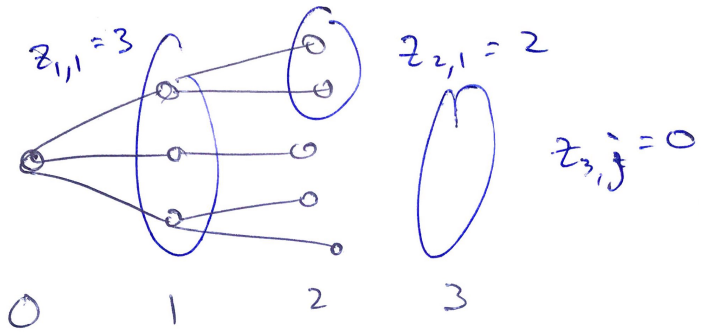
Lecture 12: Branching processes

For the exploration process in lecture eleven, at each step we add (uncover) a random number of new vertices in the component, where the random number depends on the number of already uncovered vertices. A branching process has a similar structure, but there is more independence in the added vertices.

Definition 12.1. Let $(Z_{t,i})_{t,i\geq 1}$ be i.i.d. non-degenerate and non-negative integer-valued random variables with finite mean. A Galton-Watson Branching Process (BP) $(X_t)_{t\geq 1}$ with offspring distribution $\mathscr{L}(Z_{1,1})$ is a Markov chain with $X_0 = 1$ and, given X_{t-1},

$$X_t = \sum_{i=1}^{X_{t-1}} Z_{t,i}.$$

We think of the BP as the generation sizes of a population over time, where every individual of a generation independently gives birth to a number of offspring that is distributed as $\mathscr{L}(Z_{1,1})$, and these births constitute the next generation. The following figure gives an example of a BP.



Example branching process with $Z_{1,1} = 3$, $Z_{2,1} = 2$, $Z_{2,2} = 1$, and $Z_{2,3} = 2$, and $Z_{3,j} = 0$ for $j = 1, 2, 3, 4, 5$. $(X_0, X_1, \ldots) = (1, 3, 5, 0, 0, \ldots)$

The first property of BPs we study, is their chance of extinction $\eta := \mathbb{P}(\cup_{t\geq 1}\{X_t = 0\})$. Let $(p_k)_{k\geq 0} := (\mathbb{P}(Z_{1,1} = k))_{k\geq 0}$ be the offspring point probabilities, set $m := \mathbb{E}Z_{1,1}$, and denote the generating function of $Z_{1,1}$ by

$$f(z) = \sum_{k\geq 0} p_k z^k.$$

Proposition 12.2. *The extinction probability η is the smallest non-negative solution to*

$$z = f(z).$$

If $m \leq 1$, then $\eta = 1$, and if $m > 1$, then $\eta < 1$.

Proof. Let $\eta_t = \mathbb{P}(X_t = 0) = \mathbb{P}(\cup_{s=1}^{t}\{X_s = 0\})$ be the probability the BP is extinct by time t. Then continuity of probabilities implies $\eta_t \nearrow \eta$. Denoting the generating function of X_t by $f_t(x) = \mathbb{E}[x^{X_t}]$, note that $\eta_t = f_t(0)$. For $t \geq 1$,

$$\begin{aligned} f_t(x) &= \mathbb{E}\big\{\mathbb{E}[x^{X_t}|X_{t-1}]\big\} \\ &= \mathbb{E}\big\{\mathbb{E}[x^{\sum_{j=1}^{X_{t-1}} Z_{t,j}}|X_{t-1}]\big\} \\ &= \mathbb{E}\big[f(x)^{X_{t-1}}\big] \\ &= f_{t-1}\big(f(x)\big) \\ &= f^{[t]}(x), \end{aligned}$$

where $f^{[t]}$ denotes t-fold composition of f with itself. We state the key properties of f, which follow from its power series representation with radius of convergence at least one, and generating function facts.

- $f(0) = p_0$ and $f(1) = 1$,
- f is increasing on $[0, \infty)$,
- $f'(1) = m$,
- $f''(x) \geq 0$ for $x \in (0, \infty)$, and equals zero only if $p_0 + p_1 = 1$.

If $p_0 + p_1 = 1$, then $m < 1$ (since we are assuming the offspring distribution is non-degenerate) and the process dies out after a geometrically distributed amount of time, so $\eta = 1$, which is the only solution to $z = f(z)$, and the result follows. Otherwise, let z_0 be the minimal non-negative solution to $z = f(z)$; the following picture exemplifies the situation.

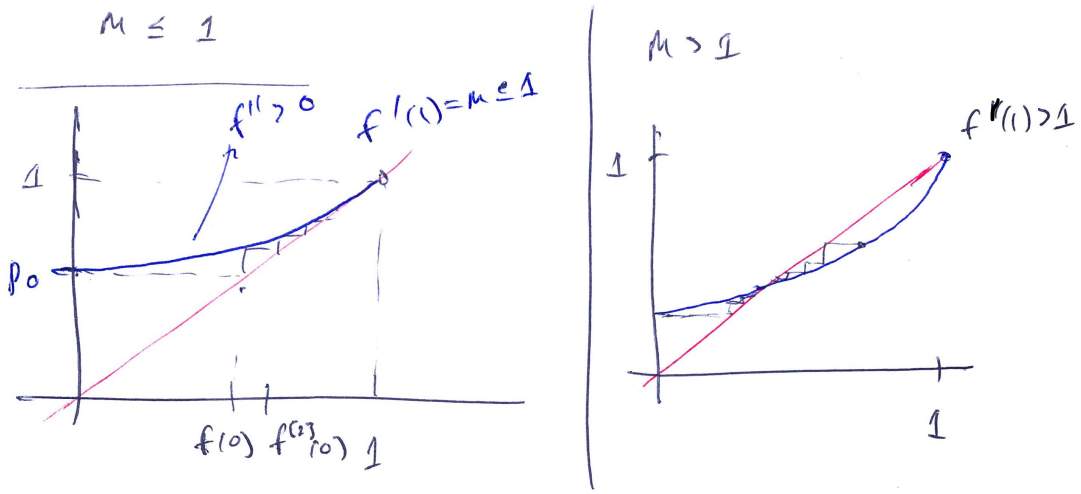


Illustration of iterating the generating function f for $m \leq 1$ and $m > 1$.

Since f is (strictly) convex and increasing, for $0 \leq x < z_0$, $x \leq f(x) \leq f(z_0) = z_0$, and so $f^{[t]}(x) \nearrow y$ for some limit $y \leq z_0$, but the continuity of f implies that

$$y = \lim_{t \to \infty} f^{[t]}(x) = f\Big(\lim_{t \to \infty} f^{[t-1]}(x)\Big) = f(y).$$

And so $y = z_0$. (A similar argument implies $f^{[t]}(x) \searrow z_0$ for $z_0 < x < 1$.) Thus the first assertion follows.

The second assertion follows by properties of $f(x) - x$ and its derivatives. □

An important special case that will figure prominently when the offspring distribution is Poisson distributed.

Example 12.3. If $Z_{1,1} \sim \mathrm{Po}(\lambda)$, then $f(z) = e^{-\lambda(1-z)}$. If $\lambda > 1$, then denote the extinction probability by $\eta_\lambda < 1$, which is the minimal non-negative solution to $z = e^{-\lambda(1-z)}$, and write $\zeta_\lambda = 1 - \eta_\lambda$ for the chance of survival, which satisfies $\zeta_\lambda = 1 - e^{-\lambda\zeta_\lambda}$.

Connection with exploration process: To connect with the exploration process of an Erdős-Rényi random graph, we view the BP as a random tree with root $(\mathcal{T}, 0)$, where individuals are vertices and we put an edge between children and their parents, and the root corresponds to the individual at generation zero; see the introductory picture above.

We define the exploration process for $(\mathcal{T}, 0)$ very similarly to that for an Erdős-Rényi random graph. The process is a random sequence of the partitions of the vertices $(\mathcal{A}_t, \mathcal{E}_t, \mathcal{N}_t)_{t=0,1,2,\ldots}$; we refer to these respectively as active, explored, and neutral. If $\mathcal{T}$ is finite, then we label its vertices as $\{0, 1, 2, \ldots, T-1\}$ (so there are T total vertices), and otherwise label them as $\{0, 1, 2, \ldots\}$. We initiate the process as

$$(\mathcal{A}_0, \mathcal{E}_0, \mathcal{N}_0) = \big(\{0\}, \emptyset, \{1, \ldots, T-1\}\big).$$

Given $(\mathcal{A}_{t-1}, \mathcal{E}_{t-1}, \mathcal{N}_{t-1})$, we generate the tth step according to the following rules.

- If $\mathcal{A}_{t-1} \neq \emptyset$, choose an element of $\mathcal{A}_{t-1}$ from among those *closest to the root* 0, call it a_t, and denote the set of vertices in $\mathcal{N}_{t-1}$ that are attached to a_t by

  $$\mathcal{Z}_t = \{u \in \mathcal{N}_{t-1} : \{u, a_t\} \text{ is an edge in } \mathcal{T}\}.$$

  Then set

  $$(\mathcal{A}_t, \mathcal{E}_t, \mathcal{N}_t) = \big(\mathcal{A}_{t-1} \setminus \{a_t\} \cup \mathcal{Z}_t,\, \mathcal{E}_{t-1} \cup \{a_t\},\, \mathcal{N}_{t-1} \setminus \mathcal{Z}_t\big).$$

  In other words, at each step, choose an active vertex, label it as explored, and its "children" as active, and label the active vertex as explored.

- If $\mathcal{A}_{t-1} = \emptyset$, then set $(\mathcal{A}_t, \mathcal{E}_t, \mathcal{N}_t) = (\mathcal{A}_{t-1}, \mathcal{E}_{t-1}, \mathcal{N}_{t-1})$.

The basic idea is that we explore the tree by revealing the children of vertices one at a time; note that the number of these children revealed at each step are i.i.d. distributed as the offspring distribution. To continue to make connections to Erdős-Rényi exploration process, we define the analogous notation:

$$A_t := |\mathcal{A}_t|, \quad E_t := |\mathcal{E}_t|, \quad N_t := |\mathcal{N}_t|, \quad Z_t := |\mathcal{Z}_t|.$$

Note that E_t is non-decreasing, N_t is non-increasing, and when $\mathcal{T}$ is finite,

$$A_t + E_t + N_t = T$$

(if $\mathcal{T}$ is infinite, then N_t is infinite for all t), and, setting

$$\tau_0 := \inf\{t \geq 0 : A_t = 0\},$$

for $t \leq \tau_0$, $E_t = t$ and $A_t + t + N_t = T$. Note also that for $t \geq \tau_0$,

$$(A_t, E_t, N_t) = (A_{\tau_0}, E_{\tau_0}, N_{\tau_0}) = (0, \tau_0, 0) = (0, T, 0).$$

Thus, the analog of Lemma 11.1 from Lecture 11 is obvious.

Lemma 12.4. *In the notation and definitions above*

$$\tau_0 = T.$$

The important difference between the BP exploration process and the Erdős-Rényi exploration, is that the number of revealed vertices at each step is i.i.d. for the former, but not for the latter. In particular, the analog of Lemma 11.2 from Lecture 11 for BPs is as follows; the proof is basically the same.

Lemma 12.5. *In the notation and definitions above, the process $(A_t)_{t\geq 0}$ is a Markov chain given by setting $A_0 = 1$, and for $t \geq 1$, setting*

$$A_t = A_{t-1} + \mathbb{I}[A_{t-1} > 0](-1 + Z_t), \tag{12.1}$$

where the Z_t are i.i.d. distributed as the offspring distribution $\mathscr{L}(Z_{1,1})$.

The key point is that for the Erdős-Rényi exploration process, the analog of (12.1) is (11.1), which says given A_{t-1} (positive)

$$Z_t \sim \mathrm{Bi}\big(n-(t-1)-A_{t-1}, \tfrac{\lambda}{n}\big) \approx \mathrm{Po}\big(\lambda\big(1-\tfrac{t-1+A_{t-1}}{n}\big)\big) \approx \mathrm{Po}(\lambda),$$

where the last approximation is good when $t-1+A_{t-1} = n - N_{t-1} \ll n$. Since $N_0 = n-1$, for any fixed λ and large n, the initial steps of the Erdős-Rényi exploration process looks like a BP with Poisson offspring distribution. As we will see, if $\lambda < 1$, then with high probability, the BP dies out quickly enough that the approximation is valid for the whole exploration process, and all components are "small". If $\lambda > 1$, then some of the exploration processes "die out" quickly and form small compents, while others live long enough to form one "giant" component. We will need a different approximation for the later stages of the exploration in the latter case. Note also that when the approximation is good, $C_v \approx T$, since these are both equal to their respective τ_0's.

A topical aside: You can think of the spread of a communicable disease among a population with a finite number of individuals as an exploration process. One individual initially has the disease, they give it to a random number of other people, then each of those people give it to a random number of new people, and so on. The randomness generating the process won't necessarily be like the Erdős-Rényi exploration process, but the beginning of the process will look like a BP with some distribution that depends on the social behaviour of the population (e.g., number and type of interactions), and properties of the disease (e.g., infectiousness, length of illness). If the mean number of people infected by a given individual is $m > 1$, then after t "generations" the number who have the disease is around the mean number of a BP with that mean, of order m^t, which is the reason for the exponential growth of cases at the start of the "epidemic curve". The purpose of social distancing measures is to lower the mean, which makes a huge difference in the size of m^t for even moderate values of t. Also note that epidemic curves eventually decrease due to the finite population. The decrease does not look like a BP, even one that goes extinct, and this corresponds to the approximation not being very good in later stages of the exploration process due to finite population, similar to the discussion in the previous paragraph. Approximating the spread of the disease by BPs is an important tool in epidemic modelling.