Poisson & Birth/Death Processes
Karlin & Taylor, *A First Course in Stochastic Processes, pp 117–150*

Continuous Time, Discrete State Markov Processes

Let $T = \mathbb{R}^+ = [0, \infty)$, $\mathcal{S} = \mathbb{N} = \{0, 1, 2, ...\}$, and consider a stationary-increment $\mathcal{S}$-valued continuous-time Markov process X_t. Such a process is determined by its initial state x_0 or distribution μ_0 and its transition probability function,

$$P_{ij}(s) = \mathsf{P}[X_{s+t} = j \mid X_t = i],$$

which is independent of t by assumption. In a very short time ϵ, say, from t to $t + \epsilon$, what can happen? For most processes, only one of two things: the process stays put at $X_t = i$, or it jumps from i to some j. All other possibilities (*i.e.*, two or more jumps) sum to only $o(\epsilon)$. This suggests that we consider the "infinitesimal jump rate" Q_{ij} of jumping from i to j, with $P_{ij}(\epsilon) = Q_{ij}\epsilon + o(\epsilon)$, $i \neq j$.

The simplest such process is the *Poisson Process*, which starts at $X_0 = 0$ and jumps only from i to $i + 1$, with a constant rate λ not depending upon i:

$$P_{ij}(\epsilon) = \mathsf{P}[X_{t+\epsilon} = j \mid X_t = i] = \begin{cases} 1 - \lambda\epsilon + o(\epsilon) & \text{if } j = i \\ \lambda\epsilon + o(\epsilon) & \text{if } j = i + 1. \end{cases}$$

Karlin and Taylor use difference equations to show that this equation has a unique solution, the familiar $P_{ij}(t) = \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t}$ for $i \le j$, $P_{ij}(t) = 0$ for $i > j$. Here is a Martingale proof:

It will be enough to show that $P_{ij}(t)$ has the right generating function, $\mathsf{E}[r^{(X_t - X_0)}]$; for a Poisson random variable we must show that this is the Poisson generating function, $e^{\lambda t(r-1)}$. But for *any* function $f(x)$ on $\mathcal{S}$, $\mathsf{E}[f(X_{t+\epsilon})|\mathcal{F}_t] = f(X_t) + \epsilon\lambda\big(f(X_t + 1) - f(X_t)\big) + o(\epsilon)$, so

$$M_t \equiv f(X_t) - \int_0^t \lambda\big(f(X_s + 1) - f(X_s)\big)\, ds$$

is a martingale. In particular, for the power function $f(x) = r^{(x - X_0)}$,

$$M_t \equiv r^{X_t - X_0} - \lambda \int_0^t (r - 1) r^{(X_s - X_0)}\, ds$$

is a martingale and so $g(t) \equiv \mathsf{E}[r^{X_t - X_0}]$ satisfies $g(0) = 1$ and $g'(t) = -\lambda(r - 1)t$, with familiar and unique solution

$$\mathsf{E}[r^{X_t - X_0}] = g(t) = e^{\lambda t(r-1)} = \sum_{j=i}^{\infty} \frac{(\lambda t)^{j-i} e^{-\lambda t}}{(j - i)!}\, r^{(j-i)},$$

the Poisson generating function.

Pure Birth

The Poisson process can be thought of as modeling the size of a population with no death rate, but constant birth (or immigration) rate λ, irrespective of the population size X_t. The simplest generalization is to allow the birth rate λ_i to depend upon the state $X_t = i$, perhaps even to be proportional to it:

$$P_{ij}(\epsilon) = \mathsf{P}[X_{t+\epsilon} = j \mid X_t = i] = \begin{cases} 1 - \lambda_i\epsilon + o(\epsilon) & \text{if } j = i \\ \lambda_i\epsilon + o(\epsilon) & \text{if } j = i + 1. \end{cases}$$

The solution $P_{ij}(t)$ or just the occupation probabilities $P_j(t) = P_{1j}(t) = \mathsf{P}[X_t = j]$ can be found using difference equations (as we did for the Poisson process) or by using martingales, but the easiest way is to notice that the sojourn times τ_j in state j are independent exponentially-distributed RV's with means $\frac{1}{\lambda_j}$, and $P_{ij}(t) = \mathsf{P}[\tau_i + ... + \tau_{j-1} \le t < \tau_i + ... + \tau_j]$. The expected time to travel from i to j is $\mathsf{E}[\tau_i + ... + \tau_{j-1}] = \sum_{i \le k < j} \frac{1}{\lambda_k}$. What happens if $\lambda_i \to \infty$ so fast that the total expected waiting time for all jumps is finite, *i.e.*, $\sum_i \frac{1}{\lambda_i} < \infty$? Then the process explodes! The exploding process is well-defined for times less than $\tau_\infty = \sum_{i<\infty} \tau_i$, but we must specify what happens at and after time τ_∞ to completely specify the process.

Yule-Furry (Linear Birth) Process

If $X_0 = x > 0$ and $\lambda_j = j\beta$ for some $\beta > 0$ we have a model for a population of X_t individuals, each giving birth at rate β; this is the stochastic analogue of exponential growth. Starting with $x = 1$ individuals one can show that $P_n(t) = P_{1n}(t) = e^{-\beta t}(1 - e^{-\beta t})^{n-1}$, *i.e.*, the population has an exponential distribution with mean $e^{\beta t}$ but a finite probability $e^{-\beta t}$ of never increasing at all.

Birth and Death

If both births and deaths are possible, the simplest generalization of the pure-birth process is:

$$P_{ij}(\epsilon) = \mathsf{P}[X_{t+\epsilon} = j \mid X_t = i] = \begin{cases} \lambda_i\epsilon + o(\epsilon) & \text{if } j = i+1 \\ 1 - (\lambda_i + \mu_i)\epsilon + o(\epsilon) & \text{if } j = i \\ \mu_i\epsilon + o(\epsilon) & \text{if } j = i-1. \end{cases}$$

For example, the Yule-like process with $X_0 = x$, $\lambda_j = j\beta$, and $\mu_j = j\delta$ will grow at *about* an exponential rate, with $X_t \sim x\, e^{(\beta-\delta)t}$ if birth rate exceeds death rate, but (unlike in the deterministic case) *could* die out! The waiting time τ_i at a state i again has an exponential distribution, this time with mean $\frac{1}{\lambda_i+\mu_i}$; when X_t leaves site i, it goes up to $i+1$ with probability $\frac{\lambda_i}{\lambda_i+\mu_i}$ and down to $i-1$ with probability $\frac{\mu_i}{\lambda_i+\mu_i}$, independent of τ_i. This suggests how to simulate X_t, a continuous-time random walk.

Some Martingales

Fix a birth/death process X_t and any real number $\alpha \in \mathbb{R}$. What would the process $G(t)$ have to be like for $M_t = \exp[\alpha X_t - G(t)]$ to be a martingale? If $G(t)$ is differentiable then, by Taylor's theorem, $M_{t+\epsilon} = M_t \exp[\alpha X_{t+\epsilon} - \alpha X_t] e^{G(t)-G(t+\epsilon)} \approx M_t e^{\alpha(X_{t+\epsilon}-X_t)-\epsilon G'(t)}$ for small ϵ, so

$$\begin{aligned} \mathsf{E}[M_{t+\epsilon}|X_t = i] &= M_t \times \big(\epsilon\lambda_i e^{\alpha} + \epsilon\mu_i e^{-\alpha} + (1 - \epsilon\lambda_i - \epsilon\mu_i)e^0\big)e^{-\epsilon G'(t)} + o(\epsilon) \\ &= M_t \times \big(1 + \epsilon(e^{\alpha} - 1)[\lambda_i - e^{-\alpha}\mu_i]\big)e^{-\epsilon G'(t)} + o(\epsilon) \\ &= M_t + o(\epsilon^2) \qquad \text{if} \quad G'(t) = (e^{\alpha} - 1)[\lambda_i - e^{-\alpha}\mu_i]. \end{aligned}$$

Thus for any α, the process $M_t = \exp\big[\alpha X_t - (e^{\alpha} - 1)\int_0^t (\lambda_{X_s} - \mu_{X_s} e^{-\alpha})\, ds\big]$ is a martingale. In case of either constant rates $(\lambda_j, \mu_j) \equiv (\beta, \delta)$ or proportional rates $(\lambda_j, \mu_j) \equiv (j\beta, j\delta)$, we can arrange for $G(t)$ to be constant (*i.e.*, $G'(t) \equiv 0$) by making sure that $[\beta - \delta e^{-\alpha}] \equiv 0$, *i.e.*, $e^{\alpha} = \delta/\beta$. In this case $M_t = \big(\delta/\beta\big)^{X_t}$ is a martingale, so for integers $a \le x \le b$, $F_{ab}(x) = \mathsf{P}[\tau_a < \tau_b \mid X_0 = x] = \mathsf{P}[X_t = a \text{ before } X_t = b \mid X_0 = x]$ is just

$$\mathsf{P}[\tau_a < \tau_b \mid X_0 = x] = \frac{(\delta/\beta)^x - (\delta/\beta)^b}{(\delta/\beta)^a - (\delta/\beta)^b} \longrightarrow \begin{cases} (\delta/\beta)^{x-a} & \text{if } \delta < \beta, \text{ as } b \to \infty \\ 1 & \text{if } \delta > \beta, \text{ as } b \to \infty. \end{cases}$$

The expectation and generating function for X_t can be found in the same way by considering expectations of the martingales

$$M_t^* \equiv X_t - \int_0^t (\lambda_{X_s} - \mu_{X_s})\, ds \qquad M_t^{(\alpha)} \equiv e^{\alpha X_t} - (e^{\alpha} - 1)\int_0^t e^{\alpha X_s}\big(\lambda_{X_s} - e^{-\alpha}\mu_{X_s}\big)\, ds;$$

for the Yule-like process with $\lambda_j = j\beta$ and $\mu_j = j\delta$, these are simply $M_t^* = X_t - (\beta - \delta)\int_0^t X_s\, ds$ and $M_t^{(\alpha)} = e^{\alpha X_t} - (e^{\alpha} - 1)(\beta - e^{-\alpha}\delta)\int_0^t X_s e^{\alpha X_s}\, ds$, so the expectations $f_x(t) \equiv \mathsf{E}[X_t \mid X_0 = x]$ and $g_x(t,\alpha) \equiv \mathsf{E}[e^{\alpha X_t} \mid X_0 = x]$ can be found by solving the integral equations $x = f_x(t) - (\beta - \delta)\int_0^t f_x(s)\, ds$ and $e^{\alpha x} = g_x(t,\alpha) - (e^{\alpha} - 1)(\beta - e^{-\alpha}\delta)\int_0^t \frac{\partial}{\partial\alpha} g_x(s,\alpha)\, ds$, or the differential equations (with initial values $f_x(0) = x$, $g_x(0,\alpha) = e^{\alpha x}$):

$$\frac{\partial}{\partial t} f_x(t) = (\beta - \delta) f_x(t) \qquad \frac{\partial}{\partial t} g_x(t,\alpha) = (e^{\alpha} - 1)(\beta - e^{-\alpha}\delta)\frac{\partial}{\partial\alpha} g_x(t,\alpha).$$

The solutions are $f_x(t) \equiv \mathsf{E}[X_t|X_0 = x] = x e^{(\beta-\delta)t}$, $g_x(t,\alpha) \equiv \mathsf{E}[e^{\alpha X_t}|X_0 = x] = \Big(\frac{\delta(1-e^{\alpha})e^{(\beta-\delta)t} + (\beta e^{\alpha} - \delta)}{\beta(1-e^{\alpha})e^{(\beta-\delta)t} + (\beta e^{\alpha} - \delta)}\Big)^x$; the probability $P_{ij}(t) = \mathsf{P}[X_t = j \mid X_0 = i]$ is just the coefficient of r^j in the power-series for $g_i(t, \log r) = \mathsf{E}[r^{X_t}|X_0 = i] = \sum_j P_{ij}(t) r^j = \Big(\frac{\delta(1-r)e^{(\beta-\delta)t} + (\beta r - \delta)}{\beta(1-r)e^{(\beta-\delta)t} + (\beta r - \delta)}\Big)^i$.