# MATH 56A: STOCHASTIC PROCESSES
# CHAPTER 3

## Plan for rest of semester

(1) st week (8/31, 9/6, 9/7) Chap 0: Diff eq's and linear recursion
(2) nd week (9/11...) Chap 1: Finite Markov chains
(3) rd week (9/18...) Chap 1: Finite Markov chains
(4) th week (9/25...) Chap 2: Countable Markov chains
(5) th week (oct 3,4,5) Chap 3: Continuous time Markov chains
(6) th week (oct 9,11,12) Ch 4: Stopping time
(7) th week (oct 16,18,19) Ch 5: Martingales
(8) th week (oct 23,25,26) Ch 6: Renewal processes
(9) th week (oct 30,1,2) Ch 7: Reversible Markov chains
(10) th week (nov 6,8,9 ) Ch 8: Weiner process
(11) th week (nov 13,15,16) Ch 8: more
(12) th week (nov 20,22) (short week) Ch 9: Stochastic integrals
(13) th week (nov 27,29,30,4) (extra day) Ch 9: more

## 3. Continuous Markov Chains

The idea of continuous Markov chains is to make time continuous instead of discrete. This idea only works when the system is not jumping back and forth at each step but rather moves gradually in a certain direction.

3.1. **making time continuous.** On the first day I discussed the problem of converting to continuous time. In the discrete Markov chain we have the transition matrix $P$ with entries $p(i,j)$ giving the probability of going from $i$ to $j$ in one unit of time. The $n$-th power, say $P^5$ has entries

$$p_5(i,j) = \mathbb{P}(X_5 = j \mid X_0 = i)$$

We want to interpolate and figure out what happened for all positive time $t$. (Negative time is discussed in Chapter 7.) We already know how to do that. You write:

$$P = QDQ^{-1}$$

*Date*: October 9, 2006.

where $D$ is a diagonal matrix whose diagonal entries are the eigenvalues of $P$ and $Q$ is the matrix of right eigenvectors of $P$. The first eigenvector of $P$ is 1 and the first right eigenvector is the column vector having all 1's.

If the eigenvalues are all positive then we can raise them to arbitrary values:

$$P^t = QD^tQ^{-1}$$

Usually you take logarithms. For example, if there are 3 states:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} = \begin{pmatrix} e^0 & 0 & 0 \\ 0 & e^{\ln d_2} & 0 \\ 0 & 0 & e^{\ln d_3} \end{pmatrix}$$

Then $P^t = e^{tA}$ where

$$A = Q(\ln D)Q^{-1} = Q \begin{pmatrix} 0 & 0 & 0 \\ 0 & \ln d_2 & 0 \\ 0 & 0 & \ln d_3 \end{pmatrix} Q^{-1}$$

This uses:

**Theorem 3.1.** $P = QDQ^{-1} = Qe^{\ln D}Q^{-1} = e^{Q \ln D Q^{-1}}$

*Proof.* Let $L = \ln D$ then

$$D = e^L := I + L + \frac{L^2}{2} + \frac{L^3}{3!} + \cdots$$

Conjugate by $Q$:

$$Qe^LQ^{-1} = QQ^{-1} + QLQ^{-1} + \frac{QL^2Q^{-1}}{2} + \frac{QL^3Q^{-1}}{3!} + \cdots$$

This is equal to $e^{QLQ^{-1}}$ since $QL^nQ^{-1} = (QLQ^{-1})^n$. □

The other theorem I pointed out was:

**Theorem 3.2.**

$$\frac{d}{dt}P^t = P^tA = AP^t$$

*Proof.* This is just term by term differentiation.

$$\frac{d}{dt}P^t = \sum \frac{d}{dt}\frac{Qt^nL^nQ^{-1}}{n!} = \sum \frac{Qnt^{n-1}L^nQ^{-1}}{n(n-1)!}$$

$$= QLQ^{-1} \sum \frac{Qt^nL^nQ^{-1}}{n!} = AP^t$$

□

3.2. **Poisson processes.** On the second day I explained continuous Markov chains as generalizations of Poisson processes.

A *Poisson process* is

- an event which occurs from time to time
- is time homogeneous (i.e., the probability that it will occur tomorrow is the same as the probability that it will occur today)
- and the occurrences are independent

The independence of occurrences of a Poisson event means that the probability of future occurrence is independent of both past and present. Markov processes are independent of the past. They depend only on the present. We will transform a Poisson processes so that it looks more like a Markov process.

Here is an example where a Poisson event occurs three times in a time interval $\Delta t = t_1 - t_0$. (We put $t_0 = 0$ in class so that $\Delta t = t_1$.)

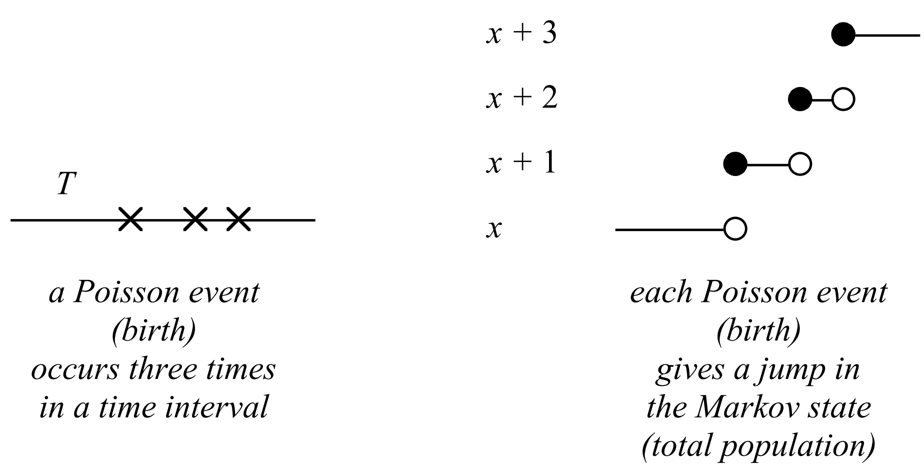


FIGURE 1. Poisson to Markov

The Poisson process has one parameter $\lambda$ called the *rate*. This is measured in inverse time units (number of occurrences per unit time). Thus $\lambda \Delta t$ is the expected number of occurrences in any time interval of length $\Delta t$.

3.2.1. *variables associated to a Poisson process.* There are two random variables associated with a Poisson process:

| Poisson variable (nonnegative integer) | Exponential variable (positive real) |
|---|---|
| $X =$ number of occurrences in $\Delta t$ <br> $\mathbb{P}(X = k) = e^{\lambda t_1} \dfrac{\lambda^k t_1^k}{k!}$ | $T =$ time until 1st occurrence |
| $\mathbb{P}(\text{event does not occur}) =$ <br> $\mathbb{P}(X = 0) = e^{\lambda t_1} = 1 - \lambda t_1 + \dfrac{\lambda^2 t_1^2}{2} - \cdots$ | $\mathbb{P}(\text{event does not occur}) =$ <br> $\mathbb{P}(T > t_1) = e^{\lambda t_1}$ |
| $\mathbb{P}(\text{event occurs}) =$ <br> $\mathbb{P}(X \geq 1) = 1 - e^{\lambda \Delta t} = \lambda \Delta t + o(\Delta t)$ | $\mathbb{P}(\text{event occurs in time } \Delta t) =$ <br> $\mathbb{P}(T \leq \Delta t) = 1 - e^{\lambda \Delta t} \approx \lambda \Delta t$ |

Here the book uses the "little oh" notation $o(\Delta t)$ to denote anything which vanishes faster than $\Delta t$:

$$\frac{o(\Delta t)}{\Delta t} \to 0 \quad \text{as} \quad \Delta t \to 0$$

3.2.2. *Poisson to Markov.* There are two changes we have to make to transform a Poisson process into a continuous time Markov process.

a) Every time an event occurs, you need to move to a new state in the Markov process. Figure 1 shows an example where the state is the total population:

$$X_t := \text{ population at time } t = X_0 + \#\text{births} - \#\text{deaths}$$

b) The rate $\alpha(x)$ depends on the state $x$. For example, the rate at which population grows is proportional to the size of the population:

$$\alpha(x) = \lambda x, \quad \lambda : \text{constant}$$

Notice that, when the rate increases, the events will occur more frequently and the waiting time will decrease. So, there is the possibility of *explosion*, i.e., an infinite number of jumps can occur in a finite amount of time.

3.3. **Definition of continuous Markov chain.** This is from Lawler, p. 69.

We need to start with an *infinitesmal generator* $A$ which is a matrix with entries $\alpha(x, y)$ for all states $x, y \in S$ so that $\alpha(x, y) \geq 0$ for $x \neq y$ and $\alpha(x, x) \leq 0$ and so that the sum of the rows is zero:

$$\sum_{y \in S} \alpha(x, y) = 0$$

We use the notation

$$\alpha(x) = -\alpha(x,x) = \sum_{y \neq x} \alpha(x,y)$$

**Definition 3.3.** *A continuous time Markov chain with infinitesmal generator $A = (\alpha(x,y))$ is a function $X : [0,\infty) \to S$ so that*

1. *$X$ is right continuous. I.e., $X_t$ is equal to the limit of $X_{t+\Delta t}$ as $\Delta t$ goes to zero from the right (the positive side).*
2. *$\mathbb{P}(X_{t+\Delta t} = x \,|\, X_t = x) = 1 - \alpha(x)\Delta t + o(\Delta t)$*
3. *$\mathbb{P}(X_{t+\Delta t} = y \,|\, X_t = x) = \alpha(x,y)\Delta t + o(\Delta t)$ for $y \neq x$.*
4. *$X_t$ is time homogeneous*
5. *$X_t$ is Markovian ($X_{\Delta t}$ depends on $X_t$ but is independent of the state before time $t$.)*

I pointed out that the numbers $\alpha(x), \alpha(x,y)$ were necessarily $\geq 0$ and that

$$\alpha(x) = \sum_{y \neq x} \alpha(x,y)$$

since $X_{t+\Delta t}$ must be in some state. The $(x,x)$ entry of the matrix $A$ is $\alpha(x,x) = -\alpha(x)$. So, the rows of the matrix $A$ add up to zero and all negative numbers lie on the diagonal.

3.4. **probability distribution vector.** At any time $t$ we have a *probability distribution vector* telling what is the probability of being in each state.

$$p_x(t) := \mathbb{P}(X_t = x)$$

This should not be confused with the time dependent probability transition matrix:

$$p_t(x,y) := \mathbb{P}(X_t = y \,|\, X_0 = x)$$

**Theorem 3.4.** *The time derivative of the probability distribution function $p_x(t)$ is given by*

$$\frac{d}{dt} p_x(t) = \sum_{y \in S} p_y(t)\alpha(y,x)$$

*In matrix notation this is*

$$\frac{d}{dt} p(t) = p(t)A$$

The unique solution of this differential equation is:

$$p(t) = p(0)e^{tA}$$

This implies that $P_t := e^{tA}$ is the time $t$ probability transition matrix.

*Proof.* The difference $p_x(t+\Delta t) - p_x(t)$ is equal to the probability that the state moves into $x$ minus the probability that it will move out of $x$ in the time period from $t$ to $t + \Delta t$. So,

$$p_x(t+\Delta t) - p_x(t) = \mathbb{P}(X_{t+\Delta t} = x, X_t = y \neq x) - \mathbb{P}(X_{t+\Delta t} = y \neq x, X_t = x)$$

$$= \sum_{y \neq x} \mathbb{P}(X_t = y)\mathbb{P}(X_{t+\Delta t} = x \mid X_t = y) - \sum_{y \neq x} \mathbb{P}(X_t = x)\mathbb{P}(X_{t+\Delta t} = y \mid X_t = x)$$

$$\approx \sum_{y \neq x} p_y(t)\alpha(y,x)\Delta t - \sum_{y \neq x} p_x(t)\alpha(x,y)\Delta t$$

$$= \sum_{y \neq x} p_y(t)\alpha(y,x)\Delta t - p_x(t)\alpha(x)\Delta t$$

$$= \sum_{y} p_y(t)\alpha(y,x)\Delta t$$

So,

$$\frac{p_x(t+\Delta t) - p_x(t)}{\Delta t} \approx \sum_y p_y(t)\alpha(y,x)$$

Take the limit as $\Delta t \to 0$ to get the theorem. $\square$

3.5. **example.** What is the probability that $X_4 = 1$ given that $X_0 = 0$ if the infinitesmal generator is

$$A = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}?$$

The answer is given by the $(0,1)$ entry of the matrix $e^{4A}$. The given information is that $p(0) = (1,0)$ and the question is: What is $p_1(4)$? The solution in matrix terms is the second coordinate of

$$p(4) = (p_0(4), p_1(4)) = (1,0)e^{4A}$$

We worked out the example: Since the trace of $A$ is $-1 + -2 = -3 = d_1 + d_2$ and $d_1 = 0$ we must have $d_2 = -3$. So, $A = QDQ^{-1}$ where

$$D = \begin{pmatrix} 0 & 0 \\ 0 & -3 \end{pmatrix}$$

and $Q$ is the matrix of right eigenvectors of $A$:

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{pmatrix}$$

The time 4 transition matrix is

$$e^{4A} = Qe^{4D}Q^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{-12} \end{pmatrix} \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{pmatrix}$$

$$= P_4 = \begin{pmatrix} \frac{2+e^{-12}}{3} & \frac{1-e^{-12}}{3} \\ \frac{2-2e^{-12}}{3} & \frac{1+2e^{-12}}{3} \end{pmatrix}$$

So, the answer is

$$p_1(4) = \frac{1-e^{-12}}{3}$$

3.6. **equilibrium distribution, positive recurrence.** An equilibrium distribution does not change with time. In other words, the time derivative is zero:

$$\frac{d}{dt}\pi(t) = \pi(t)A = 0$$

So, $\pi(t) = \pi(0)$ is the left eigenvector of $A$ normalized by:

$$\sum_{x \in S} \pi(x) = 1$$

Since $\pi$ does not change with time, we forget the $t$ and write $\pi(x)$ for $\pi_x(t)$. Recall that irreducible Markov chains are *positive recurrent* if and only if there is an equilibrium distribution.

The example above is irreducible and finite, therefore positive recurrent. The equilibrium distribution is $\pi = (2/3, 1/3)$.

3.7. **birth-death chain.**

**Definition 3.5.** *A* birth-death chain *is a continuous Markov chain with state space* $S = \{0, 1, 2, 3, \cdots\}$ *(representing population size) and transition rates:*

$$\alpha(n, n+1) = \lambda_n, \quad \alpha(n, n-1) = \mu_n, \quad \alpha(n) = \lambda_n + \mu_n$$

*representing births and deaths which occur one at a time.*

Notice that the total flow between the set of states $\{0, 1, 2, \cdots, n\}$ to the states $\{n+1, n+2, \cdots\}$ is given by the birth rate $\lambda_n$ and the death rate $\mu_{n+1}$. So, $\pi(n)$ is an equilibrium if and only if

$$\pi(n)\lambda_n = \pi(n+1)\mu_{n+1}$$

Solving for $\pi(n+1)$ gives:

$$\pi(n+1) = \frac{\lambda_n}{\mu_{n+1}}\pi(n) = \frac{\lambda_n \lambda_{n-1}}{\mu_{n+1}\mu_n}\pi(n-1) = \cdots = \frac{\lambda_n \lambda_{n-1} \cdots \lambda_0}{\mu_{n+1}\mu_n \cdots \mu_1}\pi(0)$$

If we can normalize these numbers we get an equilibrium distribution. So,

**Theorem 3.6.** *The birth-death chain is positive recurrent if and only if*

$$\sum_{n=0}^{\infty} \frac{\lambda_n \lambda_{n-1} \cdots \lambda_0}{\mu_{n+1} \mu_n \cdots \mu_1} < \infty$$

3.8. **birth and explosion.** If there is no death, the birth-death chain is obviously transient. The population is going to infinity but how fast? Suppose that $T_n$ is the time that the population stays in state $n$. Then (when $\mu_n = 0$) $T_n$ is exponential with rate $\lambda_n$. So,

$$\mathbb{E}(T_n) = \frac{1}{\lambda_n}$$

**Theorem 3.7.** *a) If $\sum_{n=0}^{\infty} \frac{1}{\lambda_n} < \infty$ then explosion occurs with probability one.*

*b) If $\sum_{n=0}^{\infty} \frac{1}{\lambda_n} = \infty$ then the probability of explosion is zero.*

For example, in the Yule process with $\lambda_n = n\lambda$, explosion will not occur since

$$\sum \frac{1}{\lambda_n} = \sum \frac{1}{\lambda n} = \frac{1}{\lambda} \sum \frac{1}{n} = \infty$$

3.9. **transient birth-death chains.** Recall that an irreducible Markov chain is transient if and only if there is a right eigenvector of $P$ with entries converging to zero corresponding to eigenvector 1. In the continuous time case, this is the same as a right eigenvector of $A$ corresponding to eigenvalue 0. So, we want numbers $a(n)$ such that

$$a(n-1)\mu_n + a(n)(-\lambda_n - \mu_n) + a(n+1)\lambda_n = 0$$

This equation can be rewritten as

$$[a(n+1) - a(n)]\lambda_n = [a(n) - a(n-1)]\mu_n$$

$$[a(n+1) - a(n)] = \frac{\mu_n}{\lambda_n}[a(n) - a(n-1)] = \frac{\mu_n \mu_{n-1} \cdots \mu_1}{\lambda_n \lambda_{n-1} \cdots \lambda_1}[a(1) - a(0)]$$

$a(k+1)$ is the sum of these numbers:

$$a(k+1) = a(0) + \sum_{n=0}^{k} [a(n+1) - a(n)] = a(0) + \sum_{n=0}^{k} \frac{\mu_n \mu_{n-1} \cdots \mu_1}{\lambda_n \lambda_{n-1} \cdots \lambda_1}[a(1) - a(0)]$$

**Theorem 3.8.** *A birth-death chain is transient if and only if*

$$\sum_{n=0}^{k} \frac{\mu_n \mu_{n-1} \cdots \mu_1}{\lambda_n \lambda_{n-1} \cdots \lambda_1} < \infty$$

*Proof.* Let $L$ be this limit. Let $a(0) = 1$ and $a(1) = 1 - 1/L$. Then $a(k+1)$ given by the above formula will converge to zero. Conversely, if the $a(k+1)$ goes to zero, the infinite sum must converge. □