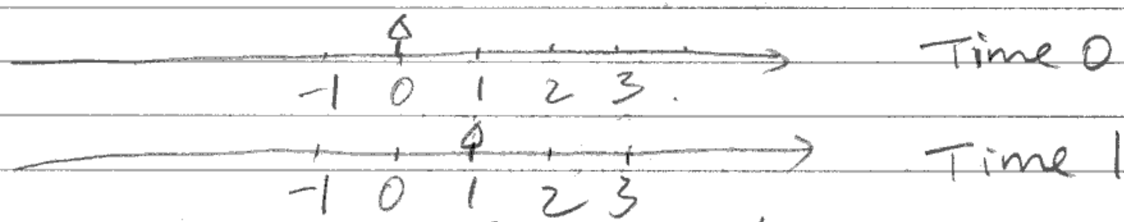
5652

9/8

Stochastic Processes

Randomness in Time

Random Walk



i.i.d sequence of Bernoulli random variable

$\xi_1, \xi_2, \cdots \xi_n, \cdots$ independent

$P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$

The position of the walker at time n:

$$S_n = \xi_1 + \cdots + \xi_n$$

We are thinking of S_n.

$$P\left((\underbrace{+1, -1, \cdots, +1}_{n})\right) = \frac{1}{2^n}$$

Suppose we have l steps that gives us $+1$

then we have $n-l$ steps that gives us -1.

$$l - (n-l) = 2l - n = k$$

$$l = \frac{n+k}{2}$$

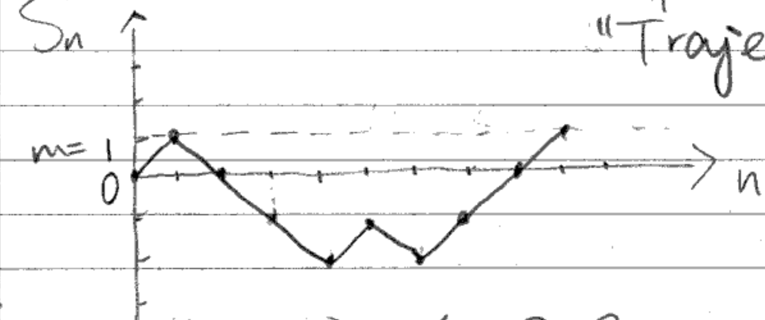
If $n+k$ is odd, $P_0(S_n = k) = 0$

If $n+k$ is even, $P_0(S_n = k) = \binom{n}{\frac{n+k}{2}} \frac{1}{2^n}$

binomial

discrete time
continuous function
"Trajectory" = "Sample path".
random walk

future depends on now not past



$(\xi_0, \xi_1, \cdots)$ $\quad \xi_k = S_k - S_{k-1}$

$(S_0, S_1, S_2 \cdots) = (S_n)_{n \geq 0}$

sequence of random variable describes the positions of walker

Def: A stochastic (random) process $(X_t)_{t \geq 0}$ is a sequence of r.v.'s X_t parametrized by time.

Sample path of a random walk

$$\omega = (i_0, i_1, i_2 \cdots, i_n, *, *, \cdots)$$

$$|i_k - i_{k-1}| = 1$$

For each fixed ω, $P(\{\omega_n\}) = \frac{1}{2^n}$

$\tau_1 = \inf\{n \geq 0, S_n = 1\}$ first hitting time

$\{\tau_1 = n\} = \{S_n = 1, S_0 \neq 1, S_1 \neq 1, \cdots, S_{n-1} \neq 1\}$

events are reasonable Trajectories

$$(\underbrace{i_0, i_1, \cdots,}_{\neq 1} \underset{\parallel}{i_n})$$

P_0: probability, starts from 0

Lem: Let $\tau_m = \inf(n \geq 0, S_n = m)$

If $m > 0$ and $n+m$ is even,

Then $P_0(\tau_m = n) = \frac{m}{n}\binom{n}{\frac{m+n}{2}} 2^{-n}$

Pf. If $n+m$ odd. then $P_0(\tau_m = n) \leq P(S_n = m) = 0$

If $n+m$ even, $P_0(S_n = m) = P_0(\tau_m = n) + P_0(S_n = m, \tau_m < n)$

hit m but not the first time

$P_0(S_n = m) = \binom{n}{\frac{n+m}{2}} \frac{1}{2^n}$

$P(S_n = m, \tau_m < n) = P(S_n = m, \tau_m < n, S_{n-1} = m-1)$

$+ P(S_n = m, \tau_m < n, S_{n-1} = m+1)$

(two possibilities in $n-1$ step)

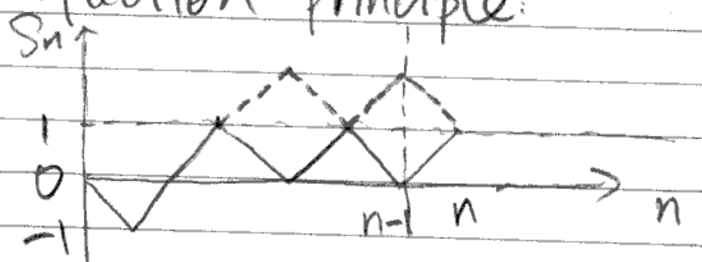
$P_0(S_n = m, \tau_m < n, S_{n-1} = m+1) = P_0(S_n = m, S_{n-1} = m+1)$

$= P_0(S_{n-1} = m+1, \xi_n = -1) = \frac{1}{2} P_0(S_{n-1} = m+1)$

$= \binom{n-1}{\frac{n+m}{2}} \frac{1}{2^n}$ by reflection principle

$= P_0(S_n = m, \tau_m < n, S_{n-1} = m-1)$

reflection principle:



$|-|$

reflect from the first hitting point

Therefore, $P_0(\tau_m = n) = P_0(S_n = m) - P_0(S_n = m, \tau_m < n)$

$$= P_0(S_n = m) - 2P_0(S_n = m, S_{n-1} = m+1)$$

$$= P_0(S_n = m) - 2 \cdot \frac{1}{2} P_0(S_{n-1} = m+1)$$

$P_0(S_{n-1} = m+1, S_n = m)$

$= P_0(S_{n-1} = m+1, \xi = -1)$

$$= P_0(S_n = m) - P_0(S_{n-1} = m+1)$$

$$= \binom{n}{\frac{n+m}{2}} 2^{-n} - \binom{n-1}{\frac{n+m}{2}} 2^{-(n-1)}$$

$$= \frac{m}{n} \binom{n}{\frac{n+m}{2}} 2^{-n}$$

$$P_0(\tau_1 = 2k-1) = \frac{1}{2k-1} \binom{2k-1}{k} 2^{-(2k-1)}, \quad k = 1, 2 \cdots$$

$\to$ discrete distribution

$$\begin{cases} P_0(\tau_1 < \infty) = 1 & \text{(gambling game)} \\ E(\tau_1) = \infty \end{cases}$$

random variable is finite.

Sooner or later, the gambler wins 1 dollar, but the expected time is infinite.

9/10 Random Walk

a family of random variables parametrized by Time

"Sample path" (Trajectory) $(X_t)_{t \in T}$

pick up a trajectory randomly (random process)

$S_n = \xi_1 + \cdots + \xi_n$

$S_{n-1} = \xi_1 + \quad + \xi_{n-1}$

$S_{n+1} = \xi_1 + \cdots + \xi_n + \xi_{n+1} = S_n + \xi_{n+1}$

event $\{\tau_m = n\}$ = step n, hits certain level m for first time

$m \geq 1$, $\tau_m = \inf\{n \geq 1, S_n = m\}$ Markov Time

reflect with respect to the piece after the first hitting time ☆ reflection principle

gambling game

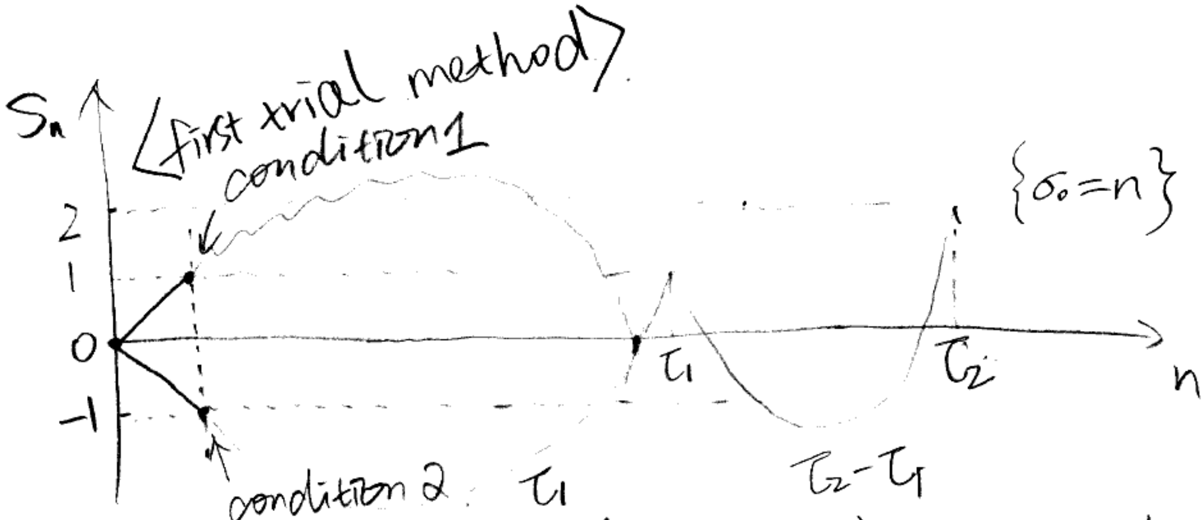
$$\begin{cases} P_0(\tau_1 < \infty) = 1 \\ E_0(\tau_1) = \infty \end{cases} \qquad \begin{cases} P_0(\tau_{-1} < \infty) = 1 \\ E_0(\tau_{-1}) = \infty \end{cases}$$

$\sigma_0 = \inf\{n \geq 1, S_n = 0\}$

☆ Law of total probability formula:

$$P(A) = \sum_{i=1}^{k} P(A|B_i) P(B_i)$$

B_i is a decomposition of Ω.



"cycle"
① $0 \to 1$
② $1 \to 2$
same in distribution

$P_0(\sigma_0 < \infty) = 1$ (Recurrent) one dimensional

transient (3-D) high degree of freedom

8 different directions.

$$E_0(\tau_1) = \sum_{n=1}^{\infty} n P(\tau_1 = n) = \infty$$

Start from 0, then hit 1 or -1, use first trial method,

$$E_0(\tau_1) = E_0(\tau_1 | \xi_1 = 1) P(\xi_1 = 1) + E_0(\tau_1 | \xi_1 = -1) P(\xi_1 = -1)$$

$$= 1 \cdot \frac{1}{2} + [1 + E_0(\tau_2)] \cdot \frac{1}{2} = \frac{1}{2} + \frac{1}{2}(1 + 2E_0\tau_1)$$

random walk from "0" to "2"

In order to reach 2, you must cross 1

Start from 1, then hit 2 (same as from 0 to 1) probabilically

Claim: τ_1 and $\tau_2 - \tau_1$ are independently identically distributed

expected value of first hitting time

Markov property is crucial for cycle method to work

time-space duality

future depends on current. only depends on starting position, not past.

pf of cycle method:

observations: (1) $\tau_1 < \tau_2$

(2) We can view $\tau_2 - \tau_1$ as the first hitting time of a random walk to start from 1 and hit 2 due to Markov property. This is the same thing as saying we start from 0 and hit 1 for the first time (shift invariance)

(3) The piece of RW from 1 to 2 is independently the same in distribution as the piece of RW from 0 to 1.

("The RW does not remember past")

Pf: Fix $k, l \geq 0$,

$$S_{k+r} = \xi_1 + \cdots + \xi_k + \underbrace{\xi_{k+1} + \cdots \xi_{k+r}}_{\tilde{\xi}_1 + \cdots + \tilde{\xi}_r}$$

Define $\tilde{\xi}_r = \xi_{k+r}$,

$\tilde{S}_0 = 0$, $\tilde{S}_r = \tilde{\xi}_1 + \cdots + \tilde{\xi}_r$

We see $S_{k+r} = S_k + \tilde{S}_r$ (duality)

$\{\tilde{S}_n\}_{n\geq 1}$ is i.i.d. as $\{S_n\}_{n\geq 1}$

time behavior

$$P_0(\tau_1 = k, \tau_2 - \tau_1 = l) = P_0(\tau_1 = k, \tau_2 = k + l)$$

space behavior

$$= P_0(S_1 \neq 1, \cdots, S_{k-1} \neq 1, S_k = 1; \underbrace{S_{k+1}}_{S_k + \tilde{S}_1} \neq 2, \cdots, \underbrace{S_{k+l-1}}_{S_k + \tilde{S}_{l-1}} \neq 2, \underbrace{S_{k+l}}_{S_k + \tilde{S}_l} = 2)$$

$$= P_0(S_1 \neq 1, \cdots, S_{k-1} \neq 1, S_k = 1; \tilde{S}_1 \neq 1, \cdots, \tilde{S}_{l-1} \neq 1, \tilde{S}_l = 1)$$

$= P_0(S_1 \neq 1, \cdots, S_{k-1} \neq 1, S_k = 1) P_0(\tilde{S}_1 \neq 1, \cdots, \tilde{S}_{\ell-1} \neq 1, \tilde{S}_\ell = 1)$

$= P_0(\tau_1 = k) P_0(\tau_1 = \ell)$ shift from 2 to 1

In summary, $P_0(\tau_1 = k, \tau_2 - \tau_1 = \ell) = P_0(\tau_1 = k) P_0(\tau_1 = \ell)$

Final step,

$$P_0(\tau_2 - \tau_1 = \ell) = \sum_{k=1}^{\infty} P_0(\tau_1 = k, \tau_2 - \tau_1 = \ell)$$

$\tau_2 - \tau_1$ has same distribution as τ_1

$$= \sum_{k=1}^{\infty} P_0(\tau_1 = k) P_0(\tau_1 = \ell)$$

$$= P_0(\tau_1 < \infty) P_0(\tau_1 = \ell) = P_0(\tau_1 = \ell)$$

so $P_0(\tau_1 = k, \tau_2 - \tau_1 = \ell) = P_0(\tau_1 = k) P_0(\tau_2 - \tau_1 = \ell) \Rightarrow \tau_1 \perp\!\!\!\perp \tau_2 - \tau_1$

τ_1 and $\tau_2 - \tau_1$ are independent

$m \geq 1$, $\tau_m = \inf\{n \geq 1, S_n = m\}$

$\tau_2 = \tau_1^{(1)} + \tau_1^{(2)}$

first trial method

$\tau_m = \tau_1^{(1)} + \cdots + \tau_1^{(m)} \Rightarrow E_0(\tau_m) = m E_0 \tau_1 = \infty$

9/15

$\tau_m = \inf\{n \geq 1, S_n = m\}$ discrete random variables

$\uparrow$

first hitting time

"reflection principle"

"first trial method"

"cycle method" — $\{\tau_1^{(k)}\}_{k=1}^{n}$ i.i.d. sequences

decompose time, spacial behavior of random walk

$\tau_1^{(k)} \overset{d.}{=\!=} \tau_1$

$X \overset{d.}{=\!=} Y$ means that r.v.'s X and Y have the same dist.

Gambler's Ruin Problem

Two gamblers A and B.

Initially, A has \$m

B has \$(M−m)

In each gambling, A/B will have probability 1/2 to win/loss \$1. (fair game)

If either A or B is out of money, the game stops.

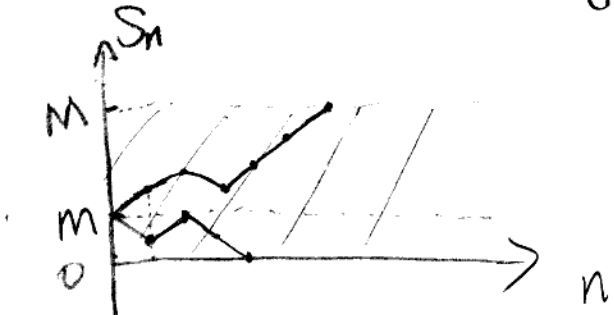
"Ruin" $\longleftrightarrow$ "complete success"

(1) Probability of ruin/complete success?

(2) Expected amount of money that each gambler is going to have at the end of game?

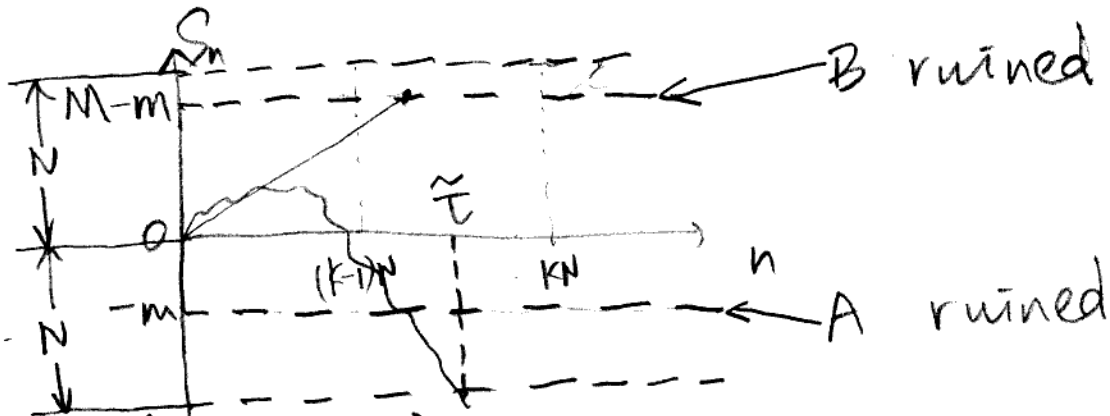
(3) Expected time we need to finish the game ?

model this problem by random walk.



restrict the random walk between 0 and M.

Consider a random walk, $S_n = \xi_1 + \cdots + \xi_n$



$\tau = \min\{\tau_{-m}, \tau_{M-m}\}$ game stops

$= \tau_{-m} \wedge \tau_{M-m}$

$a \wedge b = \min\{a, b\}$

$a \vee b = \max\{a, b\}$

$\{\text{Ruin of } A\} = \{S_\tau = -m\}$

$\{\text{Ruin of } B\} = \{S_\tau = M-m\}$

$P(\text{A ruins in the end}) = P_0(S_\tau = -m)$

Random walk $S_n = m + \xi_1 + \cdots + \xi_n$

$\tau = \tau_M \wedge \tau_0$

$P_m(\text{Ruin for } A) = P_m(S_\tau = 0)$

$$P_m(S_\tau=0)=P_m(S_\tau=0,\ S_1=m+1)+P_m(S_\tau=0,\ S_1=m-1)$$

first trial method $\rightarrow$ $= P_m(S_\tau=0\mid S_1=m+1)P_m(S_1=m+1)+P_m(S_\tau=0\mid S_1=m-1)P_m(S_1=m-1)$

$$=\frac{1}{2}P_m(S_\tau=0\mid S_1=m+1)+\frac{1}{2}P_m(S_\tau=0\mid S_1=m-1)$$

$$P_m(S_\tau=0\mid S_1=m+1)=\sum_{k=1}^{\infty}P_m(S_k=0,\ \tau=k\mid S_1=m+1)$$

$$=\sum_{k=1}^{\infty}P_m(S_1,\ \ ,S_{k-1}\notin\{0,M\},S_k=0\mid S_1=m+1)$$

$$=\sum_{k=1}^{\infty}P_{m+1}(S_1,\cdots,S_{k-2}\notin\{0,M\},\ S_{k-1}=0)$$

$$=P_{m+1}(S_\tau=0)$$

$$P_m(S_\tau=0)=\frac{1}{2}P_{m+1}(S_\tau=0)+\frac{1}{2}P_{m-1}(S_\tau=0)$$

By first trial method, we get this recursive formula.

Let $a_m=P_m(S_\tau=0)$ then $a_m=\frac{1}{2}a_{m+1}+\frac{1}{2}a_{m-1}$

$$P_m(S_\tau=0)-P_{m-1}(S_\tau=0)=P_{m+1}(S_\tau=0)-P_m(S_\tau=0)=\Delta$$

boundary conditions

$$\begin{cases}P_0(S_\tau=0)=1\\ P_M(S_\tau=0)=0\end{cases}=1+M\Delta\ \Rightarrow\ \Delta=-\frac{1}{M}$$

$$P_m(S_\tau=0)=m\cdot\Delta+P_0(S_\tau=0)=m\cdot\Delta+1=1-\frac{m}{M}\ \text{(A ruins)}$$

$$P_m(S_\tau=M)=\frac{m}{M}\quad\text{(A has a complete success)}$$

(2) $E_mS_\tau = M\cdot P_m(S_\tau=M)+0\cdot P_m(S_\tau=0)=m$

(3) $E_m\tau$?

claim: $E_m\tau < \infty$ pf later

$$E_m\tau = \frac{1}{2}(1+E_{m+1}\tau) + \frac{1}{2}(1+E_{m-1}\tau)$$
$$= 1 + \frac{1}{2}E_{m+1}\tau + \frac{1}{2}E_{m-1}\tau$$

boundary conditions

$$E_0\tau = 0 = E_M\tau$$

Answer: $E_m\tau = m(M-m)$ (HW)

$$\le \frac{M^2}{4} \qquad ab \le \left(\frac{a+b}{2}\right)^2$$

Lem: $E_m\tau < \infty$

Pf: by using cycle method

First, we let $N = (M-m) \vee m + 1$

$$\tau \le \tau_N \wedge \tau_{-N} = \tilde{\tau}$$

$$E_m\tau = \sum_{n=0}^{\infty} P_m(\tau > n)$$
$$= \sum_{k=0}^{\infty} \sum_{n=kN}^{(k+1)N-1} P_m(\tau > n) \qquad \text{group size } N$$

claim: $\le \sum_{k=0}^{\infty} N P_m(\tau > kN)$

$0 \cdots N-1$
$N \cdots 2N-1$
$2N \cdots 3N-1$
$\vdots$

$P_m(\tau > kN) \ge P_m(\tau > kN + l)$

$$P_m(\tau \le N) \ge \frac{1}{2^N}$$
$$P_m(\tau > N) \le 1 - \frac{1}{2^N}$$

We then show $P_m(\tau > kN) \le \left(1-\frac{1}{2^N}\right)^k$ ← cycle method

$$P_m(\tau > kN) \le P_m(\tau > (k-1)N) \sup_m P_m(\tilde{\tau} > N)$$

$$\le \left(1-\frac{1}{2^N}\right) P_m(\tau > (k-1)N)$$

Then $E_m\tau \le \sum_{k=0}^{\infty} N\, P_m(\tau > kN) \le \sum_{k=0}^{\infty} N \left(1-\frac{1}{2^N}\right)^k < \infty$

9/17

Markoving process: future does not depend on past.

Wald method (A. Wald)

The first Wald Lemma

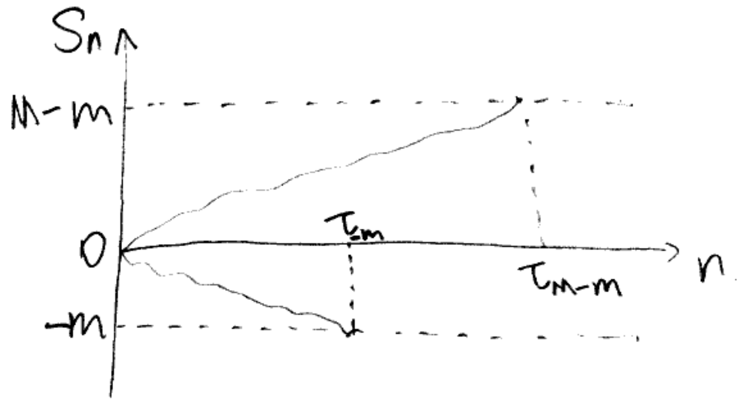
Let $\xi_1, \dots, \xi_n$ be a sequence of i.i.d r.v.'s. $E|\xi_i| < \infty$

Consider a random walk $S_n = \xi_1 + \dots + \xi_n$, $(ES_n = nE\xi)$

Let τ be a first hitting time with $E_0\tau < \infty$ (random walk bdd)

Then $E_0 S_\tau = E_0\tau \cdot E\xi$ ← same distribution with ξ_1

(change fixed-time to a random time, take Expectation)



Apply first Wald Lem,

$$\tau = \tau_{M-m} \wedge \tau_{-m}$$

$$P_0(S_\tau = -m) + P_0(S_\tau = M-m) = 1$$

$$E_0 S_\tau = (-m)P_0(S_\tau = -m) + (M-m)P_0(S_\tau = M-m) = E_0\tau \cdot E\xi = 0$$

$$\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot(-1) = 0$$

$$\Rightarrow \begin{cases} P_0(S_\tau = -m) = 1 - \frac{m}{M} \\ P_0(S_\tau = M-m) = \frac{m}{M} \end{cases}$$

The Second Wald Lemma

Let $\xi_1, \dots, \xi_n$ be an i.i.d. sequence. $P(\xi=1)=P(\xi=-1)=\frac{1}{2}$

Let $\tau = \tau_{-m} \wedge \tau_{M-m}$. Then $E_0 S_\tau^2 = E_0 \tau$

$S_n = \xi_1 + \cdots + \xi_n$

$E_0(S_n^2) = E_0[(\xi_1 + \cdots + \xi_n)^2] = E_0 \xi_1^2 + \cdots + E_0 \xi_n^2 = n$. ($\xi = 1$ or -1)

n is replaced by τ, τ is a random time.

Apply second Wald Lem,

$$E_0\tau = E_0 S_\tau^2 = P_0(S_\tau = -m)(m^2) + P_0(S_\tau = M-m)(M-m)^2$$

$$= \left(1 - \frac{m}{M}\right)(m^2) + \left(\frac{m}{M}\right)(M-m)^2 = m(M-m)$$

Pf of the first Wald Lem

$S_\tau = \xi_1 + \cdots + \xi_\tau = \sum_{n=1}^{\infty} \xi_n \mathbb{1}_{\{\tau \geq n\}}$ indicator

$$\mathbb{1}_{\tau \geq n} = \begin{cases} 1, & \text{if } \tau \geq n \\ 0, & \text{o.w.} \end{cases}$$

In the first $n-1$ steps, we do not hit. we may hit in the n^{th} step.

$E_0 S_\tau = E_0 \sum_{n=1}^{\infty} \xi_n \mathbb{1}_{\{\tau \geq n\}}$

<u>formally</u> $\sum_{n=1}^{\infty} E_0(\xi_n \mathbb{1}_{\{\tau \geq n\}}) = \sum_{n=1}^{\infty} E\xi P_0(\tau \geq n)$

$= E\xi \sum_{n=1}^{\infty} P(\tau \geq n) = E\xi E\tau$

① The event $\{\tau \geq n\}$ is independent of ξ_n.

② $\{\tau \geq n\} = \{S_0 \notin A, \dots, S_{n-1} \notin A\}$ $A = \{-m, M-m\}$ in Gambler's Ruin

③ Expectation of indicator function is a probability

④ Note that: $\sum_{n=1}^{\infty} P(X \geq n) = EX$

Absolute convergence:

$$E_0|S_\tau| \leq \sum_{n=1}^{\infty} E_0(|\xi_n| \mathbb{1}_{\{\tau \geq n\}}) = \sum_{n=1}^{\infty} E_0|\xi_n| P(\tau \geq n) = \sum_{n=1}^{\infty} E_0|\xi_n| \underbrace{E_0\tau}_{<\infty}$$

Markov Chain

R.W. $S_n = \xi_1 + \cdots + \xi_n$. ξ_i i.i.d.

Def'n: (Markov chain)

Let S be a countable set, $\{X_n\}_{n\geq 0}$ is a sequence of random variables taking values in S (state space)

If $\forall n \geq 1$, $\forall i_0, i_1, \ldots, i_{n+1} \in S$,

$$P(\overset{\text{future}}{X_{n+1}} = i_{n+1} \mid \overset{\text{current}}{X_n} = i_n, \overbrace{X_{n-1} = i_{n-1}, \ldots, X_0 = i_0}^{\text{past}}) = P(X_{n+1} = i_{n+1} \mid X_n = i_n)$$

Then we say $\{X_n\}_{n\geq 0}$ is a Markov chain.

- △ independent: future does not depend on past or current.
- △ Markov chain is weaker: future depends on current but does not depend on the past.

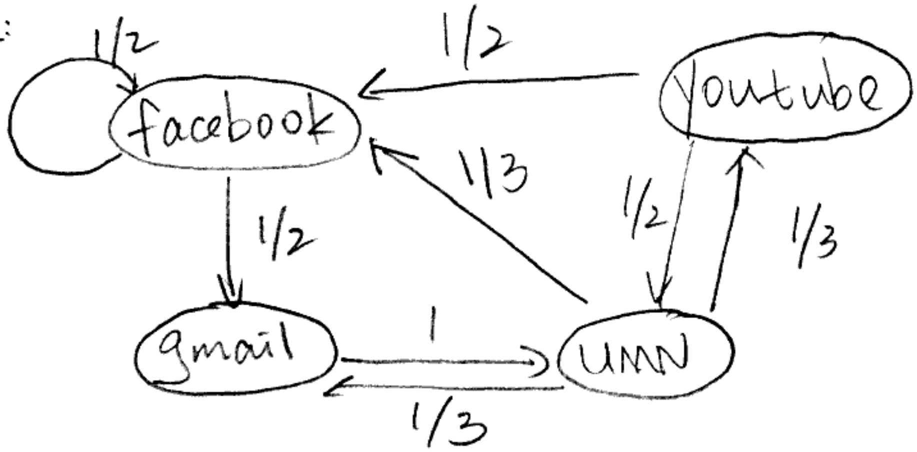
If for any $n \geq 1$ and any $i, j \in S$.

$$P(X_{n+1} = j \mid X_n = i) = p_{ij}$$

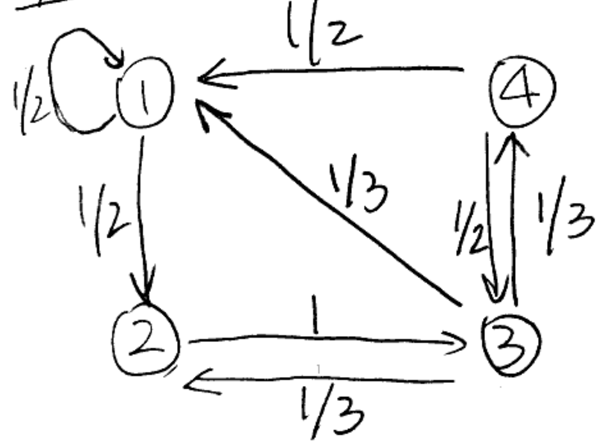
Then we say $\{X_n\}_{n\geq 0}$ is <u>time-homogeneous</u> (nothing to do with time

and p_{ij} is the one step <u>transition probability</u> from state i to state j. $P = (p_{ij})_{i,j\in S}$ is the one step <u>transition matrix</u>.

Ex:



9/22.



1 — facebook

2 — gmail

3 — umn

4 — youtube

transition matrix: (row to column) $\begin{cases}\text{row vectors}\\ \text{column vectors}\end{cases}$

$$P = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 1/2 & 1/2 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 3 & 1/3 & 1/3 & 0 & 1/3 \\ 4 & 1/2 & 0 & 1/2 & 0 \end{array} = (p_{ij})_{1\le i,j\le 4}$$

$p_{ij} = P(X_{n+1}=j \mid X_n=i)$

transition probability

time homogeneous Markov chain.

$$\sum_{j\in S} p_{ij} = \sum_{j\in S} P(X_{n+1}=j \mid X_n=i) = 1.$$

summation of each row is 1

① → ④ in 2 steps.

$1\to3\to4$

$1\to2\to4$

$1\to1\to4$

$1\to4\to4$

$$P(X_{n+m}=j \mid X_n=i) = p_{ij}^{(m)}$$

$$p_{ij}^{(2)} = P(X_{n+2}=j \mid X_n=i) = \frac{P(X_{n+2}=j, X_n=i)}{P(X_n=i)} = \frac{\sum_{k\in S}(X_{n+2}=j, X_{n+1}=k, X_n=i)}{P(X_n=i)}$$

$$= \sum_{k\in S} \frac{P(X_{n+2}=j \mid X_{n+1}=k, X_n=i)P(X_{n+1}=k, X_n=i)}{P(X_n=i)}$$

$$= \sum_{k\in S} P(X_{n+2}=j \mid X_{n+1}=k, X_n=i)\, P(X_{n+1}=k \mid X_n=i)$$

$$\overset{\text{Markov}}{\underset{\text{property}}{=}} \sum_{k\in S} P(X_{n+2}=j \mid X_{n+1}=k)\, P(X_{n+1}=k \mid X_n=i) = \sum_{k\in S} p_{ik}p_{kj}$$

both are transition probabilities

$$\boxed{p_{ij}^{(2)} = \sum_{k\in S} p_{ik}p_{kj}}$$

$$p_{ij}^{(2)} = \sum_{k=1}^{4} p_{ik}p_{kj} = (P^2)_{ij}$$

In general, $p_{ij}^{(m)} = \sum_{k\in S} p_{ik}^{(l)}\, p_{kj}^{(m-l)}$ (fix $l \le m+1$)

$$= (P^m)_{ij}$$

$\uparrow$ transition matrix

<u>Theorem</u>: The following properties are equivalent

(i) Markov $P(X_{n+1}=i_{n+1} \mid X_n=i_n, X_{n-1}=i_{n-1}, \cdots, X_1=i_1)$

$$= P(X_{n+1}=i_{n+1} \mid X_n=i_n) = p_{i_n, i_{n+1}}$$

(ii) joint $P(X_n=i_n, X_{n-1}=i_{n-1}, \dots, X_1=i_1, X_0=i_0)$

$$= P(X_0=i_0)\, p_{i_1,i_2} p_{i_2,i_3} \cdots p_{i_{n-1},i_n}$$

(iii) $\forall n, m \geq 1$, $P(X_{n+m}=j \mid X_n=i_n, \dots, X_0=i_0) = P(X_{n+m}=j \mid X_n=i_n)$
$\forall i_0, \cdots, i_n, j \in S$.

(iv) $\forall r \geq 1$, $\forall n_1 < n_2 < \cdots < n_r$, $\forall i_1, \cdots, i_r \in S$,

$$P(X_{n_1}=i_1, \cdots, X_{n_r}=i_r) = P(X_{n_1}=i_1)\prod_{l=1}^{r-1} \underbrace{P(X_{n_{l+1}}=i_{l+1} \mid X_{n_l}=i_l)}_{p_{i_l, i_{l+1}}^{n_{l+1}-n_l}}$$

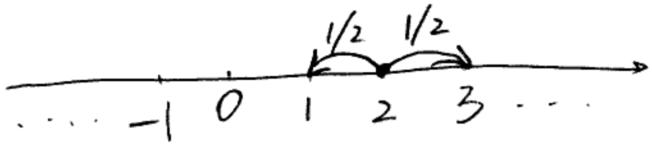
(v) $\forall r, s \geq 1$, $\forall n_1 < \cdots < n_r < n < m_1 < \cdots < m_s$, $\forall i_1, \cdots, i_r, i, j_1, \cdots, j_s \in S$,

$$P(X_{m_1} = j_1, \cdots X_{m_s} = j_s \mid X_n = i, X_{n_1} = i_1, \cdots X_{n_r} = i_r)$$
$$= P(X_{m_1} = j_1, \cdots, X_{m_s} = j_s \mid X_n = i)$$

(vi) $\forall r, s \geq 1$, $\forall n_1 < \cdots < n_r < n < m_1 < \cdots < m_s$, $\forall i_1, \cdots, i_r, i, j_1, \cdots, j_s \in S$,

$$P(X_m = j_1, \cdots, X_{m_s} = j_s, X_{n_1} = i_1, \cdots, X_{n_r} = i_r \mid X_n = i)$$
$$= P(X_{m_1} = j_1 \cdots X_{m_s} = j_s \mid X_n = i) P(X_{n_1} = i_1, \cdots, X_{n_r} = i_r \mid X_n = i)$$

Random Walk



$S = \mathbb{Z}$

$$P_{i,i+1} = P_{i,i-1} = \frac{1}{2}$$

$$P = \begin{array}{c|c} & \cdots\ -2 \quad -1 \quad 0 \quad 1 \quad 2 \\ \hline \begin{matrix} \vdots \\ -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} & \begin{pmatrix} \ddots & & & & \\ & 1/2 & 0 & 1/2 & \\ & & 1/2 & 0 & 1/2 \\ & 0 & & 1/2 & 0 & 1/2 \\ & & & & \ddots \end{pmatrix} \end{array}$$

Gambler Ruin = Random Walk w. absorbing boundary.

$$S_n = m + \xi_1 + \cdots + \xi_n$$

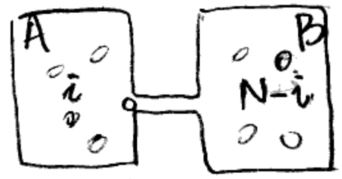
$$P_{i,i+1} = P_{i,i-1} = 1/2 \quad \text{when} \quad 1 \leq i \leq M-1$$

$$P_{0,0} = 1, \quad P_{M,M} = 1$$

$$P = \begin{array}{c|c} & 0 \quad \cdots \quad m \quad \cdots \quad M \\ \hline \begin{matrix} 0 \\ \vdots \\ m \\ \vdots \\ M \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & & & 0 \\ 1/2 & 0 & 1/2 & & & \\ & 1/2 & 0 & 1/2 & & \\ & 0 & & \ddots & & \\ & & & 0 & & 1 \end{pmatrix} \end{array}$$

Ehrenfest chain

We have N balls labelled $1 \cdots N$,



Initially, these balls are distributed randomly in boxes A&B

Suppose # of balls in A $= X_0$.

Each time we pick randomly and independently a ball from the N balls.

And put this ball into the opposite box.

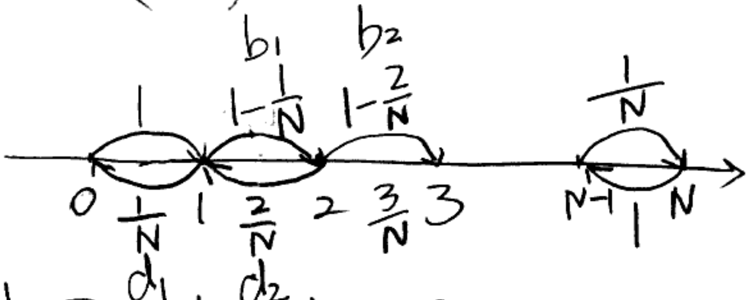
$X_n =$ # of balls in A

$p_{i,i+1} = \frac{N-i}{N} = 1 - \frac{i}{N}$

$p_{i,i-1} = \frac{i}{N}$

$$P = \begin{array}{c|c} & 0 \quad \cdots \quad N \\ \hline \begin{matrix} 0 \\ \vdots \\ \\ N \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{N} & 0 & \frac{N-1}{N} & \cdots & 0 \\ 0 & \frac{2}{N} & 0 & \frac{N-2}{N} \cdots & 0 \\ 0 & & & & \end{pmatrix} \end{array}$$

a special case of Birth and Death chains



Birth and Death Chains

State space $\{0, 1, 2, \cdots, N, \cdots\} = \mathbb{Z}_+$

$p_{01} = 1$ absorbing state

$p_{i,i-1} = d_i$

$p_{i,i+1} = b_i$ $\forall i \geq 1$, st, $d_i + d_i = 1$.

Stationary Distributions:

If $\pi = \{\pi_j, j \in S\}$ st

(1) $\pi_i \geq 0 \quad i \in S$ (2) $\sum_{j \in S} \pi_j = 1$ (3) $\sum_i \pi_i P_{ij} = \pi_j$

9/24 $S_n = \xi_1 + \cdots + \xi_n$

$S_{n+1} = S_n + \xi_{n+1} = S_{n-1} + \xi_n + \xi_{n+1}$

Markov $\Leftrightarrow$ (6) $\forall\, r, s \geq 1$, $\forall\, n_1 < \cdots < n_r < n < m_1 < \cdots < m_r$

$\forall\, i_1, \cdots, i_r, j_1, \cdots, j_s \in S$

$$P(X_{m_1} = j_1, \cdots, X_{m_s} = j_s, X_{n_1} = i_1, \cdots, X_{n_r} = i_r \mid X_n = i)$$

$$= P(X_{m_1} = j_1, \cdots, X_{m_s} = j_s \mid X_n = i) P(X_{n_1} = i_1, \cdots, X_{n_r} = i_r \mid X_n = i)$$

$$\text{LHS} = P(X_{m_1} = j_1, \cdots, X_{m_s} = j_s; X_n = i, X_{n_1} = i_1, \cdots, X_{n_r} = i_r) / P(X_n = i)$$

$$= P(X_{m_1} = j_1, \cdots, X_{m_s} = j_s \mid X_n = i, X_{n_1} = i_1, \cdots, X_{n_r} = i_r)$$

$$P(X_n = i, X_{n_1} = i_1, \cdots, X_{n_r} = i_r) / P(X_n = i)$$

$$= P(X_{m_1} = j_1, \cdots, X_{m_s} = j_s \mid X_n = i) P(X_{n_1} = i_1, \cdots, X_{n_r} = i_r \mid X_n = i)$$

conditional independence

Conditioning that we know current state, the future does not depend on the past.

$P(A \cap B \mid C) = P(A \mid C) P(B \mid C)$ conditional independency

Stationary Distribution

equilibrium state.

"Ergodicity" Spacial Average = Temporal Average.

$P_i(X_n=j) = P(X_n=j \mid X_0=i)$

$X_0 \sim \mu = (\mu_1, \cdots, \mu_n, \cdots)$ $\qquad P(X_0=i) = \mu_i$

$\sum_{i\in S} \mu_i = 1$ probability weight

$P_\mu(X_n=j) = \sum_{i\in S} P(X_n=j \mid X_0=i) P(X_0=i)$

$= \sum_{i\in S} \mu_i P(X_n=j \mid X_0=i)$ (Law of Total probability)

$\mu_j = P_\mu(X_1=j) = \sum_{i\in S} \mu_i p_{ij} = (\mu_1, \mu_2, \cdots)\begin{pmatrix} p_{ij} \end{pmatrix} = (\mu P)_j$

row vector with left multiplication

<u>Def'n</u>: If $\pi = \{\pi_i, i\in S\}$ is such that

(1) $\pi_i \geq 0 \quad \forall i \in S$,

(2) $\sum_{j\in S} \pi_j = 1$

(3) $\pi_j = \sum_{k\in S} \pi_k p_{kj}, \forall j \in S$

Then we say π is a stationary (invariant) distribution of X_n

$\pi = (\pi_i)_{i\in S}$ is a row vector (left eigenvector)

$\pi P = \pi$

2 steps: $\pi P^2 = (\pi P)\cdot P = \pi P = \pi$
(stationary)

(1) When does stationary distribution exist?

(2) Suppose exist, when is it unique? If unique, whether $\mu P^n \to \pi$? or not?

(3) If $\mu P^n \to \pi$, what is the speed?

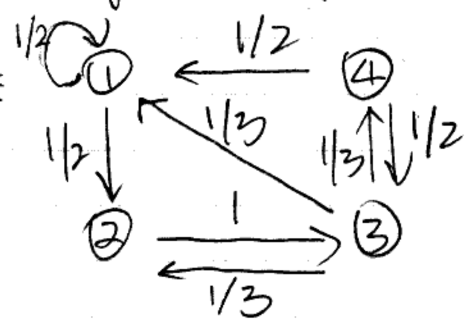
<u>Theorem</u>: (Perron – Frobenius)

Let $S = \{1, 2, \cdots, n\}$ If $P_{ij} > 0 \ \forall i, j \in S$.

Then there is a unique $\pi = (\pi_k)_{k \in S}$ s.t. π is a stationary distribution. Moreover, $\pi_k > 0 \ \forall k \in S$.

(Strong assumption $P_{ij} > 0$)

<u>Example</u>:



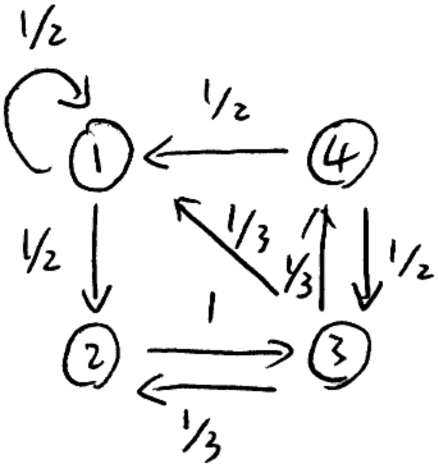
$$P = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \end{array}$$

$\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$

$$\pi P = \pi \iff \begin{cases} 1/2\, \pi_1 + 1/3\, \pi_3 + 1/2\, \pi_4 = \pi_1 \\ 1/2\, \pi_1 + 1/3\, \pi_3 = \pi_2 \\ \pi_2 + 1/2\, \pi_4 = \pi_3 \\ 1/3\, \pi_3 = \pi_4 \end{cases}$$

Website Problem

Stationary Distribution $\quad \pi = (\pi_1, \pi_2, \pi_3, \pi_4)$



$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \end{array}$$

$$\pi P = \pi$$

So

$$(\pi_1, \pi_2, \pi_3, \pi_4) \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$$

$$= (\pi_1, \pi_2, \pi_3, \pi_4)$$

$$\Rightarrow \begin{cases} \frac{1}{2}\pi_1 + \frac{1}{3}\pi_3 + \frac{1}{2}\pi_4 = \pi_1 \\ \frac{1}{2}\pi_1 + \frac{1}{3}\pi_3 = \pi_2 \\ \pi_2 + \frac{1}{2}\pi_4 = \pi_3 \\ \frac{1}{3}\pi_3 = \pi_4 \end{cases}$$

So $\quad \pi_3 = 3\pi_4, \quad \pi_2 = \frac{5}{2}\pi_4, \quad \pi_1 = 3\pi_4$

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \Rightarrow \left(3 + \frac{5}{2} + 3 + 1\right)\pi_4 = 1 \Rightarrow \pi_4 = \frac{2}{19}$$

So $\pi_1 = \frac{6}{19}$ $\pi_2 = \frac{5}{19}$ $\pi_3 = \frac{6}{19}$ $\pi_4 = \frac{2}{19}$

$$\pi = (\pi_1, \pi_2, \pi_3, \pi_4) = \left(\frac{6}{19}, \frac{5}{19}, \frac{6}{19}, \frac{2}{19}\right)$$

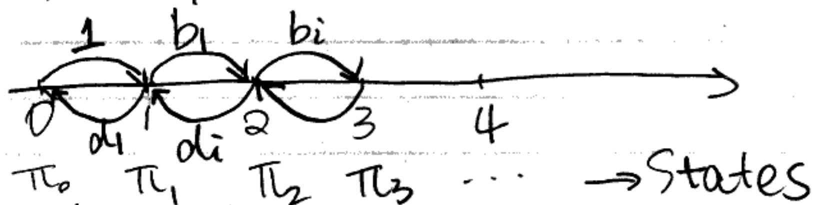
Example:

$$P = \begin{pmatrix} 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 0 & 0 & 1/3 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \end{pmatrix}$$

$$\pi = (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6)$$

$$\pi P = \pi \qquad \pi = \left(\frac{1}{8}, \frac{3}{16}, \frac{1}{8}, \frac{3}{16}, \frac{1}{8}, \frac{1}{4}\right)$$

Example: Birth / Death Chains



$$d_i + b_i = 1$$

$$\pi_0 1 = d_1 \pi_1$$

$$\pi_1 b_1 = \pi_2 d_2$$

$$\cdots$$

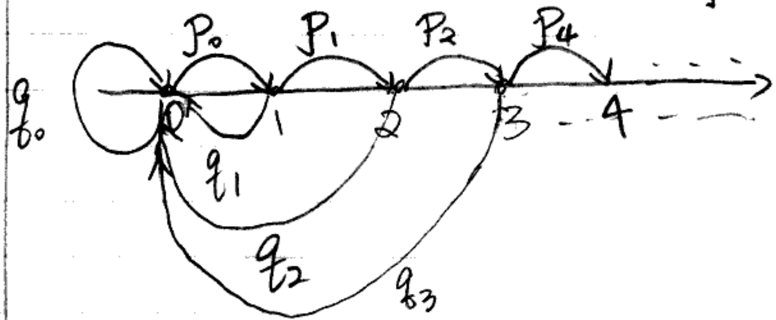
$$\pi_i b_i = \pi_{i+1} d_{i+1}$$

$$\pi_i = \frac{b_1 \cdots b_{i-1}}{d_1 \cdots d_i} \pi_0 \quad , \forall i \geq 1$$

$$1 = \sum_{i=0}^{\infty} \pi_i = \pi_0 \Big(1 + \underbrace{\sum_{i=1}^{\infty} \frac{b_1 \cdots b_{i-1}}{d_1 \cdots d_i}}_{<\infty} \Big)$$

If $\sum_{i=1}^{\infty} \frac{b_1 \cdots b_{i-1}}{d_1 \cdots d_i} < \infty$, then Invariant/stationary dist. exist

Example: $S = \mathbb{Z}_+$, $\forall i \geq 0$, ① $p_{i0} = q_i$ ② $p_{i,i+1} = p_i$ ③ $p_i + q_i = 1$.



summation only contributes 1 item to π_i

$$\boxed{\pi_i = \sum_{j \in S} \pi_j p_{ji}}$$

fix state i,

$$= \pi_{i-1} p_{i-1} = \pi_{i-2} p_{i-2} p_{i-1} = \cdots = \pi_0 p_0 p_1 p_2 \cdots p_{i-1}$$

$$\sum_{i \in S} \pi_i = 1 \Rightarrow \pi_0 \left(1 + \sum_{i=1}^{\infty} p_0 \cdots p_{i-1}\right) = 1$$

If $\sum_{i=1}^{\infty} p_0 \cdots p_{i-1} < \infty$ (convergence) Then $\exists$ stationary dist.

9/29.

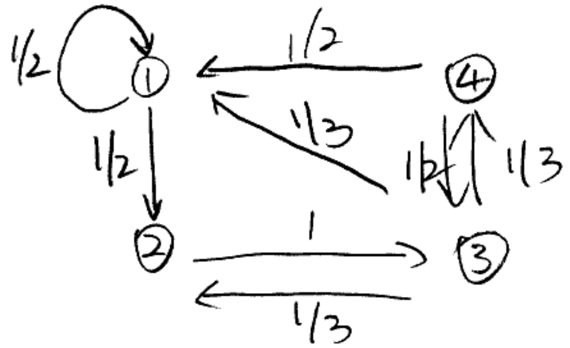
Definition $P_i(\exists n \geq 0 \text{ st } X_n = j) > 0$

(start from i, with positive probability)

Then we say that i communicates with j.

and we denote it as $i \to j$

If $i \to j$ and $j \to i$, then we say i and j mutually communicates, and we denote $i \leftrightarrow j$.



④ communicates with ①

④ communicates with ②

(sequence of $\to$'s)

④ .. with ③ }
③ .. with ④ } different probability

Observation:

(1) $i \leftrightarrow i$

(2) if $i \leftrightarrow j$ then $j \leftrightarrow i$

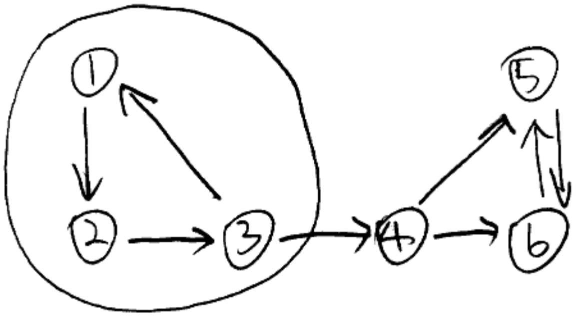
(3) if $i \leftrightarrow j$, $j \leftrightarrow k$, then $i \leftrightarrow k$ transitivity

Thus, mutually communicating is an equivalent relation

Thus, the state space S is decomposed into a disjoint union of mutually communicating classes each of these classes is called a mutually communicating class, denoted by C.

$S = C_1 \cup C_2 \cup \cdots \cup C_l$

Example.



communicating classes

$C_1 = \{1, 2, 3\}$

$C_2 = \{4\}$

$C_3 = \{5, 6\}$

$$\boxed{C_1} \longrightarrow \boxed{C_2} \longrightarrow \boxed{C_3}$$

Def'n: If $A \subseteq S$ is such that $\forall i \in A, \sum_{j \in A} p_{ij} = 1$.
Then we say A is a closed set.

$\boxed{C_3}$ is a closed set.

Remarks: ① If A is closed, then $\forall i \in A, (i \to j) \Rightarrow (j \in A)$

② If C is a mutually communicating class, A closed, then $C \cap A = \phi$ or $C \subseteq A$.

③ If A is closed, then $P|_A = (p_{ij})_{i,j \in A}$ is a transition matrix.

$$P = \begin{array}{cc} & \overbrace{\qquad}^{A} \\ A\{ & \begin{pmatrix} \boxed{P|_A} & 0 \\ * & \boxed{} \end{pmatrix} \end{array}$$

④ If $\{i\}$ is closed, then we say i is an absorbing state.

In this case. $p_{ii} = 1$

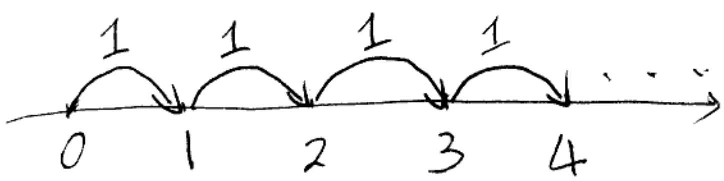
⑤ Notions of mutually communicating and closedness depends only on transition matrix $P = (p_{ij})$

<u>Def'n</u>: If state space S has a non-empty closed strict subset, Then we say S is reducible, also the corresponding transition matrix is reducible. Otherwise, we say S is irreducible and the transition matrix P is irreducible.

$$\cancel{\boxed{C_1}} \longrightarrow \boxed{C_2} \longrightarrow \boxed{C_3}$$

keep on throwing out largest mutually communicating class.

<u>Remark</u>: Is it always the case that we can keep on "throwing out" mutually communicating classes. So we end up in a closed irreducible smaller state space?



infinite.

mutually communicating classes are singleton sets

Let $B \subseteq S$. We denote the first hitting time to B by

$\tau_B = \{ n \geq 0, X_n \in B \}$.

$\tau_B = \tau_B(\omega)$ is a random variable

$\{\tau_B \leq n\}$ only depends on $X_0, \cdots, X_n$.

In this case, we say τ_B is a stopping time/Markov time.
(not depends on the future, only depends on $X_0, \cdots, X_n$)

hitting probability (distribution) (sooner or later, you will hit A)

Take $A \subseteq S$ and $i \notin A$, Then we want $P_i(\tau_A < \infty)$.

hitting time $E_i \tau_A$

First step trial method

$$\boxed{P_i(\tau_A<\infty)=\sum_{j\in S} P_{ij} P_j(\tau_A<\infty)}$$

$$P_i(\tau_A<\infty) = \sum_{j\in S} P(X_1=j, \tau_A<\infty \mid X_0=i)$$

$$= \sum_{j\in S} P(X_1=j \mid X_0=i)\, P(\tau_A<\infty \mid X_1=j, X_0=i)$$

$$\frac{P(X_1=j, X_0=i)}{P(X_0=i)} \cdot \frac{P(\tau_A<\infty, X_1=j, X_0=i)}{P(X_1=j\ X_0=i)}$$

$$= \sum_{j\in S} P_{ij} P_j(\tau_A<\infty)$$

$$E_i\tau_A = \sum_{j\in S} P(X_1=j \mid X_0=i)\left[E(\tau_A \mid X_1=j, X_0=i)+1\right]$$

$$= \sum_{j\in S} P_{ij} E_j \tau_A + 1 \qquad \boxed{E_i\tau_A = 1+\sum_{j\in S} P_{ij} E_j \tau_A}$$

Example: (Two year College)

1st year 60% freshman become sophemores

25% remain freshman

15% drop out

2nd year 70% sophermores graduate

20% remain sophermores

10% drop out

① What fractions of new students eventually graduate?

② On the average, how many years does a student take to graduate or drop out?

1 — freshmen

2 → sophermore

G → graduate

D → drop out

$$P = \begin{array}{c|cccc} & 1 & 2 & G & D \\ \hline 1 & 0.25 & 0.6 & 0 & 0.15 \\ 2 & 0 & 0.2 & 0.7 & 0.1 \\ G & 0 & 0 & 1 & 0 \\ D & 0 & 0 & 0 & 1 \end{array}$$

① $P_1(\tau_{\{G\}} < \infty) = 0.7$

② $E_1 \tau_{\{G,D\}} = 2.333\ldots$

☆ first trial method

Let $P_i(\tau_{\{G\}} < \infty) = h_i$, $i = 1, 2$.

$$\begin{cases} h_1 = 0.25h_1 + 0.6h_2 \\ h_2 = 0.2h_2 + 0.7 \times 1 \end{cases} \Rightarrow \begin{cases} h_1 = \frac{7}{10} \\ h_2 = \frac{7}{8} \end{cases}$$

Let $g_i = E_i \tau_{\{G,D\}}$.

$$g_G = g_D = 0.$$

$$\begin{cases} g_1 = 1 + 0.25g_1 + 0.6g_2 \\ g_2 = 1 + 0.2g_2 \end{cases} \Rightarrow \begin{cases} g_1 = 2.333\ldots \\ g_2 = 1.25 \end{cases}$$

Example: (Tennis game)

In tennis, winner of a game is the first player to win 4 points unless 4 − 3.

in this case, we continue until one player is ahead by 2 points and win the game.

The server win the point with probability 0.6
win successive points independently.

① What is the probability win the game if score is 3-3
What is the probability win the game if she is ahead by 1 point?
·· if she is behind by 1 point?

② What is the expected time to <u>complete</u> the game in the above cases?

State: difference of the scores

$$P = \begin{array}{c|ccccc} & 2 & 1 & 0 & -1 & -2 \\ \hline 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 & 0 \\ -1 & 0 & 0 & 0.6 & 0 & 0.4 \\ -2 & 0 & 0 & 0 & 0 & 1 \end{array}$$

① $P_0(\tau_{\{2\}}<\infty)$ $P_1(\tau_{\{2\}}<\infty)$ $P_{-1}(\tau_{\{2\}}<\infty)$

② $E_0\,\tau_{\{2,-2\}}$ $E_1\,\tau_{\{2,-2\}}$ $E_{-1}\,\tau_{\{2,-2\}}$

10/1

Example: (Tennis)

$$P = \begin{array}{c} \\ 2 \\ 1 \\ 0 \\ -1 \\ -2 \end{array} \begin{array}{c} \begin{array}{ccccc} 2 & 1 & 0 & -1 & -2 \end{array} \\ \left(\begin{array}{c|ccc|c} 1 & 0 & 0 & 0 & 0 \\ \hline 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array}$$

transition matrix.

① probabilities of success if score is $\underset{h_0}{3\text{-}3}$, $\underset{h_1}{4\text{-}3}$, $\underset{h_{-1}}{3\text{-}4}$.

② Expected time to <u>complete</u> the game (run into state 2 or -2)

Let $h_i = P_i(\tau_{\{2\}} < \infty)$

$$\begin{cases} h_1 = 0.6 \cdot 1 + 0.4 h_0 \\ h_0 = 0.6 h_1 + 0.4 h_{-1} \\ h_{-1} = 0.6 h_0 + 0.4 \times 0 \text{ (the other part get success)} \end{cases} \Rightarrow \begin{cases} h_1 = 0.8769 \\ h_0 = 0.6923 \\ h_{-1} = 0.4154 \end{cases}$$

Let $g_i = E_i(\tau_{\{2,-2\}})$

Boundary conditions: $g_2 = g_{-2} = 0$

$$\begin{cases} g_1 = 1 + 0.4 g_0 \\ g_0 = 1 + 0.6 g_1 + 0.4 g_{-1} \\ g_{-1} = 1 + 0.6 g_0 \end{cases} \Rightarrow \begin{cases} g_1 = 33/13 \\ g_0 = 50/13 \\ g_{-1} = 43/13 \end{cases}$$

$\{A_n, \text{ infinitely often}\} = \{A_n, \text{ i.o.}\} = \bigcap_{N=1}^{\infty} \bigcup_{n \geq N} A_n$ <u>event</u>

$A_1, A_2, \ldots A_n, \ldots \boxed{A_N \ldots}$ tail.

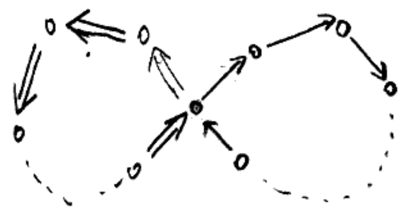
Def'n: If $P_i(X_n = i \text{ i.o.}) = P_i\left(\bigcap_{N\geq 1}\bigcup_{n\geq N}\{X_n = i\}\right) = 1$

Then we say that state i is a recurrent state

Or we say state i is recurrent. (always come back)

Otherwise, we say state i is a transient state or

we say state i is non-recurrent. $P_i < 1$



cycle method

$\sigma_i^{(1)} = \sigma_i = \inf\{n \geq 1, X_n = i\}$ first time to comeback to i

Now we recursively define $\sigma_i^{(k)} = \begin{cases} \inf\{n \geq \sigma_i^{(k-1)}+1, X_n = i\}, & \text{if } \sigma_i^{(k-1)} < \infty \\ \infty, & \text{if } \sigma_i^{(k-1)} = \infty \end{cases}$

Thus $\sigma_i^{(k)}$ is the time for the k-th visit of X_n to state i.

Claim: For any $k \geq 1$, $n \geq 1$

$$P_i(\sigma_i^{(k+1)} - \sigma_i^{(k)} = n \mid \sigma_i^{(k)} < \infty) = P_i(\sigma_i = n)$$

Moreover, $\sigma_i^{(1)}, \sigma_i^{(2)} - \sigma_i^{(1)}, \dots, \sigma_i^{(k+1)} - \sigma_i^{(k)} \dots$ are independent.

Pf. $P_i(\sigma_i^{(2)} - \sigma_i^{(1)} = n, \sigma_i^{(1)} = m)$

$= P_i(\sigma_i^{(2)} = n+m, \sigma_i^{(1)} = m)$

$= P_i(X_m = X_{m+n} = i, X_k \neq i, 1 \leq k \leq m+n-1, k \neq m)$

$= P_i(X_{m+n} = i, X_k \neq i, m+1 \leq k \leq m+n-1 \mid X_m = i, X_l = i, 1 \leq l \leq m-1)$

$\times P_i(X_m = i, X_l = i, 1 \leq l \leq m-1)$

$$P_i = (X_{n+m}=i, X_k \neq i,\ m+1 \le k \le n+m-1 \mid X_m = i)$$
$$\times P_i(X_m = i, X_l \neq l, 1 \le l \le m-1)$$
$$= P_i(X_n = i, X_k \neq i, 1 \le k \le n-1)$$
$$\times P_i(X_m = i, X_l \neq i, 1 \le l \le m-1)$$
$$= P_i(\sigma_i^{(1)} = n) P_i(\sigma_i^{(1)} = m)$$

$\rho_{ij} = P_i(\sigma_j < \infty)$ first time come back to j.

$\rho_{ii} = P_i(\sigma_i < \infty)$

Claim: $P_i = (\sigma_i^{(k)} < \infty) = \rho_{ii}^k$

Pf: $P_i(\sigma_i^{(k)} < \infty) = P_i(\sigma_i^{(k-1)} < \infty, \sigma_i^{(k)} - \sigma_i^{(k-1)} < \infty)$

$$= P_i(\sigma_i^{(k-1)} < \infty) P_i(\sigma_i^{(k)} - \sigma_i^{(k-1)} < \infty \mid \sigma_i^{(k-1)} < \infty)$$
$$= P_i(\sigma_i^{(k-1)} < \infty) P_i(\sigma_i < \infty)$$

Let $V_i = \sum_{n=1}^{\infty} 1_{\{X_n = i\}}$ number of visits to state i

We see ① $V_i \ge k \Leftrightarrow \sigma_i^{(k)} < \infty$

② $\{V_i = \infty\} = \{X_n = i, \text{ i.o.}\}$

So $P_i(V_i \ge k) = \rho_{ii}^k$ and V_i has geometric distribution with success probability ρ_{ii}

Proposition: (0-1 law)

$$P_i(V_i = \infty) = 0 \Leftrightarrow E_i V_i < \infty \Leftrightarrow \rho_{ii} < 1$$
$$P_i(V_i = \infty) = 1 \Leftrightarrow E_i V_i = \infty \Leftrightarrow \rho_{ii} = 1$$

Conclusion: The probability that state i will be visited infinitely many times is 0 or 1.

$p_{ij}^{(n)}$ n-step transition probability

Let $G_{ij} = \sum_{n=0}^{\infty} p_{ij}^{(n)}$ (Green function)

$$E_i V_i = E_i \sum_{n=1}^{\infty} 1_{\{X_n = i\}} = \sum_{n=1}^{\infty} P_i(X_n = i) = \sum_{n=1}^{\infty} p_{ii}^{(n)} = \underline{\underline{G_{ii} - 1}}$$

Corollary:

(1) $\rho_{ii} = 1 \Leftrightarrow i$ recurrent $\Leftrightarrow G_{ii} = \infty$

(2) $\rho_{ii} < 1 \Leftrightarrow i$ transient $\Leftrightarrow G_{ii} < \infty$

} one and only one of (1), (2) will happen.

Statement: If $i \leftrightarrow j$ mutually communicate.

Then i is recurrent $\Leftrightarrow$ j is recurrent.

Pf: Suppose i is transient, $G_{ii} < \infty$, then we can choose $k, l > 0$ st $p_{ij}^{(k)}, p_{ji}^{(l)} > 0$

Then $p_{ii}^{(k+r+l)} \geq p_{ij}^{(k)} p_{jj}^{(r)} p_{ji}^{(l)} \xRightarrow[r]{\text{sum over}} G_{ii} \geq G_{jj} p_{ij}^{(k)} p_{ji}^{(l)}$

$\Rightarrow G_{jj} < \infty \Rightarrow j$ is transient and vice versa.

(mutually communicating transient class)

Thus, recurrence / transient is a property of mutually communicating class and we can talk about recurrent class / transient class.

Moreover, if the markov chain is irreducible, then all states are either recurrent / transient. In this case, we talk about recurrent / transient of the Markov chain.

10/6

$P_i(V_i<\infty)=1 \Leftrightarrow i$ is transient $\Leftrightarrow E_iV_i<\infty \Leftrightarrow \rho_{ii}<1 \Leftrightarrow G_{ii}<\infty$

$P_i(V_i=\infty)=1 \Leftrightarrow i$ is recurrent $\Leftrightarrow E_iV_i=\infty \Leftrightarrow \rho_{ii}=1 \Leftrightarrow G_{ii}=\infty$

$G_{ij}=\sum_{n=0}^{\infty} p_{ij}^{(n)} \qquad E_iV_i = G_{ii}-1$

$i \longleftrightarrow j \quad i$ recurrent $\Leftrightarrow$ j recurrent

irreducible $\Leftrightarrow$ all states are either transient or recurrent.

Lem: Let π be a stationary distribution

and the state j is transient Then $\pi_j=0$

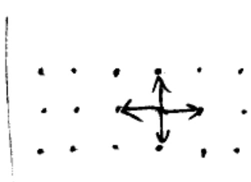
Pf: $\forall i,j\in S$, claim: $\sum_{n=0}^{\infty} p_{ij}^{(n)} \le G_{jj}=\sum_{n=0}^{\infty} p_{jj}^{(n)}<\infty$ (HW)

Therefore, $\sum_{i\in S}\pi_i \sum_{n=0}^{\infty} p_{ij}^{(n)} \le \sum_{i\in S}\pi_i G_{jj} = G_{jj}<\infty$

$$\sum_{n=0}^{\infty}\sum_{i\in S}\pi_i p_{ij}^{(n)} = \sum_{n=0}^{\infty}\pi_j<\infty \Rightarrow \pi_j=0$$

Fact: d-dimensional RW is on $\mathbb{Z}^d$

$P=\frac{1}{2d}$



$d=1,2$ it is recurrent

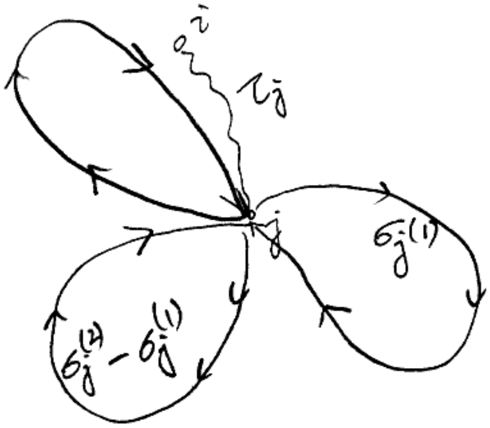
$d\geq 3$ it is transient.
(degree of freedom)

How often you are going to come back?

$V_j(n)=\sum_{k=0}^{n} 1_{\{X_k=j\}}$ visit j

"frequency" $\lim_{n\to\infty}\frac{V_j(n)}{n}$ (Law of large numbers)

Let $X_0=i$, Suppose it is on irreducible chain.



From the picture.

$V_j(n) = \#$ of excursion $j \to j+1$

By cycle argument, each excursion is independent.

Length of excursion $= \sigma_j^{(1)}, \sigma_j^{(2)}-\sigma_j^{(1)}, \sigma_j^{(3)}-\sigma_j^{(2)} \cdots$ are independent. also i.i.d.

Let us consider $\lim_{n\to\infty} \frac{n}{V_j(n)}$ should be the period

$$\frac{\text{number of steps}}{\text{number of excursions}}$$

Roughly speaking, $\lim_{n\to\infty} \frac{n}{V_j(n)} = \lim_{k\to\infty} \frac{\sigma_j^{(1)}+(\sigma_j^{(2)}-\sigma_j^{(1)})+\cdots+(\sigma_j^{(k)}-\sigma_j^{(k-1)})}{k}$

$\xrightarrow{\text{Law of large numbers}}$ $E_j\sigma_j^{(1)}$ $\quad \sigma_j^{(1)} = \sigma_j$ for short

Thus, "frequency" $\approx \lim_{n\to\infty} \frac{V_j(n)}{n} = \frac{1}{E_j\sigma_j^{(1)}}$

<u>Theorem</u>: Let $V_j(n) = \sum_{k=0}^{n} 1_{\{X_k=j\}}$

Suppose P is irreducible

then $\forall j \in S$, and any initial distribution μ.

Then $P_\mu\left(\lim_{n\to\infty} \frac{V_j(n)}{n} = \frac{1}{E_j\sigma_j}\right) = 1$

(frequency is the reciprocal of Time)

Pf: It suffices to prove the case $\mu = \delta_i = \mathbb{1}_{\{i\}}$

If P is transient then $\sum_{k=0}^{\infty} 1_{\{X_k = j\}} < \infty$

Then $\lim_{n\to\infty} \frac{V_j(n)}{n} = 0 = \frac{1}{E_j \sigma_j}$

Since j is transient, we know $P_j(\sigma_j < \infty) < 1 \Rightarrow E_j \sigma_j = \infty$

Now assume P is recurrent.

Then $1 = P_j(\sigma_j < \infty) = P_i(\sigma_j < \infty)$ (HW)

Thus $P_i\left(\sum_{k=0}^{\infty} 1_{\{X_k = j\}} = \infty\right) = 1$ number of visits.

Let $\{\tilde{X}_n = X_{\tau_j + n}, n \geq 0\}$

contains excursions, start from the first hitting time.

$\{V_j(n) = k+1\} = \{\tau_j + \tilde{\sigma}_j^{(k)} \leq n \leq \tau_j + \tilde{\sigma}_j^{(k+1)}\}$

where $\tilde{\sigma}_j^{(k)}$ is the time $\{\tilde{X}_n, n \geq 0\}$ comes back to j at the k-th excursion.

Therefore, $\tilde{\sigma}_j^{(1)}, \tilde{\sigma}_j^{(2)} - \tilde{\sigma}_j^{(1)}, \cdots, \tilde{\sigma}_j^{(k)} - \tilde{\sigma}_j^{(k-1)}$ are i.i.d by cycle method.

Thus, by Strong Law of Large numbers,

$$P_i\left(\lim_{k\to\infty} \frac{\tilde{\sigma}_j}{k} = E_j \sigma_j\right) = 1$$

weak LLN

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - EX\right| > \varepsilon\right) \to 0$$

strong LLN

$$P\left(\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = EX\right) = 1.$$

We know, $\{V_j(n) = k+1\}$

$$= \{T_j + \tilde{\sigma}_j^{(k)} \le n \le T_j + \sigma_j^{(k+1)}\}$$

and as $n \to \infty$, we know $V_j(n) \to \infty$

$$\overline{\lim_{n\to\infty}} \frac{n}{V_j(n)} \le \overline{\lim_{k\to\infty}} \frac{T_j + \tilde{\sigma}_j^{(k+1)}}{k+1} = \overline{\lim_{k\to\infty}} \frac{\tilde{\sigma}_j^{(k+1)}}{k+1} = E_j\sigma_j$$

$$\underline{\lim_{n\to\infty}} \frac{n}{V_j(n)} \ge \underline{\lim_{k\to\infty}} \frac{T_j + \tilde{\sigma}_j^{(k)}}{k+1} = \underline{\lim_{k\to\infty}} \frac{k}{k+1} \cdot \frac{\tilde{\sigma}_j^{(k)}}{k} = E_j\sigma_j$$

Thus, $P_i\left(\lim_{n\to\infty} \frac{n}{V_j(n)} = E_j\sigma_j\right) = 1$ which implies the Theorem.

$$V_j(n) = \sum_{k=0}^{n} 1\{X_k = j\}$$

By our Theorem on frequencies,

$$P_\mu\left(\lim_{n\to\infty} \frac{V_j(n)}{n} = \frac{1}{E_j\sigma_j}\right) = 1$$

Suppose $\mu = \pi$ is the stationary distribution.

$$E_\pi \frac{V_j(n)}{n} = E_\pi \frac{\sum_{k=0}^{n} 1\{X_k = j\}}{n} = \frac{1}{n}\sum_{k=0}^{n} P_\pi(X_k = j) = \pi_j$$

Thus $P_\pi\left(\lim_{n\to\infty} E\frac{V_j(n)}{n} = \frac{1}{E_j\sigma_j}\right) = 1$

Therefore $\boxed{\pi_j = \frac{1}{E_j\sigma_j}}$

- The stationary distribution is the frequency.
- If j is transient, $\pi_j = 0$, $E_j\sigma_j = \infty$
- time average (frequency) $=$ spatial average (π's)

$P_j(\sigma_j < \infty) < 1$ Transient

which $\Leftrightarrow$ $P_j(\sigma_j = \infty) > 0 \Rightarrow E_j\sigma_j = \infty \Rightarrow \pi_j = 0$.

$P_j(\sigma_j < \infty) = 1$ Recurrent

$\not\Rightarrow$ $E_j\sigma_j < \infty$

Def'n Let j be recurrent. If $E_j\sigma_j < \infty$ then j is called positive recurrent. (time is finite)

If $E_j\sigma_j = \infty$, Then j is called null recurrent. (time is infinite)

Theorem: If P is irreducible, Then following statements are equivalent.

(1) All states are positive recurrent.

(2) There exist positive recurrent state.

(3) There exist stationary distribution so that

$$\pi_j = \frac{1}{E_j\sigma_j} \quad \forall j.$$

Theorem: (Ergodic Theorem)

Suppose P is irreducible and positively recurrent. Let f be a bounded function on S let $\pi = (\pi_i)_{i\in S}$ be stationary distribution

Then (MCMC method)

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} f(X_k) = \sum_{i\in S}\pi_i f(i)\right) = 1$$

Time average Space average

Let the chain P be irreducible.

$$P_\mu\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} f(X_k) = \sum_{j\in S}\pi_j f(j)\right) = 1$$

weighted sum

$$\pi_j = \frac{1}{E_j \sigma_j}$$

<u>Example</u>: <u>(Repair chain)</u>

A machine has 3 critical parts that are subject to failure but can function as long as two are working. When two are broken then they are replaced and machine is back to working order next day.

Let <u>State space</u> be parts that are broken:

$S = \{0, 1, 2, 3, 12, 23, 13\}$ 7 states

Parts 1, 2, 3 fail with probabilities 0.01, 0.02, 0.04

No two parts fail on the same day.

<u>transition matrix</u>

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 12 \\ 23 \\ 13 \end{array}\begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & 3 & 12 & 23 & 13 \end{array} \\ \begin{pmatrix} 0.93 & 0.01 & 0.02 & 0.04 & 0 & 0 & 0 \\ 0 & 0.94 & 0 & 0 & 0.02 & 0 & 0.04 \\ 0 & 0 & 0.95 & 0 & 0.01 & 0.04 & 0 \\ 0 & 0 & 0 & 0.97 & 0 & 0.02 & 0.01 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

fix the machine

Q: We operate the machine for 1800 days. (approx. 5 years) Then how many parts of 1, 2, 3 are we going to use?

invarient distribution:

solve for $\pi = (\pi_0, \pi_1, \pi_2, \pi_3, \pi_{12}, \pi_{23}, \pi_{13})$

s.t. $\pi P = \pi$

$\pi_0 = \frac{3000}{8910}$ $\quad \pi_1 = \frac{500}{8910}$ $\quad \pi_2 = \frac{1200}{8910}$ $\quad \pi_3 = \frac{4000}{8910}$

$(E_{12}\sigma_{12})^{-1} = \pi_{12} = \frac{22}{8910}$ $\quad \pi_{23} = \frac{128}{8910}$ $\quad \pi_{13} = \frac{60}{8910}$

By the Ergodic theorem

$\#$ of part 1 $= 1800 \times (\pi_{12} + \pi_{13}) = 1800 \times \frac{82}{8910} \approx 16.56$

$\#$ of part 2 $= 1800 \times (\pi_{12} + \pi_{23}) \approx 30.30$

$\#$ of part 3 $= 1800 \times (\pi_{13} + \pi_{23}) \approx 37.98$

$P_i(X_n = j) = p_{ij}^{(n)} \xrightarrow[?]{n\to\infty} \pi_j, \forall i$, Each p_{ij} converges to π_j?

n-step transition probability

<u>Ergodic theorem</u> tells us that with probability 1

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n} 1\{X_k = j\}}{n} = \pi_j$$ we take expectation on both sides

Then $$\lim_{n\to\infty} \frac{\sum_{k=0}^{n} P_i(X_k = j)}{n} = \pi_j \qquad E 1_A = P(A)$$

i.e. $$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} p_{ij}^{(k)} = \pi_j$$ (average converges to π_j)

<u>Example:</u>

$$\begin{tikzcd} ① \arrow[r, "1"] & ② \arrow[d, "1"] \\ ④ \arrow[u, "1"] & ③ \arrow[l, "1"] \end{tikzcd}$$

irreducible

$\pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4}$

$P_{12}^{(1)} = 1$. $P_{12}^{(2)} = 0$ $P_{12}^{(3)} = 0$ $P_{12}^{(4)} = 0$

$P_{12}^{(5)} = 1$. $P_{12}^{(6)} = 0$ $P_{12}^{(7)} = 0$ $P_{12}^{(8)} = 0$. ... periodically

$\lim_{n\to\infty} P_{12}^{(n)}$ does not exist! But $\lim_{n\to\infty} \frac{\sum_{k=1}^{n} P_{12}^{(k)}}{n} = \frac{1}{4}$

average converges.

(The process has to be essentially random).

<u>Def'n</u>: We call the greatest common divisor of numbers $\{n \geq 1,\ P_{ii}^{(n)} > 0\}$ be the <u>period</u> of state i.

If i has period 1, then i is called <u>aperiodic</u>.

<u>ex:</u> $\{4, 8, 12, \cdots\} \rightarrow$ period $= 4$

periodic chains do not converge!

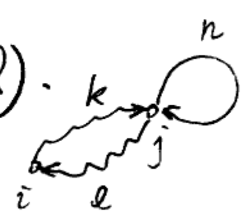
"loss of memory"

<u>Proposition</u>: If $i \leftrightarrow j$ then i and j have the same period.

<u>Pf.</u> Assume the periods of i and j are d_i, d_j

Since $i \leftrightarrow j$. we can pick up $k, l \geq 1$ st.

$P_{ij}^{(k)}, P_{ji}^{(l)} > 0$. Therefore, $P_{ii}^{(k+l)} > 0$ so $d_i \mid (k+l)$.



On the other hand, $\forall\, n \in R_j = \{m \geq 1,\ P_{jj}^{(m)} > 0\}$

We have $P_{ii}^{(k+n+l)} \geq P_{ij}^{(k)} P_{jj}^{(n)} P_{ji}^{(l)} > 0$. So $d_i \mid (k+n+l)$.

Thus, d_i / n Then $d_i \mid d_j$

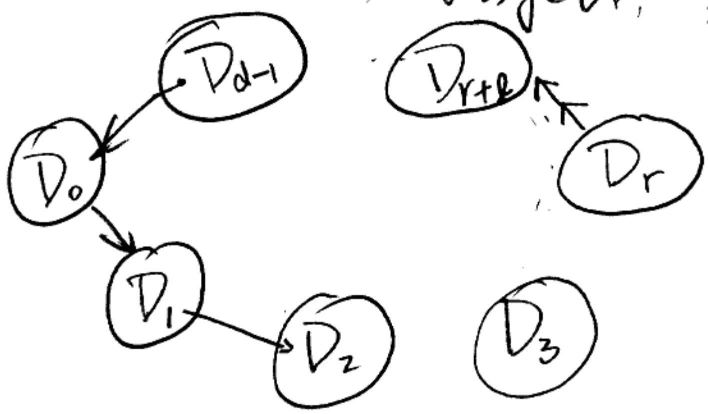
Conversely, d_j / d_i so $d_i = d_j$

Thus, we can talk about the period of a mutually communicating class, the period of an irreducible Markov chain, etc.

<u>Theorem</u>: Let P be irreducible. Then $\exists d \geq 1$ and a partition $D_0, D_1, \cdots, D_{d-1}$ of S (we define $D_{nd+r} = D_r$)

s.t. (i) $\forall r \geq 0$, $\forall i \in D_r$, $\forall l \geq 0$, $\sum_{j \in D_{r+l}} P_{ij}^{(l)} = 1$

(ii) $\forall r \geq 0$, $\forall i, j \in D_r$, $\exists n_0 \geq 0$ s.t. $\forall n \geq n_0$, $P_{ij}^{(nd)} > 0$



* Deterministically, from D_r, after l steps, reach D_{r+l}.

* after d steps, must come back to j from i (i, j in same set)

d is the period.

<u>Conclusion</u>: If P is aperiodic, Then $\forall i, j \in S$.

$\exists n_0 \geq 0$. s.t. $\forall n \geq n_0$, $P_{ij}^{(n)} > 0$ $(d=1)$

(arbitrary many steps)

<u>Theorem</u>: If P is irreducible and <u>aperiodic</u>.

mixing)

Let π be the stationary distribution.

Then $\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j$. Moreover, $\lim_{n\to\infty} \sum_{n=0}^{\infty} |P_{ij}^{(n)} - \pi_j| = 0$ $\forall i \in S$

10/13

period of state i is the greatest common divisor of the set $\{n \geq 1,\ p_{ii}^{(n)} > 0\}$.

aperiodic. Goal: $p_{ij}^{(n)} \to \pi_j$ as $n \to \infty$.

Theorem: Suppose P is irreducible and aperiodic. Then if π is the stationary distribution, we have

$$\lim_{n\to\infty} \sum_{j\in S} |p_{ij}^{(n)} - \pi_j| = 0 \quad \forall\, i \in S.$$

In particular, $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$ $\quad p_{ij}^{(n)} = P_i(X_n = j)$

"Coupling"

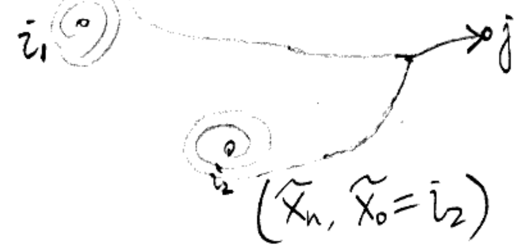
X_n $\quad Y_n$ $\quad S$ $\quad$ We don't require X_n and Y_n are Markov

$Z_n = (X_n, Y_n)$ $\quad$ State space $S \times S$

* No matter where you start the chain, you will end up with the same distribution (π_j)

* loss of memory. $\leftarrow$ aperiodicity.
  we don't know where we come from.

$(X_n, X_0 = i_1)$



Same process
Start from different states.

$(\tilde{X}_n, \tilde{X}_0 = i_2)$

$(Y_n, Y_n = i_2)$

Pf: (Details will not be on the exam)

Let $\{X_n, n\geq 0\}$ and $\{Y_n, n\geq 0\}$ be two independent Markov chains on S with transition probability matrix P.

Initial distributions μ and ν

Let $Z_n = \{(X_n, Y_n): n\geq 0\}$ be the coupled process.

The state space of Z_n is $S\times S$.

Therefore, the transition probability of Z_n is

$P_{(i_1,i_2),(j_1,j_2)} = P_{i_1j_1}P_{i_2j_2}$ and we denote the trasition matrix $\bar{P}$.

Initial distribution: $\mu\times\nu$

Step 1: Since P is irreducible. By the conclusion we have on Lecture 10 before we state the Theorem, we see that we can find n sufficiently large, so that $P^{(n)}_{i_1j_1}>0$, $P^{(n)}_{i_2j_2}>0$

$i_1 \overset{n}{\rightsquigarrow} j_1$

$i_2 \rightsquigarrow j_2$

This means that the coupling chain $Z_n = (X_n, Y_n)$ is irreducible

(Remark: In case P is periodic it can happen that Z_n is reducible

$$\begin{tikzcd} 1 \arrow[r, "1"] & 2 \arrow[d, "1"] \\ 4 \arrow[u, "1"] & 3 \arrow[l, "1"] \end{tikzcd}$$

state $(2,4)$ never reached by Z_n)

And Z_n is also aperiodic

for i, $\{n, P_{ii}^{(n)} > 0\}$

$\gcd\{n_1, n_2, \cdots\} = 1$

for j, $\{\tilde{n}, P_{jj}^{(\tilde{n})} > 0\}$

$\gcd\{\tilde{n}_1, \tilde{n}_2, \cdots\} = 1$

$\{\bar{n}, \bar{P}_{(i,j)(i,j)}^{(\bar{n})} > 0\}$ $\quad \gcd\{\bar{n}_1, \bar{n}_2, \cdots\} = 1$

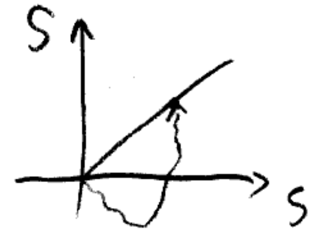
Step 2: The stationary distribution of $\bar{P}$ is $\pi \times \pi$.

Step 3: Since $\bar{P}$ is irreducible, and (admits a stationary dist.) by HW 5.1.(3), we know that all states are recurrent. (in $S \times S$)

Let $\tau = \inf\{n \geq 0, X_n = Y_n\}$

Then, $P(\tau < \infty) = 1$

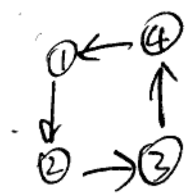
coupled process hit the diagonal.



After that, their behaviors are going to be the same.

* Coupled chain should be irreducible.

* periodicity may not give us irreducibility.



Step 4: We know that if we start X_n and Y_n after they first meet, then the distribution of X_n and Y_n shall be the same.

Therefore, $P(X_n = j, \tau \leq n) = \sum_{m=0}^{n} \sum_{i \in S} P(X_n = j, \tau = m, Z_m = (i,i))$

$$= \sum_{m=0}^{n} \sum_{i \in S} P(\tau = m, Z_m = (i,i)) \, P(X_n = j \mid \tau = m, Z_m = (i,i))$$

Since $\{\tau = m, Z_m = (i,i)\} = \{Z_0, \cdots, Z_{m-1} \notin \text{diagonal}, Z_m = (i,i)\}$

Then $P(X_n=j \mid \tau=m, Z_m=(i,i)) = P(X_n=j \mid Z_m=(i,i))$

$= P(X_n=j \mid X_m=i, Y_m=i) = P(X_n=j \mid X_m=i) = p_{ij}^{(n-m)}$

Therefore. $P(X_n=j, \tau\le n) = \sum_{m=0}^{n}\sum_{i\in S} P(\tau=m, Z_m=(i,i))\, p_{ij}^{(n-m)}$

Similarly, $P(Y_n=j, \tau\le n) = \sum_{m=0}^{n}\sum_{i\in S} P(\tau=m, Z_m=(i,i))\, p_{ij}^{(n-m)}$

Therefore, $P(X_n=j, \tau\le n) = P(Y_n=j, \tau\le n)$

Thus, $\sum_{j\in S}\Big|\underbrace{P(X_n=j)}_{=p_{ij}^{(n)}} - \underbrace{P(Y_n=j)}_{=\pi_j}\Big|$

$= \sum_{j\in S}|P(X_n=j, \tau>n) - P(Y_n=j, \tau>n)|$

$\le \sum_{j\in S} P(X_n=j, \tau>n) + \sum_{j\in S} P(Y_n=j, \tau>n)$ by triangle inequality

$= 2P(\tau>n) \to 0$ as $n\to\infty$

(aperiodicity)

Let $\mu=\delta_i$. $\nu=\pi$.

We get $\lim_{n\to\infty}\sum_{j\in S}|p_{ij}^{(n)}-\pi_j| = 0$. Done!

<u>Summary</u>: P is irreducible and aperiodic and stationary distribution π.

Then $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$

We know that $\pi_j = \frac{1}{E_j\sigma_j}$ by Ergodic Theorem.

Thus. if j is null-recurrent $(E_j\sigma_j=\infty)$

then $\pi_j=0$.

(positive recurrency $\Rightarrow$ stationary distribution)

Theorem: Suppose P is irreducible and null-recurrent

Then $\forall i, j \in S$ $\lim_{n\to\infty} P_{ij}^{(n)} = 0$

Classification of Markov Chains

irreducible
- recurrent
  - positive recurrent: There exist stationary distribution $\boxed{\pi_j = \frac{1}{E_j \sigma_j}}$
    - aperiodic $\boxed{\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j}$
    - periodic $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} P_{ij}^{(k)} = \pi_j$
  - null recurrent: There does not exist stationary distribution and $\lim_{n\to\infty} P_{ij}^{(n)} = 0$
- transient: There doesnot exist a stationary distribution, $\lim_{n\to\infty} P_{ij}^{(n)} = 0$

(Since $G_{ii} = \sum_{i=0}^{\infty} P_{ii}^{(n)} < \infty \Rightarrow P_{ii}^{(n)} \to 0$

$\Rightarrow P_{ij}^{(n)} \to 0$

(HW5 Problem 1(1))

10/15

In our coupling construction $Z_n = (X_n, Y_n)$

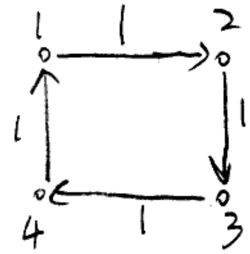
$P_{(i_1,i_2)} P_{(j_1,j_2)} = P_{i_1 j_1} P_{i_2 j_2}$

If P is aperiodic. Then $Z_n = (X_n, Y_n)$ is irreducible.

This is because we apply the conclusion in lecture 10.

$\exists\, n_0 \geq 0 \ \forall\, n \geq n_0$, $P_{i_1 j_1}^{(n)} > 0$ $\quad P_{i_2 j_2}^{(n)} > 0 \Rightarrow P^{(n)}_{(i_1,i_2)} P_{(j_1,j_2)} > 0$

If P is periodic. Then $Z_n = (X_n, Y_n)$ can be reducible.

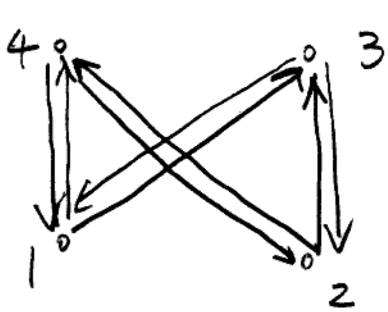


$P^{(n)}_{(1,2)(1,3)} = 0$

Ex: Consider a transition matrix.

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0.2 & 0.8 \\ 0.25 & 0.75 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix} \end{array}$$

Calculate $\lim\limits_{n \to \infty} P^{2n}$



Let X_n be a Markov chain with transition matrix P. P irreducible.

$P_{11}^{(2)} > 0 \quad P_{11}^{(3)} = 0 \quad P_{11}^{(4)} > 0$

$P_{11}^{(5)} = 0 \ \cdots$ period $= 2$

$$P^2 = \begin{pmatrix} 0.35 & 0.65 & 0 & 0 \\ 0.45 & 0.55 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 \\ 0 & 0 & 0.4 & 0.6 \end{pmatrix}$$

Let $A = \begin{pmatrix} 0.35 & 0.65 \\ 0.45 & 0.55 \end{pmatrix}$

$B = \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix}$

Thus. $P^{2n} = \begin{pmatrix} A^n & 0 \\ 0 & B^n \end{pmatrix}$

Both matrices A and B are irreducible & aperiodic.

Stationary distribution

$(\alpha, 1-\alpha) A = (\alpha, 1-\alpha)$ $\qquad \alpha = \frac{9}{22}$

$(\beta, 1-\beta) B = (\beta, 1-\beta)$ $\qquad \beta = \frac{4}{11}$

$$\lim_{n\to\infty} P^{2n} = \begin{pmatrix} \lim\limits_{n\to\infty} A^n & 0 \\ 0 & \lim\limits_{n\to\infty} B^n \end{pmatrix}$$

$A^n = \begin{pmatrix} P_{11}^{(n)} & P_{12}^{(n)} \\ P_{21}^{(n)} & P_{22}^{(n)} \end{pmatrix}$ By Theorem, $\lim\limits_{n\to\infty} P_{ij}^{(n)} = \pi_j$

$$\lim_{n\to\infty} P^{2n} = \begin{pmatrix} \frac{9}{22} & \frac{13}{22} & 0 & 0 \\ \frac{9}{22} & \frac{13}{22} & 0 & 0 \\ 0 & 0 & \frac{4}{11} & \frac{7}{11} \\ 0 & 0 & \frac{4}{11} & \frac{7}{11} \end{pmatrix}$$

$\lim\limits_{n\to\infty} P_{11}^{(n)} = \pi_1 = \alpha$.

$\lim\limits_{n\to\infty} P_{12}^{(n)} = \pi_2 = 1-\alpha$

$\lim\limits_{n\to\infty} P_{21}^{(n)} = \beta$

$\lim\limits_{n\to\infty} P_{22}^{(n)} = 1-\beta$.

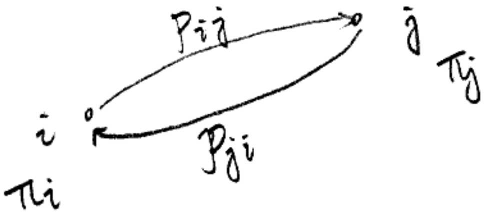
Let X_n be a Markov Chain with transition matrix P.

Let $X_0 \sim \mu = (\mu_1, \ \ , \mu_m)$ s.t. $P(X_0 = i) = \mu_i \quad 1 \le i \le m$

$$P(X_1 = j) = \sum_{i \in S} P(X_1 = j \mid X_0 = i) P(X_0 = i) = \sum_{i \in S} \mu_i P_{ij} = \mu P$$

μ is stationary iff $\mu P = \mu$

<u>Def'n</u>: The distribution $\pi = (\pi_i)_{i \in S}$ is called <u>detailed balance</u> iff $\forall i, j \in S$, $\pi_i P_{ij} = \pi_j P_{ji}$



<u>Remark</u>: (i) Detailed balance does not always exist.

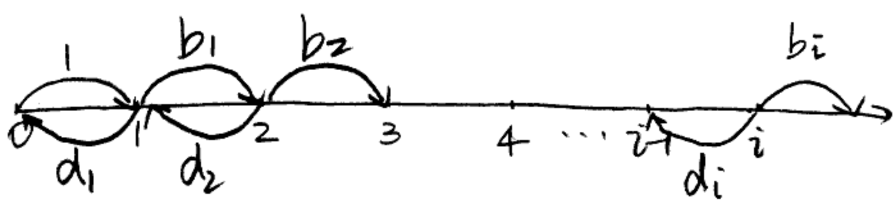
(ii) If Detailed balance exists,

Stationary requires $\pi_i = \sum_{j \in S} \pi_j P_{ji}$

But $\sum_{j \in S} \pi_j P_{ji} = \sum_{j \in S} \pi_i P_{ij} = \pi_i \sum_{j \in S} P_{ij} = \pi_i$

Then Detailed balance $\Rightarrow$ Stationary.

<u>Ex</u>: <u>Birth-Death Chain</u>



$P_{01} = 1$, $P_{i,i+1} = b_i$, $P_{i,i-1} = d_i$

$d_i + b_i = 1$, $i \geq 1$.

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & \cdots \\ d_1 & 0 & b_1 & \\ & d_2 & 0 & b_2 & \ddots \\ & & \ddots & 0 & \ddots \\ & & & \ddots & \ddots \end{pmatrix} \end{array}$$

Suppose $(\pi_i)_{i=0,1,2,\cdots}$ is detailed balance

$\pi_i P_{i,i+1} = \pi_{i+1} P_{i+1,i}$ so $\pi_i b_i = \pi_{i+1} d_{i+1} \Rightarrow \pi_{i+1} = \frac{b_i}{d_{i+1}} \pi_i$

In particular, $\pi_0 = \pi_1 d_1$, $b_0 = 1$

$$\pi_i = \frac{b_{i-1}}{d_i} \pi_{i-1} = \frac{b_{i-1} b_{i-2}}{d_i d_{i-1}} \pi_{i-2} = \cdots = \frac{b_{i-1} b_{i-2} \cdots b_1}{d_i d_{i-1} \cdots d_1} \pi_0, \ i \geq 1$$

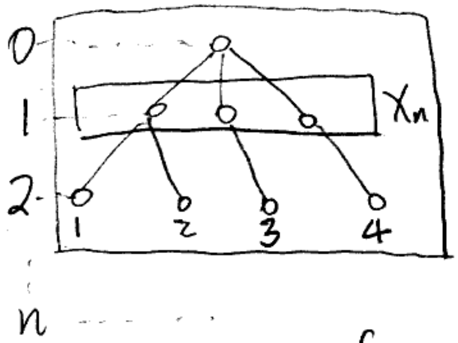
$$\sum_{i=0}^{\infty} \pi_i = 1 \iff \pi_0 \left(1 + \sum_{i=1}^{\infty} \frac{b_{i-1} \cdots b_0}{d_i \cdots d_1}\right) = 1.$$

therefore, if $\sum_{i=1}^{\infty} \frac{b_{i-1} \cdots b_0}{d_i \cdots d_1} < \infty$, then detailed balance exists. $\Rightarrow$ stationary distribution exists and equals to detailed balance.

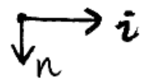
(stationary is a global balance)

10/27

Branching Process



population



X_n = # of people in the population

X_n Galton-Watson process (GW process)

n — generation

i — labeling of people in a generation

$\{\xi_{ni} : n \geq 0, i \geq 1\}$ i.i.d. non-negative integer valued random variables

ξ_{ni} — # of children in nth generation

denote in general $\xi_{ni} \stackrel{d}{=} \xi$

$P(\xi = k) = P_k$. children (offspring) distribution $(k = 0, 1, 2, \cdots)$

Branching process

$\{X_n : n \geq 0\}$ $X_0 = 1$ absorbing state $\boxed{X_{n+1} = \sum_{i=1}^{X_n} \xi_{ni}}$

$\begin{cases} \text{extinction} & X_n = 0 \\ \text{blow up} & \end{cases}$

$\tau_0 = \inf\{n \geq 0, X_n = 0\}$ extinction time

If $\tau_0 < \infty$, we say process extincts.

Otherwise, we say process survive.

$\boxed{P_1(\tau_0 < \infty) = \rho}$ extinction probability.

Remark: 0 is absorbing state

All other states are transient.

Generating Function

X - random variable

$f_X(s) = Es^X = Ee^{(\log s)X} = \psi_X(\log s)$

$Es^{\xi} = f(s)$

$X_0 = 1, \quad X_1 = \xi$

$f_{X_1}(s) = f(s)$

$f_{X_{n+1}}(s) = Es^{X_{n+1}} = Es^{\sum_{i=1}^{X_n} \xi_{ni}} = \sum_{k=0}^{\infty} P(X_n = k) E(s^{\sum_{i=1}^{k} \xi_{ni}} \mid X_n = k)$

Let us notice that X_n is independent of $\{\xi_{mi}, m \geq n, i \geq 1\}$

Therefore $f_{X_{n+1}}(s) = \sum_{k=0}^{\infty} P(X_n = k) (\underbrace{Es^{\xi}}_{\tilde{s}})^k$ (ξ is i.i.d)

$= E(Es^{\xi})^{X_n} = f_{X_n}(f(s))$ offspring ξ.

generating function of X_n

$f_{X_n}(s) = \underbrace{f(f(f \cdots (f(s) \cdots))}_{n \text{ iterations}}$

Theorem: (1) Let $m = E\xi = \sum_{k=0}^{\infty} k p_k$. Then $EX_n = m^n$.

(2) If $P_1 \neq 1$, then ρ is the minimum non-zero solution of $s = f(s)$ and $\rho < 1 \Leftrightarrow m > 1$

(ρ is the extinction probability)

fixed point equation

解: (1) $EX_{n+1} = E\sum_{j=1}^{X_n}\xi_{nj} = \sum_{k=0}^{\infty}P(X_n=k)\,E\left(\sum_{j=1}^{k}\xi_{nj}\,\middle|\,X_n=k\right)$

$= \underbrace{\sum_{k=0}^{\infty}P(X_n=k)k}\,E\xi$

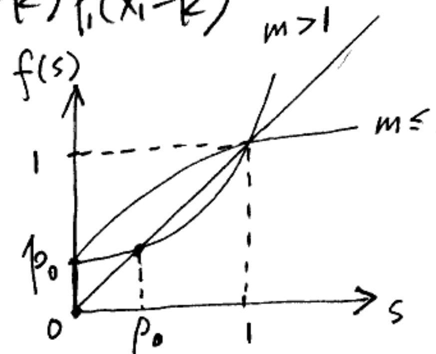
$= mEX_n$

$\Rightarrow EX_n = m^n$

(2) $\rho = \sum_{k=0}^{\infty}P_1(\tau_0<\infty, X_1=k) = \sum_{k=0}^{\infty}P_1(\tau_0<\infty\,|\,X_1=k)P_1(X_1=k)$

$= p_0 + \sum_{k=1}^{\infty}p_k P_k(\tau_0<\infty)$

$= p_0 + \sum_{k=1}^{\infty}p_k\left(P_1(\tau_0<\infty)\right)^k = f(\rho)$



Suppose $p_1 \neq 1$, $m \le 1$.

Then $f(s) = s$ has only one solution $s=1$ on $[0,1]$

$f(s) = Es^{\xi}$

$f(s) = p_0 + \sum_{k=1}^{\infty}p_k s^k = p_0 + p_1 s + p_2 s^2 + \cdots$ power series

$P(\xi=k) = p_k \quad k=0,1,2,\ldots$

$f'(1) = m$

$m>1$. $\rho = f(\rho)$ two solutions $\rho=\rho_0$ and $\rho=1$.

$\rho_0 = \rho = P_1(\tau_0<\infty) = \lim_{n\to\infty}P_1(\tau_0\le\infty) = \lim_{n\to\infty}P_1(X_n=0)$

$= \lim_{n\to\infty}f_{X_n}(0) = \lim_{n\to\infty}f^{(n)}(0) \le \rho_0$

$m>1$ supercritical — survive, extinction probability ρ_0 is the minimum solution of $f(\rho)=\rho$

$m=1$ ($p_1\neq1$) critical
$m<1$ subcritical $\Big\}$ extincts with probability 1.

Example: $P_k = p(1-p)^k$ Want: ρ

$$m = E\xi = \sum_k k \cdot P_k = \sum_k k\, p(1-p)^k = \frac{1-p}{p}$$

$\uparrow$ average # of offsprings

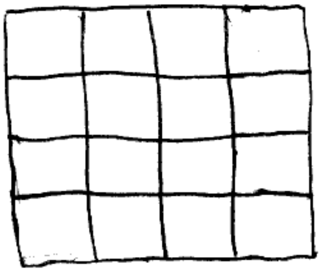
$p \geq \frac{1}{2} \Rightarrow m \leq 1$ extinct.

$p < \frac{1}{2} \Rightarrow m > 1$, $\rho = f(\rho) = \sum_k p(1-p)^k \rho^k = \frac{p}{1-\rho(1-p)}$

$$\rho = \frac{p}{1-p}$$

10/29

Hw.2



Et.

X_n – position of the rook

Y_n – position of the bishop

X_n is a random walk on $G_1 = (V_1, E_1)$

$|V_1| = 16$ each vertex of G_1 has out degree 6.

$\pi_i P_{ij} = \pi_j P_{ji} \Rightarrow \pi_i = \frac{1}{16}$

$\pi_i^X = 1/16$

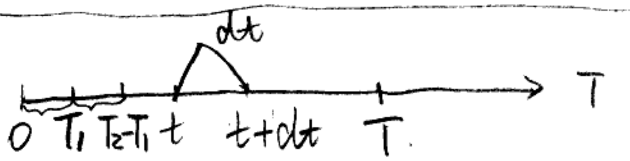
$|V_2| = 8$ 6 vertices with out degree 3.

2 vertices with out degree 5.

$|E_2| = 28$

$\pi_i P_{ij} = \pi_j P_{ji}$

$$\pi_i = \begin{cases} \frac{3}{28} & \text{if out degree of } i \text{ is } 3 \\ \frac{5}{28} & \text{if out degree of } i \text{ is } 5 \end{cases}$$



λ – rate, number of unit-time arrivals.

$P(\text{There will be a customer in } [t, t+dt]) = \lambda dt + o(dt)$

$P(\text{There will be no customer in } [t, t+dt]) = 1 - \lambda dt + o(dt)$

$P(\text{There will be} \geq 2 \text{ customers in } [t, t+dt]) = o(dt)$

Goal: • arrival time / waiting time $T_1, T_2 - T_1, \cdots$

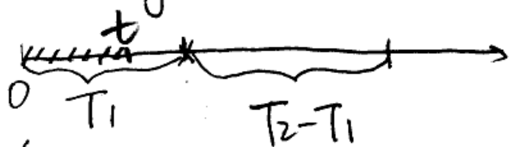
• # of customers in a fixed amount of time $[0,t] = N_t$

— poisson time

Remark: if I_1, I_2 are two time intervals so that $I_1 \cap I_2 = \phi$

then customer arrival in I_1 and I_2 are independent of each other.

<u>Waiting time</u> $T_1, T_2-T_1, T_3-T_2, \cdots$ are independent (i.i.d)

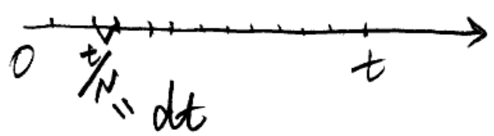


homogeneous

(let the common distribution be $\xi \overset{d}{=} T_1$)

Suppose $T_1, T_2-T_1, \cdots$ share the same distribution with a random variable ξ

$P(\xi > t) = P$(no customer arrival during time $[0,t]$)

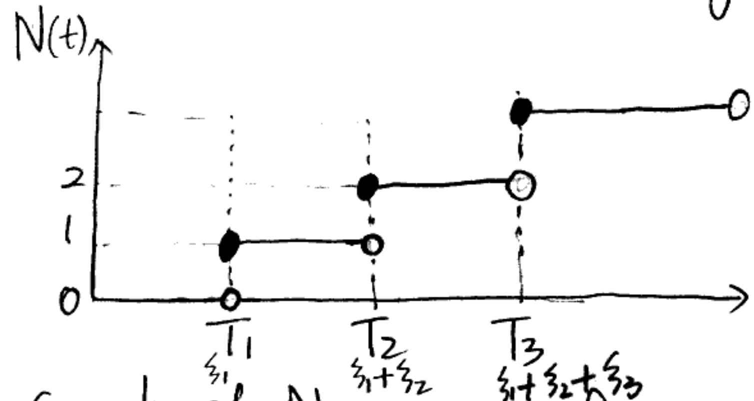


$= \lim_{N\to\infty} P$(no customer arrived during time $[0, \frac{t}{N}]$)^N

$= \lim_{N\to\infty} (1 - \frac{\lambda t}{N})^N = e^{-\lambda t}$

<u>Conclusion</u>: $\xi \sim \exp(\lambda)$ $E\xi = \frac{1}{\lambda}$ waiting time

N_t = Total # of customers during time $[0,t]$



jump process

states: $\{0, 1, 2, 3, \cdots\}$

wait for an exponential time

Graph of N_t as a function of t is continuous from the right having limit from the left.

N_t is a Markov process.

$$\{N_t = k\} = \{\xi_1 + \dots + \xi_k \le t < \xi_1 + \dots + \xi_k + \xi_{k+1}\}$$

(Space - Time Duality)

Let $S_k = \xi_1 + \dots + \xi_k$ Then $\{N_t = k\} = \{S_k \le t < S_{k+1}\}$

$$\mathbb{P}(N_t = k) = \mathbb{P}(\xi_1 + \dots + \xi_k \le t < \xi_1 + \dots + \xi_k + \xi_{k+1})$$

$$= \underset{\substack{t_1 > 0 \dots t_k > 0 \ t_{k+1} > 0 \\ t_1 + \dots + t_k < t \le t_1 + \dots + t_k + t_{k+1}}}{\int \cdots \int} \lambda^{k+1} e^{-\lambda(t_1 + \dots + t_k + t_{k+1})} dt_1 \dots dt_k \, dt_{k+1}$$

$$= \underset{\substack{t_1 > 0 \dots t_k > 0 \\ t_1 + \dots + t_k < t}}{\int dt_1 \cdots \int dt_k} \lambda^{k+1} e^{-\lambda(t_1 + \dots + t_k)} \int_{t-(t_1 + \dots + t_k)}^{\infty} e^{-\lambda t_{k+1}} dt_{k+1}$$

$$= \underset{\substack{t_1 > 0 \dots t_k > 0 \\ t_1 + \dots + t_k < t}}{\int dt_1 \cdots \int dt_k} \lambda^{k+1} e^{-\lambda(t_1 + \dots + t_k)} \frac{1}{\lambda} e^{-\lambda(t-(t_1 + \dots + t_k))}$$

$$= \lambda^k e^{-\lambda t} \int\limits_{\substack{t_1 > 0 \; \cdots \; t_k > 0 \\ t_1 + \cdots + t_k \le t}} dt_1 \cdots \int dt_k$$

$$= (\lambda t)^k e^{-\lambda t} \int\limits_{\substack{\tilde{t}_1 > 0 \; \cdots \; \tilde{t}_k > 0 \\ \tilde{t}_1 + \cdots + \tilde{t}_k \le 1}} d\tilde{t}_1 \cdots \int d\tilde{t}_k = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

<u>Conclusion</u> $N_t \sim \text{Poisson}(\lambda t)$

$N_t = \#$ of arrivals in $[0, t]$

<u>Definition</u> Let $\xi_1, \xi_2, \ldots$ be i.i.d. $\exp(\lambda)$

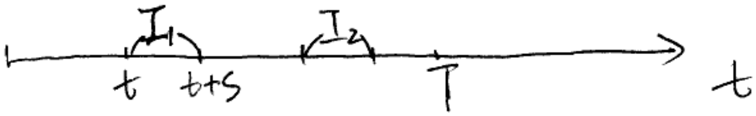
$S_n = \xi_1 + \cdots + \xi_n$

Let $\{N_t = k\} = \{S_k \le t < S_{k+1}\}$

So that $N_t = \max\{n : S_n \le t\}$

Then $\{N_t\}_{t \ge 0}$ is called a <u>Poisson process with rate / parameter</u> λ and we write $N_t \sim PP(\lambda)$

11/3



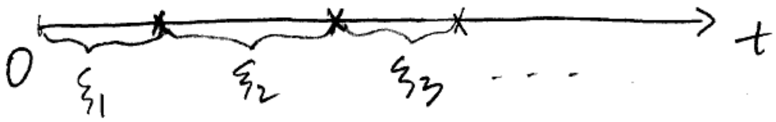
$I_1 \cap I_2 = \phi$

$\Rightarrow N_t \in I_1 \perp\!\!\!\perp N_t \in I_2$

$P(N_{t+s} - N_t = 1) = \lambda s + o(s)$

$P(N_{t+s} - N_t = 0) = 1 - \lambda s + o(s)$

$P(N_{t+s} - N_t \geq 2) = o(s)$

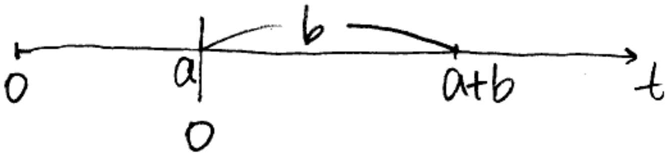


Waiting time $\sim \exp(\lambda)$

$\xi_1 = T_1,\ \xi_2 = T_2 - T_1,\ \cdots,\ \xi_n = T_n - T_{n-1}$

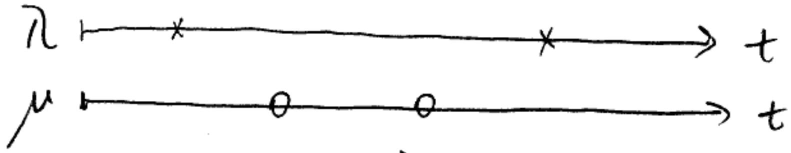
Basic Properties of $\xi \sim Exp(\lambda)$

1. Memoriless : $P(\xi - a > b \mid \xi > a) = P(\xi > b)$



2. Exponential Race :

$\xi \sim \exp(\lambda)$ $\quad \eta \sim \exp(\mu)$ $\quad \xi \perp\!\!\!\perp \eta$



$P(\xi < \eta) = \dfrac{\lambda}{\lambda+\mu}$

$P(\xi > \eta) = \dfrac{\mu}{\lambda+\mu}$

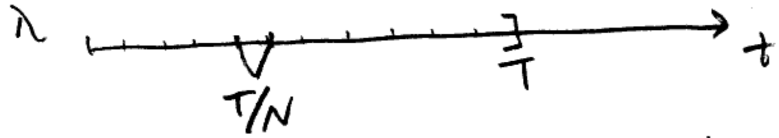
and $\xi \wedge \eta \sim \exp(\lambda+\mu)$

waiting time : either a man or a woman

$N_t = \#$ of customer arrivals in $(0,t]$.

$$P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \cdots$$

$$\{N_t = k\} = \{\xi_1 + \xi_2 + \cdots + \xi_k \le t < \xi_1 + \cdots + \xi_k + \xi_{k+1}\}$$



We divide $(0,T]$ into N intervals $I_i = (\frac{T}{N}(i-1), \frac{T}{N} i]$, $i = 1, \cdots, N$.

$$P(\text{there will be a customer in } I_i) \approx \lambda \frac{T}{N} + o(\tfrac{1}{N})$$

$N_T = X_1 + \cdots + X_N \qquad X_i \sim \text{Bernoulli}(\frac{\lambda T}{N})$

as N large, $N_t \sim B(N, \frac{\lambda T}{N})$

<u>Conclusion</u>: $PP(\lambda) = \lim_{N \to \infty} B(N, \frac{\lambda}{N})$

<u>Theorem</u>: Let $X_{n,m}$, $1 \le m \le n$, be independent r.v.'s with $P(X_{n,m} = 1) = p_m$ $\quad P(X_{n,m} = 0) = 1 - p_m$. Let $S_n = X_1 + \cdots + X_n$, $\lambda_n = ES_n = p_1 + \cdots + p_n$ and $Z_n \sim \text{Poisson}(\lambda_n)$

Then for any set A,

$$|P(S_n \in A) - P(Z_n \in A)| \le \sum_{m=1}^{n} p_m^2$$

<u>Proposition</u>: Suppose $t, s \geq 0$, n, m are non-negative integers, Then $P(N_t = k, N_{t+s} - N_t = l) = P(N_t = k) P(N_{t+s} - N_t = l)$

Moreover, $N_{t+s} - N_t$ has the same distribution as N_s. (Stationary increment)

independent increment

Summary: A stochastic process $\{N_t\}_{t \geq 0}$ is a $PP(\lambda)$ iff

(1) $N_0 = 0$, N_t takes integer values.

(2) Trajectory N_t is monotonically non-decreasing. continuous from right, having limit from left.

(3) N_t has independent increments.

(4) As $s \to 0+$, $P(N_{t+s} - N_t = 1) = \lambda s + o(s)$,

$$P(N_{t+s} - N_t = 0) = 1 - \lambda s + o(s)$$

(time homogeneity with rate λ)

Proposition: Suppose $0 \leq t_1 < \cdots < t_k < t < t+s$.

$$0 \leq n_1 \leq \cdots \leq n_k < n < n+m$$

Then $P(X_{t+s} = n+m \mid X_t = n, X_{t_1} = n_1, \ldots, X_{t_k} = n_k)$

$$= P(X_{t+s} = n+m \mid X_t = n) = P(X_s = m)$$

(markov property)

$\Uparrow$

Stationary independent increment.

Conditioned on $N_t = 1$. Let the arrival time of this customer be $S_1 \in [0, t]$,

$$P(S_1 \in [s, s+\Delta s] \mid N_t = 1) = \frac{P(S_1 \in [s, s+\Delta s], N_t = 1)}{P(N_t = 1)}$$

$$= \frac{P(\xi_1 \in [s, s+\Delta s], \xi_2 > t - \xi_1)}{\lambda t e^{-\lambda t}} = \frac{\int_s^{s+\Delta s} dt_1 \int_{t-t_1}^{\infty} dt_2 \, \lambda^2 e^{-\lambda(t_1+t_2)}}{\lambda t e^{-\lambda t}}$$

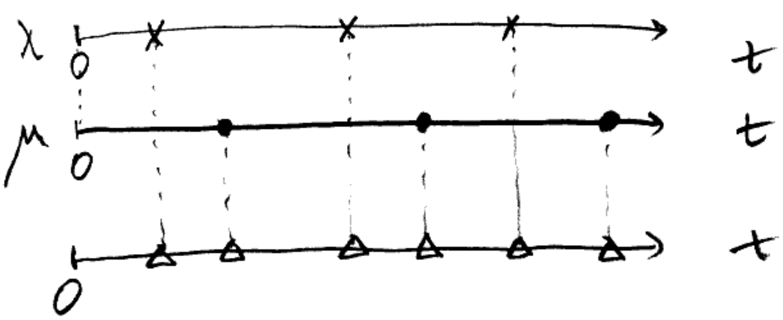
$$= \left[\int_s^{s+\Delta s} dt_1 \, \lambda^2 e^{-\lambda t_1} \frac{1}{\lambda} e^{-\lambda(t-t_1)}\right] / \left[\lambda t e^{-\lambda t}\right] = \frac{\Delta s \cdot \lambda e^{-\lambda t}}{\lambda t e^{-\lambda t}} = \frac{\Delta s}{t}$$

Therefore, $S_1 \mid N_t = 1 \sim U([0,t])$

Order statistic: $X_1, \dots, X_n$ i.i.d.

$Y_1 \le \dots \le Y_n$ so that $\{Y_1, \dots, Y_n\} = \{X_1, \dots, X_n\}$

Proposition: Conditioned on $N_t = k$, the arrival times for customers are $S_1, \dots, S_k$. is the order statistic of k independent r.v.'s uniformly distributed on $[0, t]$



Proposition: (Superposition). Let N_t and M_t be two independent poisson processes. N_t has rate λ and M_t has rate μ. Then $N_t + M_t \sim PP(\lambda + \mu)$.

$\xi \sim \exp(\lambda)$, $\eta \sim \exp(\mu)$; $\xi \wedge \eta \sim \exp(\lambda + \mu)$

regardless of gender.

Proposition: (Thinning): Let $\{N_t\}_{t \geq 0} \sim PP(\lambda)$,

$\varepsilon_1, \dots, \varepsilon_n$ i.i.d, independent of N_t.

$P(\varepsilon_1 = 1) = 1 - P(\varepsilon_1 = 0) = p$.

Then $\{M_1(t) = \sum_{n=1}^{X_t} 1\{\varepsilon_n = 1\}, t \geq 0\} \sim PP(\lambda p)$

$\{M_2(t) = \sum_{n=1}^{X_t} 1\{\varepsilon_n = 0\}, t \geq 0\} \sim PP(\lambda(1-p))$

are independent.

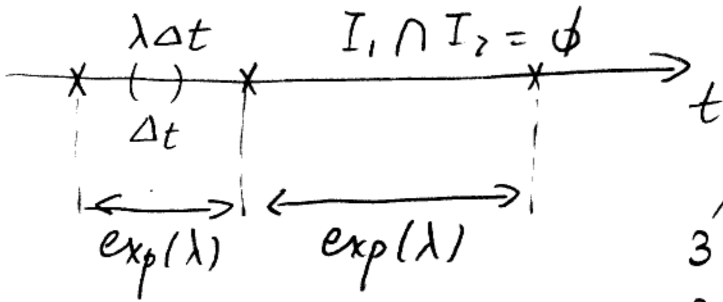
Ex: Given a Poisson process with red arrivals with rate λ and independent poisson process with green arrivals with rate μ.

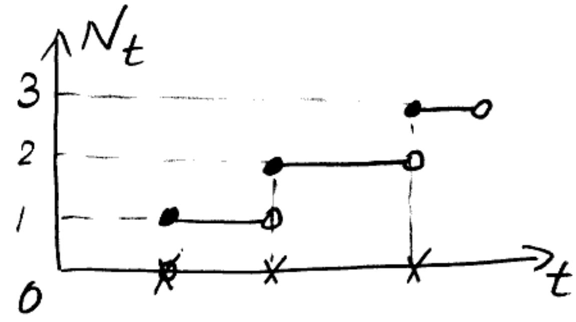
What's probability we see 6 red arrivals before a total of 4 green ones?

$$\sum_{k=6}^{9}\binom{9}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(\frac{\mu}{\lambda+\mu}\right)^{9-k}$$

at least 6 red in the first 9 arrivals.

Review





(1). Stationary & independent increment

$\uparrow$

arrival rate is always λ

(2) Markov property

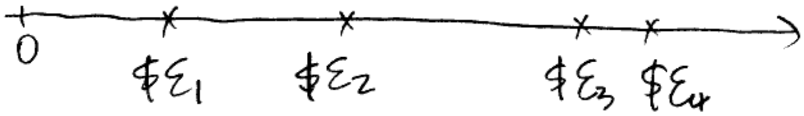
(3) Conditioning on $N_t = 1$, the arrival time $S_1 \sim U[0, t]$

(4) Thinning & Superposition

Compound Poisson Process

Let $\{N_t\}_{t \geq 0}$ be a Poisson process with rate λ. Let $\varepsilon_1\ \varepsilon_2 \dots$ i.i.d. independent of N_t

Let $S_t = \sum_{k=1}^{N_t} \varepsilon_k$ Then $\{S_t\}_{t \geq 0}$ is called a compound Poisson process.



$N_t \sim PP(\lambda)$ $\quad$ $\varepsilon_1, \varepsilon_2, \cdots$ i.i.d. $\perp\!\!\!\perp N_t$

$Y_t = \sum_{j=1}^{N_t} \varepsilon_j$ $\quad$ group sizes

$P(\varepsilon_i = k) = p_k$ $\quad$ $k = 1, 2, \cdots$

$Y_t = \sum_{k=1}^{\infty} k N_t^{(k)}$ $\quad$ $N_t^{(k)} = \sum_{j=1}^{N_t} 1_{\{\varepsilon_j = k\}} \sim PP(\lambda p_k)$ (Thinning)

groups with size k.

Ex: (Insurance) k-th type of claim is worth $\$a_k$

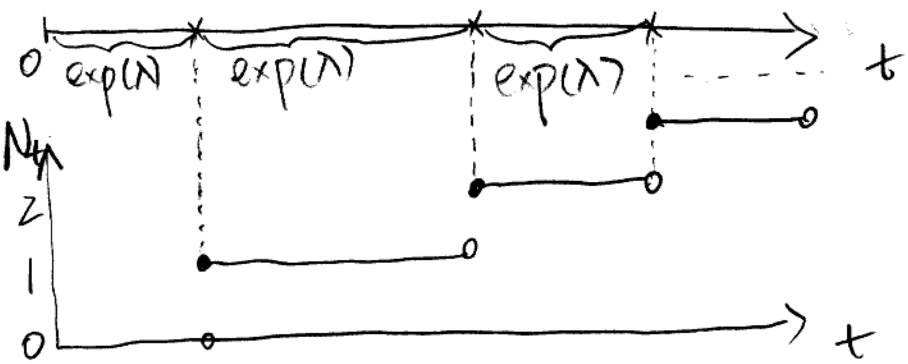
$Y_t = \sum_{k=0}^{\infty} a_k N_t^{(k)} = \sum_{j=1}^{N_t} a_{\varepsilon_j}$

M $\quad$ $I = \{k : a_k \geq M\}$ $\quad$ Let $Z_t = \sum_{k \in I} a_k N_t^{(k)} = \sum_{j=1}^{N_t} a_{\varepsilon_j} 1_{\varepsilon_j \in I}$

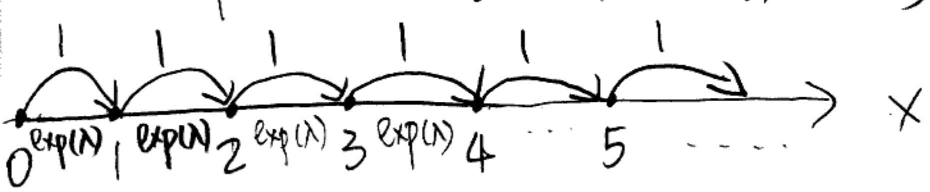
is again a compound poisson process

Let $N_t^I = \sum_{j=1}^{N_t} 1_{\varepsilon_j \in I} \sim PP(\lambda P(I))$

Thus. $Z_t = \sum_{l=1}^{N_t^I} a_{\varepsilon_l}$



State space : $N_t \in \{0, 1, 2, 3, \cdots\}$ $\quad$ $N_t \sim PP(\lambda)$



Continuous time Markov chain

We can imagine there is a clock at each site, ringing at $\exp(\lambda)$ time. ("exponential clock") and the process N_t jumps from i to $i+1$ after the clock rings.

In order to have independent increment (Markov property), waiting time has to be exponential.

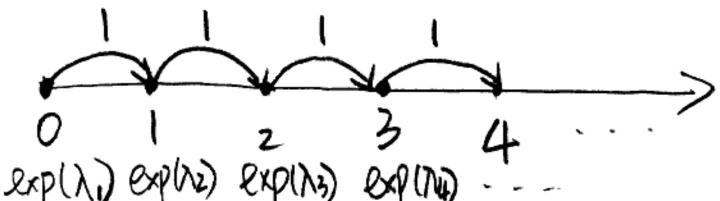
Suppose $\xi_1, \xi_2, \cdots$ are independent

$$\xi_k \sim \exp(\lambda_k).$$

$S_0 = 0, \quad S_n = \xi_1 + \xi_2 + \cdots + \xi_n$

Let $X_t = \sup\{n : S_n \le t\}$ people you can see before t.

Then X_t is called a (pure) <u>birth process</u>.



Regardless of the waiting time, the process is a Markov chain.

The λ_k's are called rate (speed) of the process.

(Ex:) $\xi \sim \exp(\lambda)$ $\quad \lim_{t\to 0} \frac{P(\xi < t)}{t} = \lambda.$

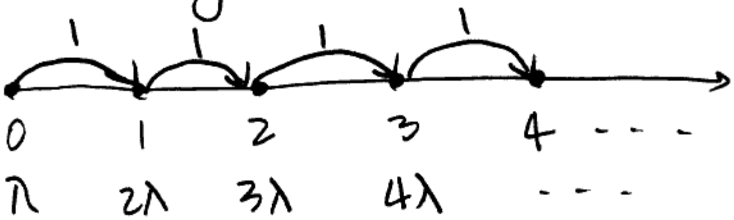
$\tau = \xi_1 + \xi_2 + \cdots$

We can imagine the bigger λ_k is, the smaller τ will be. If $\tau < \infty$, we say X_t blows up.

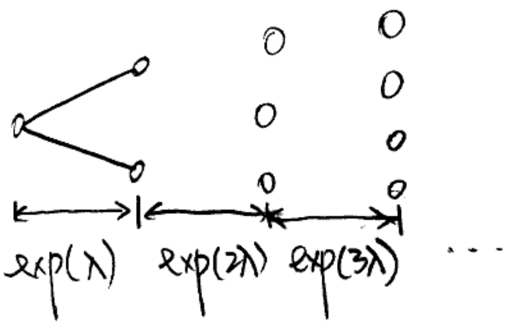
<u>Fact:</u> $\sum_{k=1}^{\infty} \frac{1}{\lambda_k} < \infty \iff P(\tau < \infty) = 1$

$\sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \infty \iff P(\tau < \infty) = 0$ "0-1" laws.

Specifically, $\lambda_k = k\lambda$. Then X_t is called <u>Yule process</u>.



$\xi \sim \exp(\lambda)$ $\quad$ $\eta \sim \exp(\mu)$ $\quad$ $\xi \wedge \eta \sim \exp(\lambda + \mu)$

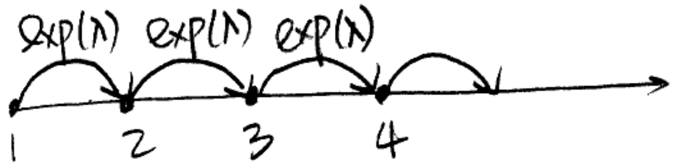


We know Yule process is not going to blow up

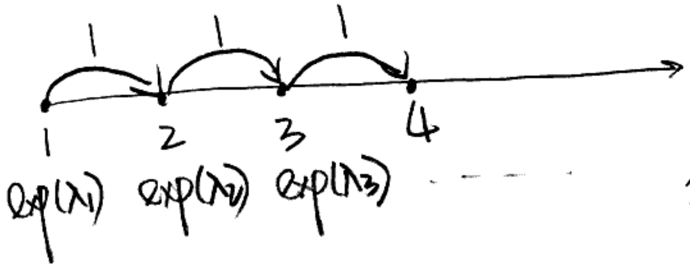
$\sum \frac{1}{k} < \infty$

<u>proposition</u>: Suppose $\{X_t : t \geq 0\}$ is a Yule process starting from $X_0 = 1$. Then $P(X_t = k \mid X_0 = 1) = (1 - e^{-\lambda t})^{k-1} e^{-\lambda t}$ for any $k \geq 1$

11/10



Chain of states

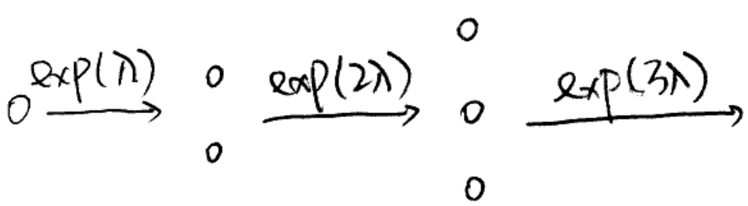


jump markov chain

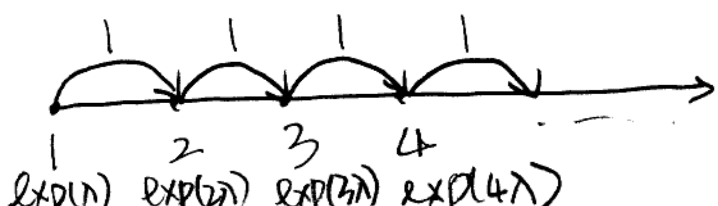
exponential clock

Continuous time Markov chain

$$\lim_{t \to 0} \frac{P(\xi < t)}{t} = \lambda \quad \text{speed of the process}$$



Yule process



Yule process: $\lambda_k = k\lambda$ not a blow-up process

Blow up: In finite time, process runs to infinity.

Proposition: Let $\{X_t\}_{t \geq 0}$ be a Yule process starting from 1.

Then $P_1(X_t = k) = (1 - e^{-\lambda t})^{k-1} e^{-\lambda t} \quad \forall k \geq 1$

transition probability for Yule process

$P_t(1,k) \qquad P_{1k}(t) \qquad p(1,k;t) \qquad p(t;1,k)$

Pf: Let $S_n = \xi_1 + \cdots + \xi_n$ where $\xi_k \sim \exp(k\lambda)$ independently.

Then $P_1(X_t = k) = P(S_{k-1} < t \leq S_k)$

$$= P(S_{k-1} \leq t) - P(S_k \leq t)$$

We claim $P(S_k \leq t) = (1-e^{-\lambda t})^k, \quad k \geq 1$

prove the claim by induction,

(1) $P(S_1 \leq t) = P(\xi_1 \leq t) = 1 - e^{-\lambda t}$.

(2) Suppose that $P(S_n \leq t) = (1-e^{-\lambda t})^n$.

Then since $\xi_{n+1} \sim \exp((n+1)\lambda)$,

Therefore, $P(S_{n+1} \leq t) = P(S_n + \xi_{n+1} \leq t)$

$$= \int_0^t (n+1)\lambda e^{-\lambda(n+1)x} P(S_n \leq t-x)\, dx \quad \text{independency}$$

$$= \int_0^t (n+1)\lambda e^{-(n+1)\lambda x} (1-e^{-\lambda(t-x)})^n\, dx$$

$$= e^{-(n+1)\lambda t} \int_0^t (n+1)\lambda e^{\lambda(t-x)} (e^{\lambda(t-x)} - 1)^n\, dx$$

$$= e^{-(n+1)\lambda t} (e^{\lambda(t-x)} - 1)^{n+1} \Big|_{x=t}^{x=0}$$

$$= (1-e^{-\lambda t})^{n+1}$$

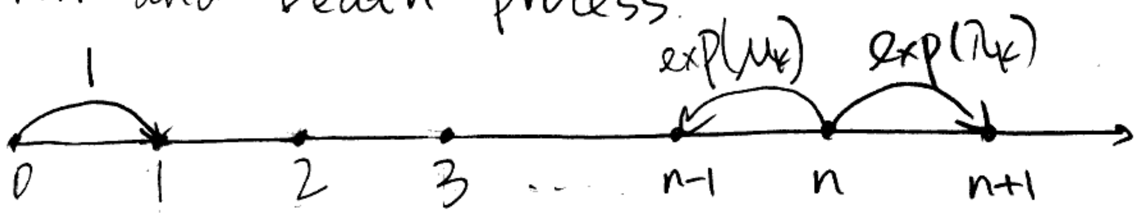
Thus, $P_1(X_1 = k) = (1-e^{-\lambda t})^{k-1} - (1-e^{-\lambda t})^k$

More generally we have $= e^{-\lambda t}(1-e^{-\lambda t})^{k-1}$

$$P_i(X_t = j) = P_{ij}(t) = \binom{j-1}{i-1} (e^{-\beta t})^i (1-e^{-\beta t})^{j-i} \quad (\text{Exercise!})$$

Negative Binomial Distribution

Birth and Death process.



Imagine at each site, there are two clocks.

Birth clock & Death clock

Birth clock at site $k \sim \exp(\lambda_k)$

Death clock at site $k \sim \exp(\mu_k)$

and they are independent.

whichever rings first, the process jumps from $k \to k+1$ or $k \to k-1$ respectively.

Such a process X_t on $\{0,1,2,\cdots\}$ is a birth-death process.

<u>By thinning</u>, we can think of the following equivalent picture, at each site k, there is a clock & a coin

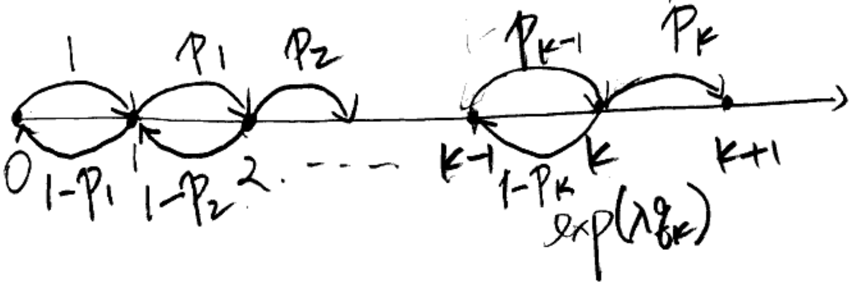
clock $\sim \exp(\lambda_k + \mu_k)$ Denote $q_k = \lambda_k + \mu_k$

Coin is a Bernoulli $\left(\frac{\lambda_k}{\lambda_k+\mu_k}\right)$

Coin $=1 \Rightarrow$ when clock rings, we have birth.

Coin $=0 \Rightarrow$ when clock rings, we have death.

Denote $p_k = \dfrac{\lambda_k}{\lambda_k+\mu_k}$



This gives us a continuous time birth and death chain.

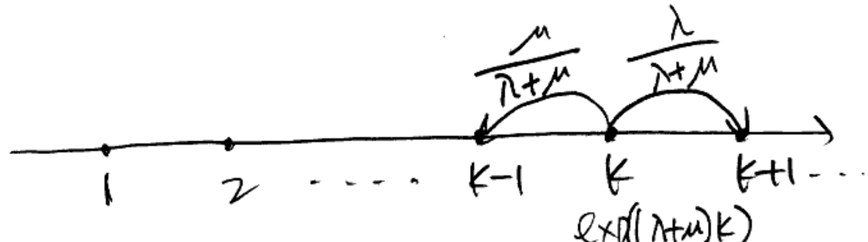
We denote the birth & death chain Y_t as follows.

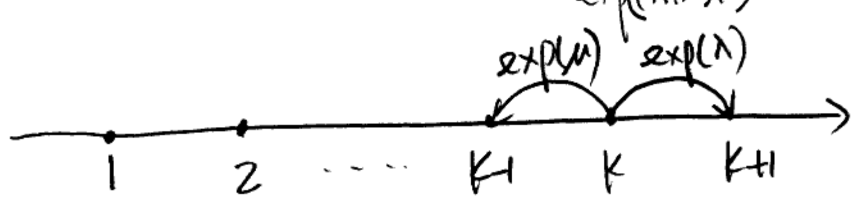
$P(Y_{n+1}=k+1 \mid Y_n=k) = p_k$, $P(Y_{n+1}=k-1 \mid Y_n=k) = 1-p_k$.

The Markov chain $\{Y_n\}_{n\geq 1}$ is called the embedding chain / skeleton chain of the birth & death process.

<u>Ex:</u> Yule process.

Suppose in the population, each individual is going to produce an offspring with rate λ, and it is going to die with rate μ. Therefore, the total population X_t is a Birth & Death process.





We can also think of X_t as a branching process with offspring distribution $P(\xi=0) = \frac{\mu}{\lambda+\mu}$, $P(\xi=2) = \frac{\lambda}{\lambda+\mu}$ and continuous time.

Suppose $X_0 = 1$. Let $T = \inf\{t : X_t = 0\}$

$T < \infty \Leftrightarrow$ extinction.

Suppose $\lambda > \mu$. Let $f(k) = P(T<\infty \mid X_0 = \underset{\text{population size}}{k})$ extinction probability.

$f(k) \downarrow$ as $k \uparrow$ and $\lim_{k\to\infty} f(k) = 0$

By first trial method,

$$f(k) = \frac{\lambda}{\lambda+\mu} f(k+1) + \frac{\mu}{\lambda+\mu} f(k-1)$$

which gives $\frac{f(k)-f(k+1)}{f(k-1)-f(k)} = \frac{\mu}{\lambda}$

Thus $f(k) = \left(\frac{\mu}{\lambda}\right)^k = P(T<\infty \mid X_0 = k)$ (minimal solution)

(2) If $\lambda < \mu$. Then $P(T<\infty \mid X_0 = 1) = 1$

Since $E\xi$ = expected value of offspring

$$= 0 \cdot \frac{\mu}{\lambda+\mu} + 2 \cdot \frac{\lambda}{\lambda+\mu} < 1.$$

Then let $g(k) = E_k T$.

By first trial method,

$$g(n) = \frac{1}{n(\lambda+\mu)} + \frac{\lambda}{\lambda+\mu} g(n+1) + \frac{\mu}{\lambda+\mu} g(n-1)$$

Therefore

$$\lambda g(n+1) - \lambda g(n) + \frac{1}{n} = \mu g(n) - \mu g(n-1)$$

So $$g(n) - g(n-1) = \frac{\lambda}{\mu}\left[g(n+1) - g(n)\right] + \frac{1}{\mu}\cdot\frac{1}{n}$$

Thus since $g(0) = 0$ (boundary condition)

$$g(1) = \frac{\lambda}{\mu}\left[g(2) - g(1)\right] + \frac{1}{1}\cdot\frac{1}{\mu}$$

$$= \left(\frac{\lambda}{\mu}\right)^2\left[g(3) - g(2)\right] + \frac{1}{2}\frac{\lambda}{\mu^2} + \frac{1}{1}\frac{1}{\mu}$$

$$= \cdots$$

$$= \left(\frac{\lambda}{\mu}\right)^n\left[g(n+1) - g(n)\right] + \frac{1}{n}\left(\frac{\lambda}{\mu}\right)^{n-1}\frac{1}{\mu} + \cdots + \frac{1}{2}\left(\frac{\lambda}{\mu}\right)\frac{1}{\mu} + \frac{1}{1}\frac{1}{\mu}$$

Since $$\lim_{n\to\infty}\left[g(n+1) - g(n)\right] = 0$$

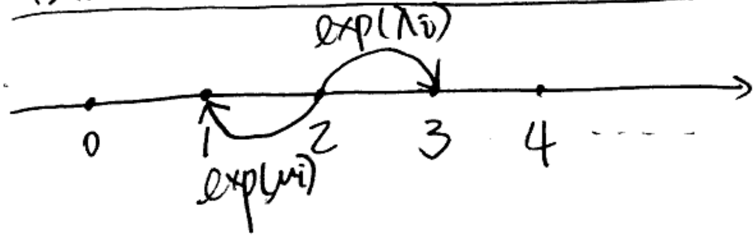
So $$\lambda g(1) = \frac{\lambda}{\mu} + \frac{1}{2}\left(\frac{\lambda}{\mu}\right)^2 + \cdots + \frac{1}{n}\left(\frac{\lambda}{\mu}\right)^n + \cdots$$

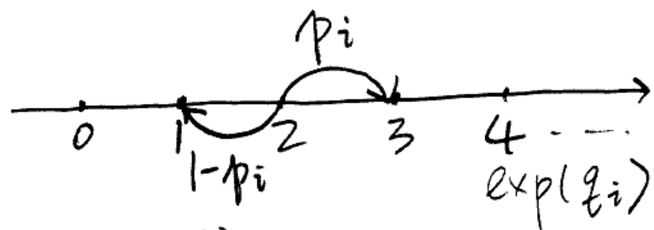
$$= \ln\frac{\mu}{\mu - \lambda}$$

Thus $\mathbb{E}_1 T = g(1) = \frac{1}{\lambda}\ln\frac{\mu}{\mu-\lambda}$. $\square$

11/12

Birth and Death Process.





$q_i = \lambda_i + \mu_i$

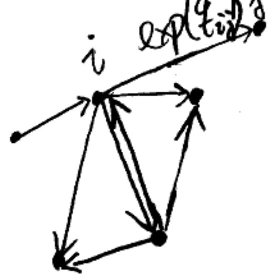
$p_i = \dfrac{\lambda_i}{\lambda_i + \mu_i}$ $\qquad 1 - p_i = \dfrac{\mu_i}{\lambda_i + \mu_i}$

---

1. Continuous time Markov Chain.
   (= Jump Markov chain = Q-process)

Model description: States $i \in S$ - statespace

Directed Graph.



At each site i, we have a family of clocks. $\sim \exp(q_{ij})$ whichever rings first, we jump from state i to state j.

At state j, there is another family of clocks $\sim \exp(q_{jk})$. independent of each other and of the clocks at site i. whichever rings first. jump $j \to k$.

By thinning, we can equivalently have the following model. at each site i, there is one clock $\sim \exp(q_i)$

$$q_i = \sum_{\substack{j \neq i \\ j \in S}} q_{ij}$$

We wait until the clock rings and then we jump from i to j with probability $p_{ij} = \frac{q_{ij}}{q_i} = \frac{q_{ij}}{\sum_{\substack{j \neq i \\ j \in S}} q_{ij}}$

That is, we roll a multiface dice with face j having probability p_{ij}.

Regardless of waiting time, the behavior of the process X_t described above is the same as a Markov chain Y_n with transition matrix $P = (p_{ij})_{i,j \in S}$

The Markov chain Y_n is called the <u>embedding chain</u>.

Or we can call it <u>skeleton chain</u>.

We append $p_{ii} = 0$ if $q_i > 0$.

$p_{ii} = 1$ if $q_i = 0$

i is an absorbing state

Let $\xi_j \sim \exp(q_j)$ independently.

$\tau = \xi_{Y_1} + \xi_{Y_2} + \cdots$ total time spent in the chain

If $P(\tau < \infty) = 1$, we say X_t blow up.

We will restrict ourselves to jump Markov Chains that do not blow up.

2. Markov Property

<u>Heuristically</u>. Markov property says that conditioned on knowing the current state, future is independent of the past. In the case of continuous time Markov chain, it means that the time we have to wait before we jump depends only on where we are, independent of how long we have been waiting. the only choice for such a waiting time is exponential distribution

Memoriless. $\xi \sim \exp(\lambda)$ $\quad P(\xi > a+b \mid \xi > b) = P(\xi > a)$

$P_i(X_t = j) = P(X_t = j \mid X_0 = i)$

We have $\{X_t = j\}$ embedding chain jumps to j.

$$= \bigcup_{n=0}^{\infty} \bigcup_{\substack{i_0, i_1, \cdots, i_{n-1} \in S \\ i_n = j}} \{Y_0 = i_0, \cdots, Y_n = i_n ; \xi_{Y_0} + \cdots + \xi_{Y_{n-1}} \le t < \xi_{Y_0} + \cdots + \xi_{Y_n}\}$$

$\xi_j \sim \exp(q_j)$ independent

$$P_i(X_t = j) = \sum_{n=0}^{\infty} \sum_{\substack{i_1 \cdots i_{n-1} \in S \\ i_n = j}} P\left(Y_0 = i_0, \ , Y_n = i_n ; \xi_{Y_0} + \cdots + \xi_{Y_{n-1}} \le t < \xi_{Y_0} + \cdots + \xi_{Y_n}\right)$$

$$= \sum_{n=0}^{\infty} \sum_{\substack{i_1 \cdots i_{n-1} \in S \\ i_n = j}} \prod_{k=0}^{n-1} P_{i_k i_{k+1}} \int \cdots \int_{t_k \geq 0,\ \sum_{k=1}^{n} t_k \le t < \sum_{k=1}^{n+1} t_k} \prod_{k=0}^{n} q_{i_k} e^{-q_{i_k} t_{k+1}} \, dt_1 \cdots dt_{n+1}$$

Make use of this formula, we can prove

$$P_i(X_t=j, X_{t+s}=k) = P_i(X_t=j)\,P_j(X_s=k) \quad (*) \quad (\underline{Ex})$$

$$P_i(X_t=j) = p_{ij}(t) = p(t; i,j) = p_t(i,j)$$

p_{ij} $\quad p_{ij}^{(n)}$: n-step transition probability.

$p_{ij}(t)$: function of t

transition matrix $\quad P(t) = \left(p_{ij}(t)\right)_{i,j\in S}$

Sum over all $j\in S$ in $(*)$ on both sides.

$$p_{ik}(t+s) = \sum_{j\in S} p_{ij}(t)\,p_{jk}(s)$$

chapmann-kolmogorv identity.

$$P(t+s) = P(t)\,P(s)$$

<u>proposition</u>: (Markov property)

(1) $\forall\, k\geq 2$, $\forall\, 0 < t_1 < \cdots < t_k$, $\forall\, i_1, \cdots, i_k \in S$

$$P_j(X_{i_1}=i_1, \cdots, X_{i_k}=i_k) = P_j(X_{i_1}=i_1)\prod_{s=1}^{k+1} P_{i_s}(X_{t_{s+1}-t_s}=i_{s+1})$$

(2) $\forall\, 0\leq t_1 < t_2 < \cdots < t_k < t$, $s\geq 0$. $\forall\, i_1, \cdots, i_k, i, j\in S$

$$P(X_{t+s}=j \mid X_t=i, X_{t_1}=i_1, \cdots, X_{t_k}=i_k) = P_i(X_s=j)$$

3. Infinistesimal description of continuous time Markov Chain

$$P(t) = \left(p_{ij}(t)\right)_{i,j\in S}$$

$$P'(t) = \frac{dP}{dt}$$

Next time !

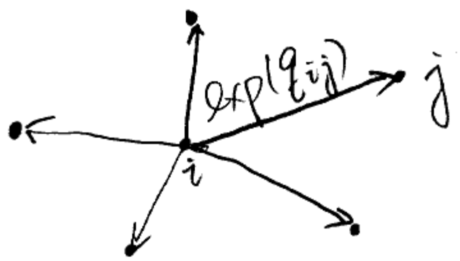
Formula

$$\mathbb{P}_i(X_t = j) = \sum_{n=0}^{\infty} \sum_{\substack{i_1 \dots i_{n-1} \in S \\ i_n = j}} \mathbb{P}\left(Y_0 = i_0, \dots, Y_n = i_n;\ \begin{array}{l} \xi_{Y_0} + \dots + \xi_{Y_{n-1}} \le t \\ < \xi_{Y_0} + \dots + \xi_{Y_n} \end{array} \right)$$

$$= \sum_{n=0}^{\infty} \sum_{\substack{i_1 \dots i_{n-1} \in S \\ i_n = j}} \prod_{k=0}^{n-1} p_{i_k i_{k+1}} \int \cdots \int_{\substack{t_k \geqslant 0 \\ \sum_{k=1}^{n} t_k \le t < \sum_{k=1}^{n+1} t_k}} \prod_{k=0}^{n} q_{i_k} e^{-q_{i_k} t_{k+1}} \, dt_1 \dots dt_{n+1}$$

11/17

3. Infinitesimal description of the process.



$$q_i = \sum_{\substack{j\in S \\ j\neq i}} q_{ij}$$

waiting time at site i.

When jump, probability $i \to j$, $p_{ij} = \frac{q_{ij}}{q_i}$

embedding chain $\{Y_n\}_{n\geq 0}$, transition matrix $(p_{ij})_{i,j\in S}$

$P(X_t = j \mid X_0 = i) = P_i(X_t = j) = p_{ij}(t)$

↳ transition probability from i to j in time t

transition matrix $P(t) = (p_{ij}(t))_{i,j\in S}$

$P(t+s) = P(t)P(s) = P(s)P(t)$

$p_{ij}(t) = P(X_t = j \mid X_0 = i)$

$i \neq j$, $\lim_{t\to 0} p_{ij}(t) = 0$.

$p_{ij}'(0) = \lim_{t\to 0} \frac{p_{ij}(t)}{t} = q_{ij}$ transition rate.

$\left[\lim_{t\to 0} \frac{p_{i,i+1}(t)}{t} = \lambda = q_{i,i+1}, \quad \lim_{t\to 0} \frac{p_{i,i+k}(t)}{t} = 0 \ (k\geq 2)\right]$

Pf: $p_{ij}(t) = P_i(N(t)=1, Y_1 = j)$

$$= p_{ij} \int_{\substack{t_1\geq 0,\, t_2\geq 0 \\ t_1+t_2 > t \geq t_1}} q_i e^{-q_i t_1} q_j e^{-q_j t_2} \, dt_1 dt_2.$$

$$= q_{ij} \int_0^t e^{-q_i t_1 - q_j(t-t_1)}\, dt_1$$

$$= q_{ij}\, e^{-q_j t} \int_0^t e^{q_i - q_j}\, dt_1$$

Thus, $\lim\limits_{t\to 0} p_{ij}(t) = q_{ij}$ ∎

When $i=j$, $\lim\limits_{t\to 0} p_{ii}(t) = 1$

$$p_{ii}'(0) = \lim_{t\to 0} \frac{p_{ii}(t)-1}{t} = \lim_{t\to 0} \frac{-\sum\limits_{\substack{j\neq i\\ j\in S}} p_{ij}(t)}{t}$$

$$= -\sum_{\substack{j\neq i\\ j\in S}} q_{ij} \qquad \left[p_{ii}(t) = 1 - \sum_{\substack{j\neq i\\ j\in S}} p_{ij}(t) \right]$$

Summary $\quad p(t) = (p_{ij}(t))_{i,j\in S}$

$$P'(0) = \begin{pmatrix} \sum\limits_{j\neq 1} -q_{1j} & q_{12} & q_{13} & \cdots & q_{1n} \\ q_{21} & \sum\limits_{j\neq 2} -q_{2j} & q_{23} & \cdots & q_{2n} \\ \vdots & & & & \vdots \end{pmatrix} = Q\text{-generator}$$

Row sum of Q-matrix $= 0$

$$P'(t) = \lim_{h\to 0} \frac{P(t+h) - P(t)}{h} = \lim_{h\to 0} \frac{P(t)P(h) - P(t)}{h}$$

$$= \lim_{h\to 0} \frac{P(t)(P(h) - I)}{h} = P(t) \lim_{h\to 0} \frac{P(h) - I}{h} = P(t)\cdot Q$$

(Forward Kolmogorov equation)

$P'(t) = QP(t)$ (Backward kolmogorov equation)

Solution: $P(0) = I$

$P(t) = e^{tQ}$

(Reference: Kolmolgorov, A.N. On analytical fundations of probability theory, Math. Ann. 104 (1931) pp. 415-458)

Ex: 2-State Markov Chain

$$\begin{tikzcd} 0 \arrow[r, bend right, "\lambda"'] & 1 \arrow[l, bend right, "\mu"'] \end{tikzcd}$$

$q_{01} = \lambda$

$q_{10} = \mu$

$$Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \qquad P(t) = \begin{pmatrix} P_{00}(t) & P_{01}(t) \\ P_{10}(t) & P_{11}(t) \end{pmatrix}$$

$$P'(t) = QP = PQ$$

$$\begin{cases} P_{00}'(t) = -\lambda P_{00}(t) + \lambda P_{10}(t) \\ P_{11}'(t) = -\mu P_{11}(t) + \mu P_{01}(t) \end{cases}$$

$P_{10}(t) = 1 - P_{11}(t)$ $\qquad$ $P_{01}(t) = 1 - P_{00}(t)$

Solution:
$$P_{00}(t) = \frac{\mu + \lambda e^{-(\mu+\lambda)t}}{\mu + \lambda}$$

$$P_{01}(t) = \frac{\lambda - \lambda e^{-(\mu+\lambda)t}}{\mu + \lambda}$$

$$P_{11}(t) = \frac{\lambda + \mu e^{-(\mu+\lambda)t}}{\mu + \lambda}$$

$$P_{10}(t) = \frac{\mu - \mu e^{-(\mu+\lambda)t}}{\mu + \lambda}$$

as $t \to \infty$,

$$\lim_{t\to\infty} P_{00}(t) = \lim_{t\to\infty} P_{10}(t) = \frac{\mu}{\mu+\lambda}.$$

$$\lim_{t\to\infty} P_{11}(t) = \lim_{t\to\infty} P_{01}(t) = \frac{\lambda}{\mu+\lambda}$$

<u>Solve the Differential Equation</u>

$$\left.\begin{aligned} p_{00}'(t) &= -\lambda p_{00}(t) + \lambda p_{10}(t) \\ p_{10}(t) &= 1 - p_{11}(t) \end{aligned}\right\}$$

$$\Rightarrow p_{00}'(t) = -\lambda p_{00}(t) + \lambda(1 - p_{11}(t))$$
$$= \lambda - \lambda(p_{00}(t) + p_{11}(t))$$

Similarly $p_{11}'(t) = \mu - \mu(p_{00}(t) + p_{11}(t))$

Let $V(t) = p_{00}(t) + p_{11}(t)$ we get $V(0) = 2$

$$V'(t) = (\lambda+\mu) - (\lambda+\mu)V(t)$$

$$\Rightarrow V(t) = e^{-(\lambda+\mu)t} + 1$$

Thus $p_{00}'(t) = \lambda - \lambda V(t) = -\lambda e^{-(\lambda+\mu)t}$

$$\Rightarrow p_{00}(t) = \frac{\mu + \lambda e^{-(\mu+\lambda)t}}{\mu+\lambda}$$

and we can then solve similarly for $p_{01}(t)$, $p_{11}(t)$, $p_{10}(t)$

12/1

Ergodic properties of continuous time Markov chain

(1) irreducibility. aperiodicity

(2) Recurrent, Transient, positive recurrent

(3) stationary distribution, ergodic theorem. detailed balance.

Assume $\{X_t\}_{t\geq 0}$ is a Q-process with Q-matrix. Q

State space S transition matrix $P(t) = e^{tQ}$.

Let $\{Y_n\}$ be the embedding chain

$S_n = \xi_{Y_0} + \xi_{Y_1} + \cdots + \xi_{Y_{n-1}}$ = jump time

N_s = # of jumps of X_t during $[0, s]$.

$\tau_j = \inf\{t \geq 0.\ X_t = j\}$ is the first time X_t jumps to j.

Def'n: If $P_i(\tau_j < \infty) > 0$, then we say i communicates with j, and we write $i \to j$. if $i \to j$ and $j \to i$, then we say i and j mutually communicate. and write $i \leftrightarrow j$. If $\forall\, i, j \in S$, $i \to j$. then we say Q is irreducible or X_t is irreducible.

proposition: If $i \neq j$, then following equivalent.

(1) $i \to j$ (2) embedding chain $\{Y_n\}$ has $i \to j$.

(3) $\exists\ i_0, \dots, i_n,\ i_0 = i,\ i_n = j.\ q_{i_0 i_1}, \dots, q_{i_{n-1} i_n} > 0$.

(4) For all $t > 0$, $P_i(X_t = j) > 0$

(5) There exists $t > 0$ s.t. $P_i(X_t = j) > 0$.

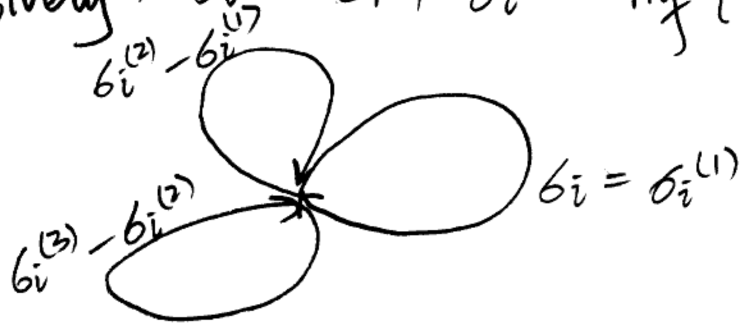
$\sigma_i = \inf\{t \geq S_1,\ X_t = i\}$

$\tau_i = \inf\{t \geq 0,\ X_t = i\}$

If $X_0 = i$ then $\sigma_i \neq \tau_i$

If $X_0 \neq i$ then $\sigma_i = \tau_i$

Recursively, $\sigma_i^{(1)} = \sigma_1$, $\sigma_i^{(k+1)} = \inf\{t \geq S_{N_{\sigma_i^{(k)}}+1},\ X_t = i\}$



<u>Proposition</u>: The following are equivalent.

(1) $P_i(\forall t > 0,\ \exists s > t,\ \text{s.t. } X_s = i) = 1$

(2) $q_i = 0$ or $\rho_{ii} = P_i(\sigma_i < \infty) = 1$

(3) $q_i = 0$ or $\forall k \geq 0$, $P_i(\sigma_i^{(k)} < \infty) = 1$

(4) $G_{ii} = \int_0^\infty p(X_t = i)\, dt = \infty$

(5) i is recurrent state for the embedding chain.

<u>Def'n</u>: If any of (1)–(5) in the previous proposition is satisfied, then we say state i is recurrent. Otherwise, i is transient.

(cycle)
<u>Lemma</u>: For any $k \geq 0$, and $A > 0$.

$$P_i(\sigma_i^{(k+1)} - \sigma_i^{(k)} < A \mid \sigma_i^{(k)} < \infty) = P_i(\sigma_i < A)$$

and conditioned on $X_0 = i$, $\sigma_i^{(k)} < \infty$, the sequence of r.v.'s $\{\sigma_i^{(k)} - \sigma_i^{(k-1)}, 1 \leq k \leq n\}$ are i.i.d. and $\sigma_i^{(0)} = 0$.

<u>Theorem</u>: Let i be recurrent, $q_i > 0$.

then $P_i\left(\lim_{t\to\infty} \frac{1}{t} \int_0^t 1_{X_s = i}\, ds = \frac{1}{q_i E_i \sigma_i} \right) = 1$

Discrete Time: $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_{X_k = i} = \pi_i = \frac{1}{E_i \sigma_i}$ frequency.

<u>Proof</u>: By LLN, $\lim_{k\to\infty} \frac{\sigma_i^{(k)}}{k} = E_i \sigma_i$

$\forall t \geq 0$, let $N_t = \sup\{k \geq 0, \sigma_i^{(k)} \leq t\}$ be the number of jumps to site i before t.

then $\frac{t}{N_t} \to E_i \sigma_i$. thus $\frac{N_t}{t} \to \frac{1}{E_i \sigma_i}$

But $\frac{1}{N_t} \int_0^t 1_{X_s = i}\, ds \to \frac{1}{q_i}$

Thus, $\frac{1}{t} \int_0^t 1_{X_s = i}\, ds = \frac{N_t}{t} \cdot \frac{1}{N_t} \int_0^t 1_{X_s = i}\, ds = \frac{1}{E_i \sigma_i} \cdot \frac{1}{q_i}$

Def'n: Let i be recurrent. If $q_i = 0$ or $E_i \sigma_i < \infty$, then we say i is positive recurrent.
If $q_i > 0$, and $E_i \sigma_i = \infty$, then we say i is null-recurrent.

Def'n: $\mu = (\mu_1, \cdots, \mu_n \cdots)$ is a probability distribution on S is stationary iff $\forall t \geq 0$, $\mu P(t) = \mu$.

$\mu P(0) = \mu$, $(\mu P(t))' = \mu P'(t) = \mu Q P(t) = 0$.

proposition: μ is stationary for $P(t) \Leftrightarrow \mu Q = 0$.

Theorem (mixing): If Q is irreducible and μ is a stationary distribution for $P(t)$, then $\lim_{t \to \infty} P_{ij}(t) = \mu_j$

proposition: If Q is irreducible then following are equivalent.

(1) All states are positive recurrent

(2) there exist positively recurrent state.

(3) there exist stationary distribution.

In particular, if μ is stationary distribution, then $\mu_i = \dfrac{1}{q_i E_i \sigma_i}$

<u>Theorem</u>: (Ergodic Theorem)

Let Q be irreducible and stationary distribution

Let $f: S \to \mathbb{R}$ be bounded.

Then $P_\nu\left(\lim_{t\to\infty}\frac{1}{t}\int_0^t f(X_s)\,ds = \sum_{i\in S}\mu_i f(i)\right) = 1$.

ν is any initial distribution.

$\pi_i p_{ij} = \pi_j p_{ji}$

<u>Def'n</u>: Let μ be a probability distribution on S.

We say μ is a detailed balance iff

$\forall i,j\in S$. $\mu_i q_{ij} = \mu_j q_{ji}$

Show then that $\mu = (\mu_i)_{i\in S}$ is a stationary distribution, this is because

$$(\mu Q)_j = \sum_{\substack{k\in S\\ k\neq j}} \mu_k q_{kj} - \mu_j \sum_{\substack{k\in S\\ k\neq j}} q_{jk} \quad \left(q_{jj} = -\sum_{\substack{k\in S\\ k\neq j}} q_{jk}\right)$$

$$\underline{\underline{\text{detailed balance}}} \quad \sum_{\substack{k\in S\\ k\neq j}} \mu_j q_{jk} - \sum_{\substack{k\in S\\ k\neq j}} \mu_j q_{jk} = 0$$

so $\mu = (\mu_j)_{j\in S}$ is a stationary distribution

Let us further show that $\mu_i p_{ij}(t) = \mu_j p_{ji}(t)$

We define $\hat{p}_{ij}(t) = \frac{\mu_j}{\mu_i} p_{ji}(t)$

and show $\hat{p}_{ij}(t) = p_{ij}(t)$ then we are done

Now we have $\hat{p}_{ij}(0) = \frac{\mu_j}{\mu_i} p_{ji}(0) = \frac{\mu_j}{\mu_i} \delta_{ji}$

$$= \delta_{ji} = \delta_{ij}$$
$$= p_{ij}(0)$$

And $\hat{p}_{ij}'(t) = \frac{\mu_j}{\mu_i} p_{ji}'(t)$

$$= \frac{\mu_j}{\mu_i} \sum_{k \in S} q_{jk} \cdot p_{ki}(t)$$

$$= \frac{\mu_j}{\mu_i} \sum_{k \in S} q_{jk} \frac{\mu_i}{\mu_k} \hat{p}_{ik}(t)$$

$$= \sum_{k \in S} \frac{\mu_j q_{jk}}{\mu_k} \hat{p}_{ik}(t) = \sum_{k \in S} q_{kj} \hat{p}_{ik}(t)$$

$$= \sum_{k \in S} \hat{p}_{ik}(t) q_{kj} \Rightarrow \hat{p}_{ij}(t) = (e^{tQ})_{ij}$$

So $\hat{p}_{ij}(t) = p_{ij}(t)$ thus $\mu_j p_{ji}(t) = \mu_i p_{ij}(t)$

12/3

$\mu_i q_{ij} = \mu_j q_{ji}$

$\Rightarrow \mu_i p_{ij}(t) = \mu_j p_{ji}(t)$ $\quad\quad P' = QP$

$\hat{p}_{ij}(t) = \frac{\mu_j}{\mu_i} p_{ji}(t)$ $\quad\quad \hat{p}'_{ij}(t) = \sum \hat{p}_{ik}(t) q_{kj}$

$\parallel$

$p_{ij}(t)$ $\quad\quad P' = \hat{P}Q$

---

1. Hitting times and hitting distributions

Continuous time Markov chain X_t state space S.

Q-matrix $\quad Q = (q_{ij})_{i,j\in S}$

$A \subset S$. $\tau_A = \min\{t : X_t \in A\}$

$a \in A$, $\quad h_i = P_i(X_{\tau_A} = a)$

boundary conditions $\quad h_a = 1$, $h_b = 0$, for all $b \in A \setminus \{a\}$.

for $i \notin A$, $\quad h_i = \sum_{j\neq i} p_{ij} h_j = \sum_{j\neq i} \frac{q_{ij}}{q_i} h_j$. $\quad q_i = -q_{ii}$

$\Leftrightarrow \sum_{j\in S} q_{ij} h_j = 0$ for all $i \notin A$

$g_i = E_i \tau_A$

boundary condition if $i \in A$, then $g_i = 0$

if $i \notin A$, $\quad g_i = \frac{1}{q_i} + \sum_{j\neq i} p_{ij} g_j = \frac{1}{q_i} + \sum_{j\neq i} \frac{q_{ij}}{q_i} g_j$

$\Leftrightarrow 1 + \sum_{j\in S} q_{ij} g_j = 0$

Queuing Theory

queuing systems.

$\boxed{\text{service}}$

(1) arrival time of customers
(2) how long does the service spend at each customer
(3) how many agents are providing service

Questions: Stability   Average queue length
Average waiting time

G (general)

M / M / 1 queue

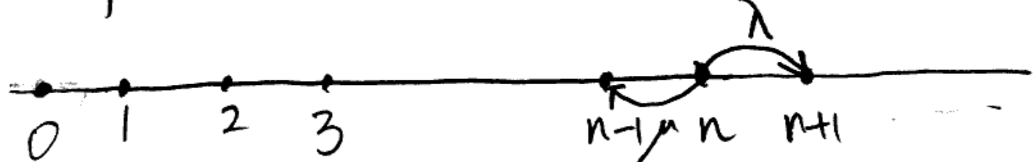
memoriless of arrival time of customer
memoriless of time at each service
1 agent

Arrival rate λ service Rate μ

$\xrightarrow{\lambda} \boxed{\text{server } \mu}$

L_t = length of the queue (birth & death chain)

State space



Birth Rate λ
Death Rate μ

Detailed balance

$\pi_i q_{i,i+1} = \pi_{i+1} q_{i+1,i}$ $\pi_i \lambda = \pi_{i+1} \mu$

$$\pi_{i+1} = \frac{\lambda}{\mu}\pi_i = \cdots = \left(\frac{\lambda}{\mu}\right)^i \pi_0$$

$$\sum_{i=0}^{\infty} \pi_i = \pi_0\left(1+\frac{\lambda}{\mu}+\left(\frac{\lambda}{\mu}\right)^2+\cdots\right)=1 \Rightarrow \pi_i = 1-\frac{\lambda}{\mu}$$

Stability criteria $\lambda < \mu$

Detailed Balance gives the stationary distribution

$$\pi_i = \left(1-\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^i \qquad i=0,1,2,\cdots$$

Average length $= EL_t = \sum_{i=0}^{\infty} i\pi_0 = \frac{\lambda}{\mu-\lambda}$

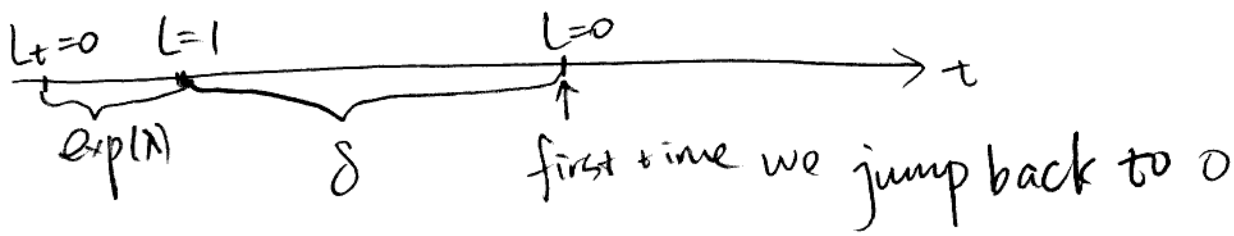
Average time a customer spend at the queue $= W$

When a person arrives at the queue at time t he is the L_t+1's person in the queue.

$$W = \sum_{i=0}^{L_t+1} \tau_i \qquad \tau_i \sim \exp(\mu)$$

$$EW = \frac{1}{\mu}(EL+1) = \frac{1}{\mu}\left(\frac{\lambda}{\mu-\lambda}+1\right) = \frac{1}{\mu-\lambda}$$

(3) proportion of time that the agent is working.



$$\frac{E\delta}{E\delta+E\xi} \qquad \sigma_0 = \delta+\xi$$

$$E_0\sigma_0 = \frac{1}{q_0\pi_0} = \frac{1}{\lambda(1-\frac{\lambda}{\mu})} = \frac{\mu}{\lambda(\mu-\lambda)}$$

$$E\delta = E_0 T_0 - E\xi = \frac{\mu}{\lambda(\mu-\lambda)} - \frac{1}{\lambda} = \frac{1}{\mu-\lambda}$$

proportion $= \frac{\lambda}{\mu}$

M/M/s queue

birth rate λ

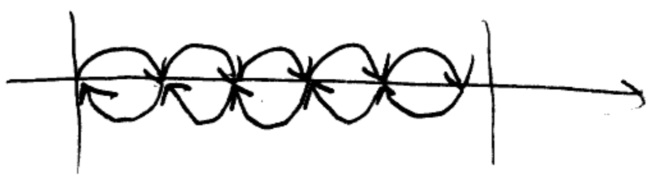
death rate $s\mu$.

M/M/∞ queue

birth rate λ, never wait

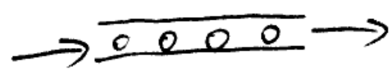
M/M/s with a finite waiting room of size N.

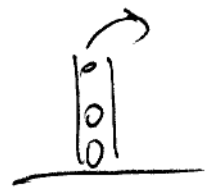
$L_t \le N$



FIFO (first in first out)

LIFO (last in first out)





Brownian Motion

X_t Einstein 1905

$X_0 = 0$

<u>Assumptions</u>

(1) X_t is space homogeneous

(2) $X_{t+s} - X_t$ is independent of X_t

$X_{t+s} - X_t$ has the same distribution as X_s

(independent and stationary increments)

(3) X_t is a continuous function of t

(with probability 1)

$X(t) = X_t$

$$\xi_i = X\left(\frac{i}{n}t\right) - X\left(\frac{i-1}{n}t\right), \quad i = 1, 2, \ldots$$

Assumption (1) $\Rightarrow$ $\mathbb{E}\xi_1 = 0$

Assumption (2) $\Rightarrow$ $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ i.i.d.

If we assume $\mathbb{E}X_t^2 < \infty$ then $\mathbb{E}X_t^2 = n\,\mathbb{E}\xi_i^2$

$$X(t) = \xi_1 + \dots + \xi_n$$

Apply Central Limit Theorem

$$\frac{\xi_1 + \dots + \xi_n - n\mathbb{E}\xi_1}{\sqrt{n \operatorname{Var}\xi_1}} \xrightarrow[n\to\infty]{d.} N(0,1)$$

Thus $\quad X(t) \overset{d.}{=\!=} N(0, \mathbb{E}X_t^2)$ as $n \to \infty$

But $\quad \mathbb{E}X_{t+s}^2 = \operatorname{Var}((X_{t+s} - X_t) + X_t)$

$$= \operatorname{Var}(X_{t+s} - X_t) + \operatorname{Var} X_t$$

$$= \mathbb{E}X_s^2 + \mathbb{E}X_t^2$$

Let $\quad m(t) = \mathbb{E}X_t^2 \quad$ Then $\quad m(t+s) = m(t) + m(s)$

and $m(t)$ is continuous in t by Assumption (3)

Thus $\quad m(t) = \sigma^2 t$

and $\quad \sigma^2 = \mathbb{E}X_1^2 = \operatorname{Var} X_1$

So $\quad X_t \sim N(0, \sigma^2 t)$

<u>Definition</u> We say a stochastic process $\{B_t : t \geq 0\}$, $B_0 = 0$ is a Brownian motion

(Wiener process) (in dimension 1) if

(i) $\forall t \geq 0$, $s > 0$, $B_{t+s} - B_t \sim N(0, \sigma^2 s)$

(ii). $\forall\, 0 < t_1 < \cdots < t_n$, B_{t_1}, $B_{t_2} - B_{t_1}$, $\ldots$, $B_{t_n} - B_{t_{n-1}}$ are independent

(iii) With probability 1 the process B_t has continuous trajectories.

If $\sigma^2 = 1$, we call B_t a standard Brownian motion. We sometime also write W_t for B_t

$B_0 = 0$ $\qquad \mathbb{P}_0$

$\tilde{B}_0 = x$ $\qquad \tilde{B}_t = x + B_t$ $\qquad \mathbb{P}_x$

<u>Proposition</u> (Finite dimensional distribution) $0 < t_1 < t_2 < \cdots < t_n$ the joint probability density of $(B(t_1), B(t_2), \ldots, B(t_n))$ is

$$p(t_1, t_2, \ldots, t_n; x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} p_{t_k - t_{k-1}}(x_{k-1}, x_k)$$

where $p_t(x,y) = \dfrac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{(y-x)^2}{2\sigma^2 t}}$

Let B_t be a Brownian motion

$\mathbb{E} B_t = 0 \qquad Cov(B_t, B_s) = t \wedge s$

Scaling Property $\dfrac{B_{ct}}{\sqrt{c}} \overset{d.}{=} B_t$

Multidimensional Brownian Motion

$B(t) = (B_1(t), \dots, B_n(t))$ $B_i(t)$ independent

Exercise Show that if B_t is a standard Brownian motion then the following are standard Brownian motions

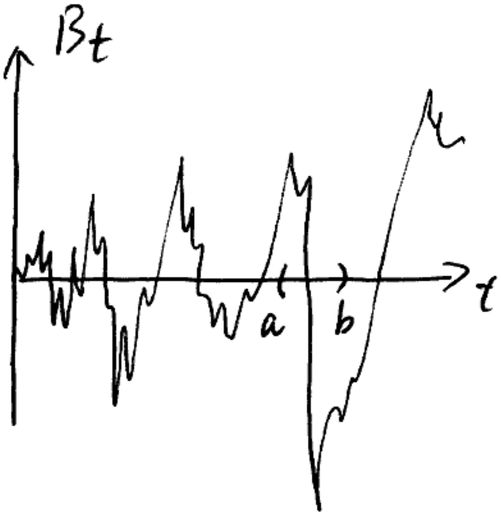
(1). $\{-B_t,\ t \geq 0\}$

(2). $\{B_{t+u} - B_u : t \geq 0\}$

(3). $\{B_{T-t} - B_T,\ 0 \leq t \leq T\}$ (for finite time)

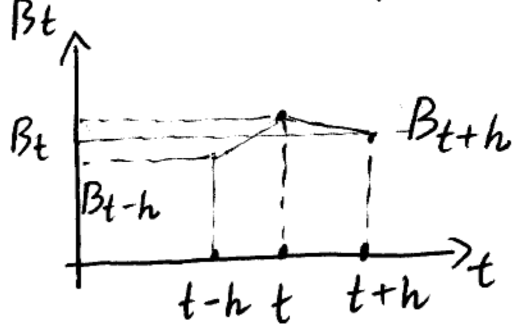
(4). $\{t B_{1/t},\ t \geq 0\}$ (define $tB_{1/t} = 0$ as $t=0$)

Sample Path property



Property 1 For any $(a,b) \subset \mathbb{R}$, B_t is not monotone in (a,b)

Property 2 B_t is not differentiable for almost all trajectories and all $t \geq 0$



Since $B_{t+h} - B_t$ is independent of $B_t - B_{t-h}$

$$\mathbb{P}(B_{t+h} - B_t = B_t - B_{t-h}) = 0$$

Markov Process : $\forall\, 0 \leq t_1 < \cdots < t_n < t,\ s > 0,\ x \in \mathbb{R}$

$$\mathbb{P}(B_{t+s} \in A \mid B_t = x,\ B_{t_1} \in B_1, \ldots, B_{t_n} \in B_n)$$

$$= \mathbb{P}(B_{t+s} \in A \mid B_t = x) = p_s(x, A)$$

where $p_s(x, A) = \int_A p_s(x, y)\, dy$

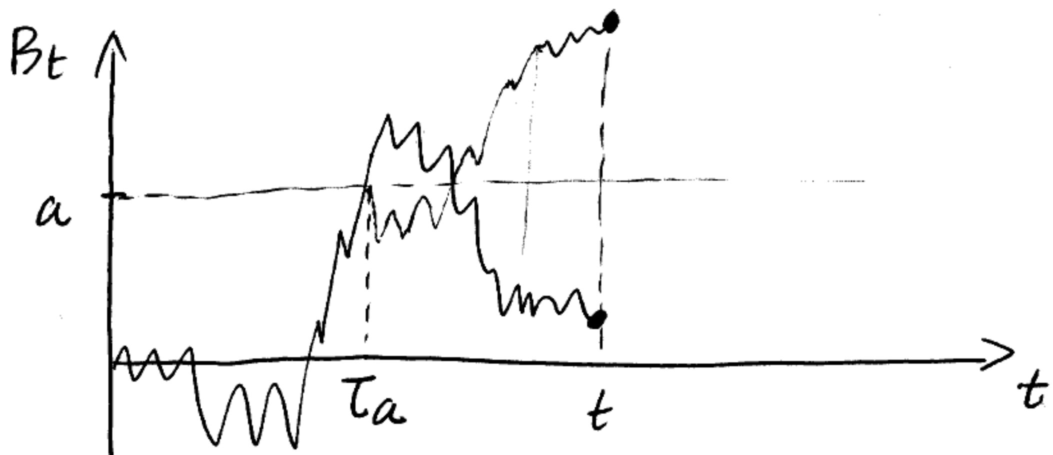
First passage time

$$\tau_A = \inf\{t \geq 0 : B_t \in A\},\ \tau_{\{a\}} = \tau_a$$

<u>Strong Markov Property</u> If τ is a first passage time, then conditioned on $\tau < \infty$ the process $\{B_{t+\tau} - B_\tau : t \geq 0\}$ is a standard Brownian motion and it is independent of $\{B_s : 0 \leq s \leq \tau\}$

<u>Reflection Principle</u>

$$\forall a > 0 \qquad \mathbb{P}_0(\tau_a \leq t) = 2\,\mathbb{P}_0(B_t > a)$$



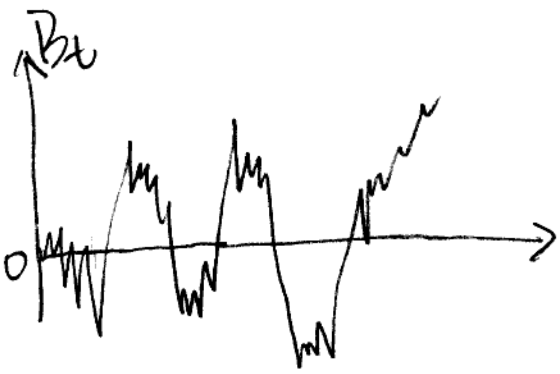
<u>Exercise</u> (1) $a > 0$ Calculate the distribution of τ_a

(2) $\mathbb{E}\, e^{-\lambda \tau_a} = e^{-a\sqrt{2\lambda}}, \ \forall \lambda, a > 0$

(3). $M_t = \sup_{0 \leq s \leq t} B_s$, Calculate the distribution of M_t

12/10 Brownian Motion (Standard)

$(B_t)_{t \geq 0}$



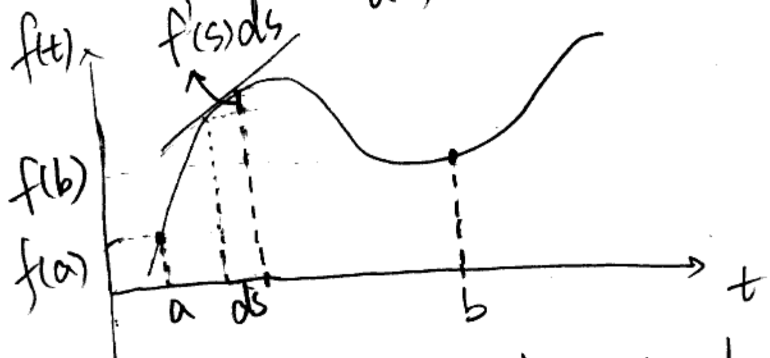
(1) $B_t \sim N(0, t)$

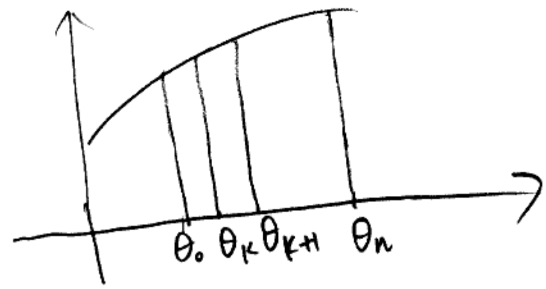
(2) $B_{t+s} - B_t$ independent of B_t

$B_{t+s} - B_t \stackrel{d}{=} B_s$

(3) continuous trajectories

$$f(b) - f(a) = \int_a^b f'(s)\, ds$$





$s = s(\theta)$ $\quad ds = s'(\theta) d\theta$

$$f(s(b)) - f(s(a)) = \int_a^b f'(s(\theta)) \underbrace{s'(\theta)\, d\theta}_{ds(\theta)} = \int_a^b f'(s(\theta))\, ds(\theta)$$

Stochastic Calculus

(K. Itô)

$$f(B_2) - f(B_1) = \int_{t_1}^{t_2} f'(B_s) \underline{B_s' \, ds} = \int_{t_1}^{t_2} f'(B_s)\, dB_s$$

$$\int_a^b f'(s(\theta))\, ds(\theta) = \lim_{\substack{n \to \infty \\ \max\limits_{0 \leq k \leq n-1} |\theta_{k+1} - \theta_k| \to 0}} \sum_{k=0}^{n-1} f'(s(\xi_k))\big(s(\theta_{k+1}) - s(\theta_k)\big)$$

$\theta_k \leq \xi_k \leq \theta_{k+1}$

$$\int_a^b f'(B_s)\,dB_s = \lim_{\substack{n\to\infty \\ \max|t_{k+1}-t_k|\to 0 \\ 0\le t\le n-1}} \sum_{k=0}^{n-1} f'(B_{\xi_k})(B_{t_{k+1}}-B_{t_k})$$

$$\int_0^t B_s\,dB_s = \lim_{\substack{n\to\infty \\ \max|t_{k+1}-t_k|\to 0}} \sum_{k=0}^{n-1} B_{t_k}(B_{t_{k+1}}-B_{t_k}) = I_1$$

$$\neq \lim_{\substack{n\to\infty \\ \max|t_{k+1}-t_k|\to 0}} \sum_{k=0}^{n-1} B_{t_{k+1}}(B_{t_{k+1}}-B_{t_k}) = I_2$$

$$I_2 - I_1 = \lim_{\substack{n\to\infty \\ \max|t_{k+1}-t_k|\to 0}} \sum_{k=0}^{n-1} (B_{t_{k+1}}-B_{t_k})^2 \xrightarrow{?} 0$$

$$\boxed{E(B_{t+h}-B_t)^2 = EB_h^2 = h} \qquad \boxed{(dB_t)^2 \approx dt} \qquad dB_t \approx \sqrt{dt}$$

$$E(I_2 - I_1) = \lim_{n\to\infty} \sum_{k=0}^{n-1} (t_{k+1}-t_k) = t_{n-1} - t_0 = t$$

differentiability deterministic

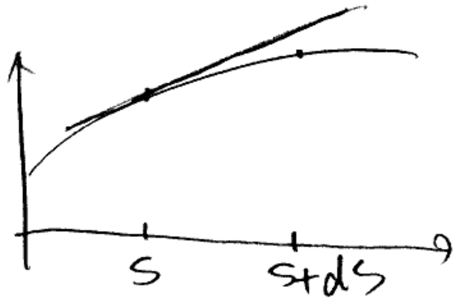
<u>Def'n:</u> (Itô's Integral)

$$\int_0^t f(B_s)\,dB_s = \lim_{\substack{n\to\infty \\ \max|t_{k+1}-t_k|\to 0}} \sum_{k=0}^{n-1} f(B_{t_k})(B_{t_{k+1}}-B_{t_k})$$

$$f(S(t_2)) - f(S(t_1)) = \int_{t_1}^{t_2} f'(S(\theta))\,dS(\theta)$$

$$f(S(t)+dS(t)) - f(S(t)) = f'(S(t))\,dS$$

$$(dS)^2 = 0$$



$$f(B_t + dB_t) - f(B_t) = f'(B_t)\,dB_t + \frac{1}{2} f''(B_t)\underbrace{(dB_t)^2}_{= dt} + \cancel{\frac{1}{3} f^{(3)}(B_s)(dB_s)^3} \cdots$$

$$f(B_t + dB_t) - f(B_t) = f'(B_t)\,dB_t + \frac{1}{2} f''(B_t)\,dt$$

$$\boxed{f(B_t) - f(B_0) = \int_0^t f'(B_s)\,dB_s + \frac{1}{2}\int_0^t f''(B_s)\,ds}$$ ← Itô's correction

$$f(\varphi_t) - f(\varphi_0) = \int_0^t f'(\varphi_s)\,d\varphi(s) \qquad d\varphi_s = ds \qquad (ds)^2 = 0$$

$f(x) = x^2$

$$B_t^2 - B_0^2 = \int_0^t 2B_s\,dB_s + \frac{1}{2}\int_0^t 2\,ds$$

$$= 2\int_0^t B_s\,dB_s + t$$

$$B_t^2 - t = 2\int_0^t B_s\,dB_s$$