LECTURE NOTES

ANDREW

10/31

Cell Proliferation and Cancer

The following setup is the motivation for a later theorem on Hitting Probabilities. Mutations can affect the rates of cell proliferation (splitting) and death. Suppose we counted the number of occurences of a particular mutation in each of many tissue samples. We wonder if the distribution of mutation numbers is consistent with a driver mutation, i.e. one with a higher than normal rate of proliferation † .

Model

Let X_t be the number of mutated cells at time t, so $X_t \in \mathbb{N}_{\geq 0}$, and assume that $X_0 = 1$ (so we start with one mutated cell).

Assume that each cell divides at rate λ and dies at rate μ. This means that

$$\begin{aligned} X_t &\to X_t + 1 \quad \text{at rate } \lambda X_t \\ X_t &\to X_t - 1 \quad \text{at rate } \mu X_t. \end{aligned}$$

Questions:

(1) What is the probability that the mutation "grows?" That is, for some large N, what is
$$\mathbb{P}\{X_t = N \text{ before } X_t \text{ hits } 0\} = ?$$

(2) If the mutation is not a driver but present due to an ongoing mutation of rate a, what is the expected distribution of the number of mutations per tissue sample?

Hitting Probabilities

To answer question 1 we need a bit more theory. Suppose $(X_t)_{t\geq 0}$ is a Continuous Time Markov Chain with generator G on a state space $\mathcal{X}$. For any subset $A \subset \mathcal{X}$, we define

$$\tau_A = \inf\{t \geq 0 \mid X_t \in A\},$$

which is the first time that the Markov chain lands in A. If $A = \{x\}$ we just write τ_x, which is the first time the Markov chain hits x. We now state the theorem:

> **0.1 Theorem** (Harmonic Functions) Suppose that $A, B \subset X$ are disjoint subsets with the property that
> $$\shortmid\{\tau_{A\cup B} < \infty\} = 1.$$

Date: 31 October 2018.

† Here, normal would mean that the rate of cell proliferation equals the rate of cell death.

For $x \in \mathcal{X}$ we define
$$h(x) = \mathbb{P}^x\{\tau_A < \tau_B\} := \mathbb{P}\{\tau_A < \tau_B \mid X_0 = x\}.$$
Then the function h is the unique function satisfying the following two conditions:

(1) $h(x) = \begin{cases} 1 & \text{for } x \in A \\ 0 & \text{for} x \in B \end{cases}$

(2) $Gh(x) = 0$ for $x \notin A \cup B$.

To prove the theorem we need two lemmas.

0.2 Lemma Let $T_1, \dots, T_n$ be independent random variables with $T_i \sim \text{Exp}(\lambda_k)$. Let $T := \min\{T_1, \dots, T_n\}$. Then $T \sim \text{Exp}\left(\sum_{k=1}^n \lambda_k\right)$ and
$$\mathbb{P}\{T_k = T\} = \frac{\lambda_k}{\lambda_1 + \cdots + \lambda_n}.$$

Proof of Lemma 1.
$$\begin{aligned} \mathbb{P}\{T > t\} &= \mathbb{P}\{T_k > t \quad \forall k\} \\ &= \prod_{k=1}^n \mathbb{P}\{T_k > t\} && \text{(by independence)} \\ &= \prod_{k=1}^n e^{-\lambda_k t} \\ &= e^{-t \sum_k \lambda_k} \end{aligned}$$
which establishes the first claim. For the second claim, first note that
$$\begin{aligned} \mathbb{P}\{T_k \in dt, T_j > t \quad \text{for } j \neq k\} &= \lambda_k e^{-\lambda_k t} \prod_{j \neq k} e^{-\lambda_j t} \\ &= \lambda_k e^{-t \sum_j \lambda_j}, \end{aligned}$$
so we now integrate over t to get
$$\mathbb{P}\{T_k = T\} = \lambda_k \int e^{-t \sum_j \lambda_j}\, dt = \frac{\lambda_k}{\lambda_1 + \cdots + \lambda_n}.$$
$\square$