STAT 870 Lecture 15

Continuous Time Markov Chains

Consider a population of single celled organisms in a stable environment.

Fix short time interval, length h.

Each organism has some probability of dividing to produce two organisms and some other probability of dying.

We might suppose:

- Different organisms behave independently.
- Probability of division (for specified organism) is λh plus $o(h)$.
- Probability of death is μh plus $o(h)$.
- Probability that an organism divides twice (or divides once and dies) in the interval of length h is $o(h)$.

Notice tacit assumption: constants of proportionality do not depend on time (that is our interpretation of "stable environment").

Notice too that we have taken the constants not to depend on which organism we are talking about. We are really assuming that the organisms are all similar and live in similar environments.

$Y(t)$: total population at time t.

$\mathcal{H}_t$: history of the process up to time t.

We generally take

$$\mathcal{H}_t = \sigma\{Y(s); 0 \le s \le t\}$$

General definition of a **history** (alternative jargon **filtration**): any family of σ-fields indexed by t satisfying:

- $s < t$ implies $\mathcal{H}_s \subset \mathcal{H}_t$.
- $Y(t)$ is a $\mathcal{H}_t$ measurable random variable.
- $\mathcal{H}_t = \bigcap_{s>t} \mathcal{H}_s$.

The last assumption is a technical detail we will ignore from now on.

Condition on event $Y(t) = n$.

Then the probability of two or more divisions (either more than one division by a single organism or two or more organisms dividing) is $o(h)$ by our assumptions.

Similarly the probability of both a division and a death or of two or more deaths is $o(h)$.

So probability of exactly 1 division by any one of the n organisms is $n\lambda h + o(h)$.

Similarly probability of 1 death is $n\mu h + o(h)$.

We deduce:

$$\begin{aligned} P(Y(t+h) &= n+1|Y(t)=n,\mathcal{H}_t) \\ &= n\lambda h + o(h) \\ P(Y(t+h) &= n-1|Y(t)=n,\mathcal{H}_t) \\ &= n\mu h + o(h) \\ P(Y(t+h) &= n|Y(t)=n,\mathcal{H}_t) \\ &= 1-n(\lambda+\mu)h + o(h) \\ P(Y(t+h) &\notin \{n-1,n,n+1\}|Y(t)=n,\mathcal{H}_t) \\ &= o(h) \end{aligned}$$

These equations lead to:

$$\begin{aligned} P(Y(t+s) = j|Y(s)=i,\mathcal{H}_s) & \\ &= P(Y(t+s)=j|Y(s)=i) \\ &= P(Y(t)=j|Y(0)=i) \end{aligned}$$

This is the **Markov Property**.

Definition: A process $\{Y(t); t \geq 0\}$ taking values in S, a finite or countable state space is a Markov Chain if

$$\begin{aligned} P(Y(t+s) = j|Y(s)=i,\mathcal{H}_s) & \\ &= P(Y(t+s)=j|Y(s)=i) \\ &\equiv \mathbf{P}_{ij}(s,s+t) \end{aligned}$$

Definition: A Markov chain Y has **stationary transitions** if

$$\mathbf{P}_{ij}(s,s+t) = \mathbf{P}_{ij}(0,t) \equiv \mathbf{P}_{ij}(t)$$

From now on: our chains have stationary transitions.

Summary of Markov Process Results

- Chapman-Kolmogorov equations:

$$\mathbf{P}_{ik}(t+s) = \sum_j \mathbf{P}_{ij}(t)\mathbf{P}_{jk}(s)$$

- Exponential holding times: starting from state i time, T_i, until process leaves i has exponential distribution, rate denoted v_i.

- Sequence of states visited, $Y_0, Y_1, Y_2, \ldots$ is Markov chain – transition matrix has $\mathbf{P}_{ii} = 0$. Y sometimes called **skeleton**.

- **Communicating classes** defined for skeleton chain. Usually assume chain has 1 communicating class.

- Periodicity irrelevant because of continuity of exponential distribution.

- Instantaneous transition rates from i to j:

$$q_{ij} = v_i \mathbf{P}_{ij}$$

- Kolmogorov backward equations:

$$\mathbf{P}'_{ij}(t) = \sum_{k \neq i} q_{ik}\mathbf{P}_{kj}(t) - v_i \mathbf{P}_{ij}(t)$$

- Kolmogorov forward equations:

$$\mathbf{P}'_{ij}(t) = \sum_{k \neq j} q_{kj}\mathbf{P}_{ik}(t) - v_j \mathbf{P}_{ij}(t)$$

- For strongly recurrent chains with a single communicating class:

$$\mathbf{P}_{ij}(t) \to \pi_j$$

- Stationary initial probabilities π_i satisfy:

$$v_j \pi_j = \sum_{k \neq j} q_{kj}\pi_k$$

- Transition probabilities given by

$$\mathbf{P}(t) = e^{\mathbf{R}t}$$

where $\mathbf{R}$ has entries

$$\mathbf{R}_{ij} = \begin{cases} q_{ij} & i \neq j \\ -v_i & i = j \end{cases}$$

- Process is a **Birth and Death** process if

$$\mathbf{P}_{ij} = 0 \text{ if } |i-j| > 1$$

In this case we write λ_i for the instantaneous "birth" rate:

$$P(Y(t+h) = i+1 | Y_t = i) = \lambda_i h + o(h)$$

and μ_i for the instantaneous "death" rate:

$$P(Y(t+h) = i-1 | Y_t = i) = \mu_i h + o(h)$$

We have

$$q_{ij} = \begin{cases} 0 & |i-j| > 1 \\ \lambda_i & j = i+1 \\ \mu_i & j = i-1 \end{cases}$$

- If all $\mu_i = 0$ then process is a **pure birth** process. If all $\lambda_i = 0$ a **pure death** process.

- Birth and Death process have stationary distribution

$$\pi_n = \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}\right)}$$

Necessary condition for existence of π is

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

- **Linear** birth and death processes have

$$\lambda_n = n\lambda \qquad \mu_n = n\mu$$

In this case

$$\pi_n = \frac{(\lambda/\mu)^n}{1-(\lambda/\mu)}$$

provided $\lambda < \mu$.

Detailed development

Suppose X a Markov Chain with stationary transitions. Then

$$\begin{aligned}
&P(X(t+s)=k|X(0)=i)\\
&\quad=\sum_j P(X(t+s)=k, X(t)=j|X(0)=i)\\
&\quad=\sum_j P(X(t+s)=k|X(t)=j, X(0)=i)\\
&\qquad\quad\times P(X(t)=j|X(0)=i)\\
&\quad=\sum_j P(X(t+s)=k|X(t)=j)\\
&\qquad\quad\times P(X(t)=j|X(0)=i)\\
&\quad=\sum_j P(X(s)=k|X(0)=j)\\
&\qquad\quad\times P(X(t)=j|X(0)=i)
\end{aligned}$$

This shows

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(t)$$

which is the Chapman-Kolmogorov equation.

Now consider the chain starting from i and let T_i be the first t for which $X(t) \neq i$. Then T_i is a stopping time.

[Technically:

$$\{T_i \le t\} \in \mathcal{H}_t$$

for each t.] Then

$$\begin{aligned}
&P(T_i > t+s|T_i > s, X(0)=i)\\
&\qquad= P(T_i > t+s|X(u)=i; 0 \le u \le s)\\
&\qquad= P(T_i > t|X(0)=i)
\end{aligned}$$

by the Markov property. Note: we actually are asserting a generalization of the Markov property: If f is some function on the set of possible paths of X then

$$\begin{aligned}\mathrm{E}(f(X(s+\cdot))|X(u) = x(u), 0 \leq u \leq s)\\ &= \mathrm{E}\left[f(X(\cdot))|X(0) = x(s)\right]\\ &= \mathrm{E}^{x(s)}\left[f(X(\cdot))\right]\end{aligned}$$

The formula requires some sophistication to appreciate. In it, f is a function which associates a sample path of X with a real number. For instance,

$$f(x(\cdot)) = \sup\{t : x(u) = i, 0 \leq u \leq t\}$$

is such a functional. Jargon: **functional** is a function whose argument is itself a function and whose value is a scalar.

FACT: Strong Markov Property – for a stopping time T

$$\mathrm{E}\left[f\left\{X(T+\cdot)\right\}|\mathcal{F}_T\right] = \mathrm{E}^{X(T)}\left[f\left\{X(\cdot)\right\}\right]$$

with suitable fix on event $T = \infty$.

Conclusion: given $X(0) = i$, T_i has memoryless porperty so T_i has an exponential distribution. Let v_i be the rate parameter.

Embedded Chain: Skeleton

Let $T_1 < T_2 < \cdots$ be the stopping times at which tranisitons occur. Then $X_n = X(T_n)$. The sequence X_n is a Markov chain by the strong Markov property. That $\mathbf{P}_{ii} = 0$ reflects the fact that $P(X(T_{n+1}) = X(T_n)) = 0$ by design.

As before we say $i \rightsquigarrow j$ if $\mathbf{P}_{ij}(t) > 0$ for some t. It is fairly clear that $i \rightsquigarrow j$ for the $X(t)$ if and only if $i \rightsquigarrow j$ for the embedded chain X_n.

We say $i$leftrightarrow$j$ if $i \rightsquigarrow j$ and $j \rightsquigarrow i$.

Now consider

$$P(X(t+h) = j|X(t) = i, \mathcal{H}_t)$$

Suppose the chain has made n transitions so far so that $T_n < t < T_{n+1}$. Then the event $X(t+h) = j$ is, except for possibilities of probability $o(h)$ the event that

$$t < T_{n+1} \leq t+h \text{ and } X_{n+1} = j$$

The probability of this is

$$(v_i h + o(h))\mathbf{P}_{ij} = v_i \mathbf{P}_{ij} + o(h)$$