Homework 3 posted, due Tuesday, November 29.

Continuing with our classification of birth-death chains on nonnegative integers. Last time we started by looking for an invariant measure and found one. Because the Markov chain is irreducible, this invariant measure is unique up to multiplicative constant. The question of whether a stationary distribution exists or not hinges on whether or not this invariant measure is normalizable.

\[ \sum_{j=0}^{\infty} v_j < \infty \]

\[ \sum_{j=0}^{\infty} \left( \sum_{k=0}^{i-1} \frac{r_k}{q_{k+1}} \right) < \infty \]

If this is true, then a stationary distribution exists, and the Markov chain is known to be positive recurrent. If this criterion fails, then the Markov chain is either null recurrent or transient, but to decide which, we need to apply the decisive for transience.

Choose \( \hat{i}_x = 0 \)

As the reference state. Study solutions to \( Qx = x \) where \( Q \) is the matrix obtained by deleting the 0th row and column from the probability transition matrix.
This system of equations can be solved recursively in the same way that we solved absorption probability equations for a finite birth-death chain

\[
\begin{align*}
    r_1 x_1 + \beta_1 x_2 &= x_1 \\
    q_2 x_1 + \gamma_2 x_2 + \beta_2 x_3 &= x_2 \\
    &\quad \vdots \\
    q_j x_{j-1} + \gamma_j x_j + \beta_j x_{j+1} &= x_j
\end{align*}
\]

Solution:

Decisive for transience: Existence (or lack thereof) of bounded nontrivial nonnegative solutions to \( Qx = x \).

Putting all these tests together, we obtain the following classification of birth-death chains on the nonnegative integers with reflecting boundary conditions at 0 based on the particular model for the birth and death probabilities:

- Positive recurrent whenever

\[
\sum_{j=1}^{\infty} \gamma_j < \infty
\]

- Transient whenever

\[
\sum_{j=1}^{\infty} \frac{\beta_j}{\gamma_j} < \infty
\]

- Positive recurrent whenever

\[
\sum_{j=1}^{\infty} \frac{\beta_j}{\gamma_j} < \infty
\]

- Transient whenever

\[
\sum_{j=1}^{\infty} \frac{\beta_j}{\gamma_j} < \infty
\]
Branching Processes

Readings: Lawler Sec. 2.4
Resnick Sec. 1.4

Branching processes are a special case of discrete-time countable-state Markov chains that have their own theory and many applications. In fact they have continuous-time versions which are also important but more complicated; we'll just introduce the basic ideas.

We will focus on the most basic branching process model which is called the Galton-Watson model:

One represents the branching process in terms of the number $X_n$ of (indistinguishable) agents present at epoch (generation) $n$. Each agent, in each epoch, gives rise to a random number of offspring in the next epoch. (If the agent survives until the next epoch, that's counted as one of its offspring.) Each of these "branching events" is independent and identically distributed in terms of the number of offspring produced.

Probability transition matrix is a pain to write down. But the stochastic update rule is fairly direct:
Probability transition matrix is a pain to write down. But the stochastic update rule is fairly direct:

$$X_{n+1} = \sum_{k=1}^{Y_{n,k}}$$

Where $Y_{n,k}$ is the number of offspring generated by the $k$th agent from the $n$th epoch. They are all iid random variables with some prescribed probability distribution

$$P_j = P(\sum_{i=1}^{Y_{n,k}} = j)$$

This basic Galton-Watson model can be extended in the following important directions:
- Multiple types of agents
- Age structure
- Continuous time

Applications of the general branching process models:
- Genealogy, including mutations and variations, phylogenetics
- Growth and activation of cancerous tumors (Rick Durrett (Duke) and Jasmine Foo (Minnesota))
- Polymerase chain reaction (PCR) (K. Mullis)
- Cell division
- Population growth (without hindrances)
- Disease spread (in early stages) (work w/ K. Bennett)
- Proliferation of viruses in an immune system model
- Photomultiplier tube cascades
- Nuclear fission
- Earthquake triggering
- Queueing models

A good reference for discussion on application of more advanced branching process models is:
**Branching Processes in Biology, Kimmel and Axelrod**

Mathematical analysis of branching processes.

The key tool is **probability generating function**

$$P_X(s) = E s^X = \sum_{j=0}^{\infty} P(X=j)s^j$$
Because:
- Generating functions work well with sums and random sums of independent random variables, and that's what the stochastic update rule looks like
- Recursive structure

We use pgf rather than characteristic function because the random variable is discrete, but one could have used the characteristic function.

Recall the following formula for the probability generating function for a random sum:

\[ Z = \sum_{j=1}^{N} X_j \]

\[ \{ X_j \}_{j=1}^{N} \text{ are independent,} \]

\[ \{ X_j \}_{j=1}^{\infty} \text{ are i.i.d.} \]

\[ P_Z(s) = \prod_{N=1}^{\infty} P(X_i(s)) \]

Applying this to the stochastic update rule for branching processes, we get:

\[ P_{\Sigma_{n+1}}(s) = P\left( \frac{P_{\Sigma_n}(s)}{P_{\Sigma_0}(s)} \right) \]

\[ P_{\Sigma_{-1}}(s) = P\left( \frac{P_{\Sigma_0}(s)}{P_{\Sigma_0}(s)} \right) \]

Induction....
Induction....

\[
P_{\Sigma_n}(s) = P_{\Sigma_0}(P^{\times(n)}(s))
\]

\[
P^{\times(n)}(s) = P_{\Sigma}(P_{\Sigma}(\cdots P_{\Sigma}(s)))
\]

\[\text{in-fold composition fun}\]

This is easy to implement on a computer with recursion.

One can get analytical formulas for some basic statistics.

\[
E[X_n] = \left. s \frac{d}{ds} P_{\Sigma_n}(s) \right|_{s=1}
\]

\[
= \left. s \frac{d}{ds} \left( P_{\Sigma_0}(P^{\times(n)}(s)) \right) \right|_{s=1}
\]

\[
= \left. s \left( P^{\times(n)}(s) \right) \left( P^{\times(n)}(s) \right) \cdots \right|_{s=1}
\]

\[
= P^{\times(n)}(1) \cdot P^{\times(n)}(1) \cdot \cdots
\]

(iterated chain rule)

\[
P_{\Sigma}(1) = 1
\]

\[
P^{\times(n)}(1) = 1
\]

\[
P'(1) \cdot s \cdot P'(s)
\]
One gets more nontrivial yet still computable formulas for higher moments through the same procedure:

\[ \mathbb{E}X_n = \mathbb{E}X_0 \left( \mathbb{E}X \right)^n \]

This is how one can compute statistics of the branching process over finite time horizons.

What about the long time properties of branching processes?

Let's look first topologically:

0 is an absorbing state. Under generic assumptions about the branching probabilities the other states form their own communication class. They must be transient states since that communication class is not open.

But the transience could take two qualitatively different forms: either the Markov chain gets absorbed at 0 or it grows without bound. Key questions:

- What is the probability to go extinct (vs. grow unboundedly)? (absorption probability)
- When the branching process goes extinct, what is the expected time for it to go extinct? (first passage time)
- If the branching process goes extinct, how much cost/reward did the branching process incur before it
went extinct?

• When the branching process grows indefinitely, what is its rate of growth?

We'll address the first question. The second and third questions can be attacked by similar means (see the texts). The fourth question requires more sophisticated tool that we'll return to later in the class.

We'll turn now to compute the probability for the branching process to go extinct.

\[ a(k) = \mathbb{P}(\exists n \geq 0 \mid X_0 = k) \]

(The complementary probability is the probability to grow unboundedly).

We will compute these extinction probabilities through a separate first-step analysis (because we don't have the probability transition matrix handy).

First, we use independence of the agents to argue that:

\[ a(k) = \left( a(1) \right)^k \]

\[ a(1) = \mathbb{P} \left( \bigcap_{i=1}^k \left( \text{descendants of initial agent } i \right) \right) \]

\[ = \prod_{i=1}^k \mathbb{P} \left( \text{descendants of initial agent } i \right) \]

\[ = \prod_{i=1}^k \mathbb{P} \left( \text{descendants of initial agent } i \right) \]

\[ = \prod_{i=1}^k a(1) \]

\[ = \left( a(1) \right)^k \]

So for simplicity of notation set

\[ a = a(1) \]

\[ q = \mathbb{P} \left( \bigcup_{n=1}^q \{ X_n > 0 \} \mid X_0 = 1 \right) \]
Now we employ first-step analysis to set up a recursive equation for $a$. Use law of total probability to insert information about what happens at the first step.

\[
    a = \sum_{j=0}^{\infty} \sum_{n=1}^{\infty} P(B \cup \{X_n > 0\} \mid X_1 = j, X_0 = 1) P(X_1 = j \mid X_0 = 1)
\]

\[
    = P(B \cup \{X_n > 0\} \mid X_1 = 0, X_0 = 1) P(X_1 = 0 \mid X_0 = 1)
\]

\[
    + \sum_{j=1}^{\infty} P(B \cup \{X_n > 0\} \mid X_1 = j, X_0 = 1) P(X_1 = j \mid X_0 = 1)
\]

\[
    = 1 \cdot p_0 + \sum_{j=1}^{\infty} P(B \cup \{X_n > 0\} \mid X_1 = j, X_0 = 1) p_j
\]

\[
    Morkov property
\]

\[
    = p_0 + \sum_{j=1}^{\infty} a(j) p_j
\]

\[
    q = p_0 + \sum_{j=1}^{\infty} a(j) p_j
\]

\[
    q = p \left( \frac{1}{a(j)} \right)
\]
What do the solutions of this equation look like? Is it unique? If not, which solution is the right one?

There are some trivial cases to dispense with:

\[ \alpha \rho_0 = 0, \; \rho_j > 1, \; \rho_j \geq 0 \quad \text{for} \; j \geq 2 \quad \Rightarrow \quad q_j (a) > q, \; a > 0, \]

\[ \alpha \rho_0 > 1, \; \rho_j > 1, \; \rho_j \geq 0 \quad \text{for} \; j \geq 2 \quad \Rightarrow \quad q_j (a) = \rho_j (1 - \rho_0) q_j, \; a < 1. \]

Next time: the nontrivial cases.