \in \mathbb{N}$ $\quad n_1 \dots n_K$.

$P((X_{n_1}, \dots X_{n_K}) \in B)$ $\quad B \in \mathbb{R}^K$

ex. Iid process

$(X_n)_{n=0}^{\infty}$ assume X_n independent and idenfically distributed.

$P((X_{n_1}, \dots X_{n_K}) \in B_1 \times \dots \times B_K)$ $\quad B_i \in \mathbb{R}$.

$= P(X_{n_1} \in B_1, \dots X_{n_K} \in B_K)$

$= P(X_{n_1} \in B_1) \times \dots \times P(X_{n_K} \in B_K)$ $\quad$ by independence.

$= P(X_0 \in B_1) \times \dots \times P(X_0 \in B_K)$ $\quad$ by idenfically distributed.

$= \prod_{i=1}^{K} P(X_0 \in B_i)$

△ Strong Law of Large Numbers.

ex. $(X_n)_{n=1}^{\infty}$ of coin flip

$$X_n = \begin{cases} 0 & \text{if } H \\ 1 & \text{if } T \end{cases}$$

Law: $\frac{1}{N}\sum_{n=1}^{N} X_n \approx E(X_1) = 0\cdot\frac{1}{2} + 1\cdot\frac{1}{2} = \frac{1}{2}$.

△ Weak Law $\varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left|\underbrace{\frac{1}{N}\sum_{n=1}^{N} X_n - E(X_1)}\right| > \varepsilon\right) = 0$$

$$= \frac{1}{N}\sum_{n=1}^{N} (\underbrace{X_n - E(X_1)}_{Y_n})$$

$E(Y_n) = 0$ $(Y_n)_{n=1}^{\infty}$ iid process.

$$P\left(\left|\frac{1}{N}\sum_{n=1}^{N} Y_n\right|^2 > \varepsilon^2\right)$$

↑ positive R.V.

Markov: $Z \geq 0$ $P(Z \geq \varepsilon^2) \leq \frac{E(Z)}{\varepsilon^2}$

$$\leq \frac{E\left|\frac{1}{N}\sum_{n=1}^{N} Y_n\right|^2}{\varepsilon^2}$$

$$= \frac{1}{\varepsilon^2 N^2}\sum_{n_1,n_2=1}^{N} E(Y_{n_1}, Y_{n_2})$$

$n_1 \neq n_2 \Rightarrow E(Y_{n_1}, Y_{n_2}) = E(Y_{n_1})E(Y_{n_2}) = 0$

$$= \frac{1}{\varepsilon^2}\frac{1}{N^2}\sum_{n=1}^{N} E(Y_n^2)$$

$$= \frac{1}{\varepsilon^2 N^2} N E(Y_1^2)$$

$$= \frac{1}{\varepsilon^2}\frac{1}{N} E(Y_1^2) = \frac{C}{N} \longrightarrow 0 \text{ (converge to } 0)$$

· Strong Law

$$P\left(\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} X_n = \frac{1}{2}\right) = 1$$

Suppose a sequence of $\mathbb{R}$ $\{a_n\}_{n=1}^{\infty}$ $a_n \geq 0$

$$\sum_{n=1}^{\infty} a_n < \infty \rightarrow \lim_{n\to\infty} a_n = 0$$

$$a_n < \sum_{n=1}^{\infty} a_n a_n < \varepsilon$$

$P\left(\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} Y_n = 0\right)$ $E(Y_n) = 0$.

Let $S_N = \sum_{n=1}^{N} Y_n$

need to choose $E\left(\sum_{n=1}^{\infty}\varphi\left(\frac{S_N}{N}\right)\right) < \infty$ $\varphi: \mathbb{R} \to \mathbb{R}_+$.

$a_N = \varphi\left(\frac{S_N}{N}\right)$

$\varphi(x) = x^2$.

$\varphi(x) = x^2 \rightarrow \sum_{N=1}^{M} \frac{C}{N} \simeq \log M \rightarrow \infty$ as $M \rightarrow \infty$ not work.

Take $\varphi(x) = x^4$

$$E\left(\frac{S_N}{N}\right)^4 = \frac{1}{N^4} \sum_{n_1,n_2,n_3,n_4=1}^{N} E(Y_{n_1} Y_{n_2} Y_{n_3} Y_{n_4})$$

Independent $E(Y_n)=0$.

$E(Y_{n_1}^3 Y_{n_4}) = 0$.

$$= \frac{1}{N^4} \sum_{n_1,n_2=1}^{N} E(Y_{n_1}^2 Y_{n_2}^2)$$

$$\leq \frac{1}{N^4} N^2 \max_{1 \leq n_1, n_2 \leq N} E(Y_{n_1}^2 Y_{n_2}^2) = \frac{1}{N^2} \max(E(Y_1^4), E(Y_1^2 Y_2^2))$$

$$= \frac{C}{N^2}$$

$$\sum_{N=1}^{\infty} E\left(\frac{S_N}{N}\right)^4 \leq \sum_{N=1}^{\infty} \frac{C}{N^2} < \infty \qquad \sum_{N=1}^{\infty} \left(\frac{S_N}{N}\right)^4 < \infty$$

$$\lim_{n\to\infty} \left(\frac{S_N}{N}\right)^4 = 0 \Rightarrow \lim_{N\to\infty} \frac{S_N}{N} = 0.$$

9/9

0 Renewal Process

- continuous time stochastic process
  each time: $(N_t)_{t\geq 0}$.

- N_t : number of occurances of a random event up until time t

ex: N_t : # of customers arriving in a store up to time t

describe in term of waiting time $(X_k)_{k\geq 1}$

X_1: time that the first 1 arrives

X_2: time elaspe b/w custome 1 & 2 arrives.

↳ 2nd customer arrives at time $X_1 + X_2$ ...

X_n: time elaspe b/w $n-1$ and n customer.

Standing assumption: waiting time random.

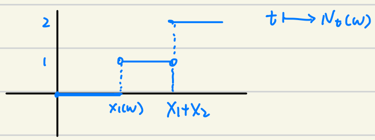
$(X_k)_{k\geq 1}$ is an iid sequence of R.V.

Basic Question: How to express N_t in terms of the waiting time $(X_k)_{k\geq 1}$

A: if $X_1 + \cdots X_n \le t$, then at least n customers have arrived by time t.
(perhaps $X_1 + \cdots X_{n+1} \le t$)
if exactly n customers arrives by time t,
then $X_1 + \cdots X_n \le t < \underset{\substack{\uparrow \\ \text{time when customer } n+1 \text{ arrives}}}{X_1 + \cdots X_{n+1}}$

. $X_n = 0$ for $t < X_1$
$N_{X_1} = 1$
for $t \geqslant X_1$, $N_t = \max\{n \geq 1 : X_1 + \cdots X_n \le t\}$

Basic Properties
① " trojectories " of the stochastic process.
· for fixed times t, N_t is a R.V. $\omega \longmapsto N_t(\omega)$
· for fixed ω, $t \longmapsto N_t(\omega)$ is a trojectary.
· a stochastic process is essentially a random function



almost surely

· N_t is integer valued
· $t \longmapsto N_t(\omega)$ is a step function with jumps at times $X_1 + \cdots + X_n$

② (long time behavior) wait long enough, expect ↓
· $\lim_{t\to\infty} N_t = \infty$ almost surely with prob 1.
· $N_t = n$ at time $X_1 + \cdots + X_n$
$X_1 + \cdots + X_n \to \infty$ consequence of the strong law of large numbers
$= \underset{\substack{\uparrow \\ \text{average converge to } E(X_1)}}{\dfrac{X_1 + \cdots + X_n}{n}} \cdot \underset{\substack{\uparrow \\ \text{goes to } \infty \text{ as } X\to\infty}}{n}$ $\Rightarrow \to \infty$ as $n \to \infty$.

△ **Strong Law for Renewal Process.**

· additional assumptions $\{X_k\}_{k=1}^{\infty}$ strictly positive $P(X_k > 0) = 1$
(arrival time, iid R.V) (two customer can not arrive at the same time)

· Thm: Let $\{X_k\}_{k=1}^{\infty}$ be iid strictly positive waiting times,
then $\lim_{t\to\infty} \frac{N_t}{t} = \frac{1}{E(X_1)}$

Intuition: $\frac{N_t}{t}$ = rate of customer arrival
→ inversely relative to the waiting time.

(Proof in Lecture Note)
Sketch: $\{X_k\}_{k=1}^{\infty}$ with SLLN
$t \to \infty$, $N_t \to \infty$
Let t be large and $N_t = n$ for n large.
SLLN: $\frac{X_1 + \cdots + X_n}{\underset{\uparrow N_t}{n}} \approx E(X_1)$

$$\Rightarrow \frac{n}{X_1 + \cdots X_n} \approx \frac{1}{E(X_1)} \approx \frac{n+1}{X_1 + \cdots + X_{n+1}}$$

If $N_t = n$, $X_1 + \cdots X_n \le t \le X_1 + \cdots + X_{n+1}$

$$\therefore \frac{1}{E(X_1)} \approx \frac{n}{X_1 + \cdots + X_n} = \frac{N_t}{X_1 + \cdots X_n} \geq \frac{N_t}{t}$$

$$\frac{1}{E(X_1)} \approx \frac{n+1}{X_1 + \cdots + X_n} = \frac{N_t + 1}{X_1 + \cdots + X_n} \leq \frac{N_t + 1}{t}$$

$$\lim_{t\to\infty} \frac{N_t + 1}{t} = \lim_{t\to\infty} \frac{N_t}{t} \Rightarrow t \to \infty \quad \frac{N_t}{t} = \frac{1}{E(X)}$$

ex. 1.20 in Lecture Note.

△ **Renewal Rewards Process**

· each customer speaks with an employee for some amounts of time

· Q: Long run fraction of time that the employee spend talking to customers?
2 sequence of R.V. $\{X_k\}_{k=1}^{\infty}$ of arrival time
$\{Y_k\}_{k=1}^{\infty}$ Y_k: amount of time customer k spend talking to an employer.

$\{R_t : t \geq 0\}$: total time spent (till time t) talking to customers.

$$R_t = \sum_{k=1}^{N_t} Y_k .$$

· Remark : Renewal process is a special case where $Y_k \equiv 1$, $\sum_{k=1}^{N_t} Y_k = N_t$.

Thm: $\lim_{t \to \infty} \frac{R_t}{t} = \frac{1}{E(X_1)} \cdot E(Y_1)$

9/11

Chap 2

o **Markov Chain**

· $(X_n)_{n=0}^{\infty}$ random sequence (Discrete time stochastic process)

state space S possible values for all X_n, assume countable ↳ finite or infinite

each X_n is discrete R.V. if S is countable

ex. joint prob dist: $P(X_0 = x_0, X_1 = x_1, X_2 = x_2)$ $x_1, x_2, x_3 \in S$

compute in terms of conditional prob.

events A, B

$P(A \cap B) = P(A|B)P(B)$

$$P(X_0 = x_0, X_1 = x_1, X_2 = x_2)$$

$$= P(X_2 = x_2 \mid X_0 = x_0, X_1 = x_1) P(X_0 = x_0, X_1 = x_1)$$

$$= \underbrace{P(X_2 = x_2 \mid X_0 = x_0, X_1 = x_1)}_{\text{looking } x_0, x_1 \text{ in the past.}} \underbrace{P(X_1 = x_1 \mid X_0 = x_0)}_{\text{looking } x_0 \text{ in the past}} P(X_0 = x_0)$$

$\Rightarrow$ complicated if X_n, $n \uparrow$

$$P(X_0 = x_0, \cdots, X_n = x_n) = P(X_n = x_n \mid X_0 = x_0, \cdots, X_{n-1} = x_{n-1}) P(X_0 = x_0, \cdots, X_{n-1} = x_{n-1})$$

$$= \prod_{k=1}^{n} P(X_k = x_k \mid X_0 = x_0, \cdots, X_{k-1} = x_{k-1}) \cdot P(X_0 = x_0)$$

↑ Markov process comes in.

only look into one in the past (nearest relevant).

reduce this to $P(X_k = x_k \mid X_{k-1} = x_{k-1})$

△ Markov Property: $P(X_k = x_k \mid X_0 = x_0, \cdots, X_{k-1} = x_{k-1}) = P(X_k = x_k \mid X_{k-1} = x_{k-1})$

for all $k \geq 1$ and $x_0 \dots x_k \in S$

△ Markov Chain: a stochastic process $(X_n)_{n=0}^{\infty}$ with a countable state space S which has the Markov property

ex. iid process. $(X_n)_{n=0}^{\infty}$ independent, identical.

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \cdots, X_n = x_n)$$

$$= P(X_{n+1} = x_{n+1}) = P(X_{n+1} = x_{n+1} \mid X_n = x_n) \quad \text{(independent)}$$

ex. Random walks.

start with iid sequence $\{\varepsilon_n\}_{n=0}^{\infty}$ take finite number of values.

$X_0 = 0$, $X_n = \sum_{k=1}^{n} \varepsilon_k$

Remark: X_2 and X_1 are dependent. $X_2 = \varepsilon_1 + \varepsilon_2$, $X_1 = \varepsilon_1$

check Markov Property:

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) \qquad X_{n+1} = \sum_{k=1}^{n+1} \varepsilon_k = \varepsilon_{n+1} + \sum_{k=1}^{n} \varepsilon_k = \varepsilon_{n+1} + X_n$$

$$= P(\varepsilon_{n+1} = x_{n+1} - x_n \mid X_0 = x_0, \dots, X_n = x_n)$$ ε_{n+1} is independent of $\varepsilon_1 \cdots \varepsilon_n$

$$= P(\varepsilon_{n+1} = x_{n+1} - x_n)$$ and hence also $X_0 \cdots X_n$

$$P(X_{n+1} = x_{n+1} \mid X_n = x_n)$$
$$= P(\varepsilon_{n+1} = x_{n+1} - x_n \mid X_n = x_n) = P(\varepsilon_{n+1} = x_{n+1} - x_n)$$

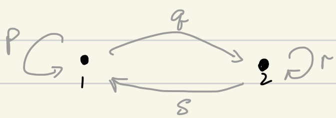
ex. Simple Random Walk.

specific case: $\varepsilon_n \in \{-1, 1\}$.

$P(\varepsilon_n = 1) = p \qquad P(\varepsilon_n = -1) = 1 - p$.

Remark: state space is infinite yet countable $S = \mathbb{Z}$ (integers)

· Two state Markov Chain $(X_n)_{n=0}^{\infty}$ $X_n \in \{1, 2\}$.



$P(X_{n+1} = 1 \mid X_n = 1) = p \qquad P(X_{n+1} = 2 \mid X_n = 2) = r$

$P(X_n = 2 \mid X_n = 1) = q \qquad P(X_{n+1} = 1 \mid X_n = 1) = s$

$p + q = 1 \qquad r + s = 1$

· Time-homogeneous Markov chains:

$P(X_{n+1} = x \mid X_n = y)$ does not depend on n

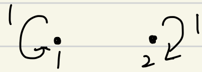
· Transition Matrix

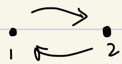
convenient representation.

$$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & p & q \\ 2 & s & r \end{array}$$

← future state.

current state

Degenerate cases: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 
(deterministic)

$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  $X_0=1 \to X_1=2 \to X_2=1 \ldots$
$X_0=2 \to X_1=1 \to \cdots$

$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ random case.

9/14

- **Transition probabilities**
- Markov chain $\{X_n\}_{n=0}^{\infty}$ $P(X_{n+1}=x_{n+1} \mid X_n=x_n, \ldots, X_0=x_0) = P(X_{n+1}=x_{n+1} \mid X_n=x_n)$

Define transition Probability.

$P: S \times S \to [0,1]$

$P(x,y) = P(X_{n+1}=y \mid X_n = x)$

starting point one step later.

$\sum_{y \in S} P(x,y) = 1$

$\hookrightarrow = \sum_{y \in S} \frac{P(X_n=x, X_{n+1}=y)}{P(X_n=x)} = \frac{P(X_n=x)}{P(X_n=x)} = 1$

Def: a transition probability is a function $p: S \times S \to [0,1]$ s.t.
$\sum_{y \in S} P(x,y) = 1$

- **Initional Distribution**

$\mu: S \to [0,1]$ $\mu(x) := P(X_0 = x)$

Have initional distribution + transition probability $\Rightarrow$ able to compute all the Markov chain

$P(X_0=x_0, \ldots, X_n=x_n) = P \bigcap_{k=0}^{n} \{X_k = x_k\} = \mu(x_0) P(x_0, x_1) P(x_1, x_2) \cdots P(x_{n-1}, x_n).$

$= \mu(x_0) \prod_{k=0}^{n-1} P(x_k, x_{k+1}).$

· **Claim:** $P(X_m = x_m, \ldots, X_{m+n} = x_{m+n})$

$= P(X_m = x_m) \prod_{k=1}^{n-1} P(X_{m+k}, X_{m+k+1})$

Proof: $P(X_{n+1} = x_{n+1} \mid X_n = x_n, \ldots, X_0 = x_0) = P(x_n, x_{n+1})$

For $P(X_{n+1} = x_{n+1} \mid X_n = x_n, \ldots X_m = x_m)$ $0 < m < n$.

intuition: only care about one step back.

Claim: $P(X_{n+1} = x_{n+1} \mid X_n = x_n, \ldots, X_m = x_m)$ $0 < m \leq n$

$= P(x_n, x_{n+1})$

$P(X_3 = x_3 \mid X_2 = x_2, X_1 = x_1) = P(x_2, x_3)$

$= \frac{P(X_1 = x_1, \ldots, X_3 = x_3)}{P(X_1 = x_1, X_2 = x_2)} = \sum_{x_0 \in S} \frac{P(X_0 = x_0, \ldots X_3 = x_3)}{P(X_1 = x_1, X_2 = x_3)}$

$= \sum_{x_0 \in S} \frac{P(X_0 = x_0, \ldots X_3 = x_3)}{P(X_0 = x_0, \ldots X_2 = x_2)} \cdot \frac{P(X_0 = x_0, \ldots X_2 = x_2)}{P(X_1 = x_1, X_2 = x_2)}$

$= P(x_2, x_3) \sum_{x_0 \in S} \frac{P(X_0 = x_0, \ldots X_2 = x_2)}{P(X_1 = x_1, X_2 = x_2)}$

$= P(x_2, x_3) \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_1 = x_1, X_2 = x_2)}$

$= P(x_2, x_3)$

Prove by induction on n check for $n=1$, $\Rightarrow n \to n+1$.

$n=1$, $P(X_m = x_m, X_{m+1} = x_{m+1}) = P(X_{m+1} = x_{m+1} \mid X_m = x_m) P(X_m = x_m)$

$= P(X_m = x_m) P(x_m, x_{m+1})$

Inductive step. assume $(*)_m$ (work at $k=m$)

claim: $(*)_{m+1}$

$P(X_m = x_m, \ldots X_{m+n+1} = x_{m+n+1})$

$= P(X_{m+n+1} = x_{m+n+1} \mid X_m = x_m, \ldots, X_{m+n} = x_{m+n}) \cdot P(X_m = x_m, \ldots, X_{m+n} = x_{m+n})$

$$= p(m+n, m+n+1) \cdot P(X_m = x_m) \prod_{k=0}^{n-1} (X_{m+k}, X_{m+k+1})$$
$$= P(X_m = x_m) \prod_{k=0}^{n} (X_{m+k}, X_{m+k+1})$$

ex. win 1 dollar with prob. p. lose 1 dollar with prob 1-p.
run out of money, stop playing
if your net worth $= 3$ stop playing.
$(X_n)_{n=0}^{\infty}$, X_n winnings at time n
$S = \{0, 1, 2, 3\}$. $P(i,j) = P(X_{n+1} = j \mid X_n = i)$
transition matrix:

$$\begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 1 & 1-p & 0 & p & 0 \\ 2 & 0 & 1-p & 0 & p \\ 3 & 0 & 0 & 0 & 1 \end{array}$$

stop playing at 0

stop playing at 3.

$\sum_{y \in S} p(x,y) = 1$
row col

9/16

Markov Chains:
ex. build with iid $\{\varepsilon_n\}_{n=1}^{\infty}$ $X_n = \sum_{k=1}^{n} \varepsilon_k$ ← random walk. markov chain
special case $X_{n+1} = \varepsilon_{n+1} + X_n$
$H(x,y) = x+y$.

ex. $X_n = \max(\varepsilon_1, \dots, \varepsilon_n)$ (markov chain)

useful fact: if $X_{n+1} = H(X_n, \varepsilon_{n+1})$
for some deterministic function $(x,y) \to H(x,y)$
Then $\{X_n\}_{n=0}^{\infty}$ is a markov chain.
only depend on one step before.

$X_{n+1} = \max(\varepsilon_1, \dots, \varepsilon_{n+1}) = \max\left(\max(\varepsilon_1, \dots, \varepsilon_n), \varepsilon_{n+1}\right)$
$H(x,y) = \max(x,y)$.

Non-ex Let $X_n = \text{medium}(\xi_1, \dots, \xi_n)$, where ξ_i is a dice roll.

Claim: $\{X_n\}_{n=0}^{\infty}$ is not a Markov chain.

. $P(X_3 = 3 \mid X_1 = 1, X_2 = 2) = \dfrac{P(X_1 = 1, X_2 = 2, X_3 = 3)}{P(X_1 = 1, X_2 = 2)}$

$\{X_1 = 1, X_2 = 2\} = \{\xi_1 = 1, \xi_2 = 3\}$.

$P(X_1 = 1, X_2 = 2) = \frac{1}{6^2}$

$\{X_1 = 1, X_2 = 2, X_3 = 3\} = \{\xi_1 = 1, \xi_2 = 3, 3 \le \xi_3 \le 6\}$.

$P(X_1 = 1, X_2 = 2, X_3 = 3) = \frac{1}{6^2} \times \frac{4}{6}$

$P(X_3 = 3 \mid X_1 = 1, X_2 = 2) = \dfrac{\frac{4}{6^3}}{\frac{1}{6^2}} = \frac{2}{3}$.

. $P(X_3 = 3 \mid X_2 = 2) = \dfrac{P(X_2 = 2, X_3 = 3)}{P(X_2 = 2)}$.

$\{X_2 = 2\} = \{\text{medium}(\xi_1, \xi_2) = 2\} = \{\xi_1 = \xi_2 = 2\} \cup \{\xi_1 = 1, \xi_2 = 3\} \cup \{\xi_1 = 3, \xi_2 = 1\}$

$P(X_2 = 2) = \frac{3}{6^2}$

$P(\xi_1 = \xi_2 = 2, X_3 = 3) = 0$

$P(\xi_1 = 1, \xi_2 = 3, X_3 = 3) = \frac{4}{6}$

$P(\xi_2 = 1, \xi_1 = 3, X_3 = 3) = \frac{4}{6}$.

$P(X_2 = 2, X_3 = 3) = \frac{8}{6^3}$

$P(X_3 = 3 \mid X_2 = 2) = \dfrac{\frac{8}{6^3}}{\frac{3}{6^3}} = \frac{4}{9} \neq \frac{2}{3} = P(X_3 = 3 \mid X_1 = 1, X_2 = 2)$

$\therefore$ We showed that $P(X_3 = 3 \mid X_1 = 1, X_2 = 2) \neq P(X_3 = 3 \mid X_2 = 2)$,

not a Markov chain.

△ Multi-step Transition Probabilities

$P(x, y) = P(X_{n+1} = y \mid X_n = x)$

transition matrix $P\{P(x,y)\}_{x,y \in S}$. (1 step transition probabilities).

· $P(X_{n+m} = y \mid X_n = x) \leftarrow$ multi-step transition probability.

can be calculated with transition matrix P.

Claim: $P(X_{n+m} = y \mid X_n = x) = p^m(x,y)$

Recall: $p^2(x,y) = \sum_z p(x,z)\,p(z,y)$

$p^3(x,y) = \sum_{z_1} p(x,z_1)\,p^2(z_1,y) = \sum_{z_1} p(x,z_1) \sum_{z_2} p(z_1,z_2)\,p(z_2,y)$

$= \sum_{z_1 z_2} p(x,z_1)\,p(z_1,z_2)\,p(z_2,y)$

$\vdots$

$p^m(x,y) = \sum_{z_1 \cdots z_{m-1}} p(x,z_1)\,p(z_1,z_2) \cdots p(z_{m-1},y)$

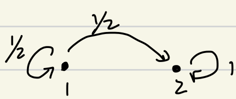
$P(X_n = x, X_{n+1} = z_1, \cdots, X_{n+m-1} = z_{m-1}, X_{n+m} = y)$

$= P(X_n = x)\,\underbrace{p(x,z_1)\,p(z_1,z_2) \cdots p(z_{n-1},y)}$

$p^m(x,y) = \sum_{z_1 \cdots z_{m-1}} \dfrac{P(X_n = x, X_{n+1} = z_1, \cdots, X_{n+m} = y)}{P(X_n = x)}$

$= \dfrac{P(X_n = x, X_{n+m} = y)}{P(X_n = x)} = P(X_{n+m} = y \mid X_n = x)$

ex. 2 states



$$P = \begin{array}{c|cc} & 1 & 2 \\ \hline 1 & \frac{1}{2} & \frac{1}{2} \\ 2 & 0 & 1 \end{array}$$

long time behavior of $(X_n)_{n=0}^{\infty}$

$\lim_{x \to \infty} P(X_n = 2 \mid X_0 = 1) = ?$

p^n ? $\quad p^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ 0 & 1 \end{bmatrix}$

p^n is a stochastic matrix, i.e. row sums up to 1.

$p^n = \begin{bmatrix} p_n & 1-p_n \\ 0 & 1 \end{bmatrix}$

$p^{n+1} = \begin{bmatrix} p_n & 1-p_n \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}p_n & 1-\frac{1}{2}p_n \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} p_{n+1} & 1-p_{n+1} \\ 0 & 1 \end{bmatrix}$

$p_{n+1} = \frac{1}{2} p_n$.

$\Rightarrow p^n = \begin{bmatrix} 2^{-n} & 1-2^{-n} \\ 0 & 1 \end{bmatrix}$.

$\lim_{n\to\infty} P(X_n = 2 \mid X_0 = 1) = \lim_{n\to\infty} p^n(1,2) = \lim_{n\to\infty} 1 - 2^{-n} = 1$.

9/18

- Remark on notations

  $(X_n)_{n=0}^{\infty}$ w/ inition distribution μ (probability distribution of X_0)

  we write P_μ associate prob measure.

  transition prob (P, μ) $\quad$ P_μ v.s. P_x for $x \in S$.

  $\uparrow$ initial condition is deterministic.

  If $P(X_0 = x) = 1$

- Given a markov chain $(X_n)_{n=0}^{\infty}$ w/ transition matrix p and i.d. μ (↙ initial distribution)

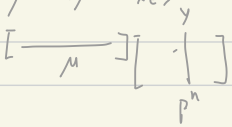
  compute $P_\mu(X_n = y) = \sum_{x_0, \dots, x_{n-1}} P_\mu(X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = y)$

  $= \sum_{x_0 \dots x_{n-1}} \mu(x_0) P(x_0, x_1) \cdots P(x_{n-1}, y)$

  $= \sum_{x_0} \mu(x_0) \sum_{x_1 \dots x_{n-1}} P(x_0, x_1) \cdots P(x_{n-1}, y)$

  $= \sum_{x_0} \mu(x_0) (P^n)(x_0, y)$

  $\mu = (\mu(x))_{x \in S}$



  $$P_\mu(X_n = y) = (\mu P^n)(y)$$

- Morel: the probability dist of X_n is encoded in μP^n

- Remark: $P_\mu(X_{n+1} = y \mid X_n = x) = P(x, y) = P_x(X_1 = y)$

  $\uparrow$ start over, fix determinist $X_0 = x$, 1 step go to y.

  $P_\mu(X_{n+1} = y_1, \dots, X_{n+m} = y_m \mid X_n = x)$

  $= \dfrac{P_\mu(X_n = x, X_{n+1} = y_1, \dots, X_{n+m} = y_m)}{P_\mu(X_n = x)}$

  $= \dfrac{P_\mu(X_n = x) P(x, y_1) \cdots P(y_{m-1}, y_m)}{P_\mu(X_n = x)}$

  $= P(x, y_1) \cdots P(y_{m-1}, y_m) = P_x(X_1 = y_1, X_2 = y_2, \cdots X_m = y_m)$

- Thm: (2.21) True also for $n = \infty$. (example 2.22)

- stopping times / strong markov

$P_\mu(X_{n+1}=y_1, \cdots, X_{n+m}=y_m \mid X_n=x) = P_x(X_1=y_1, \cdots, X_m=y_m)$

- what happens if n is random?

$P_\mu(X_{T+1}=y \mid X_T=x) = P_x(X_1=y)$ if T is a random time

ex. gambler's ruin on $\{0, \cdots, 100\}$ (similar to random walk)

- T_f = first time you have more than 99 dollars
- T_l = last time you have - - - - - - - - -

- Is $P_\mu(X_{T+1}=y \mid X_T=x) = P_x(X_1=y)$?

Yes for $T=T_f$, No for $T=T_l$.

$P_\mu(X_{T+1}=100 \mid X_T=99)$

$\to P_\mu(X_{T_l+1}=100 \mid X_{T_l}=99) = 0$

$P_{99}(X_1=100) = 1/2$ $\neq$

- stopping Time: A random variable T is a stopping time if.

informal: knowing $T=n$ requires only knowledge of $X_0, \cdots, X_n$.

rigurous: for each $n \in \mathbb{N}$, there exists $C_n \in S^{n+1}$ s.t.

$\{T=n\} = \{(X_0, X_1, \cdots, X_n) \in C_n\}$

- Thm: strong markov property.

for every stopping time T, $P_\mu(X_{T+1}=y_1, \cdots, X_{T+m}=y_m) = P_x(X_1=y_1, \cdots, X_m=y_m)$

Proof: ($m=1$ for simplicity)

Goal: show $P_\mu(X_{T+1}=y \mid X_T=x) = P_x(X_1=y)$

equivalantly, $P_\mu(X_T=x, X_{T+1}=y) = P(X_T=x)P_x(X_1=y)$

$$
\begin{aligned}
P_\mu(X_T=x, X_{T+1}=y) &= \sum_{n=0}^{\infty} P_\mu(X_{n+1}=y, X_n=x, T=n) \\
&= \sum_{n=0}^{\infty} P_\mu(X_{n+1}=y \mid X_n=x, T=n)\, P(X_n=x, T=n)
\end{aligned}
$$

$\uparrow$ depends only on $X_0, \cdots, X_{n-1}, X_n$.

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} P_\mu(X_{n+1}=y \mid X_n=x)\, P(X_n=x, T=n) = P_x(X_1=y) \sum_{n=0}^{\infty} P(X_n=x, T=n) \\
&= P(X_1=y)\, P(X_T=x).
\end{aligned}
$$

9/21

▷ Recurrance and Transience

$T_x = \inf\{\underbrace{n \geq 1}_{\text{exclude } n=0} : X_n = x\}$ for $x \in S$ "First hitting time of state x"

T_x is a stopping time. $\{T_x = n\}$ depend on process only up till n. ← stopping time.

$= \{X_1 \neq x, \cdots, X_{n-1} \neq x, X_n = x\}$.

· classify $x \in S$ either be Recurrance or Transience

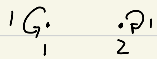
$f_{xx} = P_x(T_x < \infty)$ finite.

$= P_x(X_n = x$ for some $n \geq 1)$

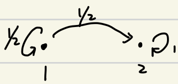
A state x in our state space ($x \in S$) is Recurrant if $f_{xx} = 1$ (start at x and eventually reach x)

Transient if $f_{xx} < 1$

· Two State Markov chain

ex.  $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ · both states 1 & 2 are recurrent.

$f_{11} = 1$, $f_{22} = 1$

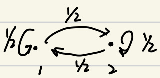
ex.  $\begin{bmatrix} 1/2 & 1/2 \\ 0 & 1 \end{bmatrix}$ · state 2 is recurrant.

state 1 is transient.

$P_1(T_1 < \infty) = 1$ iff $P_1(T_1 = \infty) = 0$

$P_1(T_1 = \infty) = P_1(X_n = 2$ for some $n \geq 1) \geq P_1(X_1 = 2) = P(1,2) = 1/2$.

Hence, $P_1(T_1 = \infty) \neq 0$

ex.  $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ · Both states 1 & 2 are Recurrant.

Claim: $P_1(T_1 = \infty) = 0$ Also true for state 2. $P_2(T_2 = \infty) = 0$

$P_1(T_1 = \infty) = P_1(X_n = 2$ for all $n \geq 1)$ always stay in state 2

$\leq P_1(X_1 = 2, \cdots, X_n = 2)$

$= P(1,2)\underbrace{P(2,2) \cdots P(2,2)}_{n-1}$

$= 1/2 \cdot (1/2)^{n-1} = 2^{-n} \to 0$ as $n \to \infty$.

$P_1(T_1 = \infty) = \infty$.

ex. Gambler's Ruin   Simple random walk. win \$1 or lose \$n, stop at 0 or n $S = \{0, \dots, N\}$.

$$\begin{bmatrix} 1, 0, \dots, 0 \\ 1/2, 0, 1/2, \dots 0 \\ \dots \\ 0, 0, 0 - 1/2, 0, 1/2 \\ 0, 0, \dots, 1 \end{bmatrix}$$

claim: 0 and N are recurrant, $1 \sim n-1$ are transient.

$j \in \{1, \dots, N-1\}$

$P_j(T_j = \infty) \geq P_j(X \in \{0, N\}$ for some $N \geq 1)$ lower bound, show $P_j(T_j = \infty) \neq 0$.

$\geq P_j(X_1 = j-1, X_2 = j-2, \dots, X_n = 0)$

$= P(j, j-1)\, P(j-1, j-2) \cdots P(1, 0) = 2^{-j} \neq 0.$

△ Successive Returns to a State

Goal: show a transience / recurrance dichotomy.

1. A markov chain started in a recurrent state visits that state infinitely often
2. A - - - - - - - - - - - - - transient - - - - - - - finitely many times.

Remark: (Gambler's Ruin) $\{0, \dots, N\}$

start at 0 or N, always be there.

start at $\{1, \dots, N\}$, stuck at 0 or N at some time.

· $T_y^k = \inf\{n > T_y^{k-1} \mid X_n = y\}$ if $T_y^{k-1} < \infty$.   first time visit y after k-1 visit to y.

· Lemma: $P(T_y^k < \infty) = \rho_{yy}^k$

Proof: $P_y(T_y^k < \infty) = P_y(T_y^{k-1} < \infty, T_y^k = \infty)$

$= \underbrace{P_y(T_y^k < \infty \mid T_y^{k-1} < \infty)}\, P(T_y^{k-1} < \infty)$

$= P_y(T_y < \infty)\, P(T_y^{k-1} < \infty)$

strong markov property: $\{X_{T_y^{k-1}+n}\}_{n=1}^{\infty}$ has the same finite dim dist as $\{X_n\}_{n=1}^{\infty}$ started from $X_0 = y$

$P_y(T_y^k < \infty) = \rho_{yy} P(T_y^{k-1} < \infty) \rightarrow P_y(T_y^k < \infty) = \rho_{yy}^k$

$P(T_y^{k-1} < \infty, T_y^k < \infty)$ $\quad T_y^{k-1} < T_y^k$

$= \sum_{n,m=1}^{\infty} P(T_y^{k-1} = n, T_y^k = n+m)$

$= \sum_{n,m=1}^{\infty} P(X_n = X_{m+n} = y, X_k \neq y \text{ for } n+1 \leq k \leq n+m-1, T_y^{k-1} = n)$

$= \sum_{n,m=1}^{\infty} P(X_{n+m} = y, X_k \neq y \text{ for } n+1 \leq k \leq n+m+1 \mid X_n = y, T_y^{k-1} = n) \cdot P(T_y^{k-1} = n, X_n = y)$

$= \sum_{n,m=1}^{\infty} P_y(X_m = y, X_k \neq y \text{ for } 1 < k < m-1) \cdot P(T_y^{k-1} = n)$

$= \rho_{yy} \cdot P(T_y^{k-1} < \infty)$

$\rightarrow P(T_y^k < \infty \mid T_y^{k-1} < \infty) = \rho_{yy}.$

Theorem: If x is recurrant, then $P_x(T_y^k < \infty \text{ for all } k \geq 1) = 1$

If x is transient, then $P_x(T_x^k < \infty \text{ for all } k \geq 1) = 0$

Proof: (Recurrant Case)

$P_x(T_y^k = \infty \text{ for some } k \geq 1) \leq \sum_{k \geq 1}^{\infty} P(T_y^k = \infty) = 0$ $\quad P(T_y^k < \infty) = \rho_{yy}^k = 1$

Proof: (Transient case)

$P_x(T_y^k < \infty \text{ for all } k \geq 1) \leq P_x(T_y^n < \infty) = \rho_{yy}^n \rightarrow 0$ since $\rho_{yy} < 1$

Hence $P_y(T_y^k < \infty \text{ for all } k \geq 1) = 0$

$P_y(T_y^k = \infty \text{ for some } k \geq 1) = 1$

9/23.

Transience v.s. Recurrance.

$\rho_{xy} = P_x(T_y < \infty)$ where T_y is the first time that the process ends up at y

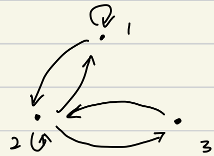
$T_y = \inf\{n \geq 1 : x_n = y\}$

- x is recurrant if $\rho_{xx} = 1$
- ... transient ... $\rho_{xx} < 1$.

△ **Accesibility** $x \to y$

y is accessible from x if $P^{(n)}(x,y) > 0$ for some $n \geq 0$.

ex. $P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1 & 0 \end{bmatrix}$ $S = \{1, 2, 3\}$.



$1 \to 2$ since $P(1,2) > 0$

$1 \to 3$ $P(1,3) = 0$ but $P^{(2)}(1,3) = \sum_z P(1,z)P(z,3) \geq P(1,2)P(2,3) > 0$

$2 \to 1$ $P(2,1) > 0$

$2 \to 3$ $P(2,3) > 0$

$3 \to 1$ $P^{(2)}(3,1) = \sum_z P(3,z)P(z,1) \geq P(3,2)P(2,1) > 0$

$3 \to 2$ $P(3,2) > 0$

△ **Finding Transient States**

Recall: Gambler's Rules on $\{0, \cdots, N\}$ 0 and N are recurrant, $1 \cdots N-1$ are transient

- Lemma: suppose $\rho_{xy} > 0$, but $\rho_{yx} < 1$, then x is transient

Proof: since $\rho_{xy} > 0$, there exists a k s.t. $p^{(k)}(x,y) > 0$.

Let $K = \min\{k \geq 1 : p^{(k)}(x,y) > 0\}$.

$p^{(k)}(x,y) = \sum_{z_1 \cdots z_{k-1}} P(x_1, z_1) P(z_1, z_2) \cdots P(z_{k-1}, y)$

Since $p^{(k)}(x,y) > 0$, $\exists\, z_1 \cdots z_{k-1} \in S$ s.t. $P(x, z_1) \cdots P(z_{k-1}, y) > 0$, $z_j \neq x$

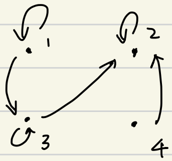
Want to show: $P_x(T_x = \infty) > 0$ (so x is transient).

$P_x(T_x = \infty) \geq P(x_1 = z_1, \cdots, X_{k-1} = z_{k-1}, X_k = y, X_n \neq x \text{ for } n > k)$

$= P_x(X_n \neq k \text{ for } n > k \mid X_1 = z_1 \cdots X_k = y) \cdot P(x, z_1) \cdots P(z_{k-1}, y)$

$= \underbrace{P_y(X_n \neq x \text{ for } n \geq 1)}_{(1 - \rho_{yx})} p(x, z_1) \cdots p(z_k, y) > 0$

ex



$$\begin{bmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Lemma: 1, 3, 4 are transient state. $x \in \{1,3,4\}$, then $\rho_{x2} > 0$, $\rho_{2x} = 0$.

△ **Closed Sets** start the process in the set, never leave the set.

$A \subset S$ is closed if $P(x,y) = 0$ for any $x \in A$, $y \notin A$

Above example: closed sets: $\{1,2,3\}$ $P(1,4) = P(2,4) = P(3,4) = 0$

$\{2\}$. $P(2,2) = 1$

not closed: $\{1,3\}$ $P(3,2) > 0$ $2 \notin \{1,3\}$

△ **Irreducible Set**

$x \leftrightarrow y$ (x communicates with y)

if x is accessible from y and y is accessible from x.

· A set $B \subset S$ is irreducible if $\forall x, y \in B$, $x \leftrightarrow y$.

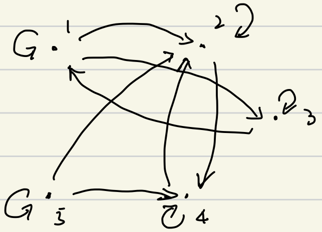
· Theorem: Let A be a closed, irreducible finite subset of S.

then all states in A are recurrant.

9/25

- closed set A $\quad P(x,y)=0 \quad \forall x\in A,\ y\notin A$
- irreducible set A
  $x,y\in A \quad P^{(n)}(x,y)>0$ for some $n\geq 0$
- Theorem: Any finite closed irreducible set of states is recurrent
- Lemma (criteria for transience)
  $\rho_{xy}>0,\ \rho_{yx}<1$, then x is transient.

ex. $P=\begin{bmatrix} 0.4 & 0.3 & 0.3 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0.3 & 0 & 0.3 & 0.4 \end{bmatrix}$



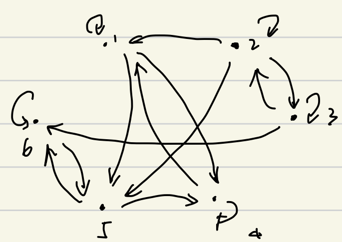
$S=\{1,2,3,4,5\}$

$\{2,4\}$ is closed + irreducible
closed : $P(x,1)=P(x,3)=P(x,5)=0$ for $x\in\{1,3\}$.
irreducible : $P(2,4)=0.5>0$, $P(4,2)=0.5>0$.

2, 4 are recurrent.
1, 3, 5 are transient. $\rho_{x2}>0$ but $\rho_{2x}=0$ for $x=1,3,5$

ex. $P=\begin{bmatrix} 0.1 & 0 & 0 & 0.4 & 0.5 & 0 \\ 0.1 & 0.2 & 0.2 & 0 & 0.5 & 0 \\ 0 & 0.1 & 0.3 & 0 & 0 & 0.6 \\ 0.1 & 0 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0 & 0.6 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}$

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



Claim: $\{1,4,5,6\}$ is closed + irreducible.

Closed: $P(x, 2) = P(x,3) = 0 \quad \forall x \in \{1,4,5,6\}$.

irreducible: $P(1,4) > 0 \quad P(1,5) > 0 \quad p^{(2)}(1,6) \geq P(1,5)P(5,6) > 0$

$P(4,1) > 0 \quad P^{(2)}(4,5) \geq P(4,1)P(1,5) > 0 \quad P^{(3)}(4,6) = P(4,1)P(1,5)P(5,6) > 0$

$P(5,4) > 0 \quad P(5,6) > 0 \quad P^{(2)}(5,1) \geq P(5,4)P(4,1) > 0$

$P(6,5) > 0 \quad P^{(2)}(6,4) \geq P(6,5)P(5,4) > 0 \quad P^{(3)}(6,1) \geq P(6,5)P(5,4)P(4,1) > 0$

Transient: 2, 3

Recurrant: 1, 4, 5, 6.

ex. ehrenfest chains

n balls $\quad N_1$ balls + N_2 balls in urn 1 & 2.

at each time, pick a random ball and move to the other urn

X_n = # balls in urn 1 at time n

Markov chain (each time have one more ball or one less ball) in each urn

$P(X_{n+1} = x+1 \mid X_n = x) = \frac{N-x}{N}$

$P(X_{n+1} = x-1 \mid X_n = x) = \frac{x}{N}$

$S = \{0, \dots, n\}$.

Claim: all the states are recurrant as the whole state space is closed + irreducible.

automatic

Irreducible: $x, y \in \{0, \dots, N\}$

Case 1: $x < y$

$n = y - x \quad P^{(n)}(x,y) \geq P(x, x+1) \cdots P(y-1, y) > 0$

Case 2: $x > y$.

$n = x - y \quad P^{(n)}(x,y) \geq P(x, x-1) \cdots P(y-1, y)$

- Proof of the theorem.

Thm: A finite closed, irreducible set of states is recurrant.

Lemma 1: A finite closed set has at least one recurrant state.

Lemma 2: If x is recurrant and $x \Leftrightarrow y$, then y is recurrant.

Key fact: A state x is recurrant iff $\sum_{n=1}^{\infty} p^{(n)}(x,x) = \infty$.

$N(x) = $ # of visits to state x

$E_x N(x) = \sum_{n=1}^{\infty} p^{(n)}(x,x)$

$N(x) = \sum_{n=1}^{\infty} 1_{\{x_n = x\}}$ $\leftarrow$ $1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{else} \end{cases}$ $\quad E 1_A = P(A)$.

$E_x N(x) = \sum_{n=1}^{\infty} E 1_{\{x_n = x\}} = \sum_{n=1}^{\infty} P_x(X_n = X)$

$\hookrightarrow p^{(n)}(x,x)$

Proof of Lemma 2:

strategy: show that $\sum_{n=N}^{\infty} p^{(n)}(y,y) = \infty$.

① choose N_1 and N_2 s.t. $p^{(N_1)}(y,x) > 0$, $p^{(N_2)}(x,y) > 0$

② Let $N = N_1 + N_2 + 1$ $\quad n \geq N$. $\quad p^{(n)}(y,y) \geq p^{(N_1)}(y,x)\, p^{(n-N_1-N_2)}(x,x)\, p^{(N_2)}(x,y)$.

$$\sum_{n=N}^{\infty} p^{(n)}(y,y) \geq p^{(N_1)}(y,x)\, p^{(N_2)}(x,y) \sum_{n=N}^{\infty} p^{(n-N_1-N_2)}(x,x)$$

$$= \text{------------} \; \underbrace{\sum_{n=1}^{\infty} p^{(n)}(x,x)}_{\infty}.$$

9/28

- Lemma 1: if x is recurrant and x communicates with y, then y is recurrant
- Lemma 2: if C is a finite closed set, then there is at least one recurrant state.

For Lemma 1

Fact: x is recurrant iff $\sum_{n=1}^{\infty} P^{(n)}(x,x) = \infty$

$N(y)$ = # of visit to state y by Markov chain

$$N(y) = \sum_{n=1}^{\infty} 1\{X_n = y\}.$$

$$E_x N(y) = \sum_{n=1}^{\infty} P_x(X_n = y) = \sum_{n=1}^{\infty} P^{(n)}(x,y)$$

Fact: $E_x N(y) = \begin{cases} \infty & \text{if } y \text{ is recurrant} \\ \frac{\rho_{xy}}{1-\rho_{yy}} & \text{if } y \text{ is transient.} \end{cases}$

$$\rho_{xy} = P_x(T_y < \infty)$$

Proof: · step one (Layer cake representation)

$$N(y) = \sum_{k=1}^{\infty} 1\{N(y) \geq k\}$$

↳ proof: If $N(y) = n$. then $\sum_{k=1}^{\infty} 1\{N(y) \geq k\} = \sum_{k=1}^{\infty} 1\{k \leq N\} = \sum_{k=1}^{n} 1 = n.$

· step two (take E_x on both sides)

$$E_x N(y) = \sum_{k=1}^{\infty} E_x 1\{N(y) \geq k\}$$
$$= \sum_{k=1}^{\infty} P_x(N(y) \geq k)$$
$$= \sum_{k=1}^{\infty} P_x(T_y^k < \infty)$$
$$= \rho_{xy} \sum_{k=1}^{\infty} \rho_{yy}^{k-1} = \rho_{xy} \underbrace{\sum_{k=0}^{\infty} \rho_{yy}^{k}}_{\text{(geometric sequences)}} = \frac{\rho_{xy}}{1-\rho_{yy}}$$

y is recurrant.

↳ $= \begin{cases} \infty & \text{if } \rho_{yy} = 1 \\ \frac{\rho_{xy}}{1-\rho_{yy}} & \text{if } \rho_{yy} < 1 \end{cases}$

For Lemma 2

key fact: since C is closed, for any $n \geq 1$, $\sum_{y \in C} P^{(n)}(x,y) = 1$ for $x \in C$

Proof by contradiction:

Suppose all states in C are transient.

Then $\sum_{y \in C} E_x N(y) < \infty$ (finite) for $x \in C$ (since C is finite set).

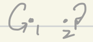
$$E_x N(y) = \sum_{n=1}^{\infty} P^{(n)}(x,y)$$

$$\Rightarrow \sum_{y \in C} E_x N(y) = \sum_{y \in C} \sum_{n=1}^{\infty} P^{(n)}(x,y) = \sum_{n=1}^{\infty} \underbrace{\sum_{y \in C} P^{(n)}(x,y)}_{=1} = \sum_{n=1}^{\infty} 1 = \infty$$

This contridict with $\sum_{y \in C} E_x N(y) < \infty$ if y is transient

- Remark. closed + irreducible $\Rightarrow$ all states in A is recurrant

  All states in A is recurrand $\not\Rightarrow$ closed, irreducible

  ex: $S=\{1,2\}$  $\{1,2\}$ is not irreducible since $P^{(n)}(1,2)=0$, $P^{(n)}(2,1)=0$

  $\{1\}, \{2\}$: closed, irreducible

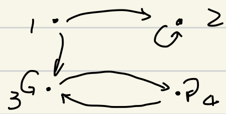
- Theorem: Each finite state Markov Chain satifies: $S = \underbrace{T \cup \bigcup_{i=1}^{k} R_i}$ Canonical decomposition of state space (k ← depend on Markov Chain)

  where T is a set of transient state

  and $\bigcup_{i=1}^{k}$ is a disjoint union of closed + irreducible sets R_i

Ex 1. $S=\{1,2,3,4\}$. $T=\{1\}$. $R_1=\{2\}$, $R_2=\{3,4\}$.

Goal: build the Markov Chain w/ canonical decomposition. $S = T \cup R_1 \cup R_2$.



$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Ex 2. Gambler's Ruin on $\{0,\dots,N\}$, $T=\{1,\dots,N-1\}$, $R=\{0,N\}$.

$\{0,N\}$ closed but not irreducible. $R_1=\{0\}$, $R_2=\{N\}$. irreducible + closed.

Canonical decomposition: $S = T \cup R_1 \cup R_2$.

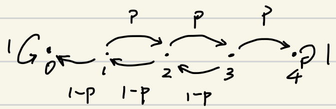
Ex 3. Ehrenfast chain on $\{0,\dots,N\}$ is closed and irreducible.

$R=\{0,\dots,N\}$ $T=\phi$ $S=R\cup T$

Termiology: the M.C is recurrant (the whole state space is closed + irreducible).

9/30

▷ More on Gambler's Ruin on $\{0,4\}$



Recurrant states : 0, 4

Transient states : 1, 2, 3

Q starting with 2 dollars, the prob. to ends up with 4 dollars?

$T = \inf\{n \geq 0 : X_n = 0 \text{ or } X_n = 4\}$.

Compute $P_2(X_T = 4)$

Strategy: vary initial condition, think T as a function(h)

$h : S \to [0,1]$

$h(x) = P_x(X_T = 4)$

Goal: compute $h(2)$

"Bondary Conditions": $h(0) = 0 \quad h(4) = 1$

$x \in \{1,2,3\}$: relate $h(x)$ and $h(y)$

$h(x) = P_x(X_T = 4) = P_x(X_T = 4, X_1 = x-1) + P_x(X_T = 4, X_1 = x+1)$

$= P_x(X_T = 4 \mid X_1 = x-1)\cdot(1-p) + P(X_T = 4 \mid X_1 = x+1)\cdot p$

Markov Property of starting again

$= P_{x-1}(X_T = 4)\cdot(1-p) + P_{x+1}(X_T = 4)\cdot p$

$= (1-p)h(x-1) + p\,h(x+1)$ for $x \in \{1,2,3\}$

$$\begin{cases} h(1) = (1-p)\,h(0) + p\,h(2) = p\,h(2) \\ h(2) = (1-p)\,h(1) + p\,h(3) \\ h(3) = (1-p)\,h(2) + p\,h(4) = (1-p)h(2) + p \end{cases}$$

$h(2) = (1-p)\,h(2) + p\,((1-p)h(2) + p)$

$= 2p(1-p)h(2) + p^2$

$h(2) = \dfrac{p^2}{1-2p(1-p)}$

Sanity check: $p = 0.2$ $\quad h(2) = \dfrac{\frac{1}{4}}{1-1\cdot\frac{1}{2}} = \dfrac{1}{2}$

· Another approach from 3 equation above

$h(1) - p\,h(2) + 0\cdot h(3) = 0$

$-(1-p)h(1) + h(2) - p\,h(3) = 0$

$0\,h(1) - (1-p)\,h(2) + h(3) = p$

$\overset{\text{matrix}}{\Rightarrow}$ $\begin{pmatrix} 1 & -p & 0 \\ -(1-p) & 1 & -p \\ 0 & -(1-p) & 1 \end{pmatrix}\begin{pmatrix} h(1) \\ h(2) \\ h(3) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ p \end{pmatrix}$

From transition matrix

$$\begin{array}{c|ccccc} & 0 & 4 & 1 & 2 & 3 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 4 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1-p & 0 & 0 & p & 0 \\ 2 & 0 & 0 & 1-p & 0 & p \\ 3 & 0 & p & 0 & 1-p & 0 \end{array}$$

inside transient Q

transient to recurrant

If invertible

$$(I-Q)\begin{pmatrix} h(1) \\ h(2) \\ h(3) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ p \end{pmatrix}$$

$$\begin{pmatrix} h(1) \\ h(2) \\ h(3) \end{pmatrix} = (I-Q)^{-1}\begin{pmatrix} 0 \\ 0 \\ p \end{pmatrix}$$

Q What is the expected duration of the game? (starting with 2)

$T = \inf\{n \geq 0 : X_n = 0 \text{ or } X_n = 4\}$.

Goal: compute $E_2(T)$

Strategy: $g(x) = E_x(T)$

Bondary Condition $g(0) = 0 \quad g(4) = 0$

$x \in \{1, 2, 3\}$. $g(x) = E_x(T)$

$= 1 + E_{x-1}(T)$

$$= E_x(T \mid X_1 = x-1)\, P_x(X_1 = x-1) + E_x(T \mid X_1 = x+1)\, P_x(X_1 = x+1)$$

$$= (1 + g(x-1))(1-p) + (1 + g(x+1))\,p$$

$$= 1 + (1-p)\,g(x-1) + p\,g(x+1)$$

$$\begin{cases} g(1) = 1 + p\,g(2) \\ g(2) = 1 + (1-p)\,g(1) + p\,g(3) \\ g(3) = 1 + (1-p)\,g(2) \end{cases}$$

$$g(2) = 1 + (1-p)(1 + p\,g(2)) + p(1 + (1-p)\,g(2)) = 2 + 2p(1-p)\,g(2)$$

$$g(2) = \frac{2}{1-2p(1-p)}$$

If $p = 1/2$ $\quad g(2) = \dfrac{2}{1-1/2} = 4$

As a matrix equation $\quad (I-Q)\begin{pmatrix} g(1) \\ g(2) \\ g(3) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} g(1) \\ g(2) \\ g(3) \end{pmatrix} = (I-Q)^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

if invertible

Reminder: geometric series $q \in (-1,1)$ $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$

Q. Can we do this with matrics ?

$$(I-Q)^{-1} = \sum_{k=0}^{\infty} Q^k$$

Abstract Setting:

T = transient state

$R = S \setminus T$

$\sigma = \inf\{n \geq 0.\ x_n \in R\}$

$P_x(X_\sigma = y)$ $\quad y \in R,\ x \in T$

ExT

$$P = \begin{array}{c|cc} & R & T \\ \hline R & P_R & 0 \\ T & S & Q \end{array}$$

Lemma: For $x, y \in T$, the series $\sum_{n=0}^{\infty} (Q^n)_{xy}$ converges

$\Rightarrow I-Q$ is invertible and $(I-Q)^{-1} = \sum_{n=0}^{\infty} (Q^n)$

10/2

△ Stationary Distribution

- Markov chain determined by initial distribution + transition matrix.

Q. a way to randomize the initial state so that the prob. distribution of the M.C. stays constant

Mathematically: want to choose an initial probability distribution

$$P_\mu(X_n = y) = \mu(y)$$

$$P_\mu(X_n = y) = \sum_x P_\mu(X_0 = x, X_n = y)$$

$$= \sum_x \mu(x) p^{(n)}(x,y)$$ ← vector · matrix, computed to an entry.

$$= (\mu P^{(n)})(y)$$

$$\underbrace{[\cdots\cdots]}_{\mu}\ \underbrace{\begin{bmatrix} & \overset{y}{\vdots} & \end{bmatrix}}_{P^{(n)}}$$

- Want: $(\mu P^{(n)})(y) = \mu(y)$ $\forall y \in S$ $\Rightarrow$ $\mu P^n = \mu$ $(*)$

Further simplification: If $(*)$ holds for $n=1$, then $(*)$ holds for all $n \geq 1$

Assume $\mu P = \mu$, then $\mu P^{(n)} = \mu P \cdot P^{n-1} = \mu \cdot P^{n-1} \cdots = \mu$.

- **Def:** given a M.C. w/ transition matrix P, an invariant prob. distribution μ satisfies $\mu P = \mu$

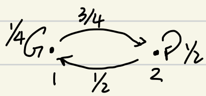
Remark (Terminology): stationary distribution = invariant probability distribution.

Q1. Do stationary distribution exists?

Q2. Are they unique?

Q3. Relation to the long time behavior of our M.C.

Ex1. $S = \{1, 2\}$ Find μ



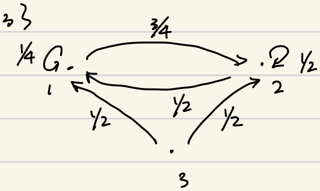
Let $\mu = [\alpha, 1-\alpha]$ $P = \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix}$

$\mu P = [\frac{1}{4}\alpha + \frac{1}{2} - \frac{1}{2}\alpha \quad \frac{3}{4}\alpha + \frac{1}{2} - \frac{1}{2}\alpha] = [\frac{1}{2} - \frac{1}{4}\alpha \quad \frac{1}{2} + \frac{1}{4}\alpha]$

$\mu P = \mu$ if $\begin{cases} \frac{1}{2} - \frac{1}{4}\alpha = \alpha \\ \frac{1}{2} + \frac{1}{4}\alpha = 1 - \alpha \end{cases} \Rightarrow \alpha = \frac{2}{5}$

Hence, $\mu = [\frac{2}{5}, \frac{3}{5}]$ is a stationary distribution. ← also unique.

Ex2. $S = \{1, 2, 3\}$



R: 1, 2

T: 3. $\Rightarrow$ $\mu(3) = 0$ as $P_3(X_n = 3) = 0$.

Intuition:

never come back to 3.

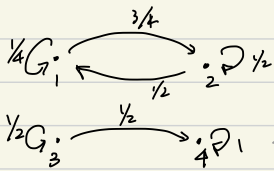
Let $\mu = (\mu_1, \mu_2, 1 - \mu_1 - \mu_2)$ $P = \begin{pmatrix} 1/4 & 3/4 & 0 \\ 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}$

Need: $\begin{cases} \frac{1}{4}\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{2}\cancel{(1-\mu_1-\mu_2)} = \mu_1 \\ \frac{3}{4}\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{2}\cancel{(1-\mu_1-\mu_2)} = \mu_2 \\ 0 = 1-\mu_1-\mu_2. \end{cases}$ $\Leftarrow$ Same equation as Ex1

$\Rightarrow \mu_1+\mu_2=1$

$\mu = (\frac{2}{5} \ \frac{3}{5} \ 0)$ unique.

Ex3. $S=\{1,2,3,4\}$



$R: \{1,2\}, \{4\}$ $T: 3.$

$\Leftarrow$ intuition: 3 is transient $\Rightarrow \mu(3)=0.$

$$P = \begin{pmatrix} 1/4 & 3/4 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow \quad (\mu_1 \ \mu_2)\begin{pmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{pmatrix} = (\mu_1 \ \mu_2)$$

$$(\mu_3 \ \mu_4)\begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix} = (\mu_3 \ \mu_4) \Rightarrow \begin{matrix} \mu_3 = 0 \\ \mu_4 = \mu_4. \end{matrix}$$

Not unique: $(\frac{2}{5}(1-\mu_4) \ \ \frac{3}{5}(1-\mu_4) \ \ 0 \ \ \mu_4)$ $\mu_4 \in [0,1].$

take $\alpha \in [0,1]$ $\alpha(\frac{2}{5}, \frac{3}{5}, 0, 0) + (1-\alpha)(0,0,0,1)$

10/5

Continue: stationary distribution.

- Stationary probability measure μ $\quad \mu(x) = P_\mu(X_n = x) \quad \forall x \in S \Leftarrow \mu = \mu P$
  $= (\mu P^{(n)})(x)$

- Imply a stronger property.

- Stationary Stochastic Process:
  $(X_n)_{n=0}^{\infty}$ is stationary stochastic process if
  $$P(X_0 = x_0, \dots X_n = x_n) = P(X_m = x_0, \dots\dots X_{n+m} = x_n)$$
  shift later m unit.

- Lemma: A M.C. $\{X_n\}_{n=0}^{\infty}$ started from a stationary dist
  is a stationary stochastic process

Proof: Let μ be the stationary prob. dist.
$$P_\mu(X_0 = x_0, \dots, X_n = x_n)$$
$$= \underbrace{\mu(x_0)}_{P_\mu(X_m = x_0)} \underbrace{P(x_0, x_1) \dots\dots P(x_{n-1}, x_n)}_{P_\mu(X_{m+1} = x_1, \dots, X_{n+m} = x_n \mid X_m = x_0)}$$
$$= P_\mu(X_m = x_0, \dots\dots, X_{m+n} = x_n)$$

- Existence & Uniqueness of stationary prob. dist.
  $S = T \cup \bigcup_{i=1}^{k} R_i$, where $R_i \cap R_j = \phi$ and R_i is closed, irreducible.

Recall: A M.C. is recurrent if S is irreducible
$\Rightarrow T$ is empty, $k=1$

- Theorem: A finite state, recurrent M.C. chain has a unique invariant prob. dist.

Proof: $\mu P = \mu \iff P^t \mu = \mu$

row vector $\quad$ col vector.

$\mu = [\quad]$ $\quad$ $\mu = [\quad]$

transfer matrix $P^t(x, y) = P(y, x)$

$$P^t \mu = \mu \text{ iff } \boxed{(P^t - I)\mu = 0} \quad \text{focus on this}$$

- Some linear algebra.
  - Null space: given an $n\times n$ matrix, we define $Null(A)=\{v\in\mathbb{R}^n \mid Av=0\}$

    Remark: this is a sub-space $\Leftarrow$ if $v_1, v_2 \in Null(A)$ and $\alpha_1,\alpha_2\in\mathbb{R}$.

    then $\alpha_1 v_1+\alpha_2 v_2 \in Null(A)$
  - Linear Independence: vectors $v_1\cdots v_n$ are linearly independent if

    $\alpha_1 v_1+\cdots+\alpha_n v_n=0$ for some $\alpha_1\cdots\alpha_n\in\mathbb{R}$. then $\alpha_1=\cdots=\alpha_n=0$
  - Dimension of $Null(A)$: $\dim(Null(A))=k$ if $\exists$ k non-zero linearly independent vectors in $Null(A)$ but not $k+1$
  - Rank of A: $Rank(A)$ is # of linearly independent rolls.
  - Rank-Nullity Theorem: given an $n\times n$ matrix A, $rank(A)+\dim(Null(A))=n$.

- Corollary: A matrix A with column sums all zero has $\dim(Null(A))\geq 1$

- Back to proof:

$\underbrace{(P^t-I)}_{A}\mu=0$

Let $\alpha_1=\cdots=\alpha_n=1$

$$\alpha_1\begin{bmatrix}A_{11}\\ \vdots\\ A_{1n}\end{bmatrix}+\cdots+\alpha_n\begin{bmatrix}A_{n1}\\ \vdots\\ A_{nn}\end{bmatrix}=\begin{bmatrix}\sum_{j=1}^n A_{j1}\\ \vdots\\ \sum_{j=1}^n A_{jn}\end{bmatrix}=0 \Rightarrow \dim(Null(A))\geq 1$$

$n=\dim(S)$

P: $n\times n$ matrix.

Step 1: $\exists\,\mu\in\mathbb{R}^n$ s.t. $\mu P=\mu$. iff $\mu\in Null(P^t-I)$

Since column sums of P^t-I are zero, $\dim(Null(P^t-I))\geq 1$

Step 2: For any μ w/ $\mu P=\mu$, all components of μ have the same sign.

Assume: $P(x,y)>0$ for $x,y\in S$

Argue by contradiction.

xs at different sign

$$|\mu(y)|=\left|\sum_x \mu(x)P(x,y)\right|<\sum_x|\mu(x)|P(x,y)$$

$$\sum_y|\mu(y)|<\sum_x|\mu(x)|\sum_y P(x,y)=\sum_x|\mu(x)|$$

both suming over μ, same $\Rightarrow$ ※.

Consider $Q=\frac{1}{2}I+\frac{1}{2}P$ $\quad \mu Q=\mu\cdot\frac{1}{2}I+\mu\frac{1}{2}P=\mu \Rightarrow \mu Q^{(n)}=\mu$.

Step 3: (uniqueness) $\dim(\text{Null}(Q^n - I)) = 1$

otherwise, there are μ_1, μ_2 linearly independent and $\mu_1 Q^n = \mu_1$, $\mu_2 Q^n = \mu_2$.

$\mu_1, \mu_2 \neq 0$, $\mu_1(y) = \sum_x \underbrace{\mu_1(x)}_{\geq 0} \underbrace{Q^n(x,y)}_{> 0} > 0$

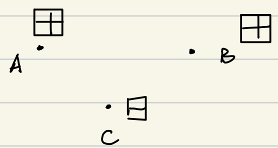
Gram-schmidt: wlog, we may assume $\mu_1 \cdot \mu_2 = 0 \Leftrightarrow \sum_i \mu_1(i)\mu_2(i) > 0$ or < 0 $\times$.

10/7

Specific type of invariant distr.: Mass Transport

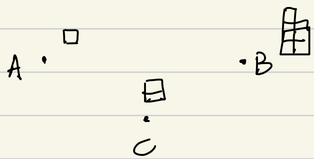
3 states: 10 blocks total $\mu = [\frac{4}{10}\ \frac{4}{10}\ \frac{2}{10}]$

$n=0$



$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$

$n=1$



distr. $[\frac{1}{10}\ \frac{2}{10}\ \frac{7}{10}]$

For an invariant prob. measure, the number of blocks at each side is the same

- Reversable measure: μ is reversable if $\mu(x)P(x,y) = \mu(y)P(y,x)$

Remark: in mass transport, this means #block $x \to y$ = #blocks $y \to x$

e.g. $\begin{tikzcd} A \arrow[r, bend left, "1"] & B \arrow[l, bend left, "1"] \end{tikzcd}$ $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then $[\frac{1}{2}\ \frac{1}{2}]$ is a reversable prob. measure.

- Lemma: reversable $\to$ invariant.

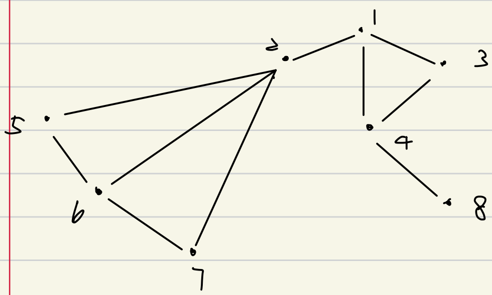
Proof: NTS. $\mu P = \mu$

proportion of mass going from x to z

$(\mu P)(z) = \sum_x \mu(x)P(x,z) = \sum_x \mu(z)P(z,x) = \mu(z)\sum_x p(z,x) = \mu(z)$

ex. Random walk on a graph.

• vertex
— edges, double connected.
degree : #neighbours



$$P = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ & & & \vdots & & & & \end{bmatrix}$$

State x and y are neighbours is denoted $x \sim y$
At a given vertex, we choose a neighbour uniformly at random
Adjacency Matrix $\quad a(x,y) = \begin{cases} 1 & \text{if } x \sim y \\ 0 & \text{otherwise} \end{cases}$

Degree = # neighbours of x
$= \sum_{y \in V} a(x,y) \qquad V$: vertices

Transition matrix: $\quad P(x,y) = \dfrac{a(x,y)}{\deg(x)}$

Q: can we find a reversable measure?
Want: $\mu(x) P(y,x) = \mu(y) P(y,x)$ $\quad$ trival if $x \nsim y$

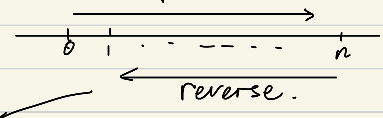
$$\mu(x) \frac{1}{\deg(x)} = \mu(y) \frac{1}{\deg(x)}$$

$$\frac{\mu(x)}{\mu(y)} = \frac{\deg(x)}{\deg(y)}$$

Set $\mu(x) = \dfrac{\deg(x)}{\sum_{x \in V} \deg(x)}$

- Time Reversal of the M.C

$(X_n)_{n=0}^{\infty}$ started from an invariant dist. μ.



Def $(Y_n)_{n=0}^{\infty}$ by $Y_n = X_{N-n}$

- **Lemma:** $(Y_n)_{n=0}^{\infty}$ has the M.C. property

Remark: assume the lemma, compute the

$$q(x,y) = P(Y_{n+1}=y \mid Y_n = x)$$

$$= \frac{P(Y_{n+1}=y, Y_n=x)}{P(Y_n=x)} = \frac{P(X_{N-(n+1)}=y, X_{N-n}=x)}{P(X_{N-n}=x)}$$

$$= \frac{P(X_{N-(n+1)}=y)\,P(y,x)}{P(X_{N-n}=x)}$$

$$= \frac{\mu(y)P(y,x)}{\mu(x)}$$

If μ is reversable, then $\mu(x)P(x,y) = \mu(y)P(y,x)$

$$P(x,y) = \frac{\mu(y)}{\mu(x)}P(y,x) = q(x,y)$$

Proof of the Lemma.

$$P(Y_{n+1}=y_{n+1} \mid Y_n = y_n, \cdots, Y_0 = y_0)$$

$$= \frac{P(X_{N-(n+1)} = y_{n+1}, \cdots\cdots, X_N = y_0)}{P(X_{N-n}=y_n, \cdots\cdots, X_N = y_0)}$$

$$= \frac{P(X_{N-(n+1)}=y_{n+1})P(y_{n+1}, y_n)\cdots P(y_1, y_0)}{P(X_{N-n}=y_n)P(y_n=y_{n-1})\cdots P(y_1,y_0)}$$

$$= \frac{P(X_{N-(n+1)}=y_{n+1})P(y_{n+1},y_n)}{P(X_{N-n}=y_n)}$$

$$= \frac{P(X_{N-(n+1)}=y_{n+1}, X_{N-n}=y_n)}{P(X_{N-n}=y_n)} = P(Y_{n+1}=y_{n+1} \mid Y_n = y_n)$$

10/9

Infinite State Space of M.C.

· Carnonical Decomposition $S = T \cup \bigcup_{i \in I} R_i$ R_i: closed, irreducible
I: either finite or countable

ex. $\underset{1}{\cdot} \rightleftarrows \underset{2}{\cdot}$ $\underset{3}{\cdot} \rightleftarrows \underset{4}{\cdot}$ - - - - . . . infinite close & irreducible sets

$R_1 = \{1, 2\}$ ⋯ $R_k = \{2k, 2k+1\}$ $\forall k \geq 1$

$S = \bigcup_{k \geq 1} R_k$ ← Canonical Decomposition.

Lots of invariant.

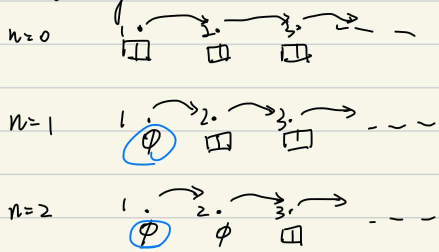
ex. Deterministic, monotone M.C.

$S = \mathbb{N} = \{1, 2, 3, \cdots\}$ $\underset{1}{\cdot} \to \underset{2}{\cdot} \to \underset{3}{\cdot}$ - - - - - ⋯ ∞

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ & & \vdots & \end{pmatrix}$$

all states are transient

Q. Are there any invariant measures ?



only solution: $\mu = 0$. $\mu(y) = \sum_x \mu(x) P(x, y) = \mu(y-1) P(y-1) = \underline{\mu(y-1)}$

$\mu(0) = \sum_x \mu(x) P(x, 0) = 0$

$\Rightarrow \mu(1) = \mu(0) = 0$

$\cdots \quad \mu(x) = 0 \quad \forall x \in S$

ex. $S = \mathbb{Z}$

$$-\infty \cdots \quad \underset{-1}{\cdot} \curvearrowright \underset{0}{\cdot} \curvearrowright \underset{1}{\cdot} \quad \cdots \infty$$

All states are transcient.

$\mu(x) = c$ for all $x \in \mathbb{Z}$ solves $\mu P = \mu$ — invariant measure, but not reversable.

- not a prob. measure $\sum_{x \in S} c = \infty$

Invariant Measure $\mu: S \to [0, \infty)$ s.t. $\mu(y) = \sum_x \mu(x) P(x,y)$

Reversible inv. Measure $\mu(x) P(x,y) = \mu(y) P(y,x)$

ex. Success Run Chain.

$(X_n)_{n=0}^{\infty}$ $\qquad S = \{0, 1, \cdots\}$

X_n: # heads in a row at time n

$p(x, x+1) = \alpha \quad p(x, 0) = 1-\alpha. \quad \forall x$

- all states are recurrant, S is irreducible

Case 1 $x \to y$ flip $y-x$ heads in a row $\to$ go from x to y.
$p^{(y-x)}(x,y) > 0$

Case 2 $y \to x$ flip 1 tail. flip x heads. $P^{(1+x)}(y,x) > 0$.

$(X_n)_{n=0}^{\infty}$ is a recurrant M.C.

Q. Invariant Measures. $\mu(y) = \sum_x \mu(x) P(x,y)$

$$\mu(0) = \sum_{k=0}^{\infty} \mu(k) p(k,0) = (1-\alpha) \sum_{k=0}^{\infty} \mu(k)$$

$$\mu(k) = \alpha \mu(k-1) \qquad k \geq 1$$

$$= \alpha^k \mu(0)$$

$$(1-\alpha) \sum_{k=0}^{\infty} \alpha^k \mu(0) = (1-\alpha) \mu(0) \frac{1}{1-\alpha} = \mu(0)$$

Want $1 = \sum_{k=0}^{\infty} \mu(k) = \sum_{k=0}^{\infty} \alpha^k \mu(0) = \mu(0) \frac{1}{1-\alpha} \Rightarrow \mu(0) = 1-\alpha \quad \mu(k) = (1-\alpha)\alpha^k$

(invariant prob. dist.)

Remark: All states are recurrant ① 0 is recurrant
② 0 is communicate with all states
$\Rightarrow$ all other states are recurrent

$T_0 = \inf\{n \geq 1 : X_n = 0\}$

NTS: $P(T_0 < \infty) = 1 \rightarrow P_0(T_0 = \infty) = 0$

$\{T_0 = \infty\} = \{$flipping infinitely many heads in a row$\}$

$= P(0) = 0$

$P(n \text{ heads in a row}) = \alpha^{\infty} = 0$ as $\alpha < 1$

Theorem: Let P be irreducible and recurrant, then there exists an invariant measure and it is unique up to multipalcation by a constant c

$\mu, \nu \rightarrow \nu = c\mu$

Q: How to encure the existence & uniqueness of an invariant prob. measure?

Recurrance: $T_x = \inf\{n \geq 1 \mid X_n = x\}$, x is recurrant if $P_x(T_x < \infty) = 1$

Positive Recurrance (stronger version)

x is positive recurrant if $E_x(T_x) < \infty$

Null Recurrance: $P_x(T_x < \infty) < 1$ but $E_x(T_x) = \infty$

10/12

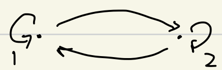
Last. M.C on countably infinite state spaces

- x is recurrent if $P_x(T_x < \infty)$
- Stronger: x is positive recurrant if $E_x(T_x) < \infty$

Thm Let P be irreducible. P has invariant prob. dist. iff all states are positive recurrent.
Futhermore, π is unique and $\pi(x) = \frac{1}{E_x(T_x)}$

Remark. For a finite state recurrent M.C. chain, this tells us that
$$E_x(T_x) = \frac{1}{\pi(x)}$$

Ex.



$\pi = (\pi(1), \pi(2))$

$E_1(T_1) = \frac{1}{\pi(1)}$

Lemma. Let π be an invariant prob. dist.
Criterian for recurrence: Let $x \in S$, if $\pi(x) > 0$, then x is recurrent
Proof: analyze the quantities $M = \sum_z \pi(z) E_z(N(x))$
argue by contradiction

1. Suppose x is transient.
$$\sum_z \pi(z) E_z(N(x)) \le \max_z E_z(N(x)) \underbrace{\sum_z \pi(z)}_{=1 \text{ as } \pi \text{ is invariant prob dist}}$$
$$E_z(N(x)) = \frac{f_{zx}}{1-f_{xx}} \le \frac{1}{1-f_{xx}} < \infty$$
$\Leftarrow$ x is transient, $f_{xx} < 1$

Hence, $\sum_z \pi(z) E_z(N(x)) < \infty$.

2. $E_z(N(x)) = \sum_{n=1}^{\infty} P^{(n)} p^{(n)}(z,x)$
$$M = \sum_z \pi(z) E_z(N(x)) = \sum_z \pi(z) \sum_{n=1}^{\infty} p^{(n)}(z,x)$$
$$= \sum_{n=1}^{\infty} \sum_z \pi(z) p^{(n)}(z,x)$$
$$= \sum_{n=1}^{\infty} P_\pi(X_n = x) = \sum_{n=1}^{\infty} \pi(x) = \infty$$
} π is invariant prob dist.

Hence, $M = \infty$. ※.

· Proof of the Thm.

1. P is irreducible
2. All states are positive recurrent

Strategy: write down an explicit invariant measure.

Fix $x \in S$, define $y \to \lambda x(y)$ $\quad \lambda x(y) = E_x[\#$ of visit to y before returning to $x]$

$$= E_x \sum_{k=1}^{T_x} 1\{x_k = y\}.$$

1. λx has finite total mass $\quad \sum_{y \in S} \lambda x(y) = E_x(T_x) < \infty$ (by def of positive recurrence)

\# visit go to all state except x before return to x

Proof:

Law of total expectation

$$E_x\left(\sum_{k=1}^{T_x} 1(x_k = y)\right) = E_x\left(E_x\left[\sum_{k=1}^{T_x} 1\{x_k = y\} \mid T_x\right]\right)$$

random variable

$$= \sum_{n=1}^{\infty} P(T_x = n)\, E_x\left[\sum_{k=1}^{T_x} 1\{x_k = y\} \mid T_x = n\right]$$

$$\sum_{y \in S} \lambda x(y) = \sum_{n=1}^{\infty} P(T_x = n)\, E_x\left[\sum_{y \in S} \sum_{k=1}^{n} 1\{x_k = y\} \mid T_x = n\right]$$

$$= \sum_{n=1}^{\infty} n\, P_x(T_x = n) = E_x(T_x)$$

Hence, λx has finite total mass.

If λx is an invariant measure, then $\pi(y) = \dfrac{\lambda x(y)}{E_x(T_x)}$ is an invariant prob dist.

total mass.

Note: $\lambda x(x) = E_x\left[\sum_{k=1}^{T_x} 1\{X_k = x\}\right] = 1$

Hence, $\pi(x) = \dfrac{\lambda x(x)}{E_x(T_x)} = \dfrac{1}{E_x(T_x)}$

Next: NTS λx is an invariant measure.

NTS: $\lambda x = \lambda x P$

$$\lambda x(y) = E_x\left(\sum_{k=1}^{T_x} 1\{x_k = y\}\right)$$

$$\sum_{k=1}^{T_x} 1\{x_k = y\} = \sum_{k=1}^{\infty} 1\{x_k = y, k \le T_x\}$$ (for all n, if $T_x = n$, both sides agree)

$$\lambda x(y) = \sum_{k=1}^{\infty} E_x 1\{x_k = y, k \le T_x\} = \sum_{k=1}^{\infty} P_x(X_k = y, T_x \ge k)$$

$k=1 \Rightarrow P(X_1=y, T_x=1) = P_x(X_1=y) = P(x,y)$

$$\lambda_x(y) = \underset{\substack{\Uparrow \\ \lambda_x(x)=1}}{\lambda_x(x)} P(x,y) + \sum_{k=2}^{\infty} P_x(X_j \neq x \text{ for } 1 \le j \le k-1,\ X_k = y)$$

$$\begin{aligned}
&\sum_{k=2}^{\infty} P_x(X_j \neq x \text{ for } 1 \le j \le k-1,\ X_k = y) \\
&= \sum_{k=2}^{\infty} \sum_{z \neq x} P_x(X_j \neq x \text{ for } 1 \le j \le k-2,\ X_{k-1} = z)\, P(z,y) \\
&= \sum_{z \neq x} P(z,y) \sum_{k=2}^{\infty} P_x(X_j \neq x \text{ for } 1 \le j \le k-2,\ k-1 = z) \\
&= \sum_{z \neq x} P(z,y) \sum_{k=2}^{\infty} P_x(X_{k-1} = z,\ T_x \geq k-1) \\
&= \sum_{z \neq x} P(z,y) \sum_{k=1}^{\infty} P_x(X_k = z,\ T_x \geq k) \qquad \text{chang } k-1 \Rightarrow k \\
&= \sum_{z \neq x} P(z,y)\, \lambda_x(z)
\end{aligned}$$

$$\begin{aligned}
\lambda_x(y) &= \lambda_x(x) P(x,y) + \sum_{z \neq x} \lambda_x(z) P(z,y) \\
&= \sum_z \lambda_x(z) P(z,y)
\end{aligned}$$

$\Rightarrow \lambda_x$ is an invariant measure

$\lambda_x(x) = 1$

$$\pi(y) = \frac{\lambda_x(y)}{E_x(T_x)} \Rightarrow \pi(x) = \frac{1}{E_x(T_x)}$$

2. Check Uniqueness. in Lecture Note.

10/14

Long Time Behavior of M.C.

1. $p^{(n)}(x,y)$ for large n.
2. $f: S \to \mathbb{R}$ $\quad \frac{1}{n}\sum_{k=0}^{n-1} f(X_k)$ converge to limit ?

Typical question: what proportion of time is spent in a recurrent state x.

$$\frac{1}{n}\sum_{k=0}^{n-1} 1\{X_k = x\} \qquad f(y) = 1\{y=x\}.$$

Recall: LLN

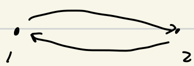
if $X_0, \dots, X_n$ iid w/ finite moments. (discrete)

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow[\text{converge}]{} E[f(X_1)] = \sum_{x\in S} f(x) P(X_1 = x)$$

"time average" "space average".

For a M.C. $(X_n)_{n=0}^{\infty}$, we don't meet the iid assumption

Ex. $S = \{1, 2\}$



$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\qquad f: S \to \mathbb{R}$

$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \approx \frac{1}{n}\left[\frac{n}{2} f(1) + \frac{n}{2} f(2)\right] \approx \frac{1}{2} f(1) + \frac{1}{2} f(2)$

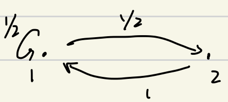
$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(X_k) = \frac{1}{2} f(1) + \frac{1}{2} f(2)$

$\pi = (\frac{1}{2} \ \frac{1}{2})$ is the stationary dist.

Thm. SLLN for M.C.

Let π be the unique invariant dist for a finite state M.C. w/ irreducible transition matrix P. Then for all $f: S \to \mathbb{R}$ $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(X_k) = \sum_{x\in S} f(x)\pi(x)$ with prob one.

Ex.



$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$

P is irreducible

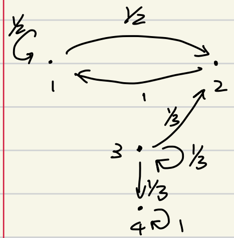
Q: long-run proportion of time spent in state 1 ?

$\pi = [\frac{2}{3} \ \frac{1}{3}]$ is an invariant distribution.

By the SLLN for M.C. $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = \pi(1) f(1) + \pi(2) f(2) = \frac{2}{3} f(1) + \frac{1}{3} f(2)$

$f(x) = 1_{x=1}$ then $f(1) = 1$, $f(2) = 0$

$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1_{X_n = 1} = \pi(1) = \frac{2}{3}$

Ex. 

$$P = \begin{pmatrix} \frac12 & \frac12 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac13 & \frac13 & \frac13 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$\{1,2\}$ closed, irreducible
$\{4\}$ closed, irreducible.
$\{3\}$ transient.

$X_0 = 3$

$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = ?$ random!

Two possibilities. $T_4 < \infty$ or $\{T_2 < \infty\} = \{T_4 = \infty\}$.

1. $T_4 < \infty$ $\quad n > T_4$

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) = \frac{1}{n}\sum_{k=1}^{T_4} f(X_k) + \frac{1}{n}\sum_{k=T_4}^{n} f(X_k)$$
$$= \frac{T_4 - 1}{n}\cdot f(3) + \frac{n - T_4}{n}\cdot f(4)$$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = f(4)$$

2. $T_4 = \infty$ $\quad n > T_2$

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) = \frac{1}{n}\sum_{k=1}^{T_2} f(X_k) + \frac{1}{n}\sum_{k=T_2}^{n} f(X_k)$$
$$= \frac{T_2 - 1}{n} f(3) + \frac{2}{3} f(1) + \frac{1}{3} f(2)$$
$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = \frac{2}{3} f(1) + \frac{1}{3} f(2)$$

Summary: $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = Y$

$$Y = \begin{cases} f(4) & \text{if } T_4 < \infty \\ \frac{2}{3} f(1) + \frac{1}{3} f(2) & \quad\; T_4 = \infty \end{cases}$$

$$P(T_4 < \infty) = \sum_{n=1}^{n} P(T_4 = n, T_2 = \infty)$$
$$= \sum_{n=1}^{n} P(X_1 \cdots = X_{n-1} = 3, X_n = 4)$$
$$= \sum_{n=1}^{n} \left(\frac{1}{3}\right)^{n-1}\cdot\frac{1}{3} = \frac{\frac{1}{3}}{1 - \frac{1}{3}} = \frac{1}{2}.$$

$$P(Y = f(4)) = P\left(Y = \frac{2}{3} f(1) + \frac{1}{3} f(2)\right) = \frac{1}{2}$$

10/16

Last: SLLN of M.C. on finite state spaces

- $(X_k)_{k=0}^{\infty}$ recurrent w/ invarient dist. π

$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = \sum_{x\in S} f(x)\pi(x)$

Reduction:

$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \quad X_k \in S$

$= \sum_{z\in S} f(z) \cdot \frac{(\text{\# of visits to state } z \text{ up till time } n)}{n}$

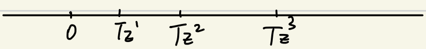
$= \sum_{z\in S} f(z) \underbrace{\frac{1}{n}\sum_{k=1}^{n} 1\{X_k = z\}}_{\text{focus on this}}$

- Goal: show for all $z\in S$, $\underbrace{\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1\{X_k = z\} = \pi(z)}_{\text{not iid.}}$ ⇐ Reduction.

$\pi(z) = \frac{1}{E_x(T_x)} \quad X_0 \sim \mu$

$T_z^1 = \inf\{n>0 : X_n = z\}$.

$T_z^2 = \inf\{n > T_z^1 : X_n = z\}$.



$T_z^1,\ T_z^2 - T_z^1,\ T_z^3 - T_z^2$ are independent.

$T_z^2 - T_z^1,\ T_z^3 - T_z^2 \ \ldots$ are iid.

$\Rightarrow E_\mu[T_z^2 - T_z^1] = E_z[T_z]$

By the SLLN, $\lim_{n\to\infty} \frac{1}{n} \underbrace{\sum_{k=1}^{n} T_z^{k+1} - T_z^k}_{\text{Telescoping sum.}} = E_z[T_z]$.

$= \lim_{n\to\infty} \frac{1}{n}(T_z^{n+1} - T_z^1) \qquad \lim_{n\to\infty} \frac{T_z^1}{n} = 0$

$= \lim_{n\to\infty} \frac{T_z^{n+1}}{n} = E_z[T_z] = \frac{1}{\pi(z)}$

$\lim_{n\to\infty} \frac{n}{T_z^{n+1}} = \pi(z)$ = limit $n\to\infty$ of propotion of time spent at z

$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1\{X_k = z\}$

$= \lim_{n\to\infty} \frac{1}{T_z^{n+1}} \sum_{k=1}^{T_z^{n+1}} 1\{X_k = z\}$.

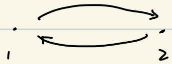
$= \lim_{n\to\infty} \frac{1}{T_z^{n+1}}(n+1)$

$= \lim_{n\to\infty} \frac{n+1}{n} \cdot \frac{n}{T_z^{n+1}} = \pi(z)$

$\Downarrow$

1

▷ Goal: analyze $p^{(n)}(x,y)$ for large n.

ex.  $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

start at 1

1-step: $1 \to 2$ 2-steps: $1 \to 2 \to 1$

3-steps: $1 \to 2 \to 1 \to 2$. ——— . . . . .

$p^{(2k+1)}(1,2) = 1$ but $p^{(2k)}(1,2) = 0$

Hence $\lim_{k\to\infty} p^{(k)}(1,2)$ DNE! (does not exist)

○ **Periodicity**

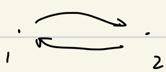
ex. $\gcd\{4,8,32\} = 4$

$\gcd\{4,8,13\} = 1$

$\gcd(A) = \sup\{n \geq 1 : \text{each } x \in A \text{ is divisible by } n\}$.

○ **Period** of a state $x \in S$

notation: $d(x) = \gcd I_x$ $I_x = \{n \geq 1 : p^{(n)}(x,x) > 0\}$.

ex.  $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$p^{(2n)}(1,1) > 0$ for all $n \geq 1$

$p^{(2n+1)}(1,1) = 0$ for all $n \geq 0$

I_1 = the even integers

$d(1) = \gcd(I_1) = 2 = d(2)$

Lemma: Two communicating states have the same period.

Def: The period of an irreducible, recurrent M.C. is the common period of all the states.

Def: Aperiodic: A M.C. is aperiodic if the period is 1.

ex.
$$\begin{tikzcd} 1 \arrow[loop left] \arrow[r, bend left] & 2 \arrow[l, bend left] \end{tikzcd}$$

$I_1 = \{n \geq 1 : P^{(n)}(1,1)\} \ni 1 \Rightarrow \gcd(I_1) = 1$

$I_2 = \{n \geq 1 : P^{(n)}(2,2)\} = \{2,3,4\dots\} \Rightarrow \gcd(I_2) = 1$

$\Rightarrow$ aperiodic.

Remark. The period of a M.C. is well-defined only for a recurrent state.

Thm: Let P be aperiodic and irreducible w/ an invariant prob. dist. then,

$$\lim_{n\to\infty} P_\mu(X_n = y) = \pi(y) \quad \text{for all } y \in S$$

Remark. In particular: $\lim_{n\to\infty} P^{(n)}(x,y) = \pi(y)$

10/21

Chap3 Martingales - class of discrete time stochastic process

Review of Condition Expectation

Setup: info in the form of a discrete random vector $Y=(Y_1, \dots, Y_k)$

random variable $X: \Omega \to \mathbb{R}$

Want to define $E[X|Y]$ ← random variable

Simpler. ordinary expectation: for X discrete $E[X] = \sum_x x P(X=x)$

Fix $y=(y_1, \dots, y_k)$ (vector) s.t $P(Y=y) = P(Y_1=y_1, \dots, Y_k=y_k)$ assume positive (>0)

$E(X|Y=y) = \sum_x x P(X=x|Y=y)$

scaler

ex. Let $(X_n)_{n=0}^{\infty}$ be a M.C. (finite state) $Y=(X_0, \dots X_{100})$ $X=X_{105}$

$E[X_{105} | (X_0, \dots, X_{100}) = (x_0 \dots, x_{100})]$ where $(x_0, \dots, x_{100}) \in S^{100}$ and $P(X_0=x_0 \dots X_{100}=x_{100}) > 0$

$= \sum_y y P(X_{105}=y | X_0=x_0 \dots X_{100}=x_{100})$

$= \sum_y y P(X_{105}=y | X_{100}=x_{100})$ by Markov Property.

$= \sum_y y P(X_5=y | X_0=x_{100})$ by time homogenouity.

$= E_{x_{100}}[X_5]$

More generally: $E[X_{n+k} | X_0=x_0, \dots, X_n=x_n] = E_{x_n}[X_k]$

$E[X|Y]$ $E[X|Y=y]$

view as a deterministic function of y s.t. $P(Y=y)>0$

Def: Let $H(y) = E[X|Y=y]$ for y s.t. $P(Y=y)>0$,

def the $E[X|Y] = H(Y)$ = "plug in" Y to the function $y \mapsto E[X|Y=y]$

Remark: The randomness of $E[X|Y]$ is all in Y

The possible values of $E[X|Y]$ are $E[X|Y=y]$

$P(E[X|Y] = E[X|Y=y]) = P(Y=y)$

r.v. real number

Alt. Def.

r.v. sample space

$\omega \in \Omega \to \mathbb{R}$

$E[X|Y](\omega) = E[X|Y=y]$ for $\omega \in \{Y=y\}$

M.C. $E[X_{n+k} \mid X_0=x_0, \dots X_n = x_n] = E_{x_n}[X_k]$

View $y \to E_y[X_k]$ as a function of the initial condition y

$E[X_{n+k} \mid X_0, \dots, X_n] = E_{X_n}[X_k]$

Law: total expectation $E[E[X|\overset{\text{r.v.}}{Y}]] = E[X]$

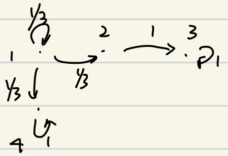
Reason: $E[E[X|Y]] = \sum_z z P(E[X|Y]=z)$ core: randomness of $E[X|Y]$ is in Y.

$$= \sum_y E[X|Y=y] \, P(E[X|Y] = E[X|Y=y])$$
$$= \sum_y E[X|Y=y] \, P(Y=y)$$
$$= \sum_{x,y} x P(X=x|Y=y) P(Y=y)$$
$$= \sum_{x,y} x P(X=x, Y=y)$$
$$= \sum_x x \sum_y P(X=x, Y=y)$$
$$= \sum_x x P(X=x)$$
$$= E(X)$$

$(X_n)_{n=0}^\infty$ M.C. X_0 is random, $X_0 = \mu$ $\quad f: S \to \mathbb{R}$

$E[f(X_n)] = E_\mu[f(X_n)]$

$$= E[E[f(X_n)|X_0]] = \sum_x E[f(X_n)|X_0=x] P(X_0=x)$$
$$= \sum_x E[f(X_n)|X_0=x] \mu(x)$$

ex. 

$\mu = [\frac12 \ \frac14 \ 0 \ \frac14]$

$E_\mu[f(X_3)] = \frac12 E_1[f(X_3)] + \frac14 E_2[f(X_3)] + \frac14 E_4[f(X_3)]$

graph $= \frac12 E_1[f(X_3)] + \frac14 f(3) + \frac14 f(X_4)$

Matrix Mul $= \sum_z z \, p^{(3)}(1,z)$

Property: conditional expectation.

1. if X and Y are independent, then $E[X|Y] = E[X]$

$E[X|Y=y] = \sum_x x P(X=x|Y=y) = \sum_x x P(X=x) = E[X]$. for all y.

$y \mapsto h(y) = E[X|Y=y]$ is constant at $E[X]$

$E[X|Y] = h(Y) = E[X]$.

2. $g: \mathbb{R}^k \longmapsto \mathbb{R}$, $Y = (Y_1, \cdots, Y_k)$, then $E(g(Y) \mid Y] = g(Y)$

$E[g(Y) \mid Y=y] = \sum_x g(x) P(g(Y)=x \mid Y=y)$ $\quad P(g(Y)=x, Y=y) = \begin{cases} 0 & \text{if } x \neq g(y) \\ P(Y=y) & \text{if } x = g(y) \end{cases}$

$= \sum_x g(y) P(Y=y) = g(y)$

$E[g(Y) \mid Y] = g(Y)$

3. $E[X g(Y) \mid Y] = g(Y) E[X \mid Y]$

$X=1$ (special case) $\rightarrow E[g(Y) \mid Y] = g(Y)$

10/23

Martingales

ex. $(\xi_k)_{k=0}^{\infty}$ of iid, mean-zero, r.v.

Define $(M_n)_{n=0}^{\infty}$ by $M_n = \sum_{k=0}^{n} \xi_k$.

- sequence of games

ξ_k : winnings in game k

M_n : total winnings up till n.

Q. given $\xi_0, \cdots, \xi_n$ (the outcomes of first n games), what is the best guess of M_{n+1}

A. M_n

$E[M_{n+1} \mid \xi_0, \dots, \xi_n] = E[\sum_{k=0}^{n+1} \xi_k \mid \xi_0, \cdots \xi_n]$

$= \sum_{k=0}^{n} E[\xi_k \mid \xi_0 \cdots \xi_k] + E[\xi_{n+1} \mid \xi_0 \cdots \xi_n]$

$g(\xi_0 \cdots \xi_k) = \xi_k$ $\quad$ iid

$= \sum_{k=0}^{n} \xi_k + E[\xi_{n+1}]$

0

$= M_n$

useful fact: ① $E[X+Y \mid Z]$
$= E[X \mid Z] + E[Y \mid Z]$
② $E[g(Y) \mid Y] = g(Y)$

Martingale Property: $E[M_{n+1} \mid \xi_0 \cdots \xi_n] = M_n$

Def: Martingale

"information process" $(X_k)_{k=0}^{\infty}$ (discrete time process)

$(M_n)_{n=0}^{\infty}$ is a martingale wrt $(X_k)_{k=0}^{\infty}$ if

① M_n can be determined from $X_0 \cdots X_n$: $M_n = g(X_0 \cdots X_n)$ for $g: \mathbb{R}^{n+1} \to \mathbb{R}$

② $E[M_n] < \infty$

③ Martingale property: $E[M_{n+1} \mid X_0 \ldots X_n] = M_n$.

Facts: 1. Martingales has constant mean for all times. (not expected to grow in time)

$E[E[M_{n+1} \mid X_0 \ldots X_n]] = E[M_{n+1}]$

$E[M_{n+1}] = E[M_n] = \cdots = E[M_0]$

2. If ① holds, then $M_n = g(X_0 \cdots X_n) = E[g(X_0 \cdots X_n) \mid X_0 \cdots X_n]$

③ is equivalent to $E[M_{n+1} - M_n \mid X_0 \cdots X_n] = 0$

increment of martingale.

ex. $(\xi_k)_{k=0}^{\infty}$ iid, mean-zero, $M_n = \left(\sum_{k=0}^{n} \xi_k\right)^2$

Q. Is M_n a martingale wrt to $(\xi_k)_{k=0}^{\infty}$?

A. No, except $\xi_0 = \cdots = \xi_n = 0$.

Property ① holds taking $g(\{\xi_0, \ldots, \xi_n\}) = (\xi_0 + \cdots + \xi_n)^2$

Property ③ fails

$E[M_0] = E[\xi_0^2]$

$E[M_n] = E\left[\left(\sum_{k=0}^{n} \xi_k\right)^2\right] = Var\left(\sum_{k=0}^{n} \xi_k\right)$ as $E\left[\sum_{k=0}^{n} \xi_k\right]^2 = 0$

$= E\left(\sum_{k=0}^{n} \xi_k^2 + \sum_{k \neq j} \xi_k \xi_j\right)$

$= \sum_{k=0}^{n} E(\xi_k^2) + \sum_{k \neq j} E(\xi_k) E(\xi_j)$ ↑ 0

$= \sum_{k=0}^{n} E[\xi_k^2]$

$= n+1\, E(\xi_0^2) \neq E(\xi_0^2)$

Q. Can we make this into a martingale ?

A. define $M_n = \left(\sum_{k=0}^{n} \xi_k\right)^2 - (n+1) E(\xi_0^2)$

Claim: $(M_n)_{n=0}^{\infty}$ is a martingale.

$$\Rightarrow M_{n+1} = \left(\sum_{k=1}^{n} \xi_k + \xi_{n+1}\right)^2 - (n+1)E(\xi_0^2)$$
$$= \underbrace{\left(\sum_{k=1}^{n} \xi_k\right)^2 - nE(\xi_0^2)}_{M_n} + \xi_{n+1}^2 - E(\xi_0^2) + 2\left(\sum_{k=1}^{n} \xi_k\right)\cdot \xi_{n+1}$$

$$E[M_{n+1} - M_n \mid \xi_0 \ldots \xi_n] = E[\xi_{n+1}^2 - E(\xi_0^2) \mid \xi_0 \ldots \xi_n] + 2\sum_{k=1}^{n} E[\xi_k \xi_{n+1} \mid \xi_0 \ldots \xi_n]$$
$$= E[\xi_{n+1}^2 - E(\xi_0^2)] + 2\sum_{k=1}^{n} \xi_k \, E[\xi_{n+1} \mid \xi_0 \ldots \xi_n]$$
$$= 0$$

$E[E[\xi_0^2]] = E[E[\xi_0^2 \mid \xi_1 \cdots \xi_n]] = E[\xi_0^2]$

ex. $(\xi_k)_{k=0}^{\infty}$, iid, $E[\xi_1] = 1$, $(M_n)_{n=0}^{\infty}$, $M_n = \xi_0 \cdots \xi_n = \prod_{k=0}^{n} \xi_k$

Property ①: $g(\xi_0 \ldots \xi_n) = \prod_{k=0}^{n} \xi_k$ holds

Property ③: $E[M_n] = E[\xi_0 \cdots \xi_n] = E[\xi_0] \cdots E[\xi_n] = 1 = E[\xi_0] < \infty$

showing constant, not showing ③

$$M_{n+1} - M_n = \xi_0 \cdots \xi_{n+1} - \xi_0 \cdots \xi_n$$
$$= \xi_0 \cdots \xi_n (\xi_{n+1} - 1)$$

$$E[M_{n+1} - M_n \mid \xi_0, \ldots, \xi_n] = E[\xi_0 \cdots \xi_n(\xi_{n+1} - 1) \mid \xi_0 \ldots \xi_n]$$
$$= \xi_0 \cdots \xi_n \, E[\xi_{n+1} - 1 \mid \xi_0 \cdots \xi_n]$$
$$= \xi_0 \cdots \xi_n \, E[\underbrace{\xi_{n+1} - 1}_{0}]$$
$$= 0$$

3 holds.

10/26 Continue - Martingale.

ex.1. Poly Urn Model

g_0 green balls

y_0 yellow balls

$\frac{g_0}{g_0 + y_n}$ = proportion of green balls.

One time step: choose a ball at random

put it back and add another of the same color.

"Info process": $(X_n)_{n=0}^{\infty}$ $X_n = (G_n, Y_n)$ G_n = # green balls, Y_n = # of yellow balls.

$(M_n)_{n=0}^{\infty}$ is defined by $M_n = \frac{G_n}{G_n + Y_n}$

Claim: $(M_n)_{n=0}^{\infty}$ is a martingale w.r.t. X_n.

Proof: NTS $E[M_{n+1} | X_0 \cdots X_n] = M_n$.

$$E[M_{n+1} | X_0, \cdots X_n] = E[M_{n+1} | X_n]$$
$$= \frac{G_n}{G_n + Y_n} \cdot \frac{G_n + 1}{G_n + Y_n + 1} + \frac{Y_n}{G_n + Y_n} \cdot \frac{G_n}{G_n + Y_n + 1}$$
$$= \frac{G_n (G_n + 1 + Y_n)}{(G_n + Y_n)(G_n + Y_n + 1)}$$
$$= \frac{G_n}{G_n + Y_n} = M_n$$

ex.2. Betting on a Martingale:

Let $(M_n)_{n=0}^{\infty}$ be a martingale w.r.t. $(X_k)_{k=0}^{\infty}$

Betting strategy: H_k is a bet on M_k given info up till $k-1$

H_k may only depend on $X_0 - X_{k-1}$

$H_k > 0$ pays out if $M_k > M_{k-1}$ (long position)

$H_k < 0$ pays out if $M_k < M_{k-1}$ (short position)

payoff in step k is $H_k (M_k - M_{k-1})$

payoff af n bets $(W_n)_{n=1}^{\infty} = \sum_{k=1}^{n} H_k (M_k - M_{k-1})$

Claim: $(W_n)_{n=1}^{\infty}$ is a martingale w.r.t. $(X_n)_{n=1}^{\infty}$

1. W_n only depends on $X_0 \sim X_n$ ✓
2. finite expectation needed
3. martingale property.

$$E[W_{n+1} - W_n | X_0 \cdots X_n] = E[H_{n+1} (M_{n+1} - M_n) | X_0 \cdots X_n]$$
$$= H_{n+1} E[(M_{n+1} - M_n) | X_0 \cdots X_n]$$
$$= H_{n+1} \cdot 0 = 0 \text{ as } (M_n)_{n=0}^{\infty} \text{ is a martingale w.r.t. } (X_n)_{n=0}^{\infty}$$

Martingales + Stopping Times

$(M_n)_{n=0}^{\infty}$ relative $(X_n)_{n=0}^{\infty}$

T is a stopping time relative to $(X_n)_{n=0}^{\infty}$

$\{T \le n\}$ depends only on $X_0 \cdots X_n$

Q. We know $E[M_n] = E[M_0]$ $\forall n \geq 1$? $E[M_T] = E[M_0]$

Stopped Martingales

Notation: $\min(a,b) = a \wedge b$ $\quad \max(a,b) = a \vee b$

Consider the process $(M_{n\wedge T})_{n=0}^{\infty}$

Claim: $(M_{n\wedge T})_{n=0}^{\infty}$ is a Martingale w.r.t. $(X_n)_{n=0}^{\infty}$

Proof: (martingale property)

$$E[M_{(n+1)\wedge T} \mid X_0, \ldots, X_n]$$
$$= E[1_{\{T\le n\}} M_{(n+1)\wedge T} \mid X_0, \cdots, X_n]$$
$$+ E[1_{\{T\ge n+1\}} M_{(n+1)\wedge T} \mid X_0, \cdots, X_n]$$
$$= \sum_{k=0}^{n} E[1_{\{T=k\}} M_k \mid X_0, \cdots, X_n]$$
$$+ 1_{\{T\ge n+1\}} E[M_{n+1} \mid X_0, \ldots, X_n]$$

$1 = 1_A + 1_{A^c}$

$1_A(w) = \begin{cases} 1 & w\in A \\ 0 & \text{~~otherwise~~} \end{cases}$

$1_{\{T\ge n+1\}} = g(X_0, \cdots X_n)$ since T is a stopping time $= 1 - 1_{\{T\le n\}}$

$$= \sum_{k=0}^{n} 1_{\{T=k\}} M_k + 1_{\{T\ge n+1\}} M_n$$
$$= E[M_{(n+1)\wedge T} \mid X_0, \cdots, X_n]$$

$M_{n\wedge T} = M_{n\wedge T} 1_{\{T\le n\}} + M_{n\wedge T} 1_{\{T\ge n+1\}}$
$= \sum_{k=0}^{n} 1_{\{T=k\}} M_k + M_n 1_{\{T\ge 1\}}$

Therefore, $E[M_{(n+1)\wedge T} \mid X_0, \cdots, X_1] = M_{n\wedge T}$

Q. $E[M_T] = E[M_0]$?

$(M_{T\wedge n})_{n=0}^{\infty}$ is a martingale

Hence, $E[M_{T\wedge n}] = E[M_0] \quad \forall n\ge 1$

$T\wedge n \to T$ as $n\to\infty$

$\lim_{n\to\infty} M_{T\wedge n} = M_T$ A.S. we might try to write $E[M_T] = E[\lim_{n\to\infty} M_{T\wedge n}]$

$\overset{?}{=} \lim_{n\to\infty} E[M_{n\wedge T}]$

$= \lim_{n\to\infty} E[M_0]$

$= M_0$

Not always true, $\Rightarrow$

but under some additional assumptions

this works : "Optimal Sampling Thm"

10/28

(relative to $(X_k)_{k=0}^{\infty}$)

Last Time $(M_n)_{n=0}^{\infty}$ a martingale w.r.t. $(X_k)_{k=0}^{\infty}$ and T is a stopping time, $P(T<\infty)=1$

know: $E[M_n] = E[M_0]$

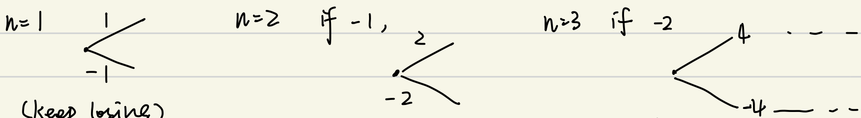
$E[M_{n\wedge T}] = E[M_0]$

Q. $n\to\infty$? $E[M_T] = E[M_{n\wedge T}]$

$\Rightarrow$? $E[M_T] = E[M_0]$

A. Sometimes, but not always.

ex. (Doubling strategy) stop at winning $E[0]=0$



(keep losing)

if the game ends at time n, total winnings $= -\Big(\underbrace{\sum_{k=0}^{n-2} 2^k}_{n-1 \text{ loss}}\Big) + \underbrace{2^{n-1}}_{1 \text{ win}}$

$= -(2^{n-1}-1) + 2^{n-1} = 1 \neq E[0]$

ex. (Betting Strategy) $(H_k)_{k=0}^{\infty}$

$(X_k)_{k=1}^{\infty}$ i.i.d. $P(X_1=1) = P(X_1=-1) = \frac{1}{2}$

bet on the simple symmetric random walk started from 0. $M_n = \sum_{k=1}^{n} X_k$.

$H_1 = 1$, $H_2 = \begin{cases} 2 & \text{if } X_1=-1 \ (M_1<M_0) \\ 0 & \text{else} \end{cases}$

$H_k = \begin{cases} 2^k & \text{if } H_{k-1}\neq 0,\ X_{k-1}=-1 \\ 0 & \text{else} \end{cases}$

$(W_n)_{n=1}^{\infty}$ defined by $W_n = \sum_{k=1}^{n} H_k(M_k - M_{k-1}) = \sum_{k=1}^{n} H_k X_k$

is a martingale w.r.t. $(X_k)_{k=1}^{\infty}$

Hence, $E[W_n] = E[W_1] = E[H_1X_1] = E[X_1] = 0$

However, let $T = \inf\{k\geq 1 : X_k = 1\}$

know: $P(T<\infty)=1$, T is a stopping time relative to $(X_k)_{k\geq 1}$

W_T is well defined

$E[W_T] = 1 \neq E[W_1]$

Optional Stopping:

Thm. if any of the a/b/c holds, then $E[M_T] = E[M_0]$

a) $P(T \le c) = 1$ for $c \in \mathbb{R}$. (the stopping T is almost surely bounded)

b) $P(|M_{T\wedge n}| \le c) = 1$ $\forall n \ge 1$, some $c \in \mathbb{R}$. (the stopped process is A.S. bounded in time)

c) (more technical) $P(|M_{T\wedge(n+1)} - M_{T\wedge n}| \le c) = 1$ $\forall n \ge 1$, some $c \in \mathbb{R}$.

ex. (Hitting Times for a simple symmetric random walk)

$(X_k)_{k=1}^{\infty}$ i.i.d. $P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}$

$M_0 = 0$, $M_n = \sum_{k=1}^{n} X_k$.

Fix $a, b > 0$, let $T = \inf\{k \ge 1 : M_n \in \{-a, b\}\}$

Goal: use the optional sampling Thm to calculate $P(M_T = -a)$

Note: $|M_{T\wedge n}| \le \max(a, b)$, so, may use (b) of the Thm to find $E[M_T] = E[M_0] = 0$.

$$E[M_T] = -a P(M_T = -a) + b P(M_T = b)$$
$$= -(a+b) P(M_T = -a) + b = 0$$
$$\Rightarrow P(M_T = -a) = \frac{b}{a+b}, \quad P(M_T = b) = \frac{a}{a+b}$$

Remark. ① if $a = b$, then $P(M_T = -a) = P(M_T = a) = \frac{1}{2}$

② if $a < b$, then $P(M_T = -a) = \frac{b}{a+b} > \frac{a}{a+b} = P(M_T = b)$

Q. ? $E[T]$ need (c)

Consider instead $M_n = \left(\sum_{k=1}^{n} X_k\right)^2 - n$, $M_0 = 0$

By optional sampling, $E[M_T] = E[M_0] = 0$

$$E\left(\left(\sum_{k=1}^{T} X_k\right)^2 - T\right) = 0$$
$$E\left[\left(\sum_{k=1}^{T} X_k\right)^2\right] = E[T]$$
$$\uparrow$$
$$a^2 \frac{b}{a+b} + b^2 \frac{a}{a+b} = \frac{ab(a+b)}{a+b} = ab$$

$E[T] = ab$.

Q. ? if the walk is asymmetric

$(X_k)_{k=1}^{\infty}$ i.i.d. $P(X_1 = 1) = p$ $P(X_1 = 0) = 1 - p$ $p \ne \frac{1}{2}$.

$M_0 = 0$, $M_n = \sum_{k=1}^{n} X_k$

$E[X_k] = E[X_1] = 1\cdot p + (-1)(1-p) = 2p-1$

$E[M_n] = \sum_{k=1}^{n} E[X_k] = n\cdot(2p-1)$

Consider, instead, $(M_n - n(2p-1))_{k=0}^{\infty}$ is a martingale

Optional Sampling implies $E[M_T - T(2p-1)] = 0$

$\rightarrow E[M_T] = (2p-1)E[T]$

$\uparrow$ not know.

Need a different martingale.

Consider, instead, $\prod_{k=1}^{n} \left(\frac{1-p}{p}\right)^{X_k} = \left(\frac{1-p}{p}\right)^{\sum_{k=1}^{n} X_k}$

10/30

Continue example last time.

Optional Stopping Thm: criteria to ensure $E[M_T] = E[M_0]$

ex. Asymetric random walk.

$S_0 = 0$, $S_n = \sum_{k=1}^{n} X_k$, where X_k is i.i.d. w/ $P(X_1=1)=p$, $P(X_1=-1)=1-p$

Fix $a, b > 0$ $T = \inf\{n \geq 0 : S_n \in \{-a, b\}\}$

Goal: compute $P(S_T = -a)$ and $P(S_T = b)$

Recall: given $(\xi_k)_{k=1}^{n}$ i.i.d. with $E(\xi_k) = 1$, then letting $M_0 = 1$ and $M_n = \prod_{k=1}^{n} \xi_k$

M_n is a martingale w.r.t. $(\xi_k)_{k=1}^{n}$

Consider $\xi_k = \left(\frac{1-p}{p}\right)^{X_k}$, note $(\xi_k)_{k=0}^{\infty}$ are i.i.d.

$\xi_k = e^{\ln\left(\frac{1-p}{p}\right)\cdot X_k}$

$E(\xi_k) = \left(\frac{1-p}{p}\right)^{1} \underbrace{P(X_k=1)}_{p} + \left(\frac{1-p}{p}\right)^{-1} \underbrace{P(X_k=-1)}_{1-p}$

$= 1-p+p = 1$

Hence, $M_n = \prod_{k=1}^{n} \left(\frac{1-p}{p}\right)^{X_k}$ is a martingale.

$= \left(\frac{1-p}{p}\right)^{X_1} \cdots \left(\frac{1-p}{p}\right)^{X_n}$

$= \left(\frac{1-p}{p}\right)^{\sum_{k=1}^{n} X_k} = \left(\frac{1-p}{p}\right)^{S_n}$

Claim: For all $n \geq 0$ $P(|M_{n\wedge T}| \leq c) = 1$ for $c = \begin{cases} \left(\frac{1-p}{p}\right)^{b} & \text{if } \frac{1-p}{p} > 1 \\ \left(\frac{1-p}{p}\right)^{-a} & \text{if } \frac{1-p}{p} < 1 \end{cases}$

Proof: $-a \le S_{n\wedge T} \le b$

if $\frac{1-p}{p} > 1$, then $g(y) = (\frac{1-p}{p})^y = e^{\ln(\frac{1-p}{p})y}$ is increasing

if $\frac{1-p}{p} < 1$, then $g(y)$ is decreasing

Hence, by optional stopping, $E[M_T] = E[M_0] = 1$

$$1 = E[M_T] = E((\frac{1-p}{p})^{S_T}) = (\frac{1-p}{p})^{-a} P(S_T=-a) + (\frac{1-p}{p})^b P(S_T=b)$$

$$= [(\frac{1-p}{p})^{-a} - (\frac{1-p}{p})^b] P(S_T=-a) + (\frac{1-p}{p})^b$$

$$P(S_T=-a) = \frac{1-(\frac{1-p}{p})^b}{(\frac{1-p}{p})^{-a} - (\frac{1-p}{p})^b}$$

Remark: Last time, we show that $S_n - (2p-1)n$ is a martingale

and $E[S_T - (2p-1)T] = 0 \rightarrow E(T) = \frac{1}{2p-1} E[S_T]$

Long Time Behavior of Martingales

ex. poly-urn model.

1 urn, G_n green balls, Y_n yellow balls

start with 1 green ball and 1 yellow ball

$M_n = \frac{G_n}{G_n+Y_n}$ is a martingale w.r.t. $(G_n, Y_n)_{n=0}^{\infty}$

Q. Does M_n has a limit as $n \to \infty$

$n=1$ $P(G_1=1) = P(G_1=2) = \frac{1}{2}$ G_1 is uniformly distributed on $\{1,2\}$.

$n=2$ $P(G_1=1) = P(G_2=1 \mid G_1=1) P(G_1=1) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$

$P(G_2=2) = P(G_2=2 \mid G_1=1) P(G_1=1) + P(G_2=2 \mid G_1=2) P(G_1=2)$

$= \frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{3}$

$P(G_3=3) = P(G_3=3 \mid G_2=2) P(G_2=2) = \frac{2}{3} \times \frac{1}{3} = \frac{1}{3}$

G_3 is uniformly distributed on $\{1,2,3\}$.

Guess: G_n is uniformly distributed on $\{1, \cdots, n\}$

Prove: by induction.

Assume for G_{n-1}, consider G_n.

Claim G_n is uniformly distributed on $\{1, \cdots, n-1\}$

n+1-k yellow, k green

n+2-k yellow, k-1 green

$2 \le k \le n+1$: $P(G_n=k) = P(G_n=k \mid G_{n-1}=k) P(G_{n-1}=k) + P(G_n=k \mid G_{n-1}=k-1) P(G_{n-1}=k-1)$

$$= \frac{n+1-k}{n+1} \cdot \frac{1}{n} + \frac{k-1}{n+1} \frac{1}{n} = \frac{1}{n+1}$$

From $G_n \to M_n$

$M_n = \frac{G_n}{G_n+Y_n} = \frac{G_n}{n+2}$, G_n is uniformly distributed on $\{1, \dots, n\}$

M_n is uniformly distributed on $\{\frac{k}{n+2} : 1 \le k \le n+1\}$

Claim: as $n \to \infty$, M_n converges to a uniform distribution, $\text{unif}[0,1]$

$F_{(n)}(t) = P(M_n \le t)$

$0 < t < 1 \quad F_n(t) = \frac{\lfloor (n+2)t \rfloor}{n+1}$

$$\lim_{t \to \infty} F_n(t) = \lim_{t\to\infty} \frac{\lfloor (n+2)t \rfloor}{n+1}$$

Hence, $F_n(t) \to F(t) \quad F(t) = \begin{cases} 0 & t<0 \\ t & 0 \le t \le 1 \\ 1 & t > 1 \end{cases}$

Thm: Martingale Convergence Thm:

Given a martingale $(M_n)_{n=0}^{\infty}$, which satisfies either

① $P(M_n \ge c) = 1 \quad \forall n \ge 0$ and some $c \in \mathbb{R}$.

② $P(M_n \le c) = 1 \quad \forall n \ge 0$

Then M_n converges a.s. to a random variable M_∞.

Remark: Given a sequence $(a_n)_{n=0}^{\infty}$, which takes values in $\mathbb{Z}$, then

$\lim_{n\to\infty} a_n = a$ iff $a_n = a$ for all $n \ge N_0$ (for some N_0)

Hence, the single r.v. (either symmetric / asymmetric)

does not converge as $n \to \infty$

11/02

Branching Processes

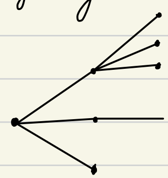
- model for population growth
- $(X_n)_{n=0}^{\infty}$ Markov chain s.t. X_n = size of population at time x

Q. 1. How do we module this? Compute the transition matrix
2. Long time behavior: extinction $\Rightarrow X_n = 0$ for $n \geq N$

① Model: picture of a sample trajectory graph: infinitely long or stop at some time

$n=0$ $n=1$ $n=2$

$X_0=1$ $X_1=3$ $X_2=4$



- offspring distribution $(\beta_k)_{k=0}^{\infty}$ β_k = prob. that a given node has k children.

Define X_n: array $\{Z_{n,j}\}_{n,j=1}^{\infty}$ of i.i.d random distribution s.t. $P(Z_{1,1}=k)=\beta_k$

$Z_{n,j}$ = # offspring off population #j at time $n-1$

starting with a population size $X_0 > 0$

$X_1 = Z_{1,1} + \cdots + Z_{1,X_0}$

At time $n+1$, either $X_{n+1} = 0$ or $X_n > 0$, consider $Z_{n+1,1}, \cdots Z_{n+1,X_n}$

$$X_{n+1} = \begin{cases} 0 & \text{if } X_n = 0 \\ \sum_{j=1}^{X_n} Z_{n+1,j} & \text{if } X_n > 0 \end{cases}$$

Remark: not all elements of the array are used (for a given trajectory w)

ex. if $X_0 = 3$, $Z_{1,1}, Z_{1,2}, Z_{1,3}$ are used, but $Z_{1,k}$ for $k \geq 4$ are not used

not sure how many will be used $\rightarrow$ need to make array Z large.

Transition matrix: $p(0,0) = 1$

$p(k,l)$ for $k>0$

$= P(X_{n+1} = l \mid X_n = k)$

Use: sum of independent r.v.

$Y_1, \cdots, Y_k$ independent

Q. dist. of $Y_1 + \cdots + Y_k$

$k=2$: $P(Y_1+Y_2 = l) = \sum_{\substack{l_1+l_2=l \\ l_1,l_2 \geq 0}}^{\infty} P(Y_1 = l_1) P(Y_2 = l_2)$

$k \geq 2$: $P(Y_1 + \cdots + Y_k = l) = \sum_{l_1+\cdots+l_k=0}^{\infty} P(Y_1 = l_1, \cdots Y_{k-1} = l_{k-1}, Y_1 + \cdots + Y_k = l)$

$= \sum_{\substack{l_1+\cdots+l_k=l \\ l_1,\cdots l_k \geq 0}} P(Y_1 = l_1) \cdots P(Y_k = l_k)$

$P(k,l) = P(X_{n+1} = l \mid X_n = k)$

$= P\left(\sum_{j=1}^{k} Z_{i+1,j} = l \mid X_n = k\right)$ depend on $Z_{n,k}$, independent of $Z_{n+1,j}$

$= P\left(\sum_{j=1}^{k} Z_{i+1,j} = l\right)$

$= \sum_{\substack{l_1+\cdots+l_k=l \\ l_1,\cdots,l_k \geq 0}} \beta_{l_1} \cdots \beta_{l_k}$

② Long Time: what is the likelyhood of distinction.

Goal: calculate $\pi = P_1(X_n = 0 \text{ for some } n \geq 1)$ initial population size is 1

$= P_1 \bigcup_{n=1}^{\infty} \{X_n = 0\}$

depending on $(\beta_k)_{k=0}^{\infty}$

Note: If $X_n = 0$ for some n, then $X_m = 0 \ \forall m \geq n$

Endpoint cases: (trival)

1. $\beta_0 = 0$ each node has at least one child.

$\to \pi = 0$

2. $\beta_0 = 1$ each node cannot reproduce

$\to P(X_1 = 0) = 1$

$\to \pi = 1$

Interesting cases: $0 < p_0 < 1$

Canonical decomposition of $S = \mathbb{Z}_+ \{0, 1, 2, \cdots\}$

$S = T \cup R$

$T = \mathbb{N} = \{1, 2, 3 \cdots\}$

$R = \{0\}$ absorbing state.

Proof: $p(0,k) = 0$ for $k \neq 0$

$$p(k,0) = \sum_{\substack{l_1+\cdots+l_k=0 \\ l_1,\cdots,l_k \geq 0}} p_{l_1} \cdots p_{l_k} = p_0^k > 0 \quad \text{for } k \geq 1$$

Extinction v.s. explosion: If $X_n \geq 0$ $\forall n \geq 1$, what happens as $n \to \infty$?

Ex. $\lim_{n\to\infty} X_n = 1{,}000{,}000$?

No: only possibility is $\lim_{n\to\infty} X_n = \infty$

Recall: transient states are only visited a finite # times

For any $M \in \mathbb{N}$, $\exists$ a random time N s.t. for $n \geq N$, $X_n = 0$ or $X_n \geq M$.

Summary: $P(\text{extinction}) + P(\text{explosion}: \lim_{n\to\infty} X_n = \infty) = 1$

Criticality and Extinction:

Consider the mean of the off-spring dist. $\mu = \sum_{k=1}^{\infty} k p_k$.

A branching process is

- sub-critical if $\mu < 1$
- critical if $\mu = 1$
- super-critical if $\mu > 1$

Thm: Let $p_0 > 0$

For a super-critical branching process, $\eta \in (0,1)$

... sub-critical or critical ..., $\eta = 1$ (extinction is certain)

11/04

branching process (ctn)

Goal: study extinction via the PGF (prob. generating function) of the off-spring distribution.

Remark: PGF is similar to MGF (moment generating function)

Setup: r.v. X w/ values in $\mathbb{Z}_+ \cup \{\infty\} = \{0,1,2 \cdots\} \cup \{\infty\}$

The PGF: $g: [-1,1] \to \mathbb{R}$ is defined by $g(s) = \sum_{k=0}^{\infty} s^k P(X=k)$

$= E[s^X 1_{X<\infty}]$

Remarks: g is well-defined, the series converges $|s| \le 1 \to |s^k P(X=k)| < P(X=k)$

$\sum_{k=1}^{\infty} |s^k P(X=k)| \le \sum_{k=1}^{\infty} P(X=k) = P(X<\infty) \le 1$

If $P(X<\infty) = 1$, then $g(s) = E[s^X]$

Ex. Binomial Dist. $X \sim B(N,p)$

$$P(X=k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & 0 \le k \le n \\ 0 & \text{else} \end{cases}$$

Note: $P(X<\infty) = 1$, so $g(s) = E[s^X] = \sum_{k=0}^{\infty} s^k P(X=k)$

$= \sum_{k=0}^{\infty} s^k \binom{n}{k} p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} \binom{n}{k} (sp)^k (1-p)^{n-k}$

$= (sp + (1-p))^n$

Ex.
$$P(X=k) = \begin{cases} \frac{1}{2} & k=0 \\ 2^{-k} & k \ge 2 \end{cases}$$
$E[X] = \infty$

$g(s) = \sum_{k=0}^{\infty} s^k P(X=k)$

$= \sum_{k=2}^{\infty} \left(\frac{s}{2}\right)^k = \sum_{k=0}^{\infty} \left(\frac{s}{2}\right)^k - \left(1 + \frac{s}{2}\right)$

$= \frac{1}{1-\frac{s}{2}} - 1 - \frac{s}{2}$

Properties: 1: $g(0) = 1 - P(X=\infty)$

pf: $g(0) = \sum_{k=0}^{\infty} P(X=k) = P(X<\infty)$

2. sum of i.i.d finite r.v.s $X_1, \cdots, X_n$ finite, i.i.d, $g_{X_1+\cdots X_n}(s) = (g_{X_1}(s))^n$

Pf: $E[s^{X_1+\cdots+X_n}] = E[s^{X_1} \cdots s^{X_n}] = (E[s^{X_1}])^n = (g_{X_1}(s))^n$

by i.i.d.

Recall: for a brenching process $(X_n)_{n=0}^{\infty}$

if $X_n > 0$, then $X_{n+1} = \sum_{k=1}^{X_n} Z_{n+1,k}$

Properties: 3 random sums of finite, i.i.d. r.v.

N r.v. w/ values in $\mathbb{Z}_+$ $\quad \{Z_k\}_{k=1}^{\infty}$ i.i.d. finite, ind. of N

$S_N = \sum_{k=1}^{n} Z_k$, consider S_N

$$\begin{aligned} g_{S_N}(s) &= E[S^{S_n}] \\ &= E[E[S^{S_n} \mid N]] \\ &= \sum_{n=0}^{\infty} E[S^{S_n} \mid N=n]\, P(N=n) \\ &= \sum_{n=0}^{\infty} E[S^{S_n}]\, P(N=n) \\ &= \sum_{n=0}^{\infty} [g_{Z_1}(s)]^n\, P(N=n) \\ &= g_N(g_Z(s)) \end{aligned}$$

Summary: $g_{S_N}(s) = g_N(g_{Z_1}(s))$

extinction b.p. $(X_n)_{n=0}^{\infty}$ w/ offspring dist. $(\beta_k)_{k=0}^{\infty}$

$\pi = P_1(\text{extinction})$

Lemma: π is a fixed point of the pgf associted to the $(\beta_k)_{k=0}^{\infty}$

$g(s) = \sum_{k=0}^{\infty} \beta_k s^k$

$g(\pi) = \pi \qquad$ (fixed point)

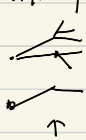
proof: consider one-step $\quad \pi = P_1(\text{extinction})$

$$= \sum_{k=0}^{\infty} P_1(\text{extinction}, X_1 = k)$$

$k=0$ $\quad P_1(\text{extinction}, X_1=0) = P_1(X_1=0) = \beta_0$

$k \geq 1$ $\quad P_1(\text{extinction}, X_1=k) = P_1(X_1=k) \cdot P_k(\text{extinction}) = \beta_k \pi^k$

eg. $k=2$:



both need to extinct

$$\pi = \beta_0 \pi^0 + \sum_{k=1}^{\infty} \beta_k \pi^k = g(\pi)$$

Difficulty: there could be many fixed points of g

Thm: π is the smallest fixed point of g in $[0,1]$

proof:

Fact: $g_{X_n} = g \circ \overset{n}{\cdots} \circ g$

$g_{X_1} = g$ since $X_0 = 1$

$X_2 = \sum_{k=1}^{X_1} \xi_{2,k}$ ← g_{X_1}; $\xi_{2,k}$ → g

$g_{X_2} = g_{X_1} \circ g = g_{X_1}(g)$

⋮

$g_{X_{n+1}} = g_{X_n} \circ g$

$\Rightarrow g_{X_n} = g \circ \overset{n}{\cdots} \circ g$

Goal: given $\sigma \in [0,1]$ s.t. $g(\sigma) = \sigma$, show $\pi < \sigma$

$\pi = P_1(\text{extinction}) = P_1 \bigcup_{n=1}^{\infty} \{X_n = 0\}$

$\{X_1 = 0\} \subset \{X_2 = 0\} \subset \cdots$

$= \lim_{n \to \infty} P_1(X_n = 0)$

$= \lim_{n\to\infty} g_{X_1}(0) \leq \lim_{n\to\infty} g_{X_n}(0)$

$= \lim_{n\to\infty} g \circ \cdots g(\sigma)$

$= \sigma$

$g \circ g(\sigma) = g(g(\sigma))$
$= g(\sigma) = \sigma$

11/06

Last: Generating Function.

$g(s) = \sum_{k=0}^{\infty} p_k s^k$

s is a fixed point of g if $g(s) = s$

π is the smallest fixed point in $[0,1]$

$\Rightarrow$ Graph of $g: [0,1] \to \mathbb{R}$

1. $g(0) = p_0$,

$g(s) = p_0 + s p_1 + s^2 p_2 \cdots$

$g'(s) = p_1 + 2s p_2 + \cdots\cdots = \sum_{k=1}^{\infty} k p_k s^{k-1}$

2. g is increasing on $[0,1]$

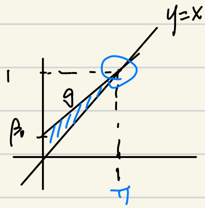
$g''(s) = 2p_2 + 3\cdot 2 p_3 s + \cdots = \sum_{k=2}^{\infty} k(k-1) p_k s^{k-2} \geq 0$

3. g is convex $(g'' > 0)$

Special Case: $p_k = 0$ for $k \geq 2$

$p_0 + p_1 = 1 \Rightarrow g(s)$ is a straight line

$\pi = 1$, extinction is certain

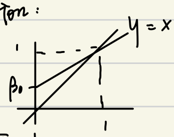


Notice: $g'(1) < 1$

$g'(1) = \mu = \sum_{k=1}^{\infty} k p_k$

Relationship with extinction:

Case 1: $g'(1) \leq 1$



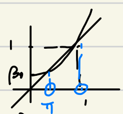
$\Rightarrow$ If $\mu \leq 1$, then $\pi = 1$

sub-critical $(\mu < 1)$ or critical $(\mu = 1)$ case.

Case 2: $g'(1) > 1$

$\mu > 1$: Super-critical

$\Rightarrow$ If $\mu > 1$, then $\pi \in (0,1)$



Poisson Process

- Renewal process $(N_t)_{t\geq 0}$ is a continuous time stochastic process.
  N_t : # of arrivals up till time t

Define, in terms of i.i.d. sequence $(X_k)_{k\geq 1}$ of waiting times,
$\sum_{k=1}^{n} X_k$ = time it takes for the first n arrivals to occur.
$N_t = \max\{n\geq 0 : \sum_{k=1}^{n} X_k \leq t\}$.

Def. Poisson Process w/ rate λ
Renewal Process where $X_k \sim Exp(\lambda)$

Review: Exponential Dist. $X\sim Exp(\lambda)$

pdf: $f(x) = \begin{cases} \lambda e^{-\lambda x} & x>0 \\ 0 & x\leq 0 \end{cases}$

$E[X] = \frac{1}{\lambda}$ $Var[X] = \frac{1}{\lambda^2}$

Q: the distribution of $\sum_{k=1}^{n} X_k$ if $X_k \sim Exp(\lambda)$

ex. $n=2$. $f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_{X_1}(y) f_{X_2}(x-y)\,dy$ $\quad y\geq 0,\ x-y\geq 0$

$= \int_0^x \lambda e^{-\lambda y}\cdot \lambda e^{-\lambda(x-y)}\,dy$

$= \lambda^2 e^{-\lambda x}\int_0^x 1\,dy$

$= \lambda^2 x e^{-\lambda x}$

$f_{X_1+X_2}(x) = \begin{cases} \lambda^2 x e^{-\lambda x} & x>0 \\ 0 & x\leq 0 \end{cases}$ (gamma dist).

Gamma Dist. $X\sim Gamma(r,\lambda)$

pdf: $f(x) = \begin{cases} \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} & x>0 \\ 0 & x\leq 0 \end{cases}$

$\Gamma(n) = (n-1)!$ $\Gamma(v)$ = integral (note)

$exp(\lambda)$ corresponds to $Gamma(1,\lambda)$

Lemma: If $(X_k)_{k=1}^{n}$ i.i.d. $Exp(\lambda)$, then $\sum_{k=1}^{n} X_k$ has $\Gamma(n,\lambda)$ distribution.

Def. Exponential Races

$(X_k)_{k=1}^{n}$ $X_k \sim \text{Exp}(\lambda_k)$ i.i.d.

Q. Dist. of ① $\min(X_1 \dots X_n)$ ② $\text{argmin}(X_1 \dots X_n) = k: X_k = \min(X_1 \dots X_n)$

Fact: $\min(X_1 \dots X_n) \sim \text{Exp}(\lambda_1 + \dots + \lambda_n)$

CDF: $X \sim \text{Exp}(\lambda)$

$$P(X \le x) = \int_{-\infty}^{x} \lambda e^{-\lambda y}\, dy = \int_0^x \frac{d(e^{-\lambda y})}{dy}\, dy = 1 - e^{-\lambda x}$$

$P(X \ge x) = e^{-\lambda x}$ survival function. $P(X = x) = 0$

$$P(\min(X_1 \dots X_n) \ge x) = P(X_1 \ge x \dots X_n \ge x) = \prod_{k=1}^{n} (X_k \ge x) = \prod_{k=1}^{n} e^{-\lambda_k x}$$
$$= e^{-\sum_{k=1}^{n} \lambda_k x}$$

$\Uparrow$

Racing Interpretation: n racers, X_k = time for racer k finishes

$\min(X_1 \dots X_n)$ = time when the winner crosses the finish line.

Fact: $P(\text{argmin}(X_1 \dots X_n) = k) = P(X_k \ge X_j \ \forall j \ne k)$

$$= \int_0^\infty \int_{X_k}^\infty \overset{k-1}{\cdots} \int_{X_k}^\infty \lambda_1 \cdots \lambda_n e^{-\lambda_1 x_1} \cdots e^{-\lambda_n x_n} \prod_{j \ne k} dx_j \, dx_k$$

$j \ne k$: $\int_{X_k}^\infty \lambda_j e^{-\lambda_j x_j} dx_j \rightarrow$ $= \int_0^\infty \lambda_k e^{-\lambda_k x_k} \prod_{j \ne k} e^{-\lambda_j x_k}\, dx_k$

$= P(X_j \ge X_k) = e^{-\lambda_j x_k}$. $= \int_0^\infty \lambda_k e^{-(\lambda_1 + \dots + \lambda_n) x_k}\, dx_k$

$$= \frac{\lambda_k}{\lambda_1 + \dots + \lambda_n} \int_0^\infty \frac{d}{dx_k}\left(e^{-(\lambda_1 + \dots + \lambda_n) x_k}\right) dx_k$$

$$= \frac{\lambda_k}{\lambda_1 + \dots + \lambda_n}$$

11/09

Last Poisson Process w/ rate λ

$(X_k)_{k=1}^{\infty}$ of i.i.d. $\text{Exp}(\lambda)$

$N_t = \max\{n \geq 0 : \sum_{k=1}^{n} X_k \leq t\}$

Recall: Poision Distribution

$X \sim \text{Poi}(\lambda)$ discrete r.v. w/ possible values $0, 1, \ldots$

$P(X=n) = e^{-\lambda}\frac{\lambda^n}{n!}$, $E[X] = \text{Var}(X) = \lambda$ λ: intensity

Last: $\sum_{k=1}^{n} X_k \sim \Gamma(n, \lambda)$

$$f_{\sum_{k=1}^{n} X_k}(s) = \begin{cases} \frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s} & s > 0 \\ 0 & s \leq 0 \end{cases}$$

Thm: given a Poisson Process, $(N_t)_{t \geq 0}$ w/ rate λ

$N_t \sim \text{Poisson}(\lambda t)$

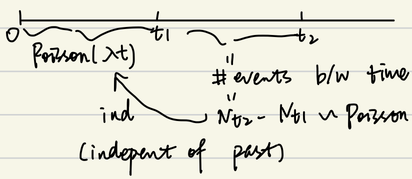
Proof: Consider the event $\{N_t = n\}$

1. $\sum_{k=1}^{n} X_k = s$ for some $s < t$
2. $X_{n+1} > t - s$

survival function last lec

$$\begin{aligned} P(N_t = n) &= \int_0^t f_{\sum_{k=1}^{n} X_k}(s)\, P(X_{n+1} > t-s)\, ds \\ &= \int_0^t \frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s}\, e^{-\lambda(t-s)}\, ds \\ &= \frac{\lambda^n e^{-\lambda t}}{(n-1)!} \int_0^t s^{n-1}\, ds \\ &= \frac{(\lambda t)^n}{n!} e^{-\lambda t} \end{aligned}$$

$\frac{d}{ds}\left(\frac{1}{n} s^n\right)$



$\overset{\shortparallel}{\#\text{events}}$ b/w time t_2 and t_1

$\overset{\shortparallel}{N_{t_2}} - N_{t_1} \sim \text{Poisson}(\lambda t_2 - \lambda t_1)$

Thm: Independent Poisson Increments

Let $0 < t_1 < \cdots < t_n$, then $N_{t_1}, N_{t_2} - N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}}$ are

1. independent
2. $N_{t_{j+1}} - N_{t_j} \sim Pois(\lambda(t_{j+1} - t_j))$

e.g. Suppose Customers arrives at starbucks at rate 30 / hour.
Model w/ poisson process w/ rate 30

Q1. Prob. that 10 customers arrive in the first hour

Sol: $N_1 \sim Poisson(30 \cdot 1) = Poisson(30)$

$P(N_1 = 10) = e^{-30} \frac{30^{10}}{10!}$

Q2. Prob. that 20 customers arrive in the first hour, 60 arrives in the first 3 hours

Sol: $P(N_1 = 20, N_3 = 60)$

N_1 and $N_3 - N_1$ are independent

$N_3 - N_1 \sim Poisson(60)$

$$\begin{aligned} P(N_1 = 20, N_3 = 60) &= P(N_1 = 20, N_3 - N_1 = 40) \\ &= P(N_1 = 20) P(N_3 - N_1 = 40) \\ &= e^{-30} \frac{30^{20}}{20!} e^{-60} \frac{30^{40}}{40!} \end{aligned}$$

Compound Poisson Process (S_t)

⇒ "Renewal Reward Process"

- $(X_k)_{k=1}^{\infty}$ arrival times i.i.d $Exp(\lambda)$
- $(Y_k)_{k=1}^{\infty}$ i.i.d. + independent from $(X_k)_{k=1}^{\infty}$
  ↳ reward from k^{th} arrival.

$(S_t)_{t=0}^{\infty}$ compound Poisson w/ rate λ

$$S_t = \begin{cases} 0 & N_t = 0 \\ \sum_{k=1}^{N_t} Y_k & N_t > 0 \end{cases}$$

e.g. starbucks continue, 30 customers arriving per hour
spend average \$6 per customer

Q. Expected revenue in the first t hours $E[S_t]$.

$E[S_t] = E[\sum_{k=1}^{N_t} Y_k]$

$= \sum_{k=1}^{\infty} E[\sum_{k=1}^{n} Y_k \mid N_t = n] P(N_t = n)$ Y_k independent from $X_k \Rightarrow N_t$.

$= \sum_{k=1}^{\infty} E[\sum_{k=1}^{n} Y_k] P(N_t = n)$

$= \sum_{n=1}^{\infty} n E[Y_1] P(N_t = n]$

$= E[Y_1] \sum_{n=1}^{\infty} n P(N_t = n)$

$= E[Y_1] E[N_t]$ $N_t \sim Pois(\lambda_t)$

$= E[Y_1] \lambda_t$

11/11

Last Poisson process w/ rate λ (intensity)
$(N_t)_{t\geq 0}$ cts. time stochastic process
· exponential waiting times $Exp(\lambda)$

Thinning a Poisson Process

e.g. starbucks continue: 30 customer per hour
N_t = # customers arriving up till time t.
Suppose there are ℓ possible types of customer.

- e.x. $\ell = 12$. "type" based on birthday month of the customer
  type $j \in \{1, \cdots, 12\}$

also independent of N_t

- P_j = prob a customer has type j $P_1 + \cdots + P_{12} = 1$ i.i.d. assumption among customers
- Consider ℓ different processes $(N_t^j)_{t \geq 0}$ for $j = 1, \cdots, \ell$
  = # customers up to time t w/ type j

Remark: Take $(Y_k)_{k=1}^{\infty}$ of i.i.d. r.v. s.t. $P(Y_1 = j) = P_j$

$$N_t^j = \begin{cases} 0 & \text{if } N_t = 0 \\ \sum_{k=1}^{N_t} 1_{\{Y_k = j\}} & \text{if } N_t > 0 \end{cases}$$

Last:

$$E[N_t^j] = E[1_{\{Y_1 = j\}}]\, E[N_t]$$
$$= P\{Y_1 = j\}\, E[N_t]$$
$$= (P_j \lambda) t$$

- Higher Moment (Variance)

$$E[(N_t^j)^2] = E[(\sum_{k=1}^{N_t} 1_{\{Y_k = j\}})^2]$$
$$= \sum_{n=1}^{\infty} E[(\sum_{k=1}^{n} 1_{\{Y_k = j\}})^2 \mid N_t = n]\, P(N_t = n)$$
$$= \sum_{n=1}^{\infty} E[(\sum_{k=1}^{n} 1_{\{Y_k = j\}})^2]\, P(N_t = n)$$
$$= \sum_{n=1}^{\infty} (nP_j + (n^2 - n)P_j^2)\, P(N_t = n)$$
$$= P_j\, E[N_t] + P_j^2\, E[n^2 - n]$$
$$= (P_j \lambda) t + P_j^2 (\underbrace{Var(N_t) - E[N_t]}_{=0}) + \underbrace{P_j^2 (E[N_t])^2}_{= E[N_t^j]^2}$$

$$E[(\sum_{k=1}^{n} 1_{\{Y_k = j\}})^2]$$
$$= E[\sum_{k=1}^{n} 1_{\{Y_k = j\}}] + E[\sum_{\substack{k_1 \neq k_2 \\ 1 \le k_i \le n}} 1_{\{Y_{k_1} = j\}} 1_{\{Y_{k_2} = j\}}]$$
$$= n E 1_{\{Y_1 = j\}} + (n^2 - n) E 1_{\{Y_1 = j\}} 1_{\{Y_2 = j\}}$$
$$= nP_j + (n^2 - n) P_j^2$$

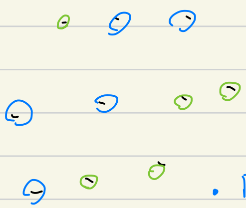
$$= P_j \lambda + E[N_t^j]^2$$

$$\Rightarrow Var(N_t^j) = (P_j \lambda) t = E[N_t^j]$$

Thm: $(N_t^j)_{t \ge 0}$ is a Poisson process w/ parameter $P_j \lambda$

- independent among types

Key Idea:



- P_b = Prob(blue)
- P_g = Prob(green)

$P_b + P_j = 1$

Y = # balls $\sim Poi(\lambda)$

Y_g, Y_b = # of green, blue balls respectively

$Y_g + Y_b = Y$

Claim: $Y_g \sim Pois(P_g\lambda)$, $Y_b \sim Pois(P_b\lambda)$
Y_g, Y_b are independent

$$P(Y_g=m, Y_b=n) = P(Y_g=n, Y_b=n, Y=m+n)$$
$$= P(Y=m+n)\cdot P(Y_j=m \mid Y=m+n)$$
$$= e^{-\lambda}\frac{\lambda^{n+m}}{(m+n)!}\cdot\binom{m+n}{n} P_j^m P_b^n$$
$$= \left(e^{-\lambda P_g}\frac{(\lambda P_g)^m}{m!} P_g^n\right)\left(e^{-\lambda P_b}\frac{\lambda^n}{n!}P_b^n\right)$$
$$= Pois(P_g\lambda)\, Pois(P_b\lambda)$$

11/13

last "thinning" a poisson process

Super-position:
$(N_t^i)_{t\geq 0}$ Poisson process
for $j=1,\cdots,l$

ex. (starbucks) $l=2$
N_t^1 counts customer at statestreet starbucks] assume independent
N_t^2 - — - - - - university ave - - - -]

Q. Is the total # of starbuck customers in madison is Poisson Process?
S. Yes.

Basic Fact: $z_1, \dots, z_l$ independent r.v. $z_i \sim Pois(\lambda_i)$
then $z_1+\cdots+z_l \sim Pois(\lambda_1+\cdots+\lambda_l)$

$m>0$, $$P(z_1+\cdots z_l=m) = \sum_{m_1+\cdots m_l=m} P(z_1=m_1)\cdots P(z_l=m_l)$$
$$= \sum_{m_1+\cdots m_l=m}\prod_{i=1}^{l} e^{-\lambda_i}\frac{\lambda_i^{m_i}}{m_i!}$$
$$= \frac{e^{-\sum_{i=1}^{l}\lambda_i}}{m!}\underbrace{\sum_{m_1+\cdots+m_l}\frac{\lambda_1^{m_1}\cdots\lambda_l^{m_l}}{m_1!\cdots m_l!}\cdot m!}_{\text{multi-normial formula}}$$
$$= \frac{(\lambda_1+\cdots+\lambda_l)^m}{m!} e^{-\sum_{i=1}^{l}\lambda_i}$$

Thm: given independent Poisson Processes $(N_t^j)_{t\geq 0}$ w/ rate λ_j for $j=1 \cdots \ell$,
it holds $(N_t)_{t\geq 0}$ is a Poisson Process w/ rate $\lambda_1 + \cdots + \lambda_\ell$
where $N_t = \sum_{j=1}^{\ell} N_t^j$

Proof sketch: $N_t^j = \sup\{n \geq 0 : \sum_{k=1}^{n} X_k^j \leq t\}$ $\quad j=1 \cdots \ell$
$(X_k^j)_{k=1}^{\infty}$ i.i.d. $Exp(\lambda_j)$ for j fixed independent for $j=1 \cdots \ell$

"first arrival": $\min(X_1^1, X_1^2) \sim Exp(\lambda_1+\lambda_2)$
"Exponential races": $\min(X_2^1, X_1^2)$ or $\min(X_1^1, X_2^2)$

ex. Customers arrive at the state street Starbuck at rate 50/hour
and at the university ave at rate 40/hour.
Assume both open at 3 am.

a) Prob. that from 8:00 ~ 8:30 Am, 60 customers go to starbucks in Madison?
N_t^1 = state street
N_t^2 = University Ave
$N_t = N_t^1 + N_t^2$ = Madison
$(N_t)_{t\geq 0}$ is a Poisson Process with rate 90/hour
$t=0$ corresponds to 3 am.
$N_{3.5} - N_3 \sim Pois(0.5 \times 90) = Pois(45)$
$P(N_{3.5} - N_3 = 60) = e^{-45} \cdot \frac{45^{60}}{60!}$

b) Prob: 40 customers at state street and 30 customers at university ave from 9~11 am.
$P(N_6^1 - N_4^1 = 40, N_6^2 - N_4^2 = 30) = P(N_6^1 - N_4^1 = 40)\, P(N_6^2 - N_4^2 = 30)$
$N_6^1 - N_4^1 \sim Pois(100)$ $\qquad = e^{-100}\frac{100^{40}}{40!} \cdot e^{-80}\frac{80^{30}}{30!}$
$N_6^2 - N_4^2 \sim Pois(80)$

Conditioning on the # of arrivals. (mid2)

ex. (Starbucks) Assume 50/hour and store opens at 5 am.
Assume that 500 customers arrive from 5~10 am.

Q. Dist. of customers from 6 ~ 7.

$$P(N_2 - N_1 = k \mid N_5 = 500)$$

$$= \frac{P(N_2 - N_1 = k,\ N_5 = 500)}{P(N_5 = 500)}$$

$$= \frac{P(N_2 - N_1 = k,\ N_1 + N_5 - N_2 = 500 - k)}{P(N_5 = 500)}$$

$$= \frac{P(N_2 - N_1 = k)\ P(N_1 + (N_5 - N_2) = 500 - k)}{P(N_5 = 500)}$$

$$= \frac{\cancel{e^{-50}} \frac{50^k}{k!}\ \cancel{e^{-200}} \frac{200^{500-k}}{(500-k)!}}{\cancel{e^{-250}} \frac{250^{500}}{500!}}$$

$$= \frac{500!}{(500-k)!\,k!} \left(\frac{50}{250}\right)^k \left(\frac{200}{250}\right)^{500-k}$$

$$= \binom{500}{k} \left(\frac{1}{5}\right)^k \left(\frac{4}{5}\right)^{500-k}$$

$\sim \mathrm{Bin}(500, \frac{1}{5})$

independent

$N_1 \perp N_2 - N_1$

$N_5 - N_2 \perp N_2 - N_1$

$N_1 + N_5 - N_2 \sim \mathrm{Pois}(200)$

$N_2 - N_1 \sim \mathrm{Pois}(50)$

$N_5 \sim \mathrm{Pois}(250)$

$k = 0, \cdots, 500.$

C6 Lec 1

Continuous Time Markov Chain

Notation: $(X_n)_{x=0}^{\infty}$ M.C. discrete time

$(X_t)_{t\geq 0}$ M.C. continuous time

e.g. 2 states M.C. $\quad 1 \rightleftarrows 2 \quad$ $\{1,2\}$.

allow: jump at any possible time. ~~to~~ stay in state 1 for a random amount of time $\to$ go to state 2.

Q. how much time should we wait to switch from $1\to 2$ or $2\to 1$.

A. One prob. dist $\to$ consistent with Markov Property.

$T = \inf\{t\geq 0 : X_t = 2\}$ dist. of T

$X_0 = 1$

Claim: T should be exponentially distributed if $(X_t)_{t\geq 0}$ is Markovian.

$s,t>0$

$$P(T>t+s) = P(T>s,\ T\geq t+s) = P(T\geq t+s \mid T>s)\,P(T>s)$$

$$= P(T>t)\,P(T>s)$$

↑ by markov property, restart at s.

↑ memoryless property of exp

e.g. $S=\{0,1,2,\cdots\}$ $\quad 0\to 1\to 2\to 3$ - always move to the right. in discrete

Continuous: wait exp dist. of time, go right

Poisson process $(N_t)_{t\geq 0}$ ↙ (e.g. of cons. M.C.)

Before: discrete $(X_n)_{n=0}^{\infty}$

$$P(X_{n+1}=y \mid X_n=x) = P(x,y) \qquad P(X_{n+m}=y \mid X_n=x) = P^{(m)}(x,y) = P^m(x,y)$$

Def: Continuous: (transition in m unit of time)

Transition Probability function

$P_t = \{P_t(x,y)\}_{t\geq 0}$ $\quad x,y\in S$

$P_t(x,y)$ = prob. of transitioning from x to y in t units of time

$\sum_{y\in S} P_t(x,y) = 1 \qquad P_0(x,y) = \begin{cases} 1 & x=y \\ 0 & x\neq y. \end{cases}$

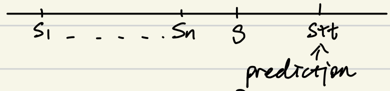
Continuous time M.C.

$(X_t)_{t \geq 0}$, state space S

Two ingredients: 1. initial prob. dist μ $\quad P(X_0 = x) = \mu(x)$

2. transition prob. func P_t.

Markov Property:



$$P(X_{t+s} = y \mid X_{s_1} = x_1, \cdots, X_{s_n} = x_n, X_s = x) = P_t(x,y)$$

Remark: $P(X_{t+s} = y \mid X_s = x) = P_t(x,y)$.

e.g. Poisson Process

$(X_t)_{t \geq 0}$ Pois Pro w/ rate λ

$S = \{0, 1, 2, \ldots\}$

$P_t(x,y) = 0$ if $x > y$

$$P(X_{t+s} = y \mid X_s = x) = \frac{P(X_{t+s} = y, X_s = x)}{P(X_s = x)} = \frac{P(X_s = x, X_{t+s} - X_s = y - x)}{P(X_s = x)}$$

$$= P(X_{t+s} - X_s = y - x) = e^{-\lambda t} \frac{(\lambda t)^{y-x}}{(y-x)!}$$

$X_{t+s} - X_s \sim \text{Pois}(\lambda t)$

$$P_t(x,y) = \begin{cases} 0 & x > y \\ e^{-\lambda t} \frac{(\lambda t)^{y-x}}{(y-x)!} & x \leq y \end{cases}$$

Check Markov Property. $\quad x_1 \leq x_2 \leq \cdots \leq x_n \leq x \leq y$.

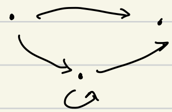
$$P(X_{t+s} = y \mid X_{s_1} = x_1, \cdots, X_{s_n} = x_n, X_s = x)$$

$$= \frac{P(X_{s_1} = x_1, \cdots, X_{s_n} = x_n, X_s = x, X_{t+s} = y)}{P(X_{s_1} = x_1, \cdots, X_{s_n} = x_n, X_s = x)}$$

$$= \frac{P(X_{s_1} = x_1, X_{s_2} - X_{s_1} = x_2 - x_1, \cdots\cdots, X_s - X_{s_n} = x - x_n, X_{t+s} - X_s = y - x)}{P(X_{s_1} = x_1, X_{s_2} - X_{s_1} = x_2 - x_1, \cdots\cdots, X_s - X_{s_n} = x - x_n)}$$

$$= P(X_{t+s} - X_s = y - x) = P_t(x,y)$$

Subordinated M.C.



$\{Y_n\}_{n=0}^{\infty}$ M.C.   Proof for M. P. in Lec Note.

w/ transition matrix $U = \{u(x,y)\}_{x,y\in S}$

switch at random exp. time

Poisson Process $(N_t)_{t\ge 0}$ ind. from the M.C. (w/ rate λ)

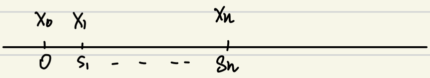
Def. $(X_t)_{t\ge0}$ continuous M.C. $X_t = Y_{N_t}$.

Transition Prob. func.

$$P_t(x,y) = P(X_{t+s}=y \mid X_s = x) = P(Y_{N_{t+s}} = y \mid Y_{N_s}=x) \qquad N_{t+s} \ge N_s$$
$$= \sum_{k=0}^{\infty} P(N_{t+s}-N_s = k,\ Y_{N_{t+s}} = y \mid Y_{N_s}=x)$$
$$= \sum_{k=0}^{\infty} P(N_{t+s}-N_s=k)\, P(Y_k = y \mid Y_0 = x)$$
$$= \sum_{k=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^k}{k!}\, u^{(k)}(x,y)$$

C6
Lec 2

Properties:



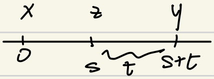
1. $P_\mu(X_0 = x_0,\ X_{s_1} = x_1, \dots, X_{s_n} = x_n)$

$$= P_\mu(X_{s_n} = x_n \mid X_0 = x_0, \dots, X_{s_{n-1}} = x_{n-1})\ P_\mu(X_0 = x_0, \dots, X_{s_{n-1}} = x_{n-1})$$
$$= P_{s_n - s_{n-1}}(x_{n-1}, x_n)\ P_\mu(X_0 = x_0, \dots, X_{s_n} = x_n)$$

- — - -

$$= P_{s_n-s_{n-1}}(x_{n-1},x_n)\cdots P_{s_2-s_1}(x_1,x_2)\ P_\mu(X_{s_1}=x_1 \mid X_0 = x_0)\ P_\mu(X_0 = x_0)$$
$$= \mu(x_0)\prod_{j=1} P_{s_j - s_{j-1}}(x_{j-1}, x_j)$$

2. Chapman-Kolmogorov equation

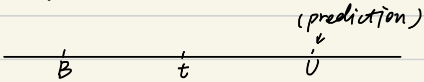


$$P_{s+t}(x,y) = \sum_{z\in S} P_s(x,z)\,P_t(z,y)$$

$\Uparrow$

$$P_\mu(X_{s+t} = y \mid X_0 = x) = \frac{P(X_0=x,\ X_{s+t}=y)}{P(X_0=x)} = \sum_{z\in S}\frac{P(X_0=x,\ X_s = z,\ X_{s+t}=y)}{P(X_0=x)}$$
$$= \sum_{z\in S}\frac{P(X_0=x)\,P_s(x,z)\,P_t(z,y)}{P(X_0=x)} = \sum_{z\in S}P_s(x,z)\,P_t(z,y)$$

Strong M.P. in to the infinite future



B depends on $\{X_s : 0 \le s \le t\}$

$P_\mu((X_{t+s})_{s\ge0} \in U \mid X_t = x, B)$ (trajectory after t)

$= P_x((X_s)_{s\ge0} \in U)$

U: collection of trajectories

e.g. (last time)

$X_0 = 1, \quad T = \inf\{t \ge 0 : X_t = 2\}$

$P_1(T > t+s \mid T > t) = P(T > s)$ $\quad B = \{X_r = 1 \text{ for } 0 \le r \le t\}$

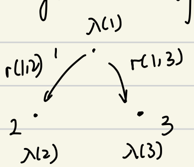
$= P_1(T > t+s \mid X_t = 1, B)$ $\quad$ trajectory after t: $U = \{X_r = 1, 0 \le r \le s\}$

$= P_1((X_{r+t})_{r\ge0} \in U \mid X_t = 1, B)$

$= P_1((X_r)_{r\ge0} \in U) = P(T > s)$

SLMC T is the stopping time, $P_\mu((X_{T+s})_{s\ge0} \in U \mid X_T = x, B) = P_x((X_s)_{s\ge0} \in U)$

Ingredients of a continuous time M.C.



(a) choose an initial state according to prob. dist. μ.

(b) $J_0 = 0$, J_1 = time of the first jump.

... J_k = time of the k^{th} jump

"holding time" = "inter-arrival" time = waiting time

$J_{k+1} - J_k \sim \exp(?)$

(c) Label the states visited as $Z_0, Z_1, \cdots, Z_k$

Z_0: initial state $\quad Z_k$: state after jump J_k

Relation: $(X_t)_{t\geq 0}$, $(J_k)_{k\geq 0}$, $(Z_k)_{k\geq 0}$

$X_t = Z_0$ for $0 \leq t < J_1$

$X_t = Z_1$ for $J_1 \leq t < J_2$

$\vdots$

$X_k = Z_k$ for $J_k \leq t < J_{k+1}$.

$(Z_k)_{k\geq 0}$ is a discrete time M.C.

$Z_k = X_{J_k}$.

w/ transition matrix $\{r(x,y)\}_{x,y}$ called the "routing matrix"

Holding time parameters $(\lambda(x))_{x\in S}$ $\lambda(x) \in [0,\infty)$ (holding rates)

(reach state x, stay $\lambda(x)$ time)

Connection: if $X \sim \exp(0)$, then $P(X=\infty) = 1$

Want: Given Z_k, $J_{k+1} - J_k \sim \operatorname{Exp}(\lambda(Z_k))$

Fact: If $X \sim \exp(1)$ and $\lambda > 0$, then $\frac{1}{\lambda} X \sim \exp(\lambda)$

$P(X>t) = e^{-1\cdot t}$, $P(\frac{1}{\lambda} X > t) = P(X > \lambda t) = e^{-\lambda t}$

Consider a sequence $(\tau_k)_{k=0}^{\infty}$ of i.i.d. $\exp(1)$ r.v.

then $J_1 = \frac{1}{\lambda(Z_0)} \tau_0$

$J_2 = \frac{1}{\lambda(Z_0)} \tau_0 + \frac{1}{\lambda(Z_1)} + \tau_1$

$\vdots$

$J_k = \sum_{j=0}^{k-1} \frac{1}{\lambda(Z_j)} \tau_j$

C6 Lec3

Last. LTMC, 2 ways to describe

① $P_t = (P_t)_{t \geq 0}$

② routing matrix $\{r(x,y)\}_{x,y \in S}$

holding rates $\{\lambda(x)\}_{x \in S}$

Connect ① & ②

jump rates: for $x \neq y$ $q(x,y) =$ "rate of jumping from x to y"

$= \lambda(x)\, r(x,y)$

Perspective 1. (algebraic) given q, we can define r and λ as follows:

$\lambda(x) =$ "rate of leaving state x"

$$= \begin{cases} \sum_{y \neq x} q(x,y) & \text{if } x \text{ is not absorbing} \\ 0 & \text{if } x \text{ is absorbing} \end{cases}$$

for $x \neq y$, $r(x,y) = \begin{cases} \dfrac{q(x,y)}{\sum_{y \neq x} q(x,y)} & \text{if } x \text{ is non-absorbing} \\ 0 & \text{if } x \text{ is absorbing} \end{cases}$

Perspective 2 (Analytic) rate = "derivative"

$x \neq y$ $q(x,y) = \lim_{h \to 0} P_h(x,y)$

Remark: $P_0(x,y) = 0$ for $x \neq y$.

$$q(x,y) = \lim_{h \to 0} \frac{P_h(x,y) - P_0(x,y)}{h} = \frac{d}{dt} P_t(x,y)\Big|_{t=0}$$

Informal Argument: $x \neq y$

$P_h(x,y) = P(X_h = y \mid X_0 = x)$

For $h \approx 0$ $\quad J_1$ = time of first leaving state x $\quad \sim Exp(\lambda(x))$

$$P_h(x,y) \approx P(J_1 \leq h) \cdot r(x,y)$$
$$= (1-e^{-\lambda(x)h})\, r(x,y)$$
$$\approx h\lambda(x)\, r(x,y)$$
$$= h\, q(x,y)$$

$e^{-\lambda(x)h} = 1 - \lambda(x)h + \cdots + \cdots$ high powers

Generator Matrix

$Q = \{q(x,y)\}_{x,y}$ where $q(x,x) = -\lambda(x)$

Note: $\sum_y q(x,y) = q(x,x) + \sum_{y\neq x} q(x,y) = -\lambda(x) + \sum_{y\neq x} \lambda(x)\, r(x,y)$
$$= \lambda(x)\left(-1 + \sum_{y\neq x} r(x,y)\right)$$
$$= 0$$

Kolmogorov Equations:

Forward: $\partial_t P_t = P_t Q$ $\quad$ (matrix equality)

Backward: $\partial_t P_t = Q P_t$

$\Uparrow$

$$P_{t+s}(x,y) = \sum_z P_t(x,z)\, P_s(z,y)$$
$$\partial_t P_t(x,y) = \lim_{h\to 0} \frac{P_{t+h}(x,y) - P_t(x,y)}{h}$$

Consider $h \approx 0$, let $s = h$ $\quad P_{t+h}(x,y) = \sum_z P_t(x,z) P_h(z,y)$
$$= P_t(x,y) P_h(y,y) + \sum_{z\neq y} P_t(x,z) P_h(z,y)$$

$\Downarrow$

$P_t(x,y)\left(1 - \sum_{y\neq z} P_h(y,z)\right)$

$$\frac{P_{t+h}(x,y) - P_t(x,y)}{h} = -P_t(x,y) \sum_{y\neq z} \frac{P_h(y,z)}{h} + \sum_{z\neq y} P_t(x,z) \frac{P_h(y,z)}{h}$$

$$\partial_t P_t(x,y) = -P_t(x,y) \underbrace{\sum_{y\neq z} q(y,z)}_{} + \sum_{y\neq z} P_t(x,z) q(y,z) = \sum_z P_t(x,z)\, q(z,y) = P_t Q$$

$P_t(x,y)\lambda(y) + \cdots$

Backward, take $t=h$, $s=t-h$.

$$P_{t+h}(x,y) = \sum P_h(x,z)P_t(z,y)$$

e.g. (Subordinated M.C.)

M.C. $(Y_n)_{n=0}^{\infty}$ w/ transition matrix U

Poisson Process $(N_t)_{t\geq 0}$ w/ rate λ

Def $X_t = Y_{N_t}$

Generator Matrix:

$$x\neq y \quad r(x,y) = P(x\to y \mid \text{leave } x) = \frac{P(x\to y)}{P(\text{leave } x)} = \frac{u(x,y)}{1-u(x,x)}$$

holding rate: $\lambda(x) = \lambda(1-u(x,x))$

comes from "thinning"

For $x\neq y$, $q(x,y) = \lambda(x)\, r(x,y) = \lambda u(x,y)$

$$q(x,x) = -\lambda \sum_{y\neq x} u(x,y) = \lambda(1-u(x,x))$$

Kolmogorov equations:

$$P_t(x,y) = \sum_{n=0}^{\infty} P(N_t=n)\, u^{(n)}(x,y) = \sum_{n=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!} u^{(n)}(x,y)$$

$$= e^{-\lambda t}\left(1+(\lambda t)u^{(1)}(x,y) + \cdots + \frac{(\lambda t)^n}{n!}u^{(n)}(x,y) \cdots\right)$$

$$\frac{d}{dt}P_t(x,y) = -\lambda e^{-\lambda t}(\cdots) + e^{-\lambda t}\left(\lambda u^{(1)}(x,y) + \cdots + \lambda^n\frac{t^{n-1}}{(n-1)!}u^{(n)}(x,y) + \cdots\right)$$

$$= -\lambda P_t(x,y) + e^{-\lambda t}\sum_{n=1}^{\infty}(\lambda t)^{n-1}u^{(n)}(x,y)$$

$$= -\lambda P_t(x,y) + \sum_z u(z,y)\sum_{n=1}^{\infty} e^{-\lambda t}(\lambda t)^{n-1}u^{(n-1)}(x,z)$$

$$= -\lambda P_t(x,y) + \sum_z u(z,y)\, \underset{\substack{\Downarrow \\ P_t(x,y)}}{P_t(x,z)} \qquad \underset{z=y}{\Downarrow}$$

$$u^{(n)}(x,y) = \sum_z u^{(n-1)}(x,z)\,u(z,y)$$

$$= \sum_{z\neq y} P_t(x,z)q(z,y) + P_t(x,y)\underbrace{\lambda(1-u(y,y))}_{q(y,y)}$$

$$\Downarrow$$

$$\frac{d}{dt}P_t = P_t Q.$$

C6 Lec4

Long time CTMC (stationary dist.)

μ : initial prob. dist.

$(P_t)_{t\ge 0}$ transition prob. func.

$$P_\mu(X_t=y) = \sum_{x\in S} P_\mu(X_0=x, X_t=y) = \sum_{x\in S} P_\mu(X_0=x) P_\mu(X_t=y \mid X_0=x)$$

$$= \sum_{x\in S} \mu(x) P_t(x,y) = (\mu P_t)(y)$$

Given $X_0 \sim \mu$, $X_t \sim \mu P_t$

stationary dist.

A probability dist. π is a stationary dist. if $\pi P_t = \pi \quad \forall t\ge 0$

Lemma: π is a stationary dist. iff $\pi Q = 0$. where Q is the generator matrix

Proof: BKE $\frac{d}{dt} P_t = Q P_t$

left multiply both sides w/ a probability dist. π

$\pi \frac{d}{dt} P_t = \pi Q P_t$

$\frac{d}{dt}(\pi P_t) = \pi Q P_t \quad (*)$

(a) if $\pi Q = 0$, then $(*)$ implies $\frac{d}{dt}(\pi P_t) = 0$ $\quad \pi P_t = \pi P_0$

Hence, π is a stationary dist.

$P_0(x,y) = \begin{cases} 1 & \text{if } x=y \\ 0 & x\ne y \end{cases}$

$= I$ (identity matrix)

(b) Assume π is a stationary dist.

then $\pi P_t = \pi$, so $\frac{d}{dt}(\pi P_t) = 0$

So $(*)$ implies $\pi Q P_t = 0$

sent $t\to 0$ $\quad \pi Q P_0 = 0 \rightarrow \pi Q = 0$

Long time behavior

Irreducibility: A CTMC is irreducible if for any $x\ne y$,

there is a path $z_0 = x, \dots, z_n = y$ s.t.

$q(z_k, z_{k+1}) > 0$.

Thm: let $(X_t)_{t\ge 0}$ an irreducible continuous time M.C. w/ stationary dist. π.

For any initial μ,

assume absolute converge.

① $\lim_{t\to\infty} P_\mu(X_t = y) = \pi(y)$

② Let $f: S \to \mathbb{R}$, then $\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X_s)\,ds = \sum_x f(x)\pi(x)$

e.g. (subordinate M.C.)

$\{Y_n\}_{n=0}^{\infty}$ w/ transition matrix $U = \{u(x,y)\}_{x,y\in S}$

irreducible, finite state space

$(N_t)_{t\geq 0} \sim$ Pois w/ rat λ

$X_t = Y_{N_t}$

stationary dist.

$$Q: q(x,y) = \begin{cases} \lambda u(x,y) & x\neq y \\ \lambda(u(x,x)-1) & x=y \end{cases}$$

$\pi Q = 0$

$\hookrightarrow \sum_x \pi(x) q(x,y) = 0$

$\sum_{x\neq y} \pi(x)\lambda u(x,y) + \pi(y)\lambda(u(y,y)-1) = 0$

$\lambda \sum_x \pi(x) u(x,y) = \lambda \pi(y)$

$\pi U = \pi$ (same stat. dist. for discrete & cons)

Remark: $(X_t)_{t\geq 0}$ is irreducible since $\{u(y,y)\}_{x,y\in S}$ is irreducible

e.g Poisson Process.

No stationary dist.

① $\lim_{t\to\infty} P_t = 0$

② Any π s.t. $\pi P_t = \pi$ $(\forall t)$ is identically 0.

② follows from ① $\pi P_t = \pi$ $\forall t\geq 0$

$\Rightarrow t\to\infty$ $\pi P_t = 0$, so $\pi = 0$

① $P_t(x,y) = e^{-\lambda t}\frac{(\lambda t)^{y-x}}{(y-x)!} = c\,e^{-\lambda t} t^{y-x}$ $\quad t\to\infty$ $P_t(x,y) \to 0$

Also not irreducible

$0 \to 1 \to 2 \to 3 \cdots$

e.g. $S = \{1, 2\}$ $\quad 1 \underset{\mu}{\overset{\lambda}{\rightleftarrows}} 2 \qquad Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$

$\lambda, \mu > 0$

$(X_t)_{t \geq 0}$ is irreducible

BKE $P_t' = Q P_t$ $\quad \begin{pmatrix} P_t'(1,1) & P_t'(1,2) \\ P_t'(2,1) & P_t'(2,2) \end{pmatrix} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \begin{pmatrix} P_t(1,1) & P_t(1,2) \\ P_t(2,1) & P_t(2,2) \end{pmatrix}$

$$P_t'(1,1) = -\lambda P_t(1,1) + \lambda P_t(2,1)$$

$$P_t'(2,1) = \mu P_t(1,1) - \mu P_t(2,1)$$

$$P_t(1,1) = \frac{\mu}{\mu+\lambda} + \frac{\lambda}{\mu+\lambda} e^{-(\mu+\lambda)t}$$

$$P_t(2,1) = \frac{\mu}{\mu+\lambda} - \frac{\lambda}{\mu+\lambda} e^{-(\mu+\lambda)t}$$

$$\lim_{t\to\infty} P_t(1,1) = \lim_{t\to\infty} P_t(2,1) = \frac{\mu}{\mu+\lambda}$$

$$\pi = \left(\frac{\mu}{\mu+\lambda}, \frac{\lambda}{\mu+\lambda} \right)$$

want $\pi Q = 0$.

$$\left(\frac{\mu}{\mu+\lambda}, \frac{\lambda}{\mu+\lambda} \right) \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} = \begin{pmatrix} \frac{-\lambda\mu}{\mu+\lambda} + \frac{\lambda\mu}{\lambda+\mu} \\ \frac{\mu\lambda}{\mu+\lambda} - \frac{\mu\lambda}{\mu+\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$