Statistics 150: Spring 2007

April 21, 2008

1 Continuous-Time Markov Chains

Consider a continuous-time stochastic process $(X_t)_{t\geq 0}$ taking values in the set of nonnegative integers. We say that the process $(X_t)_{t\geq 0}$ is a *continuous-time Markov chain* if for all sequences of times $t_0 < t_1 < \cdots < t_n$ and sequences of nonnegative integers $x_0, x_1, \ldots, x_n$,

$$\begin{aligned}
\mathbb{P}\{X_{t_n} = x_n \mid X_{t_0} = x_0, X_{t_1} = x_1, \ldots, X_{t_{n-1}} = x_{n-1}\} \\
= \mathbb{P}\{X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}\} \\
= \mathbb{P}\{X_{t_n - t_{n-1}} = x_n \mid X_0 = x_{n-1}\}.
\end{aligned}$$

This implies that for any $0 \leq s < t$, and any integer x,

$$\mathbb{P}\{X_t = x \mid X_u,\ 0 \leq u \leq s\} = \mathbb{P}\{X_t = x \mid X_s\}.$$

If we let τ_i denote the amount of time that the process stays in state i before making a transition into a different state, then the Markov property implies that

$$\mathbb{P}\{\tau_i > s + t | \tau_i > s\} = \mathbb{P}\{\tau_i > t\},$$

for all $s, t \geq 0$. Hence, the random variable τ_i is memoryless, so is exponentially distributed.

The above gives us a way of constructing a continuous-time Markov chain. Namely, it is a stochastic processes having the properties that each time it enters state i:

(i) the amount of time it spends in that state before jumping to a different state is exponentially distributed with some rate $\nu_i \geq 0$; and

(ii) when the process leaves state i, it will next enter state j with some probability P_{ij}, where $\sum_{j \neq i} P_{ij} = 1$.

A continuous-time Markov chain is said to be *regular* if, with probability 1, the number of transitions in any finite length of time is finite.

We shall, assume from now on that all Markov chains considered are regular. Let q_{ij} be defined by

$$q_{ij} = \nu_i P_{ij}, \text{ all } i \neq j.$$

It follows that q_{ij} is the rate when in state i that the process makes a transition into state j.

Finally, define the transition probabilities

$$P_{ij}(t) = \mathbb{P}\{X(t+s) = j | X(s) = i\},$$

not to be confused with the set of jump probabilities P_{ij}. The quantity does not depend on s by our assumption of homogeneity in time.

2 Birth and Death Processes

A continuous-time Markov chain with states $0, 1, \cdots$ for which $q_{ij} = 0$ whenever $|i - j| > 1$ is called a *birth and death process.* Let λ_i and μ_i be given by

$$\lambda_i = q_{i,i+1}, \quad \mu_i = q_{i,i-1}.$$

The values $\{\lambda_i, i \geq 0\}$ and $\{\mu_i, i \geq 1\}$ are called respectively the birth rate and the death rate. We see that

$$\nu_i = \lambda_i + \mu_i, \quad P_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i} = 1 - P_{i,i-1}.$$

2.1 Two Birth and Death Processes

(i) *the $M/M/s$ Queue.* Suppose that customers arrive at an s-server service station in accordance with a Poisson process having rate λ. That is, the times between successive arrivals are independent exponential random variables having mean $1/\lambda$. Each customer, upon arrival, goes directly into service if any of the servers are free, and if not, then the customer joins the queue (that is, he waits in a line). When a server finishes serving a customer, the customer leaves the system, and the next customer in line, if there are any waiting, enters the service.

The successive service times are assumed to be independent exponential random variables having mean $1/\mu$. If we let $X(t)$ denote the number in the system at time t, then $\{X(t), t \geq 0\}$ is a birth and death process with

$$\mu_n = \begin{cases} n\mu, & 1 \leq n \leq s \\ s\mu, & n > s, \end{cases}$$

$$\lambda_n = \lambda, \quad n \geq 0.$$

(ii)*A Linear Growth Model with Immigration.* A model in which

$$\mu_n = n\mu, \quad n \geq 1, \qquad \lambda_n = n\lambda + \theta, \quad n \geq 0,$$

is called a linear growth process with immigration. Such processes occur naturally in the study of biological reproduction and population growth. Each individual in the population is assumed to give birth at rate λ; in addition, there is a constant rate of increase θ due to an external source such as immigration. Hence, the total birth rate where there are n persons in the system is $n\lambda + \theta$. Deaths are assumed to occur at an exponential rate μ for each member of the population, and hence $\mu_n = n\mu$.

A birth and death process is said to be a *pure birth* process if $\mu_n = 0$ for all n (that is, if death is impossible). The simplest example of a pure birth process is the Poisson process, which has a constant birth rate $\lambda_n = \lambda, n \geq 0$.

2.2 the Yule process

A second example of a pure birth process results from a population in which each member independently gives birth at an exponential rate λ. If no one ever dies, the population size $(X_t)_{t\geq 0}$ is a pure birth process with

$$\lambda_n = n\lambda, \quad n \geq 0.$$

This pure birth process is called the *Yule* process.

Consider a Yule process starting with a single individual at time 0, and let $T_i, i \geq 1$, denote the time between the $(i-1)$th and ith birth. Now

$$\mathbb{P}(T_1 \leq t) = 1 - e^{-\lambda t},$$

$$\begin{aligned}
\mathbb{P}\{T_1 + T_2 \leq t\} &= \int_0^t \mathbb{P}\{T_1 + T_2 \leq t | T_1 = x\} \lambda e^{-\lambda x}\, d\, x \\
&= \int_0^t (1 - e^{-2\lambda(t-x)}) \lambda e^{-\lambda x}\, d\, x \\
&= (1 - e^{-\lambda t})^2,
\end{aligned}$$

We can show by induction that

$$\mathbb{P}\{T_1 + \cdots + T_j \leq t\} = (1 - e^{-\lambda t})^j.$$

Hence, as $\mathbb{P}\{T_1 + \cdots + T_j \leq t\} = \mathbb{P}\{X_t \geq j + 1 | X(0) = 1\}$, we have the transition probabilities for the Yule process:

$$P_{1j}(t) = (1 - e^{-\lambda t})^{j-1} - (1 - e^{-\lambda t})^j = e^{-\lambda t}(1 - e^{-\lambda t})^{j-1}, \quad j \geq 1$$

$$P_{ij}(t) = \binom{j - i + 1}{i - 1} e^{-\lambda t i}(1 - e^{-\lambda t})^{j-i}, \quad j \geq i \geq 1.$$

Let $S_k = \sum_{i \le k} T_i$ be the time of the kth birth. Reasoning heuristically and treating densities as if they were probabilities yields for $0 \le s_1 \le s_2 \le \cdots \le s_n \le t$

$$\begin{aligned}
&\mathbb{P}\{S_1 = s_1, S_2 = s_2, \cdots, S_n = s_n | X(t) = n+1\} \\
&\quad = \frac{\mathbb{P}\{T_1 = s_1, T_2 = s_2 - s_1, \cdots, T_n = s_n - s_{n-1}, T_{n+1} > t - s_n\}}{\mathbb{P}\{X(t) = n+1\}} \\
&\quad = \frac{\lambda e^{-\lambda s_1} 2\lambda e^{-2\lambda(s_2 - s_1)} \cdots n\lambda e^{-n\lambda(s_n - s_{n-1})} e^{-(n+1)\lambda(t - s_n)}}{\mathbb{P}\{X(t) = n+1\}} \\
&\quad = \frac{n!\, \lambda^n\, e^{-\lambda(t-s_1)} e^{-\lambda(t-s_2)} \cdots e^{-\lambda(t-s_n)}}{(1 - e^{-\lambda})^n}.
\end{aligned}$$

Hence we see that the conditional density of $(S_1, \cdots, S_n)$ given that $X(t) = n+1$ is given by $n! \prod_{i=1}^{n} f(s_i), \quad 0 \le s_1 \le s_2 \le \cdots \le s_n \le t$, where f is the density function

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda(t-x)}}{1-e^{-\lambda t}}, & 0 \le x \le t \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 2.1. *Consider a Yule process with $X(0) = 1$. Then, given that $X(t) = n + 1$, the birth times $S_1, \cdots, S_n$ are distributed as the ordered values of n i.i.d. random variables whose density is*

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda(t-x)}}{1-e^{-\lambda t}}, & 0 \le x \le t \\ 0 & \textit{otherwise.} \end{cases}$$

Consider a Yule process with $X(0) = 1$. Let us compute the expected sum of the ages of the members of the population at time t. The sum of the ages at time t, call it $A(t)$, can be expressed as

$$A(t) = a_0 + t + \sum_{i=1}^{X(t)-1} (t - S_i),$$

where a_0 is the age at $t = 0$ of the initial individual.

First we compute $\mathbb{E}[A(t)]$ conditioned on $X(t)$,

$$\begin{aligned}
\mathbb{E}[A(t)|X(t)=n+1] &= a_0 + t + \mathbb{E}[\sum_{i=1}^{n}(t-S_i)|X(t)=n+1] \\
&= a_0 + t + n\int_0^t (t-x)\frac{\lambda e^{-\lambda(t-x)}}{1-e^{-\lambda t}}\,d\,x
\end{aligned}$$

or

$$\mathbb{E}[A(t)|X(t)] = a_0 + t + (X(t)-1)\frac{1-e^{-\lambda t}-\lambda t e^{-\lambda t}}{\lambda(1-e^{-\lambda t})}.$$

Taking expectations and using the fact that $X(t)$ has mean $e^{\lambda t}$ yields

$$\mathbb{E}[A(t)] = a_0 + t + \frac{e^{\lambda t}-1-\lambda t}{\lambda} = a_0 + \frac{e^{\lambda t}-1}{\lambda}.$$

3 The Kolmogorov Differential Equations

Let $P_{ij}(t) = \mathbb{P}\{X_t = j | X_0 = i\}$ be the transition probabilities for a continuous-time Markov chain $(X_t)_{t \geq 0}$ with rate matrix $(q_{ij})_{i \neq j}$ and jump rates $\nu_i = \sum_j q_{ij}$.

Lemma 3.1. *(i)* $\lim_{t \to 0} \frac{1 - P_{ii}(t)}{t} = \nu_i$.

(ii) $\lim_{t \to 0} \frac{P_{ij}(t)}{t} = q_{ij}, \quad i \neq j$.

Lemma 3.2. *For all* $s, t \geq 0$*,*

$$P_{ij}(t + s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s).$$

Theorem 3.3 (Kolmogorov's Backward Equations). *For all i, j, and $t \geq 0$,*

$$\frac{d}{dt} P_{ij}(t) = \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t).$$

Proof: By the previous lemma, we have that

$$\begin{aligned} P_{ij}(t+h) - P_{ij}(t) &= \sum_k (P_{ik}(h) - \delta_{ik}) P_{kj}(t) \\ &= \sum_{k \neq i} P_{ik}(h) P_{kj}(t) + (P_{ii}(h) - 1) P_{ij}(t), \end{aligned}$$

where $\delta_{ik} = 1$ if $i = k$ and is zero otherwise. By the preceeding lemma then, if we can exchange the limit and the summation,

$$\frac{1}{h}(P_{ij}(t+h) - P_{ij}(t)) \to \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t) \quad \text{as } h \to 0.$$

We need only prove that we can exchange the limit and the summation. For any fixed N,

$$\begin{aligned}\liminf_{h\to 0}\sum_{k\neq i}\frac{P_{ik}(h)}{h}P_{kj}(t) &\geq \liminf_{h\to 0}\sum_{k\neq i,k<N}\frac{P_{ik}(h)}{h}P_{kj}(t)\\ &= \sum_{k\neq i,k<N} q_{ik}P_{kj}(t).\end{aligned}$$

Since the above holds for all N we see that

$$\liminf_{h\to 0}\sum_{k\neq i}\frac{P_{ik}(h)}{h}P_{kj}(t) \geq \sum_{k\neq i} q_{ik}P_{kj}(t). \tag{3.1}$$

To reverse the inequality note that for $N > i$, since $P_{kj}(t) \le 1$,

$$
\begin{aligned}
& \limsup_{h\to 0} \sum_{k\neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \\
\le \quad & \limsup_{h\to 0} \left[\sum_{k\neq i, k<N} \frac{P_{ik(h)}}{h} P_{kj}(t) + \sum_{k\ge N} \frac{P_{ik}(h)}{h} \right] \\
= \quad & \limsup_{h\to 0} \left[\sum_{k\neq i, k<N} \frac{P_{ik}(h)}{h} P_{kj}(t) + \frac{1 - P_{ii}(h)}{h} - \sum_{k\neq i, k<N} \frac{P_{ik}(h)}{h} \right] \\
= \quad & \sum_{k\neq i, k<N} q_{ik} P_{kj}(t) + \nu_i - \sum_{k\neq i, k<N} q_{ik},
\end{aligned}
$$

where we have used Lemma 3.1 and the fact that $\sum_j P_{ij}(t) = 1$.

As the above inequality is true for all $N > i$, we obtain upon letting $N \to \infty$ and using the fact $\sum_{k \neq i} q_{ik} = \nu_i$,

$$\limsup_{h \to 0} \sum_{k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \leq \sum_{k \neq j} q_{ik} P_{kj}(t).$$

We have shown that

$$\lim_{h \to 0} \sum_{k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) = \sum_{k \neq j} q_{ik} P_{kj}(t),$$

and so

$$\begin{aligned} \frac{d}{dt} P_{ij}(t) &= \lim_{h \to 0} \left(\sum_{k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) + \frac{(P_{ii}(h) - 1)}{h} P_{ij}(t) \right) \\ &= \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t). \end{aligned}$$

$\square$

The set of differential equations for $P_{ij}(t)$ given in Theorem 3.3 are known as the Kolmogorov *backward equations.* They are called the backward equations because in computing the probability distribution of the state at time $t+h$ we conditioned on the state (all the way) back at time h.

That is, we started our calculation with

$$
\begin{aligned}
&P_{ij}(t+h) \\
&\quad = \sum_k P\{X_{t+h} = j \mid X_0 = i,\, X_h = k\}\mathbb{P}\{X_h = k | X_0 = i\} \\
&\quad = \sum_k P_{ik}(h) P_{kj}(t).
\end{aligned}
$$

We may derive another set of equations, known as the Kolmogorov's *forward equations*, by now conditioning on the state at time t. This yields

$$P_{ij}(t+h) = \sum_k P_{ik}(t)P_{kj}(h)$$

or

$$\begin{aligned} P_{ij}(t+h) - P_{ij}(t) &= \sum_k P_{ik}(t)P_{kj}(h) - P_{ij}(t) \\ &= \sum_{k \neq j} P_{ik}(t)P_{kj}(h) - (1 - P_{jj}(h))P_{ij}(t). \end{aligned}$$

Therefore,

$$\lim_{h \to 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h} = \lim_{h \to 0} \left\{ \sum_{k \neq j} P_{ik}(t) \frac{P_{kj}(h)}{h} - \frac{1 - P_{ii}(h)}{h} P_{ij}(t) \right\}.$$

Assuming that we can interchange the limit with summation, we obtain by Lemma 3.1 that

$$\frac{d}{dt} P_{ij}(t) = \sum_{k \neq j} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t).$$

Theorem 3.4 (Kolmogorov forward equations)**.** *Under suitable regularity conditions,*

$$\frac{d}{dt} P_{ij}(t) = \sum_{k \neq j} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t).$$

Example 3.5. The Kolmogorov forward equations for the birth and death process are

$$\frac{d}{dt}P_{i0}(t) = \mu_1 P_{i1}(t) - \lambda_0 P_{i0}(t), \quad \text{for all } i > 0,$$

$$\frac{d}{dt}P_{ij}(t) = \lambda_{j-1}P_{i,j-1}(t) + \mu_{j+1}P_{i,j+1}(t) - (\lambda_j + \mu_j)P_{ij}(t), \quad j \neq 0.$$

Example 3.6. For a pure birth process, the forward equations reduce to

$$\frac{d}{dt}P_{ii}(t) = \lambda_i P_{ii}(t), \tag{3.2}$$

$$\frac{d}{dt}P_{ij}(t) = \lambda_{j-1}P_{i,j-1}(t) - \lambda_j P_{ij}(t), \quad j > i. \tag{3.3}$$

Integrating equation (3.2) and then using $P_{ii}(0) = 1$ yields

$$P_{ii}(t) = e^{-\lambda_i t}.$$

The above, of course, is true as $P_{ii}(t)$ is the probability that the time until a transition from state i is greater than t. The other quantities $P_{ij}(t), j > i$, can be obtained recursively from (3.2)&(3.3) as follows: we have, for $j > i$,

$$\begin{aligned} e^{\lambda_j t}\lambda_{j-1}P_{i,j-1}(t) &= e^{\lambda_j t}[\frac{d}{dt}P_{ij}(t) + \lambda_j P_{ij}(t)] \\ &= \frac{d}{dt}[e^{\lambda_j t}P_{ij}(t)]. \end{aligned}$$

Integration, using $P_{ij}(0) = 0$, yields

$$P_{ij}(t) = \lambda_{j-1} e^{-\lambda_j t} \int_0^t e^{\lambda_j s} P_{i,j-1}(s)\, d\, s, \quad , j > i.$$

In the special case of a Yule process, where $\lambda_j = j\lambda$, we can use the above to verify our previous result,

$$P_{ij}(t) = \binom{j-1}{i-1} e^{-\lambda t i} (1 - e^{-\lambda t})^{j-i}, \quad j \geq i \geq 1.$$

4 Exercises

1, A population of organisms consists of both male and female members. In a small colony any particular male is likely to mate with any particular female in any time interval of length h, with probability $\lambda h + o(h)$. Each mating immediately produces one offspring, equally likely to be male or female. Let $N_1(t)$ and $N_2(t)$ denote the number males and females in the population at t. Derive the parameters of the continuous-time Markov chain $\{N_1(t), N_2(t)\}$.

2, Suppose that a one-celled organism can be in one of two states-either A or B. An individual in state A will change to state B at an exponential rate α; an individual in state B divides into two new individuals of type A at an exponential rate β. Define an appropriate continuous-time Markov chain for a population of such organisms and determine the appropriate parameters for this model.

3, Consider a Yule process with $X(0) = i$. Given that $X(t) = i + k$, what can be said about the conditional distribution of the birth times of the k individuals born in $(0, t)$?

4, Consider a Yule process starting with a single individual and suppose that with probability $P(s)$ and individual born at time s will be robust. Compute the distribution of the number of robust individuals born in $(0, t)$.