MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070J Lecture 5 Feb 27, SP17
Lecturer: Guy Bresler Scribe notes by Justin Holmgren

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They are posted to serve class purposes.*

Branching Processes

Content.

1. The Galton-Watson Process
2. Survival Probability
3. Total Progeny
4. Random Walk Representation
5. Hitting Time Theorem

1 The Galton-Watson Process

In this lecture we study the following stochastic process, which is called the Galton-Watson process [1]. Let $X \in \mathbb{Z}_{\geq 0}$ be a random variable for which $\mu \stackrel{\text{def}}{=} \mathbb{E}[X] < \infty$. Define $p_k \stackrel{\text{def}}{=} \Pr[X = k]$

Let $\{X_{n,i}\}_{n,i\geq 1}$ be i.i.d. random variables, each of which is equal in distribution to X. Define $Z_0 = 1$ and $Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i}$ for $n \geq 1$. This formalizes a generational process with a single "founder" at generation 0, where each individual has a number of children equal in distribution to X.

This captures the dynamics of several natural processes:

- The propagation of a family name
- Nuclear chain reactions
- Percolation on the d-ary tree $\hat{\mathbb{T}}_d$.

Additionally, this is a fundamental process which appears often in the study of random graphs.

Exercise 1. *Show that* $\mathbb{E}[Z_n] = \mu^n$*, and* $M_n = \frac{Z_n}{\mu^n}$ *is a martingale with respect to* $\mathcal{F}_n = \sigma(Z_1, \ldots, Z_n)$*.*

From this we can deduce by the Martingale Convergence Theorem that $M_n \overset{\text{a.s.}}{\to} M_\infty < +\infty$ and $\mathbb{E}[M_\infty] \leq 1$ by Fatou's lemma.

2 Survival Probability

One fundamental question we can ask is: what is $\eta \overset{\text{def}}{=} \mathbb{P}\left[X_n = 0 \text{ for some } n\right]$? We will assume that $p_0 > 0$, as otherwise $\eta = 0$.

Exercise 2. *Show that almost surely, either* $Z_n \to 0$ *or* $Z_n \to \infty$*. Prove that if* $\mu \leq 1$ *then* $\eta = 1$*.*

Theorem 1. η *is the smallest fixed point of* f *in* $[0, 1]$ *(recall that* $f(s) = \mathbb{E}[s^X]$*). Furthermore,*

- *If* $\mu < 1$ *then* $\eta = 1$
- *If* $\mu > 1$ *then* $\eta < 1$*.*

Proof. We analyze the generating function of Z_n. That is, define $f_t(s) \overset{\text{def}}{=} \mathbb{E}[s^{Z_t}] = \sum_{k \geq 0} s^k p_k$. Note that $f(s)$ is at least defined when $0 < s < 1$, and if we adopt a convention that $0^0 = 1$, then $f(0) = p_0$.

$$\begin{aligned}
\eta &= \mathbb{P}\left[\exists t \text{ s.t. } Z_t = 0\right] \\
&= \mathbb{P}\left[\cup_{t=0}^{\infty} \{Z_t = 0\}\right] \\
&= \lim_{t \to \infty} \mathbb{P}\left[Z_t = 0\right] \\
&= \lim_{t \to \infty} f_t(0).
\end{aligned}$$

So it suffices to compute

$$\begin{aligned}
f_t(s) &= \mathbb{E}\left[s^{Z_t}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[s^{Z_t} \middle| \mathcal{F}_{t-1}\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[s^{\sum_{i=1}^{Z_{t-1}} X_{t,i}} \middle| \mathcal{F}_{t-1}\right]\right] \\
&= \mathbb{E}\left[\prod_{i=1}^{Z_{t-1}} \mathbb{E}\left[s^{X_{t,i}}\right]\right] \\
&= \mathbb{E}\left[f(s)^{Z_{t-1}}\right] \\
&= f_{t-1}(f(s)).
\end{aligned}$$

So by induction on t, $f_t(s) = \overbrace{f \circ \cdots \circ f}^{t}(s)$, which we will denote by $f^{(t)}(s)$.

Returning to η, we find that $\eta = \lim_{t\to\infty} f^{(t)}(0)$, which intuitively should be the first fixed point of f. Proving this rigorously uses the following properties of f, which we will not prove. On the interval $[0, 1]$, f satisfies:

- $f(0) = p_0$ and $f(1) = 1$.
- f is infinitely differentiable.
- f is convex and increasing.
- $\lim_{s\nearrow 1} f'(s) = \mu < \infty$.

$\square$

3 Total Progeny

Another quantity of interest is $W = \sum_{t=0}^{\infty} Z_t$, i.e. the total progeny.

Exercise 3. *If* $\mu < 1$, *then* $\mathbb{E}[W] = \frac{1}{1-\mu}$.

From the results of the previous section, it is easy to see that if $\mu > 1$, then $\mathbb{E}[W] = \infty$. We will see how to go further and exactly calculate the distribution of W.

3.1 Random Walk Representation

So far we have analyzed the Galton-Watson process at the granularity of generations. We can get a finer-grained view of the obtained tree T by considering one vertex at a time. Specifically, we will define sets $\mathcal{A}_t$ (the "active vertices"), $\mathcal{E}_t$ (the "explored vertices"), and $\mathcal{N}_t$ (the "neutral vertices") as follows.

1. Initialize $\mathcal{A}_0 = \{0\}$, $\mathcal{E}_0 = \emptyset$, and $\mathcal{N}_0 = V(T) \setminus \{0\}$. (Here $V(T)$ denotes the set of vertices of T).

2. For $t \geq 1$:

   - If $\mathcal{A}_{t-1} = \emptyset$, do nothing. That is, $(\mathcal{A}_t, \mathcal{E}_t, \mathcal{N}_t) = (\mathcal{A}_{t-1}, \mathcal{E}_{t-1}, \mathcal{N}_{t-1})$.
   - Otherwise, choose an arbitrary $a_t \in \mathcal{A}_{t-1}$. Let

$$\mathcal{A}_t = \mathcal{A}_{t-1} \cup \{\text{children of } a_t\} \setminus \{a_t\},$$

$$\mathcal{E}_t = \mathcal{E}_{t-1} \cup \{a_t\},$$

   and

$$\mathcal{N}_t = \mathcal{N}_{t-1} \setminus \{\text{children of } a_t\}.$$

We will write A_t, E_t, and N_t to respectively denote $|\mathcal{A}_t|$, $|\mathcal{E}_t|$, and $|\mathcal{N}_t|$. We are interested in $\tau_0 \stackrel{\mathsf{def}}{=} \inf\{t \geq 0 : A_t = 0\}$, because of the following lemma, which we will not prove.

Lemma 2. $W = \tau_0$.

Now we notice that the consecutive values of A_t are almost a random walk. Specifically, we could have just defined $A_0 = 1$ and

$$A_t = \begin{cases} A_{t-1} + X_t - 1 & \text{if } t \leq \tau_0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\{X_t\}_{t\geq 1}$ are i.i.d. random variables distributed identically to X. If we define $Y_t = X_t - 1$ and $S_t = 1 + \sum_{i=1}^{t} Y_t$, then S_t is genuinely a random walk and $A_t = S_{t \wedge \tau_0}$. Note that $Y_t \geq -1$.

τ_0 is equivalently defined in terms of S_t as

$$\begin{aligned} \tau_0 &= \inf\{t \geq 0 : S_t = 0\} \\ &= \inf\{t \geq 0 : \sum_{i=1}^{t} X_t = t - 1\}. \end{aligned}$$

The following exercise shows that when the expected number of children is greater than 1, the probability of having at least k progeny but nevertheless going extinct decreases exponentially with k.

Exercise 4 (Difficult)**.** *If $\mu > 1$, then*

$$\mathbb{P}\left[k \leq W < \infty\right] \leq \frac{e^{-kI}}{1 - e^{-I}}$$

where

$$I = \sup_{t \leq 0} \left(t - \log \mathbb{E}\left[e^{tX}\right]\right) > 0.$$

Theorem 3.

$$\mathbb{P}\left[W = n\right] = \frac{1}{n} \cdot \mathbb{P}\left[X_1 + \cdots + X_n = n - 1\right]$$

and more generally, if we have k i.i.d. copies $W_1, \ldots, W_k$ of W, then

$$\mathbb{P}\left[W_1 + \cdots + W_k = n\right] = \frac{k}{n} \cdot \mathbb{P}\left[X_1 + \cdots + X_n = n - k\right].$$

Proof. This is just a special case of the *Hitting Time Theorem*, which we prove in the Section 3.3. $\square$

Corollary 4. *When X is a Poisson distribution with intensity λ,*

$$\mathbb{P}[W = n] = \frac{(\lambda n)^{n-1} e^{-\lambda n}}{n!}$$

for $n \geq 1$.

3.2 An Aside on Duality

Suppose that we are only interested in the distribution of Z_t *conditioned on extinction*. It turns out that the resulting conditional distribution of Z_t is the same as the distribution of Z'_t in a related branching process.

Theorem 5. *Let $\{Z_t\}_{t \geq 0}$ be a Galton-Watson process defined by p_k such that $\eta < 1$. Then conditioned on extinction, $Z_t \stackrel{d}{=} Z'_t$, where $\{Z'_t\}_{t \geq 0}$ is a Galton-Watson process defined by $p'_k = \eta^{k-1} p_k$.*

Proof. First, observe that p'_k does in fact define a probability distribution, because

$$\sum p'_k = \frac{1}{\eta} \sum \eta^k p_k = \frac{1}{\eta} \mathbb{E}\left[\eta^k\right] = \frac{1}{\eta} \cdot f(\eta) = 1.$$

Now we think of $\{Z_t\}$ and $\{Z'_t\}$ as corresponding to valid "histories" $H = (X_1, \ldots, X_{\tau_0})$ and $H' = (X'_1, \ldots, X'_{\tau'_0})$. Note Z_t goes extinct iff $\tau_0 < \infty$. A (finite) history $(x_1, \ldots, x_t)$ is *valid* if:

- For every $i < t$, $X_1 + \cdots + X_i > i - 1$
- $X_1 + \cdots + X_t = t - 1$.

Now for every valid $(x_1, \ldots, x_t)$,

$$\begin{aligned}
\mathbb{P}[H = (x_1, \ldots, x_t) | \tau_0 < \infty] &= \frac{\mathbb{P}[H = (x_1, \ldots, x_t)]}{\mathbb{P}[\tau_0 < \infty]} \\
&= \frac{1}{\eta} \prod_{i=1}^{t} p_{x_i} \\
&= \frac{1}{\eta} \prod_{i=1}^{t} \eta^{1-x_i} p'_{x_i} \\
&= \prod_{i=1}^{t} p'_{x_i}
\end{aligned}$$

because $\sum_{i=1}^{t} x_i = t-1$

$$= \mathbb{P}\left[H' = (x_1, \ldots, x_t)\right].$$

□

Exercise 5. *Suppose that* $X \sim \mathrm{Pois}(\lambda)$ *for* $\lambda > 1$. *Then conditioned on extinction,* $Z_t \overset{d}{=} Z'_t$ *where* $X' \sim \mathrm{Pois}(\mu)$ *for the (unique)* $\mu < 1$ *such that* $\lambda e^{-\lambda} = \mu e^{-\mu}$.

3.3 Hitting Time Theorem

Theorem 6. *Let* Y *be any integer-valued random variable such that* $Y \geq -1$ *almost surely, let* $\{Y_t\}_{t\geq 1}$ *be i.i.d. copies of* Y, *and let* $\mathbb{P}_k$ *denote the law of* $S_n \overset{\text{def}}{=} S_0 + \sum_{i=1}^{n} Y_i$ *for* $S_0 = k$. *Define* $\tau_0 \overset{\text{def}}{=} \inf\{t \geq 0 : S_t = 0\}$. *Then for all* $n \geq 1$ *and all* $k \geq 0$,

$$\mathbb{P}_k\left[\tau_0 = n\right] = \frac{k}{n} \cdot \mathbb{P}_k\left[S_n = 0\right].$$

Proof. Our proof is by induction on n.

Base case When $n = 1$, we consider three cases for k.

- If $k = 0$, then $\tau_0 = 0$, so

$$\mathbb{P}_k\left[\tau_0 = n\right] = \frac{k}{n} \cdot \mathbb{P}_k\left[S_n = 0\right] = 0.$$

- If $k = 1$, then

$$\mathbb{P}_k\left[\tau_0 = n\right] = \frac{k}{n} \cdot \mathbb{P}_k\left[S_n = 0\right] = \mathbb{P}\left[Y = -1\right].$$

- If $k > 1$, then S_n cannot possibly be 0 (because $Y \geq -1$) so

$$\mathbb{P}_k\left[\tau_0 = n\right] = \frac{k}{n} \cdot \mathbb{P}_k\left[S_n = 0\right] = 0.$$

Inductive Step When $n > 1$, we consider two cases for k.

- If $k = 0$, then just as before $\tau_0 = 0$ so

$$\mathbb{P}_k[\tau_0 = n] = \frac{k}{n} \cdot \mathbb{P}_k[S_n = 0] = 0.$$

- The interesting case / the case where we actually use the inductive assumption is when $k \geq 1$. In this case, we have

$$\mathbb{P}_k[\tau_0 = n] = \sum_{s \geq -1} \mathbb{P}_k[\tau_0 = n | Y_1 = s] \cdot \mathbb{P}[Y_1 = s].$$

But

$$\begin{aligned}
\mathbb{P}_k[\tau_0 = n | Y_1 = s] &= \mathbb{P}_{k+s}[\tau_0 = n-1] \\
&= \frac{k+s}{n-1} \cdot \mathbb{P}_{k+s}[S_{n-1} = 0] \qquad \text{by hypothesis} \\
&= \frac{k+s}{n-1} \cdot \mathbb{P}_k[S_n = 0 | Y_1 = s]
\end{aligned}$$

So by Baye's rule,

$$\begin{aligned}
\mathbb{P}_k[\tau_0 = n] &= \frac{k}{n-1}\mathbb{P}_k[S_n = 0] + \sum_{s \geq -1} \frac{s}{n-1} \cdot \mathbb{P}_k[Y_1 = s | S_n = 0] \cdot \mathbb{P}_k[S_n = 0] \\
&= \left(\frac{k}{n-1} + \frac{1}{n-1}\mathbb{E}_k[Y_1 | S_n = 0]\right) \cdot \mathbb{P}_k[S_n = 0].
\end{aligned}$$

But since each Y_i is identically distributed, we have

$$\mathbb{E}_k[Y_1 | S_n = 0] = \frac{1}{n}\mathbb{E}_k\left[\sum_{i=1}^{n} Y_i \middle| S_n = 0\right] = -\frac{k}{n}.$$

Plugging everything in and simplifying, we get

$$\begin{aligned}
\frac{\mathbb{P}_k[\tau_0 = n]}{\mathbb{P}_k[S_n = 0]} &= \frac{k}{n-1} - \frac{k}{n(n-1)} \\
&= \frac{nk - k}{n(n-1)} \\
&= \frac{k}{n}.
\end{aligned}$$

$\square$

References

[1] Henry William Watson and Francis Galton, *On the probability of the extinction of families*, The Journal of the Anthropological Institute of Great Britain and Ireland **4** (1875), 138–144.