## Lecture notes of Stochastic Process

lectured by prof. Hsueh-I Lu

pishen

AlgoLab, CSIE, NTU

June 6, 2012

## Thank list

LeoSW, windker

pishen (AlgoLab) Stochastic Process June 6, 2012 2 / 197

### Stochastic Process

#### Definition

A Stochastic process is a set of random variables  $\{X(t)|t\in T\}$  where T is a index (time) set.

State Space: possible value of X(t) for each t, which is defined as subset of R.

pishen (AlgoLab) Stochastic Process June 6, 2012

## Markov Chain

#### Definition

A Stochastic Process X with state space S is a Markov Chain if  $\exists 0 \leq p_{ii} \leq 1 \quad \forall i, j \in S \text{ such that }$ 

$$(a) \quad \sum_{j \in S} p_{ij} = 1 \quad \forall i \in S$$

(b) 
$$P(X(t+1) = j | X(0) = i_0, X(1) = i_1, ..., X(t) = i) = p_{ij}$$
  
 $\forall t, i_0, i_1, ..., i_{t-1}$ 

 $\mathbb{P}$  denotes the matrix form of  $p_{ij}$  with sum of any row is 1.

Lemma:  $P(X(n) = i | X(0) = i) = \mathbb{P}^n[i, j]$ 

## Proof of lemma

We know statement is true for (m+n)=0. For (m+n)>0:

$$\begin{split} &P(X(m+n)=j|X(0)=i)\\ &=\sum_{k\in S}P(X(m+n)=j \text{ and } X(m)=k|X(0)=i)\\ &=\sum_{k\in S}P(X(m+n)=j|X(m)=k \text{ and } X(0)=i)\cdot\\ &P(X(m)=k|X(0)=i)\\ &=\sum_{k\in S}P(X(m+n)=j|X(m)=k)\cdot P(X(m)=k|X(0)=i)\\ &=\sum_{k\in S}P^n[k,j]\cdot P^m[i,k]\\ &=\sum_{k\in S}P^m[i,k]\cdot P^n[k,j]\\ &=\mathbb{P}^n[i,j] \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012

5 / 197

# Proof of lemma(cont)

```
: conditional on X(m)
= : definition of conditional probability
= : (see next page)
= : inductive hypothesis
```

# Proof of lemma(cont)

$$\begin{split} P(X(m+n) &= j | X(m) = k \text{ and } X(0) = i) \\ &= \sum_{r \in S} P(X(m+n) = j | \\ &X(m+n-1) = r \text{ and } X(m) = k \text{ and } X(0) = i) \cdot \\ &P(X(m+n-1) = r | X(m) = k \text{ and } X(0) = i) \\ &= \sum_{r \in S} P(X(m+n) = j | X(m+n-1) = r) \cdot \\ &P(X(m+n-1) = r | X(m) = k) \\ &= P(X(m+n) = j | X(m) = k) \end{split}$$

- =: conditional on X(m+n-1)
- =: first part by definition of Markov chain and second part by inductive hypothesis

pishen (AlgoLab) Stochastic Process June 6, 2012 7 / 197

# Absorbing State

Let  $\mathbb A$  be a set of accepting states. We would like to know the probability that  $\mathbb X$  has ever entered some state in  $\mathbb A$ . Technique: merge all state of  $\mathbb A$  into a new absorbing state a. Set matrix of  $\mathbb X$  by once enter a, then probability of a goes to a is 1.

## Recurrent & transient

### Definition

The recurrent probability of state i of Markov chain X is

$$f_i = P(\text{there exists an index } t \ge 1 \text{ with } X(t) = i | X(0) = i)$$

- State *i* of X is recurrent if  $f_i = 1$ .
- State i of X is transient if  $f_i < 1$ .

# Recurrent & transient (cont.)

- If state i is recurrent, by the property of Markov chain, once it re-enter the state i, we can take it as starting from X(0) again. Hence we know that it will keep re-entering the state i again and again in the process.
- If state i is transient, in each period it start going from i, it may have probability  $1 - f_i$  that it won't come back anymore. Hence the probability that the process will be in state i for exactly nperiods equals  $f_i^{n-1}(1-f_i)$ ,  $n \ge 1$ , which is a geometric distribution.

pishen (AlgoLab) Stochastic Process June 6, 2012 10 / 197

# Recurrent & transient (cont.)

- From the preceding page, it follows that state *i* is recurrent if and only if, starting in state *i*, the expected number of steps that the process is in state *i* is infinite.
- We can also derive that, if the Markov chain has finite states, at least one state is recurrent.

pishen (AlgoLab) Stochastic Process June 6, 2012 11 / 19

# Expected number of visits

Let

$$I(n) = \begin{cases} 1 & \text{if } X(n) = i \\ 0 & \text{if } X(n) \neq i \end{cases}$$

we have  $\sum_{n=0}^{\infty} I(n)$  represents the number of steps that the process is in state i, and

$$E\left[\sum_{n=0}^{\infty} I(n)|X(0) = i\right] = \sum_{n=0}^{\infty} E[I(n)|X(0) = i]$$
$$= \sum_{n=0}^{\infty} 1 \cdot P(X(n) = i|X(0) = i)$$
$$= \sum_{n=0}^{\infty} P_{ii}^{n}$$

We set  $T = \sum_{n=0}^{\infty} I(n)$ 

## Lemma 1

From the above statements, we prove the following

#### Lemma

State i is

recurrent 
$$\iff \sum_{n=0}^{\infty} P_{ii}^n = \infty$$
,

transient 
$$\iff \sum_{n=0}^{\infty} P_{ii}^n < \infty$$

## Proof of Lemma 1

Suppose state i is transient( $f_i < 1$ ), consider  $P(T = k) = f_i^{k-1} \cdot (1 - f_i)$ . Since T is a geometric distribution, we have

$$E[T] = \sum_{k=0}^{\infty} k \cdot f_i^{k-1} \cdot f_i$$
$$= \frac{1}{1 - f_i} < \infty$$

## Communicated states

#### Definition

State i and j communicate, denoted  $i \leftrightarrow j$ , if there exist integers  $m \ge 0$  and n > 0 such that

$$P_{ij}^m > 0$$
 and  $P_{ji}^n > 0$ 

We say a Markov chain X is irreducible if  $i \leftrightarrow j \quad \forall i, j \in S$ 

## Lemma 2

#### Lemma

If  $i \leftrightarrow j$ , then the following statements hold.

- State i is recurrent if and only if state i is recurrent.
- State i is transient if and only if state i is transient.

Corollary: X is finite and irreducible  $\implies$  all states are recurrent.

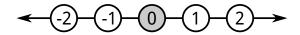
- $\mathbb{X}$  is finite  $\implies \exists i \in S$  is recurrent (proof later)
- By Lemma 2, all states are recurrent

pishen (AlgoLab) Stochastic Process June 6, 2012 16 / 197 Let m and n be nonnegative integers with  $P^m_{ij}$ ,  $P^n_{ji} > 0$ . Suppose that state j is recurrent, i.e.,  $\sum_{t=0}^{\infty} P^t_{jj} = \infty$ . We have

$$\begin{split} \sum_{t=0}^{\infty} P_{ii}^t &\geq \sum_{t=0}^{\infty} P_{ii}^{m+t+n} \\ &\geq \sum_{t=0}^{\infty} P_{ij}^m \cdot P_{jj}^t \cdot P_{ji}^n \\ &= P_{ij}^m \cdot P_{ji}^n \cdot \sum_{t=0}^{\infty} P_{jj}^t = \infty \end{split}$$

Thus, state i is also recurrent.

# Infinite drunken man problem



Let the state space consist of all integers. Let X(0) = 0 (i.e. at time 0 the drunken man is in state 0). The transition probabilities are such that

$$P_{i,(i+1)} = P_{i,(i-1)} = 0.5$$

holds for all states i of X.

Stochastic Process June 6, 2012 18 / 197

## Gambler's ruin

pishen (AlgoLab) Stochastic Process June 6, 2012 19 / 197

## Outline

- Limiting probabilities
- Stationary distribution
- 3 Long-run proportion
- 4 (Inverse of) Expected return time

Stochastic Process June 6, 2012 20 / 197

# Limiting Probabilities

#### Definition

Number  $\pi_i$  is the *limiting probability* of j if

$$\pi_j = \lim_{n \to \infty} P_{ij}^n$$

holds for all states  $i \in S$  ( $S \subseteq \mathbb{N}$  is the state space).

 $\blacksquare$   $\pi_i$  is independent of i.

$$lacksquare \lim_{n o\infty}P^n=egin{pmatrix}\pi\\\pi\\dots\\\vdots\end{pmatrix}$$
 , where  $\pi=(\pi_1,\pi_2,\ldots)$ 

pishen (AlgoLab) Stochastic Process June 6, 2012 21 / 197

# Stationary Probability Distribution

#### Definition

Non-negative row vector  $\pi = (\pi_1, \pi_2, ...)$  is a stationary probability distribution of  $\mathbb{X}$  if  $\pi \times P = \pi$  holds and  $\sum_{i \in S} \pi_i = 1$ 

- $\blacksquare$   $\pi$  is a normalized left eigenvector with eigenvalue = 1.
- If X(0) has distribution  $\pi$ , then X(t) has the same distribution  $\pi$  for all t > 1.  $\pi$  is also called as *steady-state distribution*.
- It doesn't mean that each X(t) become independent.  $\pi$  only means the distribution of X(t) when the previous random variable's value is unknown.

pishen (AlgoLab) Stochastic Process June 6, 2012 22 / 19

## Theorem 1

#### **Theorem**

Let X be an irreducible, aperiodic, positive recurrent Markov chain, then

- The limiting probability  $\pi_i$  of each state i exists.
- $\pi = (\pi_1, \pi_2, ...)$  is the unique stationary probability distribution.

■ The proof will be stated at page 38.

# Expected return time

#### Definition

The *expected return time* of state  $i \in S$  is

$$\mu_i = \sum_{n \ge 1} n \cdot f_i^{(n)}$$

where

$$f_i^{(n)} = P(\min\{t : X(t) = i, t \ge 1\} = n|X(0) = i)$$

## Positive recurrent & null recurrent

#### Definition

State *i* is positive recurrent if  $\mu_i < \infty$ 

#### Definition

State *i* is *null recurrent* if  $\mu_i = \infty$ 

- Both are recurrent states, and are class properties, which means that if state i and j communicate, they will share this property.
- If X is finite, then each recurrent state of X is positive recurrent. Proof stated at page 63.

Stochastic Process June 6, 2012 25 / 197

# Example of null recurrent

#### Example

For a Markov chain with n states  $(1, \ldots, n)$ , if

$$P(X(t+1) = i+1|X(t) = i) = 1-1/n$$

and

$$P(X(t+1) = 1|X(t) = i) = 1/n$$

According to geometric distribution (taking p = 1/n), the expectation value of "steps taken for state 1 to come back" will be 1/p = n, hence  $\lim_{n\to\infty} n = \infty$ .

Stochastic Process June 6, 2012 26 / 197

## Period of a chain

#### Definition

The *period* of state i is d if d is the largest integer such that

$$P_{ii}^n = 0$$

holds for all n which is not divisible by d.

#### Definition

If each state of X has period 1, then X is called *aperiodic*.

- If  $P_{ii} > 0$  for all  $i \in S$ , then X is aperiodic.
- Period can be seen as the gcd of all n that have  $P_{ii}^n > 0$ , note that  $P_{ii}^{\text{gcd}} > 0$  is not necessary.
- The period of drunken man problem is 2.

pishen (AlgoLab) Stochastic Process June 6, 2012 27 / 19

## Lemma 1

#### Lemma

If state i has period 1 and is positive recurrent, then

$$\pi_{ij} \equiv \lim_{n \to \infty} P_{ij}^n$$

exists and is positive for all states  $i \in S$ .

- This can be proved by the Blackwell theorem in Renewal theory.
- It doesn't promise that  $\pi_{ii} = \pi_{i'i}$  for any  $i, i' \in S$ . But they will be the same if we add the irreducible property  $(i \leftrightarrow i')$ .

Stochastic Process June 6, 2012 28 / 197

# Property of lim

■ The position of lim may not be switched arbitrarily in an equation.

### Example

$$1 = \lim_{n \to \infty} \lim_{m \to \infty} \frac{m}{m+n} \neq \lim_{m \to \infty} \lim_{n \to \infty} \frac{m}{m+n} = 0$$

lim would not influence the inequality.

#### Example

If 
$$f(n) \ge g(n)$$
, then  $\lim_{n\to\infty} f(n) \ge \lim_{n\to\infty} g(n)$ 

Stochastic Process June 6, 2012 29 / 197

# Property of lim (cont.)

lim is linear operator under finite number of functions.

### Example

For  $m < \infty$ ,

$$\sum_{i=1}^{m} \lim_{n \to \infty} f_i(n) = \lim_{n \to \infty} \sum_{i=1}^{m} f_i(n)$$

need an example of  $m = \infty$ 

# Inequality 1

## Inequality

$$\sum_{j \in S} \pi_{ij} \le 1 \quad \forall i \in S$$

pishen (AlgoLab) Stochastic Process June 6, 2012 31 / 197

## Proof

$$\lim_{m \to \infty} \sum_{j=1}^{m} \pi_{ij} = \lim_{m \to \infty} \sum_{j=1}^{m} \lim_{n \to \infty} P_{ij}^{n}$$

$$= \lim_{m \to \infty} \lim_{n \to \infty} \sum_{j=1}^{m} P_{ij}^{n}$$

$$\leq \lim_{m \to \infty} \lim_{n \to \infty} \sum_{j \in S} P_{ij}^{n} = 1$$

■ The last equation works since  $\sum_{j \in S} P_{ij}^n = 1$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 32 / 197

# Inequality 2

## Inequality

For state  $j \in S$ , we have

$$\pi_{ij} \ge \sum_{k \in S} \pi_{ik} P_{kj}$$

## Proof

For  $m \ge 1$  and  $n \ge 1$ ,

$$P_{ij}^{n+1} = \sum_{k \in S} P_{ik}^{n} P_{kj} \ge \sum_{k=1}^{m} P_{ik}^{n} P_{kj}$$

then

$$\pi_{ij} = \lim_{n \to \infty} P_{ij}^{n+1} \ge \lim_{n \to \infty} \sum_{k=1}^{m} P_{ik}^{n} P_{kj} = \sum_{k=1}^{m} \lim_{n \to \infty} P_{ik}^{n} P_{kj} = \sum_{k=1}^{m} \pi_{ik} P_{kj}$$

hence, we know

$$\lim_{m \to \infty} \pi_{ij} = \pi_{ij} \ge \lim_{m \to \infty} \sum_{k=1}^{m} \pi_{ik} P_{kj} = \sum_{k \in S} \pi_{ik} P_{kj}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 34 / 197

# Equality 1

## Equality

$$\pi_{ij} = \sum_{k \in S} \pi_{ik} P_{kj}$$

## **Proof**

Assume for contradiction  $\pi_{ij} > \sum_{k \in S} \pi_{ik} P_{kj}$ , then

$$\begin{split} \lim_{m \to \infty} \sum_{j=1}^{m} &> \lim_{m \to \infty} \sum_{j=1}^{m} \lim_{p \to \infty} \sum_{k=1}^{p} \pi_{ik} P_{kj} \\ &= \lim_{m \to \infty} \lim_{p \to \infty} \sum_{j=1}^{m} \sum_{k=1}^{p} \pi_{ik} P_{kj} \\ &= \lim_{m \to \infty} \lim_{p \to \infty} \sum_{k=1}^{p} \pi_{ik} \sum_{j=1}^{m} P_{kj} \\ &= \lim_{p \to \infty} \sum_{k=1}^{p} \pi_{ik} \lim_{m \to \infty} \sum_{j=1}^{m} P_{kj} \\ &= \lim_{p \to \infty} \sum_{k=1}^{p} \pi_{ik} \cdot 1 = \lim_{p \to \infty} \sum_{k=1}^{p} \pi_{ik} \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 36 / 197

# Proof (cont.)

- Since a value cannot be greater than itself, we got contradiction.
- In the 4th line, two lim can be switched because the value can only get larger when applying lim on it. not sure

Stochastic Process June 6, 2012

### Proof of theorem 1

- **Step 0**: existence of limiting probability.
- **Step 1**: existence of stationary probability distribution.
- Step 2: uniqueness.

## 0. Existence of limiting probability

#### Proof.

By lemma 1, we know that there exists a  $\pi_j$  for row i. Since the Markov chain is irreducible and all the states are positive recurrent, for any state i' other than i, we know that i' surely will visit i in finite steps. Therefore, the  $\pi_j$  value at row i' will equal to the  $\pi_j$  value at row i, which means that all the  $\pi_j$  for column j are the same, and is the limiting probability.  $\square$ 

still not clear enough

## 1. Existence of stationary probability distribution

We want to prove that

### **Target**

There's a vector  $s = (s_1, s_2, ...)$  such that

- $\sum_{i \in S} s_i = 1$
- $2 s \times P = s$

#### Proof.

By lemma 1, we know that there exists a  $\pi = (\pi_1, \pi_2, ...)$ . And by equality 1, we know that

$$(\pi_1, \pi_2, \ldots) \times P = (\pi_1, \pi_2, \ldots)$$

Hence  $\pi$  can satisfy the 2nd part of our target.

Then, we take  $k = \sum_{i \in S} \pi_i$ . By inequality 1, we know that  $k < \infty$ , and can get

$$\left(\frac{\pi_1}{k}, \frac{\pi_2}{k}, \ldots\right) \times P = \left(\frac{\pi_1}{k}, \frac{\pi_2}{k}, \ldots\right)$$

where  $\sum_{i \in S} \frac{\pi_i}{k} = 1$  also satisfy the 1st part of our target.

Therefore, this vector can be s, which means that it exists.

pishen (AlgoLab) Stochastic Process June 6, 2012 41 / 19

## 2. Uniqueness

### Target

If  $s = (s_1, s_2,...)$  is a stationary distribution of X, then  $s = \pi$ .

■ We'll prove this by inequality 3 & 4.

Stochastic Process June 6, 2012 42 / 197

# Inequality 3

## Inequality

$$s_j \geq \pi_j, \forall j \in S$$

#### Proof.

Let the distribution of X(0) be s, by the property of stationary distribution, we have

$$\begin{split} s_j &= P(X(n) = j) = \sum_{i \in S} P(X(n) = j | X(0) = i) P(X(0) = i) \\ &= \sum_{i \in S} P_{ij}^n \cdot s_i \\ &\geq \sum_{i = 1}^m P_{ij}^n \cdot s_i \\ &\Rightarrow s_j = \lim_{m \to \infty} \lim_{n \to \infty} s_j \\ &\geq \lim_{m \to \infty} \lim_{n \to \infty} \sum_{i = 1}^m P_{ij}^n \cdot s_i = \lim_{m \to \infty} \sum_{i = 1}^m \pi_j \cdot s_i = \pi_j \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 44 / 197

## Inequality 4

## Inequality

$$s_j \leq \pi_j, \forall j \in S$$

#### Proof.

Similar in the proof above,  $\forall m, n \geq 1$ , we have

$$s_{j} = \sum_{i \in S} P_{ij}^{n} \cdot s_{i}$$

$$\leq \sum_{i=1}^{m} P_{ij}^{n} \cdot s_{i} + \sum_{i=m+1}^{\infty} s_{i}$$

$$\Rightarrow s_{j} = \lim_{m \to \infty} \lim_{n \to \infty} s_{j}$$

$$\leq \lim_{m \to \infty} \lim_{n \to \infty} \left( \sum_{i=1}^{m} P_{ij}^{n} \cdot s_{i} + \sum_{i=m+1}^{\infty} s_{i} \right)$$

$$= \pi_{j}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 46 / 197

## An example Markov chain

### Example

$$P = \begin{pmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{pmatrix}, 0 < \alpha, \beta < 1$$
$$\pi = \left(\frac{\beta}{1 + \beta - \alpha}, \frac{1 - \alpha}{1 + \beta - \alpha}\right)$$

pishen (AlgoLab) Stochastic Process June 6, 2012 47 / 197

# Real world example: Hardy-Weinberg Law

### Example

There're two kinds of allele:

- dominant: A
- recessive: a

And three kinds of senotype with population proportion as follow:

- AA: p
- aa: *q*
- Aa: r = 1 (p + q)

### Example (cont.)

$$P = \begin{array}{cccc} AA & aa & Aa \\ AA & p + \frac{r}{2} & 0 & q + \frac{r}{2} \\ 0 & q + \frac{r}{2} & p + \frac{r}{2} \\ \frac{p}{2} + \frac{r}{4} & \frac{p}{2} + \frac{r}{4} & \frac{p+q+r}{2} \end{array}$$

we get  $\pi = (p, q, r)$  when

$$p = \left(p + \frac{r}{2}\right)^2$$

$$q = \left(q + \frac{r}{2}\right)^2$$

## Long-run proportion

#### Definition

We say that  $r_i$  is the *long-run proportion* of state  $i \in S$  if

$$r_j = \lim_{n \to \infty} \frac{1}{n} \sum_{1 \le t \le n} P_{ij}^t$$

holds for each state  $i \in S$ .

- It represents the average appearance times of state *i* in the whole process.
- We will show that (in theorem 3) if X is irreducible, then the long-run proportion of all states exist.

Stochastic Process June 6, 2012 50 / 197

### Theorem 2

### Theorem (type 1)

If  $r_i$  exists for each  $j \in S$  and  $\sum_{i \in S} r_i > 0$ , then  $r = (r_1, r_2, ...)$  is the unique stationary distribution of X.

or

### Theorem (type 2)

If  $r_i$  exists for each  $j \in S$  and a stationary distribution exists, then  $r = (r_1, r_2, ...)$  is the unique stationary distribution of X.

pishen (AlgoLab) Stochastic Process June 6, 2012 51 / 197

## **Proof**

### Existence of stationary distribution in type 1:

Let

$$R = \begin{pmatrix} r \\ r \\ \vdots \end{pmatrix} = \lim_{n \to \infty} \frac{1}{n} \sum_{1 \le t \le n} P^t$$

then

$$R \times P = \lim_{n \to \infty} \frac{1}{n} \sum_{1 \le t \le n} P^{t+1}$$
$$= \lim_{n \to \infty} \frac{1}{n} \sum_{1 \le t \le n} P^t + \lim_{n \to \infty} \frac{1}{n} (P^{n+1} - P)$$
$$= R$$

As stated later,  $\sum_{j \in S} r_j \le 1$ , hence by normalizing r, we prove that stationary distribution exist.

- $(\lim f(n)) \cdot g(n) = \lim f(n) \cdot g(n)?$
- can replace the proof on page 40?

pishen (AlgoLab) Stochastic Process June 6, 2012 52 / 197

# Proof (cont.)

#### **Uniqueness:**

Let  $\pi$  be an arbitrary stationary distribution, then

$$r = \pi \times R$$

$$= \pi \times \lim_{n \to \infty} \frac{1}{n} \sum_{1 \le t \le n} P^t$$

$$= \lim_{n \to \infty} \frac{1}{n} \sum_{1 \le t \le n} \pi \times P^t$$

$$= \lim_{n \to \infty} \frac{1}{n} \sum_{1 \le t \le n} \pi$$

$$= \pi$$

can replace the proof for page 42?

# Proof (cont.)

Prove that  $\sum_{j \in S} r_j \leq 1$ :

$$\sum_{j \in S} r_j = \lim_{m \to \infty} \sum_{j=1}^m \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^n P_{ij}^t$$

$$= \lim_{m \to \infty} \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^m P_{ij}^t$$

$$\leq \lim_{m \to \infty} \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^n \sum_{j \in S} P_{ij}^t$$

$$= \lim_{m \to \infty} \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^n 1 = 1$$

pishen (AlgoLab) Stochastic Process June 6, 2012 54 / 197

# Example 1

On a highway, if we know the probability that

- A truck is followed by a truck: 1/4
- A truck is followed by a car: 3/4
- A car is followed by a truck: 1/5
- $\blacksquare$  A car is followed by a car: 4/5

We can construct a matrix

$$\begin{array}{ccc}
T & C \\
T & (1/4 & 3/4) \\
C & (1/5 & 4/5)
\end{array}$$

and get the portion of trucks and cars on the whole highway as the eigenvector (4/19,15/19) (we will know that long-run proportion exists by Theorem 3).

## Example 2

For a system which has several good and bad states, we have a matrix P:

```
g_1 \ g_2 \ \cdots \ b_1 \ b_2 \ \cdots
g_1 \ g_2 \ \vdots \ b_1 \ b_2 \ \vdots
```

pishen (AlgoLab) Stochastic Process June 6, 2012 56 / 197

# Example 2 (cont.)

**Q1:** Breakdown rate (breakdown times / total time)
The long-run frequency of going to a bad state from a good state is

$$\sum_{i \in g} \sum_{j \in b} r_i P_{ij}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 57 / 197

# Example 2 (cont.)

**Q2:** The expected time  $\mu_G$  (resp.  $\mu_B$ ) of staying in good (resp. bad) states once we reach a good (resp. bad) state?

#### Ans:

For each t = 1, 2, ..., let  $G_t$  (resp.  $B_t$ ) be the length of the t-th good (resp. bad) phase of consecutive good (resp. bad) states. By the strong law of large numbers,

$$P\left(\lim_{t \to \infty} \frac{G_1 + B_1 + G_2 + B_2 + \dots + G_t + B_t}{t} = \mu_G + \mu_B\right) = 1$$

Since the reciprocal of above is the breakdown rate, we get equation (1):

$$P\left(\sum_{i\in G}\sum_{j\in B}\pi_i P_{ij} = \frac{1}{\mu_G + \mu_B}\right) = 1$$

pishen (AlgoLab) Stochastic Process June 6, 2012 58 / 197

# Example 2 (cont.)

Also, with probability 1, we get equation (2):

$$P\left(\sum_{i \in G} r_i = \lim_{t \to \infty} \frac{G_1 + G_2 + \dots + G_t}{G_1 + B_1 + \dots + G_t + B_t} = \frac{\mu_G}{\mu_G + \mu_B}\right) = 1$$

Then, by (2)/(1), we get that

$$P\left(\mu_G = \frac{\sum_{i \in G} r_i}{\sum_{i \in G} \sum_{j \in B} r_i P_{ij}}\right) = 1$$

 $\lim \frac{f(n)}{g(n)} = \frac{\lim f(n)}{\lim g(n)}?$ 

pishen (AlgoLab) Stochastic Process June 6, 2012 59 / 197

## Theorem 3

#### **Theorem**

If X is irreducible, then the long-run proportion  $r_i$  exists with probability 1, moreover,

- If state i is positive recurrent (i.e.  $0 < \mu_i < \infty$ ), then  $P(r_i = \frac{1}{u_i}) = 1$ .
- 2 If state i is null recurrent (i.e.  $\mu_i = \infty$ ) or transient, then  $P(r_i = 0) = 1$ .

where  $\mu_i$  is the expected return time of state i

## Proof

#### Part 1:

Suppose X(0)=i,  $T_k$  is the number of steps required for the k-th i goes to (k+1)-st i, then by the strong law of large number,

$$P\left(\lim_{k\to\infty} \frac{T_1 + T_2 + \dots + T_k}{k} = \mu_i\right) = 1$$

$$\Rightarrow P\left(r_i = \lim_{k\to\infty} \frac{k}{T_1 + T_2 + \dots + T_k} = \frac{1}{\mu_i}\right) = 1$$

 $lim(A/B) = \frac{1}{\lim(B/A)}?$ 

pishen (AlgoLab) Stochastic Process June 6, 2012 61 / 197

# Proof (cont.)

#### Part 2:

f I If i is transient, i will only appear finite times in the long-run, hence

$$r_i = \frac{finite}{\infty} = 0$$

**2** If *i* is null recurrent,  $\mu_i$  is ∞, then

$$P\left(\lim_{k\to\infty}\frac{T_1+T_2+\cdots+T_k}{k}=\infty\right)=1$$

$$P\left(r_i = \lim_{k \to \infty} \frac{k}{T_1 + T_2 + \dots + T_k} = 0\right) = 1$$

(The first equation is not promised by the strong law of large number. But if it's not  $\infty$ , we can say that  $\mu_i$  is not  $\infty$ , which is a contradiction.)

pishen (AlgoLab) Stochastic Process June 6, 2012 62 / 197

## Example 1

### Example (type 1)

If X is **irreducible** and finite, then X has no null recurrent states.

### Example (type 2)

If X is finite, then X has no null recurrent states.

Finite irreducible imply positive recurrent.

Stochastic Process June 6, 2012 63 / 197

## Proof

### Type 1:

If there's a state which is null recurrent, by irreducible, all the states will be null recurrent. Then, all states have  $P(r_i = 0) = 1$ . By changing the proof in page 54 into finite states version, we know that  $\sum r_i = 1$ . So it's impossible for finite  $r_i$ , which are all close to 0, to sum up to 1.

#### Type 2:

If it's not irreducible, the finite set of communicated null recurrent states still form an irreducible and finite Markov chain, which can fit the requirement of type 1.

Stochastic Process June 6, 2012 64 / 197

## Example 2

### Example

In the drunken man problem with infinite states, no state will be positive recurrent.

Infinite drunken man imply no positive recurrent. Note that it doesn't mean all infinite irreducible Markov chain has no positive recurrent state.

pishen (AlgoLab) Stochastic Process June 6, 2012 65 / 197

## Proof

If all the states are positive recurrent, then by theorem 3, we know that all the  $r_i>0$  and is a finite value. Since each state of drunken man problem has the same structure, all the  $r_i$  has same value. We then set  $r=\epsilon \cdot r_i$   $(0<\epsilon<1)$  such that  $r_i>r>0, \forall i$ . And get

$$\sum_{i \in S} r_i > \sum_{i \in S} r = \infty > 1$$

which is contradiction to page 54.

## Example 3: Poisson Hotel

#### Example

There's a hotel, with N representing the number of newly occupied rooms each day (N is a poisson distribution with parameter  $\lambda$ ). And the number of consecutive check-in days of each room is a geometric distribution with probability p (p is the probability of check-out). X(t) is the number of occupied rooms in day t.

pishen (AlgoLab) Stochastic Process June 6, 2012 67 / 19

# **Q1:** $P_{ij} = ?$

We set  $R_i$  as a binomial distribution with parameter (i, 1-p), which represents the number of rooms which will remain occupied in the next day, then

$$\begin{split} P_{ij} &= P(R_i + N = j) \\ &= \sum_{k \geq 0} P(R_i + N = j | R_i = k) P(R_i = k) \\ &= \sum_{k \geq 0} P(N = j - k) P(R_i = k) \\ &= \sum_{0 \leq k \leq \min(i,j)} \frac{e^{-\lambda} \cdot \lambda^{j-k}}{(j-k)!} \binom{i}{k} (1-p)^k p^{1-k} \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 68 / 197

# **Q2:** $r_i = ?$

We guess (by a dream?) there's a stationary distribution which is a poisson distribution with parameter  $\lambda_0$ . Setting X(0) with this distribution. And let R as the number of rooms in X(0) which remain check-in in the next day (R is a poisson distribution with parameter  $\lambda_0(1-p)$ ). X(1) will have distribution R+N, which is a poisson distribution with parameter  $\lambda_0(1-p)+\lambda$ . Then since X(0) is a stationary distribution, it will have the same distribution with X(1), which means that  $\lambda_0=\lambda_0(1-p)+\lambda$ , and we get  $\lambda_0=\lambda/p$ . After getting  $r_i$ , we get that with probability 1,

$$\mu_i = \frac{1}{P(X(0) = i)} = \frac{i!}{e^{-\lambda/p} \cdot (\lambda/p)^i}$$

not clear enough

pishen (AlgoLab) Stochastic Process June 6, 2012 69 / 197

## Corollary of theorem 2 & 3

### Corollary

If X is irreducible, then

X is positive recurrent  $\iff X$  admits a stationary distribution.

Stochastic Process June 6, 2012 70 / 197

# Moving to transient states

For transient states i and j, we define the following:

1 Expected steps in a transient state:

#### Definition

E is a matrix where  $E_{ij}$  is the expected number of steps t with X(t) = jwhen X(0) = i.

2 Probability of reaching a transient state:

#### Definition

F is a matrix where

$$F_{ij} = P(X(t) = j \text{ for some } t \ge 1 | X(0) = i)$$

pishen (AlgoLab) Stochastic Process June 6, 2012

# Computing E & F

#### **Theorem**

For a Markov chain X consisting finite transient states,

$$E = (I - T)^{-1}$$

where I is an identity matrix, T is the induced matrix of P by all the transient states in P. Moreover,

$$F_{ij} = rac{E_{ij} - \delta_{ij}}{E_{jj}}$$
 ,where  $\delta_{ij} = egin{cases} 1 & ext{if } i = j \ 0 & ext{if } i 
eq j \end{cases}$ 

pishen (AlgoLab) Stochastic Process June 6, 2012 72 / 197

Conditioned on X(1), we have

$$E_{ij} = \underbrace{\delta_{ij}}_{\mathsf{step}=0} + \underbrace{\sum_{k} P_{ik} \cdot E_{kj}}_{\mathsf{step} \geq 1} = \delta_{ij} + \sum_{k} T_{ik} \cdot E_{kj}$$

The 2nd equation works since the process will not go back to transient state once it enter a recurrent state. Then, we have

$$I \times E = E = I + T \times E$$

$$\implies (I - T) \times E = I$$

$$\implies E = (I - T)^{-1}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 73 / 197

## Proof (cont.)

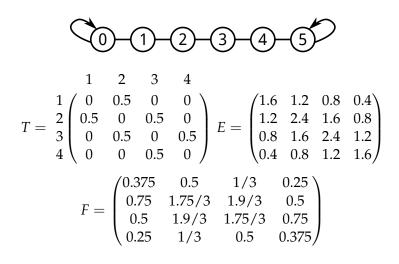
Conditioned on whether or not X(t) = j holds for some  $t \ge 1$ , we have

$$E_{ij} = \underbrace{\delta_{ij}}_{\text{step}=0} + \underbrace{F_{ij} \cdot E_{jj}}_{\text{steps} \ge \text{the first } j}$$

therefore,

$$F_{ij} = \frac{E_{ij} - \delta_{ij}}{E_{jj}}$$

#### Example: Gambler's ruin



## Branching process

In the beginning, there're X(0) life forms, each life form has probability  $p_i$  of becoming i life forms in the next step.

- state 0 is recurrent (absorbing).
- if  $p_0 > 0$ , all other states (1, 2, ...) are transient since  $P(X(t+1) = 0|X(t) = i) = p_0^i > 0$

We'll show that

$$E[X(n)] = \mu^n \cdot X(0)$$

where

$$\mu = \sum_{j \ge 1} j \cdot p_j = E[Z_k]$$

and  $Z_k$  is the number of offspring of the k-th life form, all  $Z_k$  are i.i.d.

$$\begin{split} E[X(n)] &= E[E[X(n)|X(n-1)]] \\ &= E\left[E\left[\sum_{k=1}^{X(n-1)} Z_k | X(n-1)\right]\right] \\ &= E[X(n-1) \cdot \mu] \\ &= \mu \cdot E[X(n-1)] \\ &= \mu^n \cdot X(0) \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 77 / 197

## Probability of extinction

#### **Definition**

 $e_i$  is the probability of extinction when X(0) = i.

**Case 1:**  $\mu < 1$ 

$$\begin{aligned} 1 - e_i &= \lim_{n \to \infty} P(X(n) \ge 1 | X(0) = i) \\ &= \lim_{n \to \infty} \sum_{j \ge 1} P(X(n) = j | X(0) = i) \\ &\le \lim_{n \to \infty} \sum_{j \ge 1} j \cdot P(X(n) = j | X(0) = i) \\ &= \lim_{n \to \infty} E[X(n) | X(0) = i] \\ &= \lim_{n \to \infty} \mu^n \cdot i = 0 \end{aligned}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 78 / 197

# Probability of extinction (cont.)

**Case 2:** 
$$\mu \ge 1$$

$$e_2 = e_1^2$$
,  $e_3 = e_2 \cdot e_1$ , ...
 $e_1 = P(\mathsf{extinct}|X(0) = 1)$ 
 $= \sum_{j \ge 0} P(\mathsf{extinct}|X(1) = j) \cdot P_{1j}$ 
 $= \sum_{j \ge 0} e_j \cdot p_j$ 
 $= \sum_{j \ge 0} e_1^j \cdot p_j$ 

We then solve the above equation to get  $e_1$ .

### Example

$$p_0 = p_1 = 0.25, \quad p_2 = 0.5$$
  
 $\implies \mu = 1 \cdot 0.25 + 2 \cdot 0.5 > 1$   
 $\implies e_1 = e_1^0 \cdot 0.25 + e_1^1 \cdot 0.25 + e_1^2 \cdot 0.5$   
 $\implies e_1 = \{1/2, 1\}$ 

Since  $\mu>1$ , we know  $\lim_{n\to\infty} E[X(n)]=\infty$ . But if  $e_1=1$ , we have  $\lim_{n\to\infty} P(X(n)=0)=1$ , which would not make  $\lim_{n\to\infty} E[X(n)]=\infty$ , hence  $e_1\neq 1$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 80 / 197

#### Reversed Markov chain

#### Definition

Let X (resp. Y) be a Markov chain with matrix P (resp. Q). We say that  $\mathbb{Y}$  is the reversed chain of  $\mathbb{X}$  if there exists a stationary distribution  $\pi$  of  $\mathbb{X}$  such that

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

holds for all states  $i, j \in S$ .

Stochastic Process June 6, 2012 81 / 197

#### Observation 1

#### Observation

The reversed sequence  $\mathbb{Y}$  of  $\mathbb{X}$  is a Markov chain.

pishen (AlgoLab) Stochastic Process June 6, 2012 82 / 197

$$\begin{split} &P(Y(n)=i_0|Y(n-1)=i_1,Y(n-2)=i_2,\ldots,Y(n-k)=i_k)\\ &=P(X(n)=i_0|X(n+1)=i_1,X(n+2)=i_2,\ldots,X(n+k)=i_k)\\ &=\frac{P(X(n)=i_0,X(n+1)=i_1,\ldots,X(n+k)=i_k)}{P(X(n+1)=i_1,\ldots,X(n+k)=i_k)}\\ &=\frac{P(X(n)=i_0)\cdot P(X(n+1)=i_1|X(n)=i_0)\cdot P_{i_1i_2}\cdot \cdot \cdot P_{i_{k-1}i_k}}{P(X(n+1)=i_1)\cdot P_{i_1i_2}\cdot \cdot \cdot P_{i_{k-1}i_k}}\\ &=\frac{P(X(n)=i_0,X(n+1)=i_1)}{P(X(n+1)=i_1)}\\ &=P(X(n)=i_0|X(n+1)=i_1)\\ &=P(Y(n)=i_0|Y(n-1)=i_1) \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 83 / 197

#### Observation 2

#### Observation

If  $\mathbb Y$  is the reversed sequence of Markov chain  $\mathbb X$  and  $\pi$  is a stationary distribution of  $\mathbb X$ , then

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

holds for all  $i, j \in S$ , where Q is the transition matrix of Y.

pishen (AlgoLab) Stochastic Process June 6, 2012 84 / 19

Let X and Y have distribution  $\pi$ 

$$\pi_{i} \cdot Q_{ij} = P(Y(n-1) = i) \cdot P(Y(n) = j | Y(n-1) = i)$$

$$= P(Y(n-1) = i, Y(n) = j)$$

$$= P(Y(n-1) = i | Y(n) = j) \cdot P(Y(n) = j)$$

$$= P(X(n+1) = i | X(n) = j) \cdot P(X(n) = j) = \pi_{j} \cdot P_{ji}$$

Stochastic Process June 6, 2012 85 / 197

#### Observation

Let P (resp. Q) be the transition matrix of X (resp. Y), if vector  $\pi$  satisfy the following

- $\pi_i \geq 0 \quad \forall i \in S$
- $\pi_i \cdot Q_{ij} = \pi_j \cdot P_{ji} \quad \forall i, j \in S$

then  $\mathbb{Y}$  is the reversed sequence of  $\mathbb{X}$ .

- The long-run proportion of  $i \to j$  in the sequence of Y is equal to the long-run proportion of  $i \to i$  in the sequence of X.
- Reversed Markov chain is the reversed sequence.

Stochastic Process June 6, 2012 86 / 197 From the third property, we have

$$\sum_{j \in S} \pi_i \cdot Q_{ij} = \pi_i = \sum_{j \in S} \pi_j \cdot P_{ji} \quad \forall i \in S$$

From the 2nd equation, we know that  $\pi \times P = \pi$ , hence  $\pi$  is a stationary distribution of  $\mathbb{X}$ .

Then by observation 2, we know that for any  $\pi$ , there's a reversed sequence Y', whose transition matrix Q' satisfy

$$\pi_i \cdot Q'_{ij} = \pi_j \cdot P_{ji} \quad \forall i, j \in S$$

hence  $\mathbb{Y} = \mathbb{Y}'$ , which is a reversed sequence of  $\mathbb{X}$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 87 / 197

## Example: Bulb's life



There's a room which need to be lighted by one bulb, when the bulb in use fails, it will be replaced by a new one on next day.

- X(n) = i if the bulb in use on day n is in its ith day of use.
- *L* is a random variable representing the lifetime of a bulb.

We want to know the stationary probability  $\pi_i$  of state i.

pishen (AlgoLab) Stochastic Process June 6, 2012 88 / 197

### Example: Bulb's life (cont.)

X is a irreducible, positive recurrent, aperiodic Markov chain which has the sequence like this:

$$1, 2, 3, 1, 2, 3, 4, 5, 1, 1, 2, 1, 2, 3, 4, \dots$$

We know that

$$P_{i1} = P(\text{buld, on its } i \text{th day of use, fails}) = \frac{P(L=i)}{P(L \geq i)} = 1 - P_{i(i+1)}$$

And the expected return time of state 1 is E[L], which means that the long-run proportion of state 1 is 1/E[L] by page 60.

pishen (AlgoLab) Stochastic Process June 6, 2012 89 / 197

## Example: Bulb's life (cont.)

Take  $\mathbb{Y}$  (with matrix Q) as the reversed chain of  $\mathbb{X}$ , we know that for all  $i \in S$ ,

- $Q_{(i+1)i} = 1$
- $Q_{1i} = P(L = i)$
- $\blacksquare \ \pi_1 \cdot Q_{1i} = \pi_i \cdot P_{i1}$

Hence,

$$\pi_i = \frac{\pi_1 \cdot Q_{1i}}{P_{i1}} = \frac{P(L=i) \cdot P(L \ge i)}{E[L] \cdot P(L=i)} = \frac{P(L \ge i)}{E[L]}$$

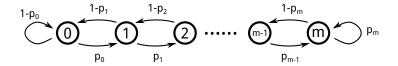
#### Time-reversible

#### Definition

X is *time-reversible* if X is the reversed chain of X.

pishen (AlgoLab) Stochastic Process June 6, 2012 91 / 197

### Example: Reversed drunken man



- $0 < p_0 \le 1$
- $0 \le p_m < 1$
- $0 < p_i < 1 \forall i = 1, ..., m-1$

The long-run proportion of transition  $i \to i+1$  and  $i+1 \to i$  are the same, since one must go back to i from i+1 in order to go to i+1 from i. Hence the drunken man problem is time-reversible.

Stochastic Process June 6, 2012 92 / 197

## Example: Reversed drunken man (cont.)

$$\pi_{0} \cdot p_{0} = \pi_{1} \cdot (1 - p_{1})$$

$$\pi_{1} \cdot p_{1} = \pi_{2} \cdot (1 - p_{2})$$

$$\vdots$$

$$\pi_{m-1} \cdot p_{m-1} = \pi_{m} \cdot (1 - p_{m})$$

Thus,

$$\pi_{1} = \pi_{0} \cdot p_{0} / (1 - p_{1})$$

$$\pi_{2} = \pi_{1} \cdot p_{1} / (1 - p_{2})$$

$$\vdots$$

$$\pi_{m} = \pi_{m-1} \cdot p_{m-1} / (1 - p_{m})$$

# Example: Reversed drunken man (cont.)

$$\pi_{i} = \underbrace{\frac{\prod_{j=0}^{i-1} p_{j}}{\prod_{j=1}^{i} (1 - p_{j})}}_{q_{i}} \cdot \pi_{0} \quad \forall i = 1, \dots m$$

$$\Longrightarrow \pi_{0} + \sum_{i=1}^{m} \pi_{i} = 1 = \pi_{0} + \sum_{i=1}^{m} q_{i} \cdot \pi_{0}$$

$$\Longrightarrow \pi_{0} = \frac{1}{1 + \sum_{i=1}^{m} q_{i}}$$

$$\Longrightarrow \pi_{k} = \frac{q_{k}}{1 + \sum_{i=1}^{m} q_{i}} \quad \forall k = 0, 1, \dots m$$

pishen (AlgoLab) Stochastic Process June 6, 2012 94 / 197

### Example: Two bukkits of balls

There're two bukkits contain total m balls.

In each step, we randomly choose one ball and put it in another bukkit. Let X(n) represent the number of balls in the first bukkit, it's the Markov chain of previous example with

$$p_0 = 1$$
,  $p_m = 0$ ,  $p_i = \frac{m-i}{m}$   $\forall i = 1, ..., m-1$ 

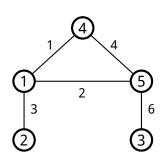
We can get that

$$q_{i} = \frac{\prod_{j=0}^{i-1} \frac{m-j}{m}}{\prod_{j=1}^{i} \frac{j}{m}} = \frac{\prod_{j=0}^{i-1} m - j}{\prod_{j=1}^{i} j} = \binom{m}{i} \quad \forall i = 1, \dots m$$

$$\implies \pi_{0} = \frac{1}{1 + \sum_{i=1}^{m} \binom{m}{i}} = \frac{1}{2^{m}} \implies \pi_{k} = \frac{\binom{m}{k}}{2^{m}} \quad \forall k = 0, 1, \dots m$$

pishen (AlgoLab) Stochastic Process June 6, 2012 95 / 197

### Example: A random walk



$$P_{ij} = \frac{w(i,j)}{\sum_{k} w(i,k)}$$

where w(a,b) is the weight of edge (a,b). To make it as a time-reversible chain, we let

$$\pi_i = \frac{\sum_k w(i, k)}{\sum_{\ell} \sum_k w(\ell, k)} \quad \forall i$$

We can see that

$$\pi_i \cdot P_{ij} = \pi_j \cdot P_{ji} \quad \forall i, j$$

## Hastings-Metropolis sampling algorithm

Design an irreducible Markov chain  $\mathbb X$  such that the unique stationary distribution of  $\mathbb X$  is the distribution of random variable Y. Since the long-run proportion of state i is P(Y=i),

$$\lim_{n \to \infty} \frac{X(1) + X(2) + \dots + X(n)}{n} = \sum_{i \in S} i \cdot P(Y = i) = E[Y] = \mu$$

While computing  $\mu$  by the law of large number is difficult (hard to sample on Y), we use this alternative method to compute  $\mu$  by generating a sequence of  $\mathbb{X}$ , which is sometime easier.

pishen (AlgoLab) Stochastic Process June 6, 2012 97 / 197

# Hastings-Metropolis sampling algorithm (cont.)

There's a random variable Y such that

$$P(Y=i) = \frac{b_i}{C}$$

for some unknown (or intractable)  $C = \sum_{i \in S} b_i$ . We then design a Markov chain X that

- $P_{ii} = Q_{ii} + \sum_{k \in S, k \neq i} Q_{ik} \cdot (1 q_{ik})$
- $P_{ij} = Q_{ij} \cdot q_{ij} \quad \forall j \neq i$

#### where

- Q is the transition matrix of an arbitrary irreducible Markov chain X which has the same state space as Y.
- q is a matrix to be determined later.

# Hastings-Metropolis sampling algorithm (cont.)

For n = 0, 1, ...

- 1 If X(n) = i, set Z such that  $P(Z = j) = Q_{ii} \quad \forall j \in S$ .
- 2 If Z = j, set X(n+1) such that
  - $P(X(n+1) = j) = q_{ii}$
  - $P(X(n+1)=i)=1-q_{ii}$

One can see that this satisfies the requirement on previous page.

# Hastings-Metropolis sampling algorithm (cont.)

Then, we let

$$q_{ij} = \min\left(\frac{b_j \cdot Q_{ji}}{b_i \cdot Q_{ij}}, 1\right)$$

$$\implies b_i \cdot Q_{ij} \cdot q_{ij} = b_j \cdot Q_{ji} \cdot q_{ji}$$

$$\implies \frac{b_i}{C} \cdot P_{ij} = \frac{b_j}{C} \cdot P_{ji}$$

By observation 3 on page 86, we know that  $(b_1/C, b_2/C,...)$  is the stationary distribution of X.

### Example: Space of permutations

#### Example

Let S consist of all the permutations  $(x_1, x_2, ..., x_n)$  of  $\{1, 2, ..., n\}$  that

$$\sum_{k=1}^{n} k \cdot x_k \ge \frac{n^3}{4}$$

- This is same as Y in page 98 with C = |S| and  $b_i = 1 \ \forall i$ .
- S is hard to compute.
- We need to design a matrix Q such that when given a permutation x, it's efficient to compute the value of  $Q_{xy} \forall y \in S$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 101 / 197

## Example: Space of permutations (cont.)

We let

$$Q_{xy} = \frac{1}{N(x)}$$
 , if  $y$  can be obtained from  $x$  by one swap

where N(x) is the number of permutations that can be obtained from x by one swap. For example:

$$\underbrace{(1,2,3,4,5)}_{y} \leftrightarrow \underbrace{(1,3,2,4,5)}_{x} \leftrightarrow \underbrace{(1,3,4,2,5)}_{y}$$

This chain is irreducible since each  $x \in S$  can go to  $(x_1, x_2, ..., x_n)$ , where  $x_1 \le x_2 \le ... \le x_n$ , by several swaps.

Also, given a x, finding all the obtainable y can be done efficiently.

pishen (AlgoLab) Stochastic Process June 6, 2012 102 / 197

#### Counting process

#### Definition

A collection  $\mathbb{N}$  of random variables is a *counting process* if N(t) denotes the total number of events that occur by time t.

- $\blacksquare N(t)$  is a nonnegative integer.
- The value of N(t) is increasing as t increase.
- N(t) N(s) is the number of events that occur between time index sand t. where t > s.

Stochastic Process June 6, 2012 103 / 197

#### Two properties

#### Independent increments:

#### Definition

A counting process is *independent increments* if the number of events in two non-overlapping time intervals are independent.

■ For example, N(s) - N(0) and N(s+t) - N(s) are independent.

#### Stationary increments:

#### **Definition**

A counting process is *stationary increments* if the number of events in any time interval depends only on the length of the interval.

■ For example,  $P(N(s_1 + t) - N(s_1) = k) = P(N(s_2 + t) - N(s_2) = k)$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 104 / 197

### Poisson process

#### **Definition**

A Poisson process with rate  $\lambda$  is a counting process with independent increments and stationary increments such that

$$P(N(s+t) - N(s) = n) = \frac{e^{-\lambda t} \cdot (\lambda t)^n}{n!}$$

holds for all nonnegative integers.

- N(s+t) N(s) is Poisson distributed with parameter  $\lambda t$ .
- The average number of events that occur in an unit time interval (t=1) is  $\lambda$  (since the expectation value of Poisson distribution with parameter  $\lambda$  is  $\lambda$ .)

Stochastic Process June 6, 2012 105 / 197

## An operational definition

#### **Theorem**

Let  $\mathbb N$  be a counting process with independent increments and stationary increments. Then  $\mathbb N$  is a Poisson process if and only if the following two conditions hold:

- $P(N(t) = 1) = \lambda \cdot t + o(t)$
- $P(N(t) \ge 2) = o(t)$ 
  - We say that f(t) = o(t) if

$$\lim_{t \to 0} \frac{f(t)}{t} = 0$$

#### Proof

 $(\Longrightarrow)$ :

Since N(t) is Poisson distributed with parameter  $\lambda t$ ,

$$P(N(t) = 1) = \frac{(\lambda t) \cdot e^{-\lambda t}}{1!} = \lambda t \cdot \left(1 - \frac{\lambda t}{1!} + \frac{(\lambda t)^2}{2!} - \cdots\right)$$
$$= \lambda t - \lambda^2 t^2 + \cdots$$
$$= \lambda t + o(t)$$

$$P(N(t) = 2) = \frac{(\lambda t)^2 \cdot e^{-\lambda t}}{2!} = \frac{(\lambda t)^2}{2!} \cdot \left(1 - \frac{\lambda t}{1!} + \frac{(\lambda t)^2}{2!} - \cdots\right)$$
  
=  $o(t)$ 

One can prove that P(N(t)=k)=o(t) for all  $k\geq 2$ , hence  $P(N(t)\geq 2)=o(t)$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 107 / 197

## Proof (cont.)

(⇐=):

The Laplace transform of a random variable X is

$$\phi(u) = E[e^{-u \cdot X}]$$

We say that two random variables have the same distribution if their Laplace transform are the same.

And if X is Poisson distributed with parameter  $\lambda t$ , then

$$E[e^{-u \cdot X}] = e^{(e^{-u} - 1) \cdot \lambda t}$$

We define  $\phi_u(t) = E[e^{-u \cdot N(t)}]$ , then we know that

$$\begin{aligned} \phi_{u}(s+t) &= E[e^{-u \cdot N(s+t)}] \\ &= E[e^{-u \cdot (N(s) - N(0))} e^{-u \cdot (N(s+t) - N(s))}] \\ &= E[e^{-u \cdot N(s)}] \cdot E[e^{-u \cdot (N(s+t) - N(s))}] \\ &= E[e^{-u \cdot N(s)}] \cdot E[e^{-u \cdot N(t)}] \\ &= \phi_{u}(s) \cdot \phi_{u}(t) \end{aligned}$$

The 3rd equation is because two independent random variables X and Y will make

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

pishen (AlgoLab) Stochastic Process June 6, 2012 109 / 197

By the two conditions in page 106, we know

$$P(N(t) = 0) = 1 - \lambda t + o(t)$$

Therefore,

$$\phi_{u}(t) = E[e^{-u \cdot N(t)}]$$

$$= e^{-u \cdot 0} \cdot (1 - \lambda t + o(t)) + e^{-u \cdot 1} \cdot (\lambda t + o(t))$$

$$+ (e^{-u \cdot 2} + e^{-u \cdot 3} + \cdots) \cdot o(t)$$

$$= 1 - \lambda t + e^{-u} \cdot \lambda t + o(t)$$

$$= 1 + (e^{-u} - 1) \cdot \lambda t + o(t)$$

And

$$\phi_u(s+t) = \phi_u(s) \cdot \phi_u(t) = \phi_u(s) \cdot (1 + (e^{-u} - 1) \cdot \lambda t + o(t))$$

pishen (AlgoLab) Stochastic Process June 6, 2012 110 / 197

Differentiate on  $\phi_u(s)$ , we can get

$$\phi_{u}'(s) = \lim_{t \to 0} \frac{\phi_{u}(s+t) - \phi_{u}(s)}{t} = \lim_{t \to 0} (\phi_{u}(s) \cdot (e^{-u} - 1) \cdot \lambda + o(t))$$
$$= \phi_{u}(s) \cdot (e^{-u} - 1) \cdot \lambda$$

By  $\frac{\phi_{u}'(s)}{\phi_{u}(s)} = (e^{-u} - 1) \cdot \lambda$ , we have

$$\ln \phi_u(s) = \int (e^{-u} - 1) \cdot \lambda \, ds = (e^{-u} - 1) \cdot \lambda s + C$$

By  $\phi_u(0) = 1$  and  $\ln 1 = 0$ , we know C = 0, hence

$$\phi_u(s) = e^{(e^{-u}-1)\cdot \lambda s} \quad \forall s, u$$

which means that N(s) is Poisson distributed for all s.

pishen (AlgoLab) Stochastic Process June 6, 2012 111 / 197

### Inter-arrival time

#### Definition

The kth inter-arrival time  $T_k$  of  $\mathbb N$  is the time interval between the (k+1)st and kth events.

 $\mathbb{T} = T_1, T_2, \dots$  is the sequence of inter-arrival times of  $\mathbb{N}$ .

• 0th event arrives at time 0.

pishen (AlgoLab) Stochastic Process June 6, 2012 112 / 19

## Observation 1: Independent & exponential distributed

#### Observation

If  $\mathbb N$  is a Poisson process with rate  $\lambda$ , then each  $T_k$  is an independent exponential distribution with parameter  $\lambda$ .

#### Proof:

The cumulative distribution function of  $T_1$  is

$$F_1(s) = P(T_1 \le s)$$
= 1 - P(T\_1 > s)  
= 1 - P(N(s) = 0)  
= 1 - e^{-\lambda s}

The 3rd equation is because  $T_1 > s \iff N(s) = 0$ . We can observe that  $T_1$  is expenential distributed.

$$P(T_2 > t | T_1 = s) = P(N(T_1 + t) - N(T_1) = 0 | T_1 = s)$$

$$= P(N(T_1 + t) - N(T_1) = 0)$$

$$= P(N(t) = 0)$$

$$= e^{-\lambda t}$$

The equations are derived by stationary increments.

Thus,  $T_2$  is also exponential distributed with parameter  $\lambda$ .

And  $T_1$ ,  $T_2$  are independent.

One can prove for  $T_k$  with  $k \ge 3$  by the same approach.

pishen (AlgoLab) Stochastic Process June 6, 2012 114 / 197

# Observation 2: Waiting time is gamma distributed

#### Observation

The waiting time  $S_k = T_1 + T_2 + \cdots + T_k$  of the kth event is gamma distributed with parameter  $(k, \lambda)$ . check gamma distribution

■ The probability density function is

$$f(t) = \lambda e^{-\lambda t} \cdot \frac{(\lambda t)^{k-1}}{(k-1)!}$$

■ It's also called Erlang distribution since  $k \in \mathbb{Z}^+$ .

### Verification

$$P(S_k \le t) = P(N(t) \ge k) = \sum_{i \ge k} P(N(t) = i) = \sum_{i \ge k} \frac{(\lambda t)^i \cdot e^{-\lambda t}}{i!}$$

So,

$$\frac{dP(S_k \le t)}{dt} = \sum_{i \ge k} \frac{\lambda \cdot (\lambda t)^{i-1} \cdot e^{-\lambda t}}{(i-1)!} - \sum_{i \ge k} \frac{(\lambda t)^i \cdot e^{-\lambda t} \cdot \lambda}{i!}$$
$$= \frac{\lambda \cdot (\lambda t)^{k-1} \cdot e^{-\lambda t}}{(k-1)!}$$

pishen (AlgoLab) Stochastic Process June 6, 2012

116 / 197

# Property 1 (Different types of events)

### Property

Let  $\mathbb N$  be a Poisson process with rate  $\lambda$ , each event is classified as type 1 with probability p or type 2 with probability 1-p.

Then the arrival of type 1 and type 2 events are both Poisson processes with rate  $p \cdot \lambda$  and  $(1-p) \cdot \lambda$ . And the two processes are independent.

- Let  $\mathbb{N}_1$  be the process of type 1 event,  $N_1(k)$  is the number of type 1 events that occur by time k. (same for  $\mathbb{N}_2$ )
- $\mathbb{N}_1$  and  $\mathbb{N}_2$  are said to be independent if  $N_1(s_1+t_1)-N_1(s_1)$  and  $N_2(s_2+t_2)-N(s_2)$  are independent for all  $s_1,t_1,s_2,t_2$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 117 / 19

### Proof

Here we prove that  $\mathbb{N}_1$  is a Poisson process with rate  $\lambda p$ .

#### **Stationary increments:**

$$P(N_1(s+t) - N_1(s) = k_1 | N(s+t) - N(s) = k) = \binom{k}{k_1} \cdot p^{k_1} \cdot (1-p)^{k-k_1}$$

Therefore,

$$P(N_1(s+t) - N_1(s) = k_1) = \sum_{k>0} {k \choose k_1} \cdot p^{k_1} \cdot (1-p)^{k-k_1} \cdot \frac{(\lambda t)^k \cdot e^{\lambda t}}{k!}$$

which has nothing to do with s.

pishen (AlgoLab) Stochastic Process 118 / 197

#### Independent increments:

Let (s, s + t) and (u, u + v) be two non-overlapping time intervals,

$$\begin{split} &P(N_1(s+t)-N_1(s)=k_1,N_1(u+v)-N_1(u)=\ell_1)\\ &=\sum_{k\geq 0}\sum_{\ell\geq 0}P(N_1(s+t)-N_1(s)=k_1,N_1(u+v)-N_1(u)=\ell_1\\ &|N(s+t)-N(s)=k,N(u+v)-N(u)=\ell)\\ &\cdot P(N(s+t)-N(s)=k,N(u+v)-N(u)=\ell)\\ &=\sum_{k\geq 0}\sum_{\ell\geq 0}P(N_1(s+t)-N_1(s)=k_1|N(s+t)-N(s)=k)\\ &\cdot P(N_1(u+v)-N_1(u)=\ell_1|N(u+v)-N(u)=\ell)\\ &\cdot P(N(s+t)-N(s)=k)\cdot P(N(u+v)-N(u)=\ell)\\ &=P(N_1(s+t)-N_1(s)=k_1)\cdot P(N_1(u+v)-N_1(u)=\ell_1) \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 119 / 197

#### Conditions on page 106:

1. 
$$P(N_1(t) \ge 2) \le P(N(t) \ge 2) = o(t)$$
  
2.  $P(N_1(t) = 1) = P(N_1(t) = 1|N(t) = 1) \cdot P(N(t) = 1) + P(N_1(t) = 1|N(t) \ge 2) \cdot P(N(t) \ge 2)$   
 $= p \cdot (\lambda t + o(t)) + o(t)$   
 $= p\lambda t + o(t)$ 

Hence we know that  $\mathbb{N}_1$  is a Poisson process with rate  $\lambda p$ . Seems like we can also derive this from the result of next page and omit this page's proof? (By Example 3.23 on textbook?)

pishen (AlgoLab) Stochastic Process June 6, 2012 120 / 197

#### $\mathbb{N}_1$ and $\mathbb{N}_2$ are independent:

$$P(N_{1}(t) = i, N_{2}(t) = j)$$

$$= P(N_{1}(t) = i, N_{2}(t) = j | N(t) = i + j) \cdot P(N(t) = i + j)$$

$$= {i + j \choose i} \cdot p^{i} \cdot (1 - p)^{j} \cdot \frac{e^{-\lambda t} \cdot (\lambda t)^{i+j}}{(i + j)!}$$

$$= \frac{e^{-\lambda pt} \cdot (\lambda pt)^{i}}{i!} \cdot \frac{e^{-\lambda(1-p)t} \cdot (\lambda(1-p)t)^{j}}{j!}$$

$$= P(N_{1}(t) = i) \cdot P(N_{2}(t) = j)$$

We only prove for two intervals that have the same length here. This also prove that  $N_1(t)$  and  $N_2(t)$  are Poisson distributed over t (Example 3.23 on textbook).

pishen (AlgoLab) Stochastic Process June 6, 2012 121 / 197

## Example 1

 $\mathbb{N}$  has rate 10 and p = 1/12,

$$P(N_1(4) = 0) = \frac{e^{-\frac{40}{12}} \cdot (\frac{40}{12})^0}{0!} = e^{-\frac{10}{3}}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 122 / 197

### Example 2: Type transitions

There're r classes of particles.

- $Y_i(k)$  is the number of class i particles at time k.
- The time is discrete in this case.
- $Y_i(0)$  is Poisson distributed with parameter  $\lambda_i$ .
- $\blacksquare$   $P_{ij}$  is the transition probability for a class i particle to class j.

We prove that  $Y_i(n)$  is Poisson distributed with parameter  $\sum_{i=1}^r P_{ii}^n \cdot \lambda_i$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 123 / 197

### Proof

Take class i for example, we consider a Poisson process  $\mathbb N$  with rate  $\lambda_i$ , where each event is classified as type k with probability  $P^n_{ik}$ .

For an arbitrary unit time interval, the number of events that occur in this interval is Poisson distributed with parameter  $\lambda_i$ . We take this Poisson distributed number as the value of  $Y_i(0)$ .

By property 1, we know that the number of type k events in this interval is Poisson distributed with parameter  $P^n_{ik} \cdot \lambda_i$ , which also means that the number of class i particles that eventually become class k at time n, which is denoted as  $C^n_{ik}$ , is also Poisson distributed with parameter  $P^n_{ik} \cdot \lambda_i$ .

 $Y_j(n) = \sum_i C_{ij}^n$ , which is Poisson distributed with parameter  $\sum_{i=1}^r P_{ij}^n \cdot \lambda_i$ . (since the Poisson parameter can be summed up.)

pishen (AlgoLab) Stochastic Process June 6, 2012 124 / 197

## Example 3: Selling a product

Consider a Poisson process with rate  $\lambda$ , where each event is an offer that has density function f(x).

A product is sold if an offer with value higher than the price y comes. Assume the accepted offer comes at time t, then the storage cost is  $c \cdot t$ , where c is a constant decided by the product.

We want to know the expected profit, which is  $E[f(x) - ct|f(x) \ge y]$ .

### Solution

The probability for each offer being accepted is

$$p(y) = P(X \ge y) = \int_{y}^{\infty} f(x) \, dx$$

The expectation of storage time t is  $1/(\lambda \cdot p(y))$ . Hence,

$$E[f(x)|f(x) \ge y] - E[ct|f(x) \ge y]$$

$$= \int_0^\infty x \cdot f_{X|X \ge y}(x) \, dx - \frac{c}{\lambda \cdot p(y)}$$

$$= \int_y^\infty x \cdot \frac{f_X(x)}{P(X \ge y)} \, dx - \frac{c}{\lambda \cdot p(y)}$$

$$= \frac{1}{p(y)} \left( \int_y^\infty x \cdot f(x) \, dx - \frac{c}{\lambda} \right)$$

pishen (AlgoLab) Stochastic Process June 6, 2012 126 / 197

### Example 4: Coupon collection



There are r types of coupons, and  $p_i$  is the probability for a collected coupon being type i.

We want to know the expectation of N, where N is the number of collected coupons so that all r types of coupons are collected.

pishen (AlgoLab) Stochastic Process June 6, 2012 127 / 19

## Solution: First attempt

Let  $N_i$  be the number of coupons collected to receive the first type i coupon. We know that

$$E[N] = E[\max(N_1, N_2, \dots, N_r)]$$

And

$$P(N \leq n) = P(N_1 \leq n, N_2 \leq n, \cdots, N_r \leq n)$$

But since each  $N_i$  are not independent, we can't go even further from here. For example, given that  $N_1=1$ ,  $P(N_2=1|N_1=1)=0$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 128 / 197

## Solution: Second attempt

Without loss of generality, we assume that the coupons arrive as a Poisson process  $\mathbb N$  with rate 1.

 $\mathbb{N}_i$  is the process of type i coupons, which has rate  $p_i$ .

 $X_i$  is the time that the first type i coupon appears, and

$$X = \max(X_1, X_2, \dots, X_r)$$

 $X_1, X_2, \dots, X_r$  are independent since  $\mathbb{N}_1, \mathbb{N}_2, \dots, \mathbb{N}_r$  are independent.

# Solution: Second attempt (cont.)

We can see that

$$P(X \le t) = P(X_1 \le t, X_2 \le t, \cdot, X_r \le t)$$
  
=  $\prod_{i=1}^{r} P(X_i \le t) = \prod_{i=1}^{r} (1 - e^{-t \cdot p_i})$ 

And

$$E[X] = \int_0^\infty P(X > t) dt = \int_0^\infty 1 - \prod_{i=1}^r (1 - e^{-t \cdot p_i}) dt$$

The first equation is from the property of probability. Surprisingly, E[X] = E[N] as explained below.

# Solution: Second attempt (cont.)

$$X = T_1 + T_2 + \dots + T_N$$

where  $T_i$  is the *j*th inter-arrival time of  $\mathbb{N}$ .

Each  $T_i$  are i.i.d. and are exponential distributed with rate 1.

Also, N is independent with each  $T_i$ .

$$E[X|N = n] = E[T_1 + T_2 + \dots + T_n] = n \cdot E[T_1] = n$$
  
 $E[X|N] = N \cdot E[T_1] = N$ 

Hence,

$$E[X] = E[E[X|N]] = E[N]$$

## Property 2-simple (Distribution of one event)

#### Property

Given that exactly one event of a Poisson process arrives in the interval [0,t], this arrival time is uniformly distributed over [0,t].

#### Proof:

$$P(T_1 \le s | N(t) = 1) = \frac{P(T_1 \le s, N(t) = 1)}{P(N(t) = 1)}$$

$$= \frac{P(N(s) = 1) \cdot P(N(t) - N(s) = 0)}{e^{-\lambda t} \cdot (\lambda t)^1 / 1!}$$

$$= \frac{e^{-\lambda s} \cdot \lambda s \cdot e^{-\lambda (t - s)}}{e^{-\lambda t} \cdot \lambda t} = \frac{s}{t}$$

## Property 2-advanced (Distribution of several events)

### Property

Given that exactly n events of a Poisson process arrive in the interval [0,t], each with arrival time  $X_1, X_2, \ldots, X_n$ .

The order statistics  $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$  of these random variables have the joint density function

$$f_{X_{(1)},X_{(2)},...,X_{(n)}}(x_1,x_2,...,x_n|N(t) = n)$$

$$= \begin{cases} \frac{n!}{t^n} & \text{if } 0 < x_1 < x_2 < \dots < x_n < t \\ 0 & \text{otherwise} \end{cases}$$

■ This implies that  $X_1, X_2, ..., X_n$  are i.i.d. and each is uniformly distributed over [0, t].

pishen (AlgoLab) Stochastic Process June 6, 2012 133 / 197

For any  $0 < x_1 < x_2 < \ldots < x_n < t$ ,

$$f_{X_{(1)},X_{(2)},...,X_{(n)}}(x_{1},x_{2},...,x_{n}|N(t) = n)$$

$$= \frac{f_{X_{(1)},X_{(2)},...,X_{(n)},N(t)}(x_{1},x_{2},...,x_{n},n)}{P(N(t) = n)}$$

$$= \frac{f_{T_{1},T_{2},...,T_{n}}(x_{1},x_{2} - x_{1},...,x_{n} - x_{n-1}) \cdot P(T_{n+1} > t - x_{n})}{P(N(t) = n)}$$

$$= \frac{\lambda e^{-\lambda x_{1}} \cdot \lambda e^{-\lambda(x_{2} - x_{1})} \cdot ... \cdot \lambda e^{-\lambda(x_{n} - x_{n-1})} \cdot e^{\lambda(t - x_{n})}}{\frac{e^{-\lambda t}(\lambda t)^{n}}{n!}}$$

$$= \frac{n!}{t}$$

 $T_i$  are the inter-arrival times, which are exponential distributed.

pishen (AlgoLab) Stochastic Process June 6, 2012 134 / 197

# Corollary of property 2 (simple version)

### Corollary

Consider a Poisson process with rate  $\lambda$ .

Each event is classified into a type i, where there are r types of event. Suppose that  $p_i(\cdot)$  is a sampling function over interval [0,t] such that each arrived event at time x has probability  $p_i(x)$  to be classified as type i. Then the number  $N_i(t)$  of type i events in [0,t] is Poisson distributed with parameter

$$\lambda \int_0^t p_i(x) \; dx = \lambda \cdot t \cdot R_i \quad ext{,where } R_i = rac{1}{t} \int_0^t p_i(x) \; dx$$

And each  $N_i(t)$  for all i are independent.

■  $N_i(\cdot)$  does not form a Poisson process here. It doesn't satisfy stationary distribution because of  $p_i(\cdot)$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 135 / 197

### Proof

Assume that N(t) = n.

Let 
$$n = n_1 + n_2 + \cdots + n_r$$
, where  $n_i \ge 0 \ \forall i = 1, \dots, r$ .

These n events arrive independent and uniformly at random over [0,t]. If an event arrives at time  $x \in [0,t]$ , then with probability  $p_i(x)$  it becomes type i. Therefore, each event is of type i with probability

$$\int_0^t P(\text{type } i|\text{arrives at time } x) \cdot \frac{1}{t} dx$$

$$= \frac{1}{t} \int_0^t p_i(x) dx$$

$$= R_i$$

which can be seen as the average of  $p_i(\cdot)$  over [0,t].

pishen (AlgoLab) Stochastic Process June 6, 2012 136 / 197

$$\begin{split} P(\bigwedge_{1 \leq i \leq r} N_i(t) &= n_i) = P(\bigwedge_{1 \leq i \leq r} N_i(t) = n_i | N(t) = n) \cdot P(N(t) = n) \\ &= \left(\frac{n!}{n_1! n_2! \cdots n_r!} \cdot \prod_{1 \leq i \leq r} R_i^{n_i} \right) \cdot \frac{e^{-\lambda t} \cdot (\lambda t)^n}{n!} \\ &= \prod_{1 \leq i \leq r} \frac{e^{-\lambda t R_i} (\lambda t R_i)^{n_i}}{n_i!} \\ &= \prod_{1 \leq i \leq r} P(N_i(t) = n_i) \end{split}$$

Hence we know that each  $N_i(t)$  is Poisson distributed with parameter  $\lambda t R_i$  and are independent. (Check Example 3.23 on textbook.)

pishen (AlgoLab) Stochastic Process June 6, 2012 137 / 197

# Corollary of property 2 (full version)

### Corollary

Consider a Poisson process with rate  $\lambda$ .

Each event is classified into a type i, where there are r types of event. Suppose that  $p_i(\cdot)$  is a sampling function over interval [s,s+t] such that each arrived event at time x has probability  $p_i(x)$  to be classified as type i. Then the number  $N_i(s+t)-N_i(s)$  of type i events in [s,s+t] is Poisson distributed with parameter

$$\lambda \int_{s}^{s+t} p_i(x) \ dx$$

And each  $N_i(s+t) - N_i(s)$  for all i are independent.

### Proof

Regarding [s, s+t] as [0, t], the sampling function becomes

$$p_i'(x) = p_i(x+s)$$

From the simple version, we know that  $N_i(s+t) - N_i(s)$  is Poisson distributed with parameter

$$\lambda \int_0^t p_i'(x) \ dx = \lambda \int_0^t p_i(x+s) \ dx = \lambda \int_s^{s+t} p_i(x) \ dx$$

pishen (AlgoLab) Stochastic Process June 6, 2012 139 / 197

## Example 1: Infinite server queue



Suppose that jobs arrive at a Poisson rate  $\lambda$ , and we have infinite number of servers. The running time of each job is independent and distributed with function  $T(\cdot)$ . We want to know

- **1** The distribution of the number X(t) of completed jobs by time t.
- **2** The distribution of the number Y(t) of running jobs by time t.
- **3** The joint distribtuion of  $Y(t_1)$  and  $Y(t_2)$ , where  $t_1 < t_2$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 140 / 197

# Solution of question 1 & 2

We classify the jobs into two types:

- **type 1**: completed by time t.
- **type 2**: not completed by time t.

If a job arrives at time x, then the sampling function is

$$p_1(x) = T(t-x)$$

$$p_2(x) = 1 - T(t - x)$$

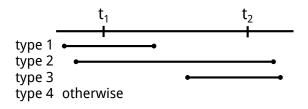
Thus,  $X(t) = N_1(t)$  is Poisson distributed with parameter

$$\lambda \int_0^t T(t-x) \ dx$$

And  $Y(t) = N_2(t)$  is Poisson distributed with parameter

$$\lambda \int_0^t (1 - T(t - x)) dx = \lambda t - \lambda \int_0^t T(t - x) dx$$

### Solution of question 3



$$p_1(x) = T(t_2 - x) - T(t_1 - x)$$
 if  $0 < x < t_1$   
 $p_2(x) = 1 - T(t_2 - x)$  if  $0 < x < t_1$   
 $p_3(x) = 1 - T(t_2 - x)$  if  $t_1 < x < t_2$ 

pishen (AlgoLab) Stochastic Process June 6, 2012 142 / 197

## Solution of question 3 (cont.)

We know that  $Y(t_1)=N_1(t_2)+N_2(t_2)$  and  $Y(t_2)=N_2(t_2)+N_3(t_2)$ , hence  $Y(t_1)$  and  $Y(t_2)$  are not independent. Therefore, we use the following method:

$$P(Y(t_1) = m_1, Y(t_2) = m_2)$$

$$= \sum_{n_2=0}^{\infty} P(N_1(t_2) = m_1 - n_2, N_2(t_2) = n_2, N_3(t_2) = m_2 - n_2)$$

$$= \sum_{n_2=0}^{\infty} P(N_1(t_2) = m_1 - n_2) \cdot P(N_2(t_2) = n_2) \cdot P(N_3(t_2) = m_2 - n_2)$$

$$= \cdots$$

The  $\infty$  in  $\Sigma$  can be replaced by  $\min(m_1, m_2)$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 143 / 197

## Example 2: Encounters on a highway



Cars enter a distance-d highway in a Poisson rate  $\lambda$ . The fixed speed of each car is i.i.d. with function  $F_S(\cdot)$ . Suppose our car enters the highway and moves at a fixed speed s, what's the distribution of the number of encountering with other cars?

pishen (AlgoLab) Stochastic Process June 6, 2012 144 / 19

### Solution

Suppose we enter at time  $t_1$  and leave at  $t_2 = t_1 + d/s$ .

Each car choose a fixed speed S according to  $F_S$ , its travel time T = d/S. The distribution function of T is

$$F_T(t) = P(T \le t) = P(S \ge \frac{d}{t}) = 1 - F_S(\frac{d}{t})$$

We classify the cars into three types:

- **type a** (overtaken by us):  $0 < t < t_1, t + T > t_2$ .
- **type b** (overtake us):  $t_1 < t < t_2$ ,  $t + T < t_2$ .
- **type c**: otherwise

pishen (AlgoLab) Stochastic Process June 6, 2012 145 / 197

$$p_a(t) = P(T > t_2 - t) = 1 - F_T(t_2 - t)$$
 if  $0 < t < t_1$   
 $p_b(t) = P(T < t_2 - t) = F_T(t_2 - t)$  if  $t_1 < t < t_2$ 

Since the Poisson parameters can be summed up,  $N_a(t_2) + N_b(t_2)$  is Poisson distributed with parameter

$$\lambda \int_0^{t_1} (1 - F_T(t_2 - t)) dt + \lambda \int_{t_1}^{t_2} F_T(t_2 - t) dt$$

which is the distribution we want.

pishen (AlgoLab) Stochastic Process June 6, 2012 146 / 197

## Example 3: HIV infection



People are infected with an unknown Poisson rate  $\lambda$ .

The incubation time for each infected person has distribution function F. At time t, we know the number  $n_1$  of people who already have the AIDS symptoms (finished incubations).

We want to estimate the value of  $\lambda$ , and the number  $n_2$  of the incubating people at time t.

pishen (AlgoLab) Stochastic Process June 6, 2012 147 / 19

### Solution

We classify the infected people into two types:

- type 1: have symptoms appear by time t.
- **type 2**: still in incubation by time t.

Then

$$E[N_{1}(t)] = \lambda \int_{0}^{t} F(t-x) dx = \lambda \int_{0}^{t} F(x) dx$$

$$E[N_{2}(t)] = \lambda \int_{0}^{t} (1 - F(t-x)) dx = \lambda \int_{0}^{t} (1 - F(x)) dx$$

$$n_{1} \approx E[N_{1}(t)] \implies \hat{\lambda} = \frac{n_{1}}{\int_{0}^{t} F(x) dx}$$

$$n_{2} \approx \hat{\lambda} \int_{0}^{t} (1 - F(x)) dx = \frac{n_{1} \int_{0}^{t} (1 - F(x)) dx}{\int_{0}^{t} F(x) dx}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 148 / 197

# Example 4: Hidden bugs



Suppose the errors come as a Poisson process.

Each error belongs to one of the m bugs in the program.

For all  $i=1,\ldots,m$ ,  $\mathbb{N}_i$  denotes the Poisson process of errors caused by bug i, which has an unknown rate  $\lambda_i$ .

At time t, we know the value of  $M_j(t)$ , which is the number of bugs causing exactly j errors.

And a bug is still hidden if it hasn't cause any error.

We want to know the expected error rate of hidden bugs, which is the expectation of the summation of all hidden bugs'  $\lambda_i$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 149 / 197

### Solution

Let

$$H_i(t)=\left\{egin{array}{ll} 1 & ext{if the $i$th bug is still hidden by time $t$} \\ 0 & ext{otherwise} \end{array}
ight.$$
  $\Lambda(t)=\sum_{i=1}^m H_i(t)\cdot \lambda_i$ 

Then

$$P(H_{i}(t) = 1) = P(N_{i}(t) = 0) = \frac{(\lambda_{i}t)^{0} \cdot e^{-\lambda_{i}t}}{0!} = e^{-\lambda_{i}t}$$

$$E[\Lambda(t)] = \sum_{i=1}^{m} \lambda_{i} \cdot E[H_{i}(t)] = \sum_{i=1}^{m} \lambda_{i} \cdot P(H_{i}(t) = 1) = \sum_{i=1}^{m} \lambda_{i} \cdot e^{-\lambda_{i}t}$$

$$= \frac{1}{t} \sum_{i=1}^{m} \lambda_{i} \cdot t \cdot e^{-\lambda_{i}t} = \frac{1}{t} \sum_{i=1}^{m} P(N_{i}(t) = 1) = \frac{1}{t} E[M_{1}(t)]$$

pishen (AlgoLab) Stochastic Process June 6, 2012 150 / 197

## Non-homogeneous Poisson process

#### Definition

 $\mathbb N$  is a non-homogeneous Poisson process with intensity function  $\lambda(\cdot)$  if  $\mathbb N$  is a counting process such that the following four conditions hold:

- N(0) = 0
- $\mathbf{2}$   $\mathbf{N}$  satisfies independent increments.
- $P(N(t+h) N(t) \ge 2) = o(h)$
- 4  $P(N(t+h) N(t) = 1) = \lambda(t) \cdot h + o(h)$ 
  - It's a Poisson process without stationary increments.
  - If  $\lambda(t) = \lambda$ , it becomes the homogeneous (normal) Poisson process.

## Proposition 1

#### Part 1:

For a Poisson process  $\mathbb N$  with rate  $\lambda$ , suppose that each arrived event has sampling function  $p_i(t)$ , then the counting process  $\mathbb N_i$  describing the arrival of type-i events is a non-homogeneous Poisson process with intensity function

$$\lambda_i(t) = \lambda \cdot p_i(t)$$

#### Part 2:

All non-homogeneous Poisson process with bounded  $\lambda(\cdot)$  can be obtained in the above way.

**Proof of part 2**: Since there's a  $\lambda$  that  $\lambda(t) \leq \lambda$  for all t, we just let  $p_i(t) = \lambda(t)/\lambda$  for a Poisson process with rate  $\lambda$ .

# Proof of part 1

Condition 1: trivial.

Condition 3:

$$P(N_i(t+h) - N_i(t) \ge 2) \le P(N(t+h) - N(t) \ge 2) = o(h)$$

Condition 4:

$$P(N_{i}(t+h) - N_{i}(t) = 1)$$

$$= \frac{P(N_{i}(t+h) - N_{i}(t) = 1 | N(t+h) - N(t) = 1)}{\times P(N(t+h) - N(t) = 1)}$$

$$+ P(N_{i}(t+h) - N_{i}(t) = 1 | N(t+h) - N(t) \ge 2)$$

$$\times \frac{P(N(t+h) - N(t) \ge 2)}{\times P(N(t+h) - N(t) \ge 2)}$$

$$= \frac{(p_{i}(t) + o(h))}{\times (\lambda h + o(h))} \cdot \frac{(\lambda h + o(h))}{\times (\lambda h + o(h))} + \frac{o(h)}{\times (\lambda h + o(h))}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 153 / 197

# Proof of part 1 (cont.)

The yellow part holds because

$$\lim_{h \to 0} P(N_i(t+h) - N_i(t)) = 1 | N(t+h) - N(t) = 1) = p_i(t)$$

$$\implies P(N_i(t+h) - N_i(t)) = 1 | N(t+h) - N(t) = 1) = p_i(t) + o(h)$$

#### Condition 2:

Consider non-overlapped intervals [s,s+t] and [u,u+v], the distribution in each interval is decided by Poisson parameter  $\lambda \int_s^{s+t} p_i(x) \ dx$  and  $\lambda \int_u^{u+v} p_i(x) \ dx$ , which doesn't influence each other.

pishen (AlgoLab) Stochastic Process June 6, 2012 154 / 197

## Distribution of non-homogeneous Poisson process

According to part 2 of proposition 1, for a non-homogeneous Poisson process with intensity function  $\lambda(\cdot)$ , we can observe that its distribution in [s,s+t] is Poisson distributed with parameter

$$\lambda \int_{s}^{s+t} p_{i}(x) \ dx = \int_{s}^{s+t} \lambda(x) \ dx$$

## Example 1: Poisson tea shop

Suppose that customers come to tea shop as a non-homogeneous Poisson process with  $\lambda(t)$  as follow:

$$\lambda(t) = \begin{cases} 0 & 0 \le t \le 8 \\ 5 + 5 \cdot (t - 8) & 8 < t \le 11 \\ 20 & 11 < t < 13 \\ 20 - 2 \cdot (t - 13) & 13 < t \le 17 \\ 0 & 17 < t < 24 \end{cases}$$

We want to know the probability that no one comes in [8.5, 9.5].

pishen (AlgoLab) Stochastic Process June 6, 2012 156 / 197

### Solution

The distribution in [8.5, 9.5] is a Poisson with parameter

$$\int_{8.5}^{9.5} (5+5\cdot(t-8)) dt = \int_{0.5}^{1.5} (5+5t) dt = 10$$

Hence the probability is  $e^{-10}$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 157 / 197

# Example 2: Infinite server queue again



Suppose that jobs arrive at a Poisson rate  $\lambda$ , and we have infinite number of servers. The running time of each job is independent and distributed with function  $F(\cdot)$ .

We want to prove that the process of jobs departing (finishing) is a non-homogeneous Poisson process.

pishen (AlgoLab) Stochastic Process June 6, 2012 158 / 197

### Proof

Condition 1: trivial.

#### Condition 2:

Consider two non-overlapping intervals [s, s+t] and [u, u+v], where u > s. Let the jobs finishing in [s, s+t] be type 1, and those finishing in [u, u + v] be type 2.

According to corollary of property 2, we know that the distribution of type 1 and type 2 jobs are independent in [0, u + v], which means that the number of jobs finishing in [s, s+t] and [u, u+v] are independent.

Stochastic Process June 6, 2012 159 / 197

# Proof (cont.)

**Condition 3 & 4**: Let  $f(\cdot)$  be the density function of  $F(\cdot)$ . The number of jobs departing in [t, t+h] is Poisson distributed with parameter

$$\lambda \int_0^{t+h} (F(t+h-x) - F(t-x)) dx$$

$$= \lambda \int_0^{t+h} (f(t+h-x) \cdot h + o(h)) dx$$

$$= \left(\lambda h \int_0^{t+h} f(x) dx\right) + o(h)$$

$$= \lambda h \cdot F(t+h) + o(h)$$

$$= \lambda h \cdot F(t) + o(h)$$

# Proof (cont.)

The first equation is because

$$\lim_{h \to 0} \frac{F(t+h-x) - F(t-x)}{h} = f(t-x) = \lim_{h \to 0} f(t-x+h)$$

Then we know that

$$\begin{split} P(N(t+h)-N(t)&=1)\\ =&(\lambda h\cdot F(t)+o(h))\cdot e^{-\lambda h\cdot F(t)+o(h)}\\ =&(\lambda h\cdot F(t)+o(h))\cdot (1-\lambda h\cdot F(t)+o(h))\\ =&\lambda h\cdot F(t)+o(h)\\ \text{and} \qquad P(N(t+h)-N(t)\geq 2)\\ =&\sum_{k\geq 2}\frac{(\lambda h\cdot F(t)+o(h))^k\cdot e^{-\lambda h\cdot F(t)+o(h)}}{k!}\\ =&o(h) \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 161 / 197

## Proposition 2

If  $\mathbb{N}_1$  and  $\mathbb{N}_2$  are independent non-homogeneous Poisson process with intensity function  $\lambda_1(\cdot)$  and  $\lambda_2(\cdot)$ ,

then  $\mathbb{N} = \mathbb{N}_1 + \mathbb{N}_2$  is also a non-homogeneous Poisson process with intensity function  $\lambda(t) = \lambda_1(t) + \lambda_2(t)$ .

**Proof** is left as exercise.

### Continuous Markov chain

#### Definition

The collection  $\mathbb{X} = \{X(t)|t \geq 0\}$  of nonnegative integral random variables is a *continuous-time Markov chain* if for all function  $x(\cdot)$ , nonnegative real numbers s and t, and integer i and j,

$$P(X(s+t) = j | X(s) = i, X(r) = x(r) \quad \forall r \in [0,s))$$
  
= $P(X(s+t) = j | X(s) = i)$ 

holds.

- The condition above is called *Markovian property*.
- Non-homogeneous Poisson process is a continuous Markov chain, where we take the values of N(t) as the states.

pishen (AlgoLab) Stochastic Process June 6, 2012 163 / 197

## Homogeneous transition probabilities

#### Definition

We say a continuous Markov chain has homogeneous transition probabilities if

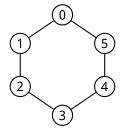
$$P(X(s+t) = j|X(s) = i)$$

is independent of s.

Our discussion of continuous Markov chain assumes this property.

pishen (AlgoLab) Stochastic Process June 6, 2012 164 / 197

## State graph



The state graph of continuous Markov chain can be taken as the state graph of Markov chain without self-links. The process will stay on a state for a certain amount of time, and go into a possible next state.

pishen (AlgoLab) Stochastic Process June 6, 2012 165 / 197

### Observation 1

#### Observation

The time for X to stay in state i is exponentially distributed.

■ Suppose that X(r) = i for some time  $r \ge 0$ . Let  $T_i$  be the waiting time for X to transit to a state other than i. We want to prove that  $T_i$  is memoryless, that is

$$P(T_i \ge s + t | T_i \ge s) = P(T_i \ge t)$$

hence  $T_i$  is exponential distributed.

■ This also means that for a Poisson process, no matter when the last event happened before, start measuring from now, the waiting time for the next event is still exponential distributed with parameter  $\lambda$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 166 / 197

$$P(T_i \ge s + t | T_i \ge s)$$

$$= P(X(v) = i, \forall r + s \le v \le r + s + t | X(u) = i, \forall r \le u \le r + s)$$

$$= P(X(v) = i, \forall r + s \le v \le r + s + t | X(r + s) = i)$$

$$= P(X(v) = i, \forall r \le v \le r + t | X(r) = i)$$

$$= P(T_i \ge t)$$

- holds because of homogeneous transition probability.
- holds because of Markovian property (more details at next page.)

Stochastic Process June 6, 2012 167 / 197

# Proof (cont.)

holds because we can use something like

$$P(A,B,C,...,Y|Z)$$

$$=P(B,C,...,Y|A,Z) \cdot P(A|Z)$$

$$=P(C,...,Y|A,B,Z) \cdot P(B|A,Z) \cdot P(A|Z)$$

$$= \cdots$$

and rearrange the condition parts?

### Observation 2

#### Observation

Given that X(s) = i, the probability that the next state other than i is j is a value  $\mathbb{P}[i,j]$  that is independent of s.

**Proof**: For any state j other than i, let  $T_i$  be the waiting time for the first transition from state i to a state other than i. We have

$$\mathbb{P}[i,j] = P(X(T_i + s) = j | X(s) = i) = P(X(T_i) = j | X(0) = i)$$

where the last equation is by homogeneous transition probability.

pishen (AlgoLab) Stochastic Process June 6, 2012 169 / 197

## Interpretation

- Starting with X(0) = i, the time staying in i is exponential distributed with parameter  $\lambda_i$ , then  $\mathbb X$  transits to a different state j with probability  $\mathbb P[i,j]$ .
- X can be characterized by

$$\lambda_i \quad \forall i \in S \quad \text{and} \quad \mathbb{P}[i,j] \quad \forall i,j \in S$$

pishen (AlgoLab) Stochastic Process June 6, 2012 170 / 197

## Example: Hair cut

- State 0: no customer
- State 1: cut hair
- State 2: wash hair

Assume customers appear as a Poisson process with rate  $\lambda'$ . When a customer appear, if the service is in state 1 or 2, then the customer just leave.

- lacksquare Staying time in state 1 is exponential distributed with parameter  ${\lambda_1}'.$
- lacksquare Staying time in state 2 is exponential distributed with parameter  ${\lambda_2}'.$

pishen (AlgoLab) Stochastic Process June 6, 2012 171 / 19

## Example: Hair cut (cont.)

Characterized as follow:

onlow. 
$$\mathbb{P}=\begin{pmatrix}1&0\\0&&1\\1&0\end{pmatrix}$$
  $\lambda_0=\lambda',\qquad \lambda_1=\lambda_1',\qquad \lambda_2=\lambda_2'$ 

pishen (AlgoLab) Stochastic Process June 6, 2012 172 / 197

## Birth and death process

#### Definition

A birth and death process with birth rate  $\{\beta_n\}$ ,  $\forall n \geq 0$  and death rate  $\{\delta_n\}$ ,  $\forall n \geq 0$  is a Markov chain that when the state is n,

- The waiting time  $B_n$  for it to go to state n+1 is exponential distributed with parameter  $\beta_n$ .
- The waiting time  $D_n$  for it to go to state n-1 is exponential distributed with parameter  $\delta_n$ .

pishen (AlgoLab) Stochastic Process June 6, 2012 173 / 19

# Birth and death process: characterize

- $\lambda_0 = \beta_0$
- $\lambda_n = \beta_n + \delta_n \quad \forall n > 0$

The last equation can be seen as a Poisson process with rate  $\lambda_n$ , where there're two types of events, one is birth and one is death. By page 117, we get the above result and

- $\mathbb{P}[n, n+1] = \frac{\beta_n}{\beta_n + \delta_n} \quad \forall n > 0$

## Example: Servers

Given s servers, where

- Jobs arrive in a Poisson process with rate  $\lambda$ .
- lacksquare Processing time for each job is exponential distributed with parameter  $\mu$ .

The number of jobs currently waiting or processing is a B & D process with

$$\beta_n = \lambda$$

$$\bullet \ \delta_n = \left\{ \begin{array}{ll} \mu \cdot n & \forall n \leq s \\ \mu \cdot s & \forall n > s \end{array} \right.$$

# Example: Linear growth with immigration

- $\beta_n = \beta \cdot n + \theta$
- $\delta_n = \delta \cdot n$

We want to know the value of E[X(t)].

#### Solution:

$$\begin{split} &P(X(t+h)=n+1|X(t)=n)=(\beta n+\theta)\cdot h+o(h) \quad \forall n\geq 0 \\ &P(X(t+h)=n-1|X(t)=n)=\delta n\cdot h+o(h) \quad \forall n\geq 1 \\ &P(X(t+h)=n|X(t)=n)=1-(\beta n+\delta n+\theta)\cdot h+o(h) \quad \forall n\geq 0 \end{split}$$

Since we make h close to 0, we ignore other probability that are all o(h).

pishen (AlgoLab) Stochastic Process June 6, 2012 176 / 197

$$\begin{split} E[X(t+h)|X(t) &= n] = (n+1) \cdot ((\beta n + \theta) \cdot h + o(h)) \\ &\quad + (n-1) \cdot (\delta n \cdot h + o(h)) \\ &\quad + n \cdot (1 - (\beta n + \delta n + \theta) \cdot h + o(h)) \\ &\quad + o(h) \\ &\vdots \\ &= n + ((\beta - \delta)n + \theta) \cdot h + o(h) \end{split}$$

Hence,

$$\begin{split} E[X(t+h)] &= E[E[X(t+h)|X(t)]] \\ &= (1 + (\beta - \delta) \cdot h) \cdot E[X(t)] + \delta h + o(h) \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 17

$$\frac{E[X(t+h)] - E[X(t)]}{h} = (\beta - \delta) \cdot E[X(t)] + \theta + \frac{o(h)}{h}$$
$$\frac{d}{dt}E[X(t)] = (\beta - \delta) \cdot E[X(t)] + \theta$$

If  $\beta = \theta$ ,

$$\frac{dE[X(t)]}{dt} = \theta$$

$$\implies E[X(t)] = X(0) + \theta t$$

pishen (AlgoLab) Stochastic Process June 6, 2012

178 / 197

If 
$$\beta \neq \theta$$
, define  $g(t) = (\beta - \delta) \cdot E[X(t)] + \theta$ , 
$$g'(t) = (\beta - \delta) \cdot \frac{d}{dt} E[X(t)] = (\beta - \delta) \cdot g(t)$$
 
$$\vdots$$
 
$$\Longrightarrow E[X(t)] = X(0) \cdot e^{(\beta - \delta)t} + \frac{\theta}{\beta - \delta} \cdot (e^{(\beta - \delta)t} - 1)$$

Note that if  $\delta > \beta$ ,

$$\lim_{t\to\infty} E[t] = \frac{\theta}{\delta - \beta}$$

which is dominated by the immigration.

pishen (AlgoLab) Stochastic Process June 6, 2012 179 / 197

#### Increment time

#### Definition

Given a birth and death process, if X(0) = i, the increment time is  $T_i = \min\{t | X(t) = i + 1\}.$ 

■ We want to know the value of  $E[T_i]$ .

Stochastic Process June 6, 2012 180 / 197

# Expectation of increment time

• If 
$$i = 0$$
,  $E[T_0] = \frac{1}{\beta_0}$ 

• If i > 0,

$$\begin{split} E[T_i] &= \text{ waiting time for the first event} \\ &+ 0 \cdot P(\text{the first event is birth}) \\ &+ (E[T_{i-1}] + E[T_i]) \cdot P(\text{the first event is death}) \\ &= \frac{1}{\beta_i + \delta_i} + 0 + (E[T_{i-1}] + E[T_i]) \cdot \frac{\delta_i}{\beta_i + \delta_i} \\ \Longrightarrow E[T_i] &= \frac{1}{\beta_i} \cdot (\delta_i \cdot E[T_{i-1}] + 1) \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 181 / 197

# Transition probability function

### Definition

The transition probability function  $P_{ij}(t) = P(X(t) = j | X(0) = i)$ .

- Note that this is different from the  $\mathbb{P}[i,j]$  mentioned before.
- There's also value for  $P_{ii}(t)$ .

Stochastic Process June 6, 2012 182 / 197

# Trnasition probability function of pure birth process

### $\mathsf{Theorem}$

If X is a pure birth process, where  $\delta_n = 0 \quad \forall n \geq 0$ , and  $\beta_i \neq \beta_i \quad \forall i \neq j$ , then

$$P_{ij}(t) = \begin{cases} e^{-\beta_i t} & \text{if } i = j \\ F(i,j) - F(i,j-1) & \text{if } i < j \end{cases}$$

where

$$F(i,j) = P(X(t) \le j | X(0) = i) = \sum_{i \le k \le j} e^{-\beta_k t} \prod_{i \le r \le j, \ r \ne k} \frac{\beta_r}{\beta_r - \beta_k}$$

Stochastic Process June 6, 2012 183 / 197 When i < j,

$$F(i,j) = P(X(t) < j+1|X(0) = i)$$

$$= P(T_i + T_{i+1} + \dots + T_j > t)$$

$$= \sum_{i \le k \le j} e^{-\beta_k t} \prod_{i \le r \le j, \ r \ne k} \frac{\beta_r}{\beta_r - \beta_k}$$

= is from equation 5.6 on textbook.

$$P(T_i + T_{i+1} + \dots + T_j > t) = 1 - P(T_i + T_{i+1} + \dots + T_j \le t)$$

where  $P(T_i + T_{i+1} + \cdots + T_j \le t)$  is the distribution of sum of j - i + 1 exponential distributions where each one must has different rate.

pishen (AlgoLab) Stochastic Process June 6, 2012 184 / 197

# Yule process

### Definition

Yule process is a pure birth process with linear rate, where  $\delta_n = 0 \ \forall n \geq 0$ and  $\beta_n = \beta \cdot n \ \forall n \geq 0$ .

According to page 183,

$$P_{ij}(t) = \dots = {j-1 \choose i-1} (e^{-\beta t})^i \cdot (1 - e^{-\beta t})^{j-i} \quad \forall 1 \le i \le j$$

which is a negative binomial distribution with parameter  $(i, e^{-\beta t})$ .

Stochastic Process June 6, 2012 185 / 197

## Instantaneous transition rate

### Definition

Given a continuous Markov chain with each state i having the exponential parameter  $\lambda_i$ , the *instantaneous transition rate* is

$$q_{ij} = \lambda_i \cdot \mathbb{P}[i,j]$$

we let  $q_{ii} = 0$  for the ease of later explanation.

■ Given all the  $q_{ij}$ , one can compute  $\lambda_i$  by  $\sum_{j \in S} q_{ij}$ , and then compute all the value of  $\mathbb{P}[i,j]$ . Hence  $q_{ij}$  also characterize the continuous Markov chain.

pishen (AlgoLab) Stochastic Process June 6, 2012 186 / 197

# Komolgorov forward equation

### **Theorem**

For most of the continuous Markov chain (ex. birth and death process), the following holds:

$$P_{ij}'(t) = \sum_{k \in S} P_{ik}(t) \cdot q_{kj} - \lambda_j \cdot P_{ij}(t)$$

# Komolgorov backward equation

### Theorem

For all the continuous Markov chain, the following holds:

$$P_{ij}'(t) = \sum_{k \in S} q_{ik} \cdot P_{kj}(t) - \lambda_i \cdot P_{ij}(t)$$

■ We only prove this equation (at page 194).

pishen (AlgoLab) Stochastic Process June 6, 2012 188 / 197

## Lemma 1

### Lemma

$$\lim_{h\to 0}\frac{1-P_{ii}(h)}{h}=\lambda_i$$

$$\lim_{h\to 0}\frac{P_{ij}(h)}{h}=q_{ij}\quad \text{if } i\neq j$$

We prove this lemma by proving the equivalence that

$$1 - P_{ii}(h) = \lambda_i \cdot h + o(h)$$

$$P_{ij}(h) = q_{ij} \cdot h + o(h)$$

Stochastic Process June 6, 2012 189 / 197

### **Equation 1**: In the interval of h,

$$P_{ii}(h) = P(\text{no transition happen}) \\ + P(\text{more than one transitions and come back to state } i) \\ \Longrightarrow 1 - P_{ii}(h) = P(\text{one transition happens}) \\ + P(\text{more than one transitions and doesn't come back}) \\ = \lambda_i \cdot h + o(h)$$

= is according to operational definition on page 106.

pishen (AlgoLab) Stochastic Process June 6, 2012 190 / 197

# Proof (cont.)

## **Equation 2**: In the interval of h,

$$\begin{split} P_{ij}(h) &= P(\text{one } i \to j \text{ transition happens}) \\ &+ P(\text{more than one transitions and the final state is } j) \\ &= (\lambda_i \cdot h + o(h)) \cdot \mathbb{P}[i,j] + o(h) \\ &= \lambda_i \cdot \mathbb{P}[i,j] \cdot h + o(h) \\ &= q_{ij} \cdot h + o(h) \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 191 / 197

## Lemma 2

### Lemma

$$P_{ij}(t+s) = \sum_{k \in S} P_{ik}(t) \cdot P_{kj}(s) \quad \forall s \ge 0, t \ge 0$$

pishen (AlgoLab) Stochastic Process June 6, 2012 192 / 197

## Proof

$$\begin{split} P_{ij}(t+s) &= P(X(t+s) = j | X(0) = i) \\ &= \sum_{k \in S} P(X(t+s) = j, X(t) = k | X(0) = i) \\ &= \sum_{k \in S} P(X(t+s) = j | X(t) = k, X(0) = i) \cdot P(X(t) = k | X(0) = i) \\ &= \sum_{k \in S} P(X(t+s) = j | X(t) = k) \cdot P(X(t) = k | X(0) = i) \\ &= \sum_{k \in S} P(X(s) = j | X(0) = k) \cdot P(X(t) = k | X(0) = i) \end{split}$$

- is by Markovian property.
- = is by homogeneous transition probability.

 $= \sum P_{ik}(t) \cdot P_{kj}(s)$ 

pishen (AlgoLab)

# Proof of backward equation

$$\begin{split} P_{ij}(h+t) - P_{ij}(t) &= \left(\sum_{k \geq 0} P_{ik}(h) \cdot P_{kj}(t)\right) - P_{ij}(t) \\ &= \left(\sum_{k \geq 0, k \neq i} P_{ik}(h) \cdot P_{kj}(t)\right) - (1 - P_{ii}(h)) \cdot P_{ij}(t) \\ &\Longrightarrow P_{ij}{}'(t) = \lim_{h \to 0} \left(\sum_{k \geq 0, k \neq i} \frac{P_{ik}(h)}{h} \cdot P_{kj}(t)\right) - \frac{1 - P_{ii}(h)}{h} \cdot P_{ij}(t) \\ &= \sum_{k \geq 0} q_{ik} \cdot P_{kj}(t) - \lambda_i \cdot P_{ij}(t) \end{split}$$

- = is by lemma 2.
- is by lemma 1 and that lim can be swap in the situation here.

pishen (AlgoLab) Stochastic Process June 6, 2012 194 / 197

# Example: Pure birth process

$$\lambda_i = \beta_i$$
 and  $\mathbb{P}[i, i+1] = 1$ 

Hence

$$P_{ij}'(t) = \beta_i \cdot P_{i+1,j}(t) - \beta_i \cdot P_{ij}(t)$$

# Example: Birth and death process

If 
$$i=0$$
, 
$$P_{0j}{}'(t)=\beta_0\cdot P_{1j}(t)-\beta_0\cdot P_{0j}(t)$$
 If  $i>0$ , 
$$P_{ij}{}'(t)=(\beta_i+\delta_i)\cdot \frac{\beta_i}{\beta_i+\delta_i}\cdot P_{i+1,j}(t)+(\beta_i+\delta_i)\cdot \frac{\delta_i}{\beta_i+\delta_i}\cdot P_{i-1,j}(t)+(\beta_i+\delta_i)\cdot P_{ij}(t)$$

pishen (AlgoLab) Stochastic Process June 6, 2012 196 / 197

# Example: A simple case

There're only two states 0 and 1, where 0 has rate  $\beta$  and 1 has rate  $\delta$ .

$$\begin{split} P_{00}'(t) &= \beta(P_{10}(t) - P_{00}(t)) \\ P_{10}'(t) &= \delta(P_{00}(t) - P_{10}(t)) \end{split} \right\} \delta \cdot P_{00}'(t) + \beta \cdot P_{10}'(t) = 0 \\ \Longrightarrow \delta \cdot P_{00}(t) + \beta \cdot P_{10}(t) = C \\ \Longrightarrow \operatorname{since} P_{00}(0) &= 1 \text{ and } P_{10}(0) = 0, \text{ we know } C = \delta \\ \Longrightarrow \beta \cdot P_{10}(t) &= \delta(1 - P_{00}(t)) \\ \Longrightarrow P_{00}'(t) &= \delta - (\delta + \beta) \cdot P_{00}(t) \\ \Longrightarrow P_{00}(t) &= \frac{\delta}{\beta + \delta} e^{-(\beta + \delta)t} + \frac{\beta}{\beta + \delta} \text{ and } \\ P_{10}(t) &= \cdots \end{split}$$

pishen (AlgoLab) Stochastic Process June 6, 2012 197 / 197