Branching process survival probability

We are close now close to proving the following Theorem about the survival probabilities of a branching process. Suppose that a branching process $\mathbf{X}$ has offspring distribution Z (meaning that each vertex has a number of children distributed as Z, independently of other vertices).

Theorem 1. *If $\mathbb{E}(Z) < 1$ then $\eta = 1$. If $\mathbb{E}(Z) > 1$ then $\eta < 1$, and if $\mathbb{E}(Z) = 1$ with $\mathbb{P}(Z = 0) > 0$ then $\eta = 1$.*

Proof. We have already proven the first statement. Suppose that $\mathbb{E}(Z) > 1$. We will show that η is the smallest nonnegative solution to the equation $x = f_Z(x)$. Recall that we let $\eta_t = \mathbb{P}(X_t = 0)$. The events $X_t = 0$ form a chain:

$$X_1 = 0 \subseteq X_2 = 0 \subseteq X_3 = 0 \subseteq \dots$$

and therefore we have

$$\eta_1 \leq \eta_2 \leq \eta_3 \leq \dots$$

Since we have

$$\eta = \mathbb{P}\left(\bigcup_{t=1}^{\infty}\{X_t = 0\}\right),$$

we have $\eta_t \to \eta$ as $t \to \infty$. Since $\eta_t = f_Z(\eta_{t-1})$ and $f_Z(x)$ is continuous in $[0,1]$, we can take the limit on both sides and obtain that $\eta = f_Z(\eta)$. Suppose that a is any nonnegative solution to $f_Z(x) = x$. Recall that $f_Z(0) = p_0$, and that $f_Z(x)$ is strictly increasing. This implies that $p_0 \leq a$. We also have that $\eta_1 = p_0$ (why?). Therefore, we have

$$\eta_2 = f_Z(\eta_1) \leq f_Z(a) = a.$$

Since $\eta_2 \leq a$ and $f_Z(x)$ is increasing, we obtain that

$$\eta_3 = f_Z(\eta_2) \leq f_Z(a) = a,$$

and so forth. We can repeat this step to show that $\eta_t \leq a$ for all t, and therefore $\eta \leq a$. □

Exercise 1. *Finish the second statement of Theorem 1.* 2

Exercise 2. *Prove the final statement of Theorem 1.* 4

Branching process random walk

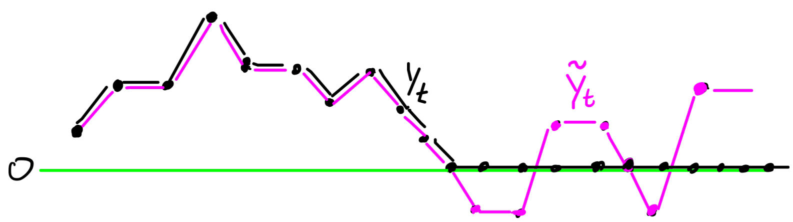
Now, we will finally use the branching process to help us bound the probabilities of connected components in $G(n,p)$ reaching a certain size. First, we describe the branching process in a slightly different way: by letting the vertices have offspring one by one, instead of per generation. Note that this does not affect the outcome of the process. We let Y_t^{bp} be the number of "live" vertices at time t. Each new vertex is live from the moment is is created until it has offspring. Let $Y_0^{bp} = 1$. Then we let

$$Y_{t+1}^{bp} = \begin{cases} Y_t^{bp} + Z_t - 1, & \text{if } Y_t^{bp} > 0, \\ 0, & \text{if } Y_t^{bp} = 0, \end{cases}$$

where Z_t is the random variable that determines the number of offspring of the vertex processed at time t. In the case of our branching process, we have that $Z_t \sim Z$ for each t. We also define a somewhat similar but simpler process $\widetilde{Y}_t^{bp}$ as follows: $\widetilde{Y}_0^{bp} = 1$, and

$$\widetilde{Y}_{t+1}^{bp} = \widetilde{Y}_t^{bp} + Z_t - 1.$$

Note that the random walks Y_t^{bp} and $\widetilde{Y}_t^{bp}$ behave the same until they hit 0, then Y_t^{bp} stays at 0 forever, while $\widetilde{Y}_t^{bp}$ is allowed to drop below 0 and then recover. As a sketch:



What is X, the total size of the process? We have $Y_0^{bp} = 1$, and at each time step t, no matter how many new vertices are created, exactly 1 is processed. Therefore, the total number of vertices in the process is

$$T^{bp} = \min\{t : \ Y_t^{bp} = 0\}.$$

Similarly, we let

$$\widetilde{T}^{bp} = \min\{t : \ \widetilde{Y}_t^{bp} = 0\},$$

and we notice that $T^{bp} = \widetilde{T}^{bp}$.

Graph exploration process

Let v be a vertex in $G(n,p)$. We are interested in the size of the connected component that contains v, which we will call C_v. We explore C_v as follows. Start at v and mark it as "live". While there are live vertices, choose a live vertex and mark it as "processed". Find it's neighbors among the vertices in $G(n,p)$ that are not yet live or processed, and mark them as live. This ensures that we find a tree in $G(n,p)$, and that each vertex is only added to the tree and processed at most once: no vertices are double counted. Continue until there are no more live vertices, and note that at this point, the processed vertices are exactly the set $V(C_v)$. We let $Y_0^{gr} = 1$, and

$$Y_{t+1}^{gr} = \begin{cases} Y_t^{gr} + R_t - 1, & \text{if } Y_t^{gr} > 0, \\ 0, & \text{if } Y_t^{gr} = 0, \end{cases}$$

where R_t represents the number of new neighbors discovered at time t. In th graph exploration process, there is a time dependence, since we have only n vertices total. We similarly define $\widetilde{Y}_0^{gr} = 1$ and

$$\widetilde{Y}_{t+1}^{gr} = \widetilde{Y}_t^{gr} + R_t - 1.$$

and

$$T^{gr} = \min\{t : \ Y_t^{gr} = 0\} = \widetilde{T}^{gr} = \min\{t : \ \widetilde{Y}_t^{gr} = 0\}.$$

Exercise 3. *Suppose that we are at time t and that $Y_{t-1}^{gr} = y_{t-1}$. Show that R_t has a binomial distribution with $n - y_{t-1} - (t-1)$ trials and success probability p.* 1

Consider a component exploration around a vertex v in $G(n,p)$. Let $1 \leq k \leq n$, and let X^1 be a branching process with offspring distribution Binomial(n,p), and let X^2 be a branching process with offspring distribution Binomial$(n-k,p)$.

Exercise 4. *Show that* 3

$$\mathbb{P}(|X^1| > k) \leq \mathbb{P}(\mathit{Binomial}(kn,p) \geq k).$$

For the next exercise, we will compare the graph exploration process to the branching process, and let the branching process dominate the graph process. More precisely, at each time t we can generate Z_t and R_t at the same time with dependence. Suppose R_t has a Binomial$(n-x,p)$ distribution (where x is just a placeholder, see Exercise 3), and Z_t has a Binomial(n,p) distribution. We can sample R_t by running $n-x$ trials and counting the number of successes, and then sample Z_t by running a further x trials, counting the successes and adding this to R_t. Verify that this gives the correct distributions for R_t and Z_t and ensures that $R_t \leq Z_t$ for all t.

Exercise 5. *Show that* 3

$$\mathbb{P}(|C_v| > k) \leq \mathbb{P}(|X^1| > k).$$

Exercise 6. *Show that* 3

$$\mathbb{P}(|C_v| > k) \geq \mathbb{P}(|X^2| > k).$$