Chapter 3

Branching processes

In this chapter we begin studying another classical process considered in probability theory: *branching*. The motivation comes from attempts to understand the dynamics of genealogical trees — as was the case for Galton and Watson who invented branching processes — but the real interest comes from applications that reach beyond this limited, and to most people little appealing, context.

3.1 Galton-Watson branching process

A branching process is uniquely determined by its *offspring distribution* which is a sequence $\mathfrak{p}_n$ of non-negative numbers such that

$$\sum_{n\geq 0} \mathfrak{p}_n = 1. \tag{3.1}$$

An informal definition of a branching process is as follows: At each time, the process has a certain number of living individuals. To get the next generation, each of the living individuals *splits* into a random number of offspring, and then *dies*. The number of offspring is sampled from the offspring distribution independently of all other neighbors. A formal definition is as follows:

Definition 3.1 [Galton-Watson branching process] *Let $(\xi_{n,m})_{m,n\geq 1}$ be i.i.d. integer-valued random variables with $\mathbb{P}(\xi_{n,m} = k) = \mathfrak{p}_k$. Define the sequence of random variables (X_n) by setting*

$$X_0 = 1 \tag{3.2}$$

and solving recursively

$$X_{n+1} = \begin{cases} \xi_{n+1,1} + \dots + \xi_{n+1,X_n}, & \text{if } X_n > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.3}$$

A Galton-Watson branching process *with offspring distribution $(\mathfrak{p}_n)$ is a sequence of random variables that has the same law as (X_n).*

The second line in the formula for X_{n+1} shows that once the sequence (X_n) hits zero — i.e., once the family has died out — it will be zero forever. We refer to this situation *extinction*; the opposite case is referred to as *survival*. The first goal is to characterize offspring distributions for which extinction occurs with probability one — or, complementarily, survival occurs with a positive probability.

Consider the moment-generating function

$$\phi_n(s) = \mathbb{E}(\mathrm{e}^{-sX_n}), \qquad s \geq 0. \tag{3.4}$$

The reason for looking at this function is that if $\mathbb{P}(X_n \geq 1)$ is bounded uniformly away from zero, say, by $\epsilon > 0$, then we have $\phi(s) \leq 1 - \epsilon\mathrm{e}^{-s}$. On the other hand, if $\mathbb{P}(X_n \geq 1)$ tends to zero, we clearly get $\phi(s) \to 1$ for all $s \geq 0$.

Lemma 3.2 *For each $s \geq 0$, let*

$$\lambda(s) := -\log \mathbb{E}(\mathrm{e}^{-s\xi_{1,1}}) = -\log\Big(\sum_{n\geq 0} \mathrm{e}^{-sn}\mathfrak{p}_n\Big). \tag{3.5}$$

Then for all $n \geq 0$,

$$\phi_{n+1}(s) = \phi_n\big(\lambda(s)\big). \tag{3.6}$$

In particular,

$$\phi_n(s) = \exp\{-\lambda_n(s)\} \tag{3.7}$$

where λ_n denotes the n-fold iteration $\lambda \circ \cdots \circ \lambda$ of function λ.

Proof. First we derive (3.6). By the definition of X_{n+1},

$$\phi_{n+1}(s) = \mathbb{E}\big(\mathrm{e}^{-s(\xi_{n+1,1}+\cdots+\xi_{n+1,X_n})}\big) \tag{3.8}$$

where we interpret the sum in the exponent as zero when $X_n = 0$. Conditioning on X_n and using the independence of $\xi_{n+1,j}$'s of themselves as well as X_n, we get

$$\begin{aligned} \mathbb{E}\big(\mathrm{e}^{-s(\xi_{n+1,1}+\cdots+\xi_{n+1,X_n})}\big) &= \sum_{k\geq 0} \mathbb{E}\big(\mathrm{e}^{-s(\xi_{n+1,1}+\cdots+\xi_{n+1,k})}1_{\{X_n=k\}}\big) \\ &= \sum_{k\geq 0} \mathbb{E}\big(\mathrm{e}^{-s(\xi_{n+1,1}+\cdots+\xi_{n+1,k})}\big)\mathbb{P}(X_n = k) \\ &= \sum_{k\geq 0} \big(\mathbb{E}(\mathrm{e}^{-s\xi_{n+1,1}})\big)^k\, \mathbb{P}(X_n = k) \\ &= \sum_{k\geq 0} \mathrm{e}^{-\lambda(s)k}\, \mathbb{P}(X_n = k) = \mathbb{E}\big(\mathrm{e}^{-\lambda(s)X_n}\big) \end{aligned} \tag{3.9}$$

The right-hand side is the moment-generating function of X_n at the point $\lambda(s)$. This is the content of (3.6). Note that $\lambda(s) \geq 0$ once $s \geq 0$ and so there is no problem with using $\lambda(s)$ as an argument of ϕ_n.

To derive the explicit formula for ϕ_n we first solve recursively for ϕ_n to get

$$\phi_n(s) = \phi_0\big(\,\lambda \circ \cdots \circ \lambda(s)\big) \tag{3.10}$$

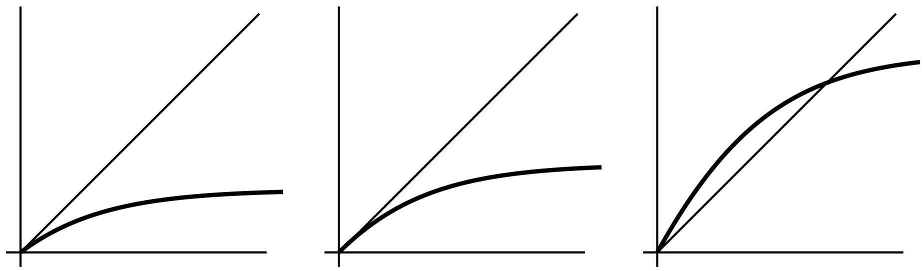

Figure 3.1: The plot of $s \mapsto \lambda(s)$ for offspring distribution with $\mathfrak{p}_0 = 1 - p$ and $\mathfrak{p}_3 = p$ for p taking values 0.25, 0.33 and 0.6, respectively. These values represent the three generic regimes distinguished — going left to right — by whether $\lambda'(0^+)$ is less than one, equal to one, or larger than one. As λ is strictly concave, non-negative with $\lambda(0) = 0$, only in the last case λ has a strictly positive fixed point. We refer to the three situations as subcritical, critical and supercritical.

(Again, we are using that λ maps $(0, \infty)$ into $(0, \infty)$ and so the iterated map is well defined.) From here (3.7) follows by noting that $\phi_0(s) = \mathrm{e}^{-s}$ due to (3.2). □

In light of (3.7), the question whether $\phi_n(s) \to 1$ or not now boils to the question whether $\lambda_n(s) \to 0$ or not. To find for the right criterion, we will need to characterize the analytic properties of λ:

Lemma 3.3 *Suppose that $\mathfrak{p}_n < 1$ for all n. Then λ is non-decreasing and continuous on $[0, \infty)$ and strictly concave and differentiable on $(0, \infty)$. In addition,*

$$\lim_{s \downarrow 0} \lambda'(s) = \sum_{n \geq 0} n \mathfrak{p}_n \tag{3.11}$$

and

$$\lim_{s \to \infty} \lambda'(s) = \inf\{n \colon \mathfrak{p}_n > 0\}. \tag{3.12}$$

Proof. Since $\mathfrak{p}_n \leq 1$, the series $\sum_{n \geq 0} \mathrm{e}^{-sn} \mathfrak{p}_n$ is absolutely summable locally uniformly on $s > 0$ and so it can be differentiated term-by-term. In particular, the first derivative $\lambda(s)$ is the expectation

$$\lambda'(s) = \sum_{n \geq 0} n \mathfrak{p}_n \, \mathrm{e}^{-ns + \lambda(s)} \tag{3.13}$$

for the probability mass function $n \mapsto \mathfrak{p}_n \mathrm{e}^{-sn} \mathrm{e}^{\lambda(s)}$ on $\mathbb{N} \cup \{0\}$, while the second derivative $\lambda''(s)$ is the negative of the corresponding variance. Under the condition $\mathfrak{p}_n < 1$ for all n, the variance is non-zero and so $\lambda''(s) < 0$ for all $s > 0$. This establishes differentiability and strict concavity $(0, \infty)$; continuity at $s = 0$ is directly checked. The limit of the derivatives (3.11) exists by concavity and equals the corresponding limit of (3.13). To prove (3.12), let k be the infimum on the right hand side. Then $\mathrm{e}^{-\lambda(s)} \sim \mathfrak{p}_k \mathrm{e}^{-sk}$ and so all terms but the k-th in (3.13) disappear in the limit $s \to \infty$. The k-th term converges to k and so (3.12) holds. □

Exercise 3.4 Compute $\lambda''(s)$ explicitly and show that it can be written as the negative of a variance. Use this to show that $\lambda''(s) < 0$ for all $s > 0$ once at least two of the $\mathfrak{p}_n$'s are non-zero.

The statements in the lemma imply that the function looks as in Fig. 3.1.

Exercise 3.5 Consider branching process with offspring distribution determined by $p_0 = p$ and $p_2 = 1 - p$ with $0 < p < 1$. Sketch the graph of $\lambda(s)$ and characterize the regime when λ has a non-zero fixed point.

We can now use these properties to control the iterations of λ:

Theorem 3.6 [Survival/extinction criteria] *Suppose* $0 < \mathfrak{p}_0 < 1$ *and let*

$$\mu = \sum_{n\geq 0} n\mathfrak{p}_n \in (0, \infty] \tag{3.14}$$

denote the mean offspring. Then we have:

(1) If $\mu \leq 1$ *then the branching process dies out,*

$$\mathbb{P}(X_n \geq 1) \underset{n\to\infty}{\longrightarrow} 0. \tag{3.15}$$

(2) If $\mu > 1$ *then the process dies out with probability* $\mathrm{e}^{-s_\star}$*,*

$$\mathbb{P}(X_n = 0) \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-s_\star}, \tag{3.16}$$

where $s_\star$ *is the unique positive solution to* $\lambda(s) = s$*, and it survives forever and, in fact, the population size goes to infinity with complementary probability,*

$$\mathbb{P}(X_n \geq M) \underset{n\to\infty}{\longrightarrow} 1 - \mathrm{e}^{-s_\star} \tag{3.17}$$

for every $M \geq 1$*.*

Proof. Suppose first that $\mu \leq 1$. The strict convexity and the fact that $\lambda'(s) \to \mu$ as $s \downarrow 0$ ensure that $\lambda(s) < s$ for all $s > 0$. This means $\lambda_{n+1}(s) = \lambda(\lambda_n(s)) < \lambda_n(s)$, i.e., the sequence $n \mapsto \lambda_n(s)$ is strictly decreasing. Thus for each $s \geq 0$ the limit

$$r(s) = \lim_{n\to\infty} \lambda_n(s) \tag{3.18}$$

exists. But $\lambda_{n+1}(s) = \lambda(\lambda_n(s))$ and the continuity of λ imply

$$\lambda(r(s)) = r(s) \tag{3.19}$$

As $\lambda(s) < s$ for $s > 0$, the only point that satisfies this equality on $[0, \infty)$ is $r(s) = 0$. We conclude that $\lambda_n(s) \to 0$ yielding

$$\lim_{n\to\infty} \phi_n(s) = 1, \qquad s \geq 0. \tag{3.20}$$

From

$$\phi_n(s) \leq 1 - (1 - \mathrm{e}^{-s})\mathbb{P}(X_n \geq 1) \tag{3.21}$$

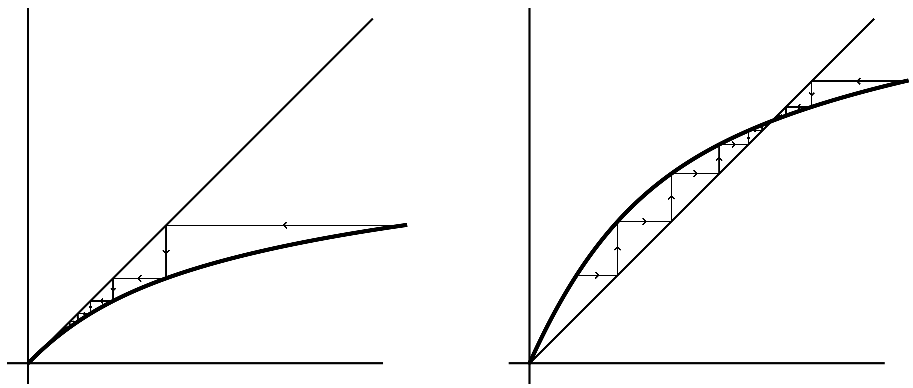

Figure 3.2: A graphical construction underlying the iterations $\lambda_n = \lambda \circ \cdots \circ \lambda$ in the proof of Theorem 3.6. For the subcritical process (left) $\lambda_n(s)$ converges to zero for any s. For the supercritical process (right), λ_n increases for $s < s_\star$ and decreases for $s < s_\star$ where $s_\star$ is the positive fixed point of λ.

it follows that

$$\mathbb{P}(X_n \geq 1) \underset{n\to\infty}{\longrightarrow} 0 \tag{3.22}$$

as we desired to prove.

Next let us assume $\mu > 1$. Then we have $\lambda(s) > s$ for s sufficiently small. On the other hand, due to $\mathfrak{p}_0 > 0$ and (3.12), we have $\lambda(s) < s$ once s is large. Thus there exists a non-zero solution to $\lambda(s) = s$. Strict convexity of λ implies that this solution is actually unique because if $s_\star$ is the least such positive solution then $\lambda'(s_\star) \leq 1$ and strict concavity tell us that $\lambda(s) < s$ for $0 < s < s_\star$ and $\lambda(s) > s$ for $s > s_\star$.

We claim that

$$\lambda_n(s) \underset{n\to\infty}{\longrightarrow} s_\star, \qquad s > 0. \tag{3.23}$$

This is proved in the same way as for $\mu \leq 1$ but we have to deal with $s > s_\star$ and $s < s_\star$ separately.

For $s > s_\star$ we have $s_\star < \lambda(s) < s$ and so the sequence $\lambda_n(s)$ is *decreasing*. The limit is again a fixed point of $\lambda(s) = s$ and since $s_\star$ is the only one available, we have $\lambda_n(s) \to s_\star$.

For $0 < s < s_\star$ we instead have $s < \lambda(s) < s_\star$, where we used that $\lambda(s)$ is increasing to get the second inequality. It follows that $\lambda_n(s)$ is *increasing* in this regime. The limit is a fixed point of λ and so $\lambda_n(s) \to s_\star$ in this case as well.

Having established (3.23), we now note that this implies

$$\phi_n(s) \underset{n\to\infty}{\longrightarrow} e^{-s_\star} \tag{3.24}$$

Since this holds for s arbitrary large, and since

$$\mathbb{P}(X_n = 0) \leq \phi_n(s) = \mathbb{P}(X_n = 0) + e^{-s} \tag{3.25}$$

we must have

$$\mathbb{P}(X_n = 0) \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-s_\star}. \tag{3.26}$$

Now fix $M \ge 1$ and suppose that $\epsilon := \liminf_{n\to\infty} \mathbb{P}(X_n = M) > 0$. Then

$$\phi_n(s) \ge \mathbb{P}(X_n = 0) + \mathrm{e}^{-sM}\mathbb{P}(X_n = M) \tag{3.27}$$

would imply $\liminf_{n\to\infty} \phi_n(s) \ge \mathrm{e}^{-s_\star} + \epsilon \mathrm{e}^{-sM} > \mathrm{e}^{-s_\star}$, a contradiction. Hence, $\mathbb{P}(X_n = M)$ tends to zero for any finite M, proving (3.17). □

Exercise 3.7 Suppose $\mathfrak{p}_0 = 0$ and $\mathfrak{p}_1 < 1$. Show that $X_n \to \infty$ with probability one.

Problem 3.8 Suppose $\mu < 1$. Show that $\mathbb{P}(X_n \ge 1)$ decays exponentially with n.

3.2 Critical process & duality

The analysis in the previous section revealed the following picture: Branching processes undergo an abrupt change of behavior when μ increases through one. This is a manifestation of what physicists call a *phase transition*. The goal of this section is to investigate the situation at the *critical point*, i.e., for generic branching processes with mean-offspring $\mu = 1$. Here is the desired result:

Theorem 3.9 *Suppose $\mu := \mathbb{E}\xi = 1$ and $\sigma^2 := \mathrm{Var}(\xi) \in (0,\infty)$. Then*

$$\mathbb{P}(X_n \ge 1) = \frac{2}{\sigma^2}\frac{1}{n}\big(1 + o(1)\big), \qquad n \to \infty. \tag{3.28}$$

Moreover,

$$\mathbb{P}\Big(\frac{X_n}{n} \ge z \,\Big|\, X_n \ge 1\Big) \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-z(2/\sigma^2)}, \tag{3.29}$$

i.e., conditional on survival the law of X_n/n converges to that of an exponential random variable with parameter ${}^2/_{\sigma^2}$.

The key to the proof are two lemmas:

Lemma 3.10 *For any $s > 0$,*

$$\frac{1-\phi_n(s)}{1-\mathrm{e}^{-s}} = \sum_{k\ge1} \mathrm{e}^{-s(k-1)}\mathbb{P}(X_n \ge k). \tag{3.30}$$

Proof. Abbreviate $a_k = \mathbb{P}(X_n \ge k)$. Then

$$\phi_n(s) = \sum_{k\ge0} \mathrm{e}^{-ks}\,(a_k - a_{k+1}). \tag{3.31}$$

Rearranging the sum, we get

$$\phi_n(s) = a_0 + \sum_{k\ge1} (\mathrm{e}^{-ks} - \mathrm{e}^{-(k-1)s})a_k. \tag{3.32}$$

Since $a_0 = 1$, subtracting both sides from one gives us

$$1 - \phi_n(s) = (1 - e^{-s}) \sum_{k \geq 1} e^{-s(k-1)} a_k \tag{3.33}$$

This is exactly (3.30). □

Lemma 3.11 *For any fixed $s > 0$,*

$$\frac{1}{\lambda_n(s)} = \frac{\sigma^2}{2} n \big(1 + o(1)\big), \qquad n \to \infty. \tag{3.34}$$

Moreover, if $\theta > 0$ is fixed, then also

$$\frac{1}{\lambda_n(\theta/n)} = \Big(\frac{1}{\theta} + \frac{\sigma^2}{2}\Big) n \big(1 + o(1)\big), \qquad n \to \infty. \tag{3.35}$$

Proof. First, let us get the intuitive idea for the appearance of $1/n$ scaling for λ_n. We know that $\lambda_n(s) \to 0$ for any $s \geq 0$. Since $\lambda_{n+1}(s) = \lambda(\lambda_n(s))$, we can thus expand λ about $s = 0$ to get a simpler recurrence relation. A computation shows

$$\lambda(s) = \mu\, s - \frac{\sigma^2}{2} s^2 + o(s^2) \tag{3.36}$$

and since $\mu = 1$, we thus have

$$\lambda_{n+1}(s) = \lambda_n(s) - \frac{\sigma^2}{2} \lambda_n(s)^2 \big(1 + o(1)\big). \tag{3.37}$$

This is solved to the leading order by setting $\lambda_n(s) = c/n$ which gets us $c = 2/\sigma^2$.

We actually do not have to work much harder in order to get the above calculation under control. First, the existence of the first two derivatives of λ tells us that, for each $\epsilon > 0$, we can find $s_0(\epsilon) > 0$ such that

$$\frac{1}{s} + (1-\epsilon)\frac{\sigma^2}{2} \leq \frac{1}{\lambda(s)} \leq \frac{1}{s} + (1+\epsilon)\frac{\sigma^2}{2}, \qquad 0 < s < s_0. \tag{3.38}$$

Thus, we have

$$\frac{1}{\lambda_n(s)} + (1-\epsilon)\frac{\sigma^2}{2} \leq \frac{1}{\lambda_{n+1}(s)} \leq \frac{1}{\lambda_n(s)} + (1+\epsilon)\frac{\sigma^2}{2} \tag{3.39}$$

whenever $\lambda_n(s) < s_0$. Define $n_0 = n_0(s)$ be the first n for which this is true for iterations started from s, i.e.,

$$n_0 = \sup\big\{n \geq 0 \colon \lambda_n(s) \geq s_0(\epsilon)\big\}. \tag{3.40}$$

Since $s_0(\epsilon) \leq \lambda_{n_0}(s) \leq \lambda(s_0(\epsilon))$, summing the inequalites from n_0 on we get

$$\frac{1}{\lambda(s_0(\epsilon))} + (1-\epsilon)\frac{\sigma^2}{2}(n - n_0) \leq \frac{1}{\lambda_n(s)} \leq \frac{1}{s_0(\epsilon)} + (1+\epsilon)\frac{\sigma^2}{2}(n - n_0) \tag{3.41}$$

From here the limit (3.34) follows by taking $n \to \infty$ followed by $\epsilon \downarrow 0$.

To prove also (3.35), we now assume that $\theta/n < s_0(\epsilon)$. Then (3.39) tells us

$$\frac{n}{\theta} + (1-\epsilon)\frac{\sigma^2}{2}k \leq \frac{1}{\lambda_k(\theta/n)} \leq \frac{n}{\theta} + (1+\epsilon)\frac{\sigma^2}{2}k \tag{3.42}$$

Setting $k = n$, we get (3.35). □

Proof of Theorem 3.9. First we prove (3.28). Fix $s > 0$. By (3.34) we then have

$$1 - \phi_n(s) = \frac{2/\sigma^2}{n}\big(1 + o(1)\big), \qquad n \to \infty. \tag{3.43}$$

The bounds $0 \leq \mathbb{P}(X_n \geq k) \leq \mathbb{P}(X_n \geq 1)$ and Lemma 3.10 imply

$$\mathbb{P}(X_n \geq 1) \leq \frac{1 - \phi_n(s)}{1 - \mathrm{e}^{-s}} \leq \frac{\mathbb{P}(X_n \geq 1)}{1 - \mathrm{e}^{-s}}. \tag{3.44}$$

From here (3.28) follows by taking $n \to \infty$ and $s \to \infty$.

Next we plug $s = \theta/n$ into the left-hand side of (3.30) and apply (3.35) to get

$$\frac{1 - \phi_n(\theta/n)}{1 - \mathrm{e}^{-\theta/n}} = \frac{1/\theta}{1/\theta + \sigma^2/2} + o(1), \qquad n \to \infty. \tag{3.45}$$

The identity

$$\mathbb{P}(X_n \geq k) = \mathbb{P}(X_n \geq 1)\,\mathbb{P}(X_n \geq k | X_n \geq 1) \tag{3.46}$$

turns (3.30) into

$$\frac{1 - \phi_n(\theta/n)}{1 - \mathrm{e}^{-\theta/n}} = n\mathbb{P}(X_n \geq 1)\mathrm{e}^{\theta/n}\,\frac{1}{n}\sum_{k\geq 1}\mathrm{e}^{-\theta k/n}\mathbb{P}(X_n \geq k | X_n \geq 1) \tag{3.47}$$

and applying (3.28) and (3.45) then gets us

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k\geq 1}\mathrm{e}^{-\theta k/n}\mathbb{P}(X_n \geq k | X_n \geq 1) = \frac{1}{\theta + 2/\sigma^2} \tag{3.48}$$

Thus, assuming that

$$\mathbb{P}(X_n \geq zn | X_n \geq 1) \underset{n\to\infty}{\longrightarrow} G(z) \tag{3.49}$$

for some non-increasing function $G\colon (0,\infty) \to (0,1)$, we can interpret the sum on the left of (3.48) as the Riemann sum of an (improper) integral to get

$$\int_0^\infty \mathrm{d}z\,\mathrm{e}^{-\theta z}\,G(z) = \frac{1}{\theta + 2/\sigma^2} \tag{3.50}$$

By the properties of the Laplace transform (whose discussion we omit) this can only be true if

$$G(z) = \mathrm{e}^{-z(2/\sigma^2)}. \tag{3.51}$$

But this shows that the limit (3.49) must exist because from any subsequence we can always extract a limit by using Cantor's diagonal argument and the fact that G is decreasing (again we omit details here). □

From Theorem 3.9 we learn that, conditional on survival up to time n, the number of surviving individuals is of order n. Next we will look at what happens when we condition on extinction. Of course, this is going to have a noticeable effect only on the supercritical processes.

Theorem 3.12 [Duality] *Consider a Galton-Watson branching process with supercritical offspring distribution $(\mathfrak{p}_n)$. Assume $0 < \mathfrak{p}_0 < 1$, let $s_\star$ be the unique positive solution to $\lambda(s) = s$, and define*

$$\mathfrak{q}_n = \mathfrak{p}_n \mathrm{e}^{-s_\star n + \lambda(s_\star)}, \qquad n \geq 0. \tag{3.52}$$

Then $\sum_{n\geq 0} \mathfrak{q}_n = 1$ and

$$\sum_{n\geq 0} n\mathfrak{q}_n = \lambda'(s_\star) < 1, \tag{3.53}$$

i.e., $(\mathfrak{q}_n)$ is an offspring distribution of a subcritical branching process. Furthermore, if T_p denotes the branching tree for $(\mathfrak{p}_n)$ and T_q is the corresponding object for $(\mathfrak{q}_n)$, then for any finite tree T,

$$\mathbb{P}\big(T_p = T \big| \text{extinction}\big) = \mathbb{P}(T_q = T) \tag{3.54}$$

i.e., the law of T_p conditioned on extinction is that of the dual process T_q.

Proof. Consider a finite tree T and let V be the set of vertices and E the set of edges. Let $n(v)$ denote the number of children of vertex $v \in T$. The probability that T occurs is then

$$\mathbb{P}(T_p = T) = \prod_{v\in T} \mathfrak{p}_{n(v)} \tag{3.55}$$

Conditioning on extinction multiplies this by $\mathrm{e}^{s_\star}$ which we can write using

$$|V| = |E| + 1 \tag{3.56}$$

and the fact that $\lambda(s_\star) = s_\star$ as

$$\mathrm{e}^{s_\star} = \prod_{e\in E} \mathrm{e}^{-s_\star} \prod_{v\in V} \mathrm{e}^{\lambda(s_\star)} = \prod_{v\in V} \mathrm{e}^{-s_\star n(v) + \lambda(s_\star)} \tag{3.57}$$

This shows

$$\mathbb{P}\big(T_p = T \big| \text{ extinction}\big) = \prod_{v\in T} \mathfrak{p}_{n(v)} \mathrm{e}^{-s_\star n(v) + \lambda(s_\star)} \tag{3.58}$$

which is exactly (3.54). □

An interesting special case of duality is the case when a process is *self-dual*. This loosely defined term refers to the situation when the dual process has a distribution "of the same kind" as the original process. Here are some examples:

Example 3.13 *Binomial distribution*: Let $\theta \in [0, 1]$ and consider the offspring distribution $(\mathfrak{p}_n)$ which is Binomial(N, θ), i.e.,

$$\mathfrak{p}_n = \binom{N}{n} \theta^n (1-\theta)^{N-n}, \qquad 0 \leq n \leq N. \tag{3.59}$$

Then the dual process is also binomial, with parameters N and $\theta^\star$ where

$$\frac{\theta^\star}{1-\theta^\star} = \frac{\theta}{1-\theta}\mathrm{e}^{-s_\star} \tag{3.60}$$

where $s_\star$ is determined from $(1-\theta+\theta\mathrm{e}^{-s_\star})^N = \mathrm{e}^{-s_\star}$.

Example 3.14 *Poisson distribution*: A limit case of the above is the Poisson offspring distribution,

$$\mathfrak{p}_n = \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}, \qquad n \ge 0, \tag{3.61}$$

with $\lambda > 1$. Here the dual $(\mathfrak{q}_n)$ is also Poisson but with parameter $\lambda^\star$ which is defined as the unique number less than one with

$$\lambda\mathrm{e}^{-\lambda} = \lambda^\star\mathrm{e}^{-\lambda^\star}. \tag{3.62}$$

The self-dual point is $\lambda = 1$.

Exercise 3.15 Verify the claims in the previous examples.

Problem 3.16 Characterize the distribution of the size of the first generation, X_1, conditioned on survival forever, i.e., the event $\bigcap_{n\ge1}\{X_n \ge 1\}$.

Problem 3.17 Apply the above analysis to prove that if $\mu > 1$, then for each $\tilde\mu < \mu$, $\mathbb{P}(X_n \ge \tilde\mu^n | X_n \ge 1) \to 1$ as $n \to \infty$.

3.3 Tree percolation

One of the special, but extremely useful, applications of branching processes is to the problem of *percolation*. We will consider percolation on two graphs for which the connection with branching is quite apparent, a regular directed tree $\mathbb{T}_n$ and the complete graph K_n on n-vertices. Here we will address the former of the two.

A regular directed tree $\mathbb{T}_b$ is a connected infinite graph without cycles where each vertex except one — called the *root* and denoted by $\varnothing$ — has exactly $b+1$ neighbors; the root has only b neighbors. One of the $b+1$ neighbors of vertex $v \ne \varnothing$ — namely, the one on the unique path from v to the root — will be referred to as the *parent*, the other b vertices are the *children*. The root still has b children but no parent.

We will now define a version of percolation called bond percolation. The definition can be made on an arbitrary graph:

Definition 3.18 [Bond percolation] *Consider a finite or infinite graph G with vertex set V and edge set E. Fix $p \in [0,1]$ and consider a collection of Bernoulli random variables ω_e indexed by edges $e \in E$ such that*

$$\mathbb{P}(\omega_e = 1) = p = 1 - \mathbb{P}(\omega_e = 0). \tag{3.63}$$

The phrase bond percolation *then generally refers to the (random) subgraph $G_\omega = (V, E_\omega)$ of G with vertex set V and edge set*

$$E_\omega = \{e \in E\colon \omega_e = 1\}. \tag{3.64}$$

Performing bond percolation on the tree $\mathbb{T}_b$ results in a *forest* — i.e., a collection of trees. The pertinent question in the context of infinite graphs is for what values of p, or whether at all, the graph G_ω contains an infinite connected component. We will narrow this question to whether the root is in an infinite connected component or not. To this end, let $C_\omega(v)$ denote the set of vertices in the connected component of G_ω containing v. Denote

$$\theta(p) = \mathbb{P}_p\big(|C_\omega(\varnothing)| = \infty\big) \tag{3.65}$$

where the index p denotes we are considering percolation with parameter p. The principal observations concerning tree percolation are summarized as follows:

Theorem 3.19 [Bond percolation on $\mathbb{T}_b$] *Consider bond percolation on $\mathbb{T}_b$ with $b \geq 2$ and parameter p. Define $p_c = 1/b$. Then*

$$\theta(p) \begin{cases} = 0, & \text{for } p \leq p_c, \\ > 0, & \text{for } p > p_c. \end{cases} \tag{3.66}$$

Moreover, $p \mapsto \theta(p)$ is continuous on $[0,1]$ and strictly increasing on $[p_c, 1]$ and it vanishes linearly at $p = p_c$,

$$\theta(p) \sim (p - p_c), \qquad p \downarrow p_c. \tag{3.67}$$

Furthermore, the mean-component size, $\chi(p) = \mathbb{E}_p(|C_\omega(\varnothing)|)$ diverges as $p \uparrow p_c$,

$$\chi(p) = \frac{p_c}{p_c - p}, \qquad p < p_c, \tag{3.68}$$

and, at $p = p_c$, the component size distribution has a power-law tail,

$$\mathbb{P}_{p_c}\big(|C_\omega(\varnothing)| \geq n\big) \sim \frac{1}{\sqrt{n}} \tag{3.69}$$

To keep the subject of branching processes in the back of our mind, we begin by noting a connection between percolation and branching:

Lemma 3.20 [Connection to branching] *Let X_n be the number of vertices in $C_\omega(\varnothing)$ that have distance n to the root. Then $(X_n)_{n\geq 0}$ has the law of a Galton-Watson branching process with binomial offspring distribution $(\mathfrak{p}_n)$,*

$$\mathfrak{p}_n = \begin{cases} \binom{b}{n} p^n (1-p)^{b-n}, & \text{if } 0 \leq n \leq b, \\ 0, & \text{otherwise.} \end{cases} \tag{3.70}$$

Proof. Let V_n be the vertices of $\mathbb{T}_b$ that have distance n to the root. Suppose that X_n is known and let $v_1, \dots, v_{X_n}$ be the vertices in $C_\omega(\varnothing) \cap V_n$. Define

$$\xi_{n+1,j} = \#\big\{u \in V_{n+1} \colon (v_j, u) \in E_\omega\big\}. \tag{3.71}$$

Then, clearly,

$$X_{n+1} = \xi_{n+1,1} + \dots + \xi_{n+1,X_n} \tag{3.72}$$

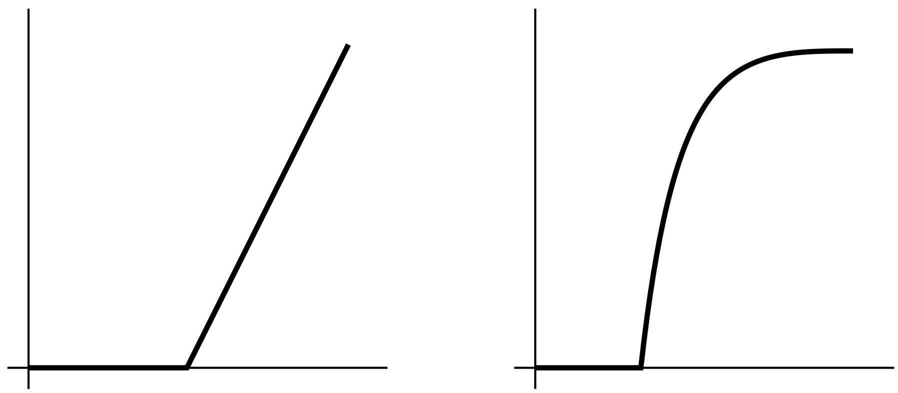

Figure 3.3: The graph of $p \mapsto \theta(p)$ for the tree with $b = 2$ and $b = 3$.

Since each $\xi_{n,j}$ is determined by different ω_e's which are in turn independent of X_n, we may think of $\xi_{n+1,1}, \dots, \xi_{n+1,X_n}$ as the first X_n terms in the sequence of i.i.d. random variables. The law of these ξ's is binomial with parameters b and p and so, inspecting Definition 3.1, the sequence (X_n) is the Galton-Watson branching process with offspring distribution as stated above. □

Corollary 3.21 *The probability $\theta(p)$ to have an infinite component at the root is the maximal positive solution in $[0,1]$ to the equation*

$$\theta = 1 - \bigl(1 - p\theta\bigr)^b \tag{3.73}$$

The generating function of the component size distribution, $\Psi_p(x) = \mathbb{E}_p(x^{|C_\omega(\varnothing)|})$, is for $x > 0$ and $0 < p < 1$ the unique solution to

$$\Psi = x\bigl(1 - p + p\Psi\bigr)^b \tag{3.74}$$

that lies strictly between zero and one.

Proof. Recall the connection with branching. The event that $C_\infty(\varnothing) = \infty$ clearly coincides with the event that the branching process (X_n) lives forever. As the probability to die out equals $q = \mathrm{e}^{-s_\star}$, where the latter quantity is the smallest solution to the equation

$$q = \sum_{n \ge 0} \mathfrak{p}_n q^n = (pq + 1 - p)^b, \tag{3.75}$$

we easily check that $\theta = 1 - q$ solves (3.73).

To derive the formula for the component moment generating function, we note that if $v_1, \dots, v_{X_1}$ are the neighbors of the root in $C_\omega(\varnothing)$, then

$$C_\omega(\varnothing) = \{\varnothing\} \cup \tilde{C}_\omega(v_1) \cup \dots \cup \tilde{C}_\omega(v_{X_1}), \tag{3.76}$$

where $\tilde{C}_\omega(v)$ denote the connected component of v in the subtree of $\mathbb{T}_b$ rooted at v. The union on the right-hand side is disjoint and the components are actually

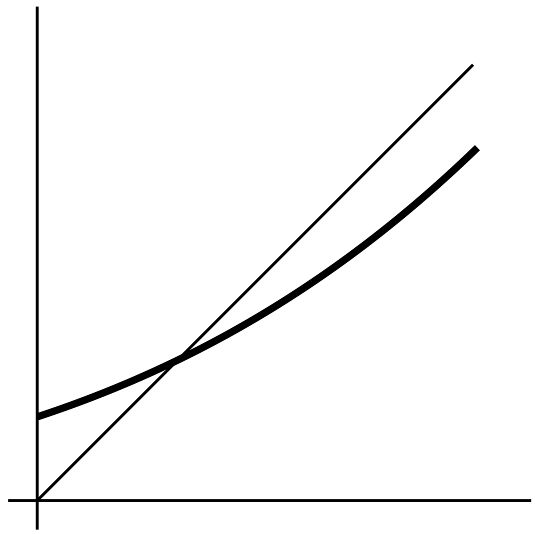

Figure 3.4: The graph of $\Psi \mapsto x(1-p+p\Psi)^b$ for the tree with $b=4$ at $x=0.8$ and $p=0.3$. The function is convex and has a unique fixed point with $\Psi \in (0,1)$.

independent copies of $C_\omega(\varnothing)$. Thus we have

$$\Psi_p(x) = x\,\mathbb{E}_p\big(x^{|C_\omega(v_1)|}\cdots x^{|C_\omega(v_{X_1})|}\big) = x\sum_{n=0}^{b}\mathfrak{p}_n\,\Psi_p(x)^n. \tag{3.77}$$

The equation now (3.74) follows from (3.70). The right-hand side of (3.74) is convex and it is strictly positive at $\Psi = 0$ and less than one at $\Psi = 1$. There is thus a unique intersection in $(0,1)$ which then has to be the value of $\Psi_p(x)$. □

Exercise 3.22 Set $b=2$ and/or $b=3$ and solve (3.73) explicitly. See Fig. 3.3.

Problem 3.23 Use the argument in the previous proof to show that $\mathbb{P}_p(|C_\omega(\varnothing)| \geq n)$ decays exponentially with n for all $p < p_c$.

Exercise 3.24 Suppose $p > p_c$. Use the duality for branching processes (Theorem 3.12) to show that the tail of the probability distribution of the *finite* components, $\mathbb{P}_p(n \leq |C_\omega(\varnothing)| < \infty)$, decays exponentially with n.

To get the asymptotic for the component size distribution at p_c, we also note:

Lemma 3.25 *For every $x < 1$,*

$$\frac{1-\Psi_p(x)}{1-x} = \sum_{n\geq 1} x^{n-1}\,\mathbb{P}_p\big(|C_\omega(\varnothing)| \geq n\big) \tag{3.78}$$

Proof. The proof is essentially identical to that of Lemma 3.10. Let us abbreviate

$a_n = \mathbb{P}_p(|C_\omega(\varnothing)| \geq n)$. Then

$$\Psi_p(x) = \sum_{n\geq 1} x^n\, \mathbb{P}_p\big(|C_\omega(\varnothing)| = n\big) = \sum_{n\geq 1} x^n\,(a_n - a_{n+1}). \tag{3.79}$$

Rearranging the latter sum, we get

$$\Psi_p(x) = a_1 + \sum_{n\geq 1}(x^n - x^{n-1})a_n. \tag{3.80}$$

The first term on the right-hand side is $a_1 = 1$ while the second term can be written as $(x-1)\sum_{n\geq 1} x^{n-1}a_n$. This and a bit of algebra yield (3.78). □

Proof of Theorem 3.19. The claims concerning $\theta(p)$ are derived by analyzing (3.73). Let us temporarily denote $\Phi(\theta) = 1-(1-p\theta)^b$. Then Φ is concave on $[0,1]$ and

$$\Phi(\theta) = pb\theta - \binom{b}{2}p^2\theta^2 + o(\theta^2) \tag{3.81}$$

This shows that for $pb \leq 1$ the only fixed point of Φ is $\theta = 0$, while for $pb > 1$ there are two solutions: one at $\theta = 0$ and one at $\theta = \theta(p) > 0$. This proves (3.66).

To get the critical scaling (3.67), we note that (3.81) can be rewritten as

$$\Phi(\theta) = \theta + b\theta\Big[(p-p_{\rm c}) - \frac{b-1}{2}p^2\big(1+o(1)\big)\theta\Big] \tag{3.82}$$

The positive solution $\theta(p)$ of $\Phi(\theta) = \theta$ will thus satisfy

$$\theta(p) = \frac{2b^2}{b-1}\big(1+o(1)\big)\,(p-p_c), \qquad p \downarrow p_{\rm c}. \tag{3.83}$$

To get the asymptotic for $\chi(p)$, we note

$$\chi(p) = \Psi_p'(1) \tag{3.84}$$

so we just need to find the good asymptotic of Ψ_p near $x=1$. This end we denote $g(x) = 1-\Psi_p(x)$ and note that (3.74) becomes

$$g(x) = 1 - x\big(1-pg(x)\big)^b \tag{3.85}$$

Differentiating at $x=1$ we get

$$g'(1) = -1 + bpg'(1) \tag{3.86}$$

and so $\Psi_p'(1) = (1-bp)^{-1} = p_{\rm c}/(p_{\rm c}-p)$.

Finally, to get the component-size distribution at $p_{\rm c}$, we will restrict ourselves only to $b=2$. In that case the equation for $g(x)$ can be solved and we get

$$g(x) = \frac{2}{x}\Big(\sqrt{1-x} - (1-x)\Big) \tag{3.87}$$

where the sign of the square root was chosen to make g positive on $(0,1)$ as it should be. By (3.78) we in turn have

$$x\,\frac{g(x)}{1-x} = \sum_{n\geq 1} x^n\,\mathbb{P}_p\big(|C_\omega(\varnothing)| \geq n\big) \tag{3.88}$$

and so to find the probabilities $\mathbb{P}_p(|C_\omega(\varnothing)| \geq n)$ we just need to expand the left-hand side into a Taylor series about $x = 0$ and use that the coefficients of this expansion are uniquely determined. This is said nearly as easily as it is done:

$$x\,\frac{g(x)}{1-x} = 2\Big(\frac{1}{\sqrt{1-x}} - 1\Big) = 2\sum_{n\geq 1}\binom{2n}{n}\Big(\frac{x}{4}\Big)^n \tag{3.89}$$

and so, by identifying the coefficients,

$$\mathbb{P}_{p_c}\big(|C_\omega(\varnothing)| \geq n\big) = 2\binom{2n}{n}4^{-n}. \tag{3.90}$$

Stirling's formula, $n! = (n/\mathrm{e})^n\sqrt{2\pi n}(1+o(1))$, shows that

$$2\binom{2n}{n}4^{-n} = \frac{2+o(1)}{\sqrt{\pi n}}, \qquad n\to\infty, \tag{3.91}$$

and so we have $\mathbb{P}_p(|C_\omega(\varnothing)| \geq n) \sim 1/\sqrt{n}$ as claimed. □

Problem 3.26 Show that for any $b \geq 2$ the limit

$$\lim_{x\uparrow 1}\frac{1-\Psi_{p_c}(x)}{\sqrt{1-x}} \tag{3.92}$$

exists and is positive and finite. Compute its value.

Some of our computations above have perhaps been somewhat unnecessarily formal. For instance, (3.73) can be derived as follows:

Let $1-\theta$ be the probability that the root is in a finite component. For that to be true, every occupied edge from the root must end up at a vertex whose component in the forward direction is also finite. The probability that the vertex v_1 is like this is $1-p\theta$, and similarly for all b neighbors of $\varnothing$. As these events are independent for distinct neighbors, this yields

$$1-\theta = (1-p\theta)^b \tag{3.93}$$

which is (3.73). Similarly, we can also compute $\chi(p)$ directly: By (3.76),

$$|C_\omega(\varnothing)| = 1 + \omega_{\varnothing,v_1}|\tilde{C}_\omega(v_1)| + \cdots + \omega_{\varnothing,v_b}|\tilde{C}_\omega(v_b)| \tag{3.94}$$

Using that $\omega_{\varnothing,v}$ is independent of $|\tilde{C}_\omega(v)|$, taking expectations we get

$$\chi(p) = 1 + bp\chi(p) \tag{3.95}$$

whereby $\chi(p) = \frac{p_c}{p_c-p}$.

The recursive nature of the tree, and of the ensuing calculations, allows us to look at some more complicated variants of the percolation process:

Problem 3.27 *k-core percolation*: Consider the problem of so called k-core percolation on the tree $\mathbb{T}_b$. Here we take $k \geq 3$ and take a sample of $C(\varnothing)$. Then we start applying the following pruning procedure: If a vertex has less than k "children" to which it is connected by an occupied edge, we remove it from the component along with the subtree rooted in it. Applying this over and over, this gives us a *decreasing* sequence $C_n(\varnothing)$ of subtrees of $C(\varnothing)$. Let $\vartheta(p)$ denote the probability that $C_n(\varnothing)$ is infinite for all n. Show that $\vartheta(p)$ is the largest positive solution to

$$\vartheta = \pi_k(\vartheta p) \quad \text{where} \quad \pi_k(\lambda) = \sum_{\ell=k}^{b} \binom{b}{\ell} \lambda^\ell (1-\lambda)^{b-\lambda}, \tag{3.96}$$

i.e., $\pi_k(\lambda)$ is the probability that Binomial(b, λ) is at least k. Explain why $\vartheta(p)$ for 1-core percolation equals to $\theta(p)$ for the ordinary percolation.

3.4 Erdős-Rényi random graph

Having understood percolation on the regular rooted tree, we can now move on to a slightly complicated setting which is percolation on the complete graph K_n. This problem has emerged independent of the development in percolation as a model of a random graph and is named after its inventors Erdős and Rényi.

The complete graph K_n has vertices $\{1, \dots, n\}$ and an (unoriented) edge between every pair of distinct vertices. Given $p \in [0, 1]$ we toss a biased coin for each edge and if it comes out heads — which happens with probability p — we keep the edge and if we get tails then we discard it. We call the resulting random graph $\mathscr{G}(n, p)$. The principal question of interest is the distribution of the largest connected component; particularly, when it is of order n.

Our main observation is that a percolation transition still occurs in the setting, even though the formulation is somewhat less clean due to the necessity to take $n \to \infty$:

Theorem 3.28 *For any $\alpha \geq 0$ and $\epsilon > 0$, let $\theta_{\epsilon,n}(\alpha)$ denote the probability that vertex "1" is in a component $\mathscr{G}(n, {}^{\alpha}/_{n})$ of size at least ϵn. Then*

$$\theta(\alpha) = \lim_{\epsilon \downarrow 0} \lim_{n \to \infty} \theta_{\epsilon,n}(\alpha) \tag{3.97}$$

exists and equals the largest positive solution to

$$\theta = 1 - e^{-\alpha\theta}. \tag{3.98}$$

In particular,

$$\theta(\alpha) \begin{cases} = 0, & \alpha \leq 1, \\ > 0, & \alpha > 1. \end{cases} \tag{3.99}$$

Exercise 3.29 Show that $\theta(\alpha)$ is the probability that the branching process with Poisson offspring distribution $\mathfrak{p}_n = \frac{\alpha^n}{n!} e^{-\alpha}$ survives forever.

The key idea of our proof is to explore the component of vertex "1" using a *search algorithm*. The algorithm keeps vertices in three classes: *explored* and *active* — called jointly “discovered” — and *undisovered*. Initially, we mark vertex “1” as active and the others as undiscovered. Then we repeat the following:

(1) Pick the active vertex v that has the least index.

(2) Find all undiscovered neighbors of v.

(3) Change the status of v to explored and its undiscovered neighbors to active.

The algorithm stops when we run out of active vertices. It is easy to check that this happens when we have explored the entire connected component of vertex “1”.

Let A_k denote the number of active vertices at the k-th stage of the algorithm; the initial marking of “1” is represented by $A_0 = 1$. If $v \in A_k$ is the vertex with the least index, we use L_k to denote the number of as of yet undiscovered neighbors of v. We have

$$A_{k+1} = A_k + L_k - 1. \tag{3.100}$$

Note that $n - k - A_k$ is the number of undiscovered vertices after the k-th run of the above procedure.

Lemma 3.30 *Conditional on A_k, we have*

$$L_k = \text{Binom}(n - k - A_k, {}^{\alpha}/_{n}). \tag{3.101}$$

Proof. The newly discovered vertices are chosen with probability ${}^{\alpha}/_{n}$ from the set of $n - k - A_k$ undiscovered vertices. □

We will use the above algorithm to design a *coupling* with the corresponding search algorithm for bond percolation on a regular rooted tree — or, alternatively, with a branching process with a binomial offspring distribution. Recall that $\mathbb{T}_b$ denotes the rooted tree with forward degree b and K_n the complete graph on n vertices. To make all percolation processes notationally distinct, we will from now on write $\mathbb{P}_{\mathbb{T}_n}$ for the law of percolation on $\mathbb{T}_n$ and $\mathbb{P}_{K_n}$ for the corresponding law on K_n. In *all* cases below the probability that an edge is occupied is ${}^{\alpha}/_{n}$.

Lemma 3.31 [Coupling with tree percolation] *For $m \le n$ and $r \le n - m$,*

$$\mathbb{P}_{\mathbb{T}_m}\big(|C(\varnothing)| \ge r\big) \le \mathbb{P}_{K_n}\big(|C(1)| \ge r\big) \le \mathbb{P}_{\mathbb{T}_n}\big(|C(\varnothing)| \ge r\big) \tag{3.102}$$

Proof. The proof is based on the observation that, as long as less than $n - m$ vertices have been discovered, we can couple the variables L_k for $\mathbb{T}_m$, K_n and $\mathbb{T}_n$ so that

$$L_k^{(\mathbb{T}_m)} \le L_k^{(K_n)} \le L_k^{(\mathbb{T}_n)} \tag{3.103}$$

Indeed, conditioning on the number of discovered vertices, we have

$$L_k^{(\mathbb{T}_b)} = \text{Binom}(b, {}^{\alpha}/_{n}) \tag{3.104}$$

Now think of the binomial random variable $\text{Binom}(b, {}^{\alpha}/_{n})$ as the sum of the first b terms in a sequence of Bernoulli random variables that are 1 with probability ${}^{\alpha}/_{n}$

and zero otherwise. To get $L_k^{(\mathbb{T}_m)}$ we then add only the first m, to sample $L_k^{(K_n)}$ we add the first $n-k-A_k$, and to get $L_k^{(\mathbb{T}_n)}$ we add the first n of these variables. Under the condition

$$m \le n-k-A_k \le n \tag{3.105}$$

we will then have (3.103). The upper bound in this condition is trivial and the lower bound will hold as long as $k+A_k \le n-m$.

This argument shows that, if the connected component $C(1)$ of vertex "1" in K_n is of size r, then so is the component $C(\varnothing)$ on $\mathbb{T}_n$, i.e., the right inequality in (3.102) holds. Similarly, thinking of adding the discovered vertices to the tree one by one, before the component $C(1)$ of K_n reaches the size $r \le n-m$, the component $C(\varnothing)$ of $\mathbb{T}_m$ will not be larger than r. This implies

$$\mathbb{P}_{K_n}\big(|C(1)|<r\big) \le \mathbb{P}_{\mathbb{T}_m}\big(|C(\varnothing)|<r\big), \qquad r \le n-m, \tag{3.106}$$

whose complement then yields the left inequality in (3.102). □

The following observation will be helpful in the proof:

Lemma 3.32 [Continuity in offspring distribution] *Let $\mathfrak{p}^{(m)} = (\mathfrak{p}_m^{(n)})$ be a family of offspring distributions and let $\mathfrak{p}$ be an offspring distribution such that $0<\mathfrak{p}_0<1$. Suppose that, for each $m \ge 0$,*

$$\mathfrak{p}_m^{(n)} \underset{n\to\infty}{\longrightarrow} \mathfrak{p}_m. \tag{3.107}$$

Then for each $m \ge 1$,

$$\mathbb{P}_{\mathfrak{p}^{(n)}}(X_m \ge 1) \underset{n\to\infty}{\longrightarrow} \mathbb{P}_{\mathfrak{p}}(X_m \ge 1). \tag{3.108}$$

where $\mathbb{P}_{\mathfrak{p}^{(n)}}$ and $\mathbb{P}_{\mathfrak{p}}$ denote the law of the branching process with offspring distributions $\mathfrak{p}^{(n)}$ and $\mathfrak{p}$, respectively.

Problem 3.33 Define

$$\lambda^{(n)}(s) = -\log \sum_{m\ge 0} \mathfrak{p}_m^{(n)} \mathrm{e}^{-sm} \tag{3.109}$$

and show that $\lambda^{(n)}(s) \to \lambda(s)$, where $\lambda(s)$ is defined using $\mathfrak{p} = (\mathfrak{p}_m)$. Then use this to prove the lemma.

Now we are ready to prove our result for the Erdős-Rényi random graph:

Proof of Theorem 3.28. We will prove upper and lower bounds on the quantity

$$\theta_{\epsilon,n} := \mathbb{P}_{K_n}(|C(1)| \ge \epsilon n). \tag{3.110}$$

By Lemma 3.31, it follows that

$$\theta_{\epsilon,n} \ge \mathbb{P}_{\mathbb{T}_{(1-\epsilon)n}}\big(|C(\varnothing)| \ge \epsilon n\big) \ge \mathbb{P}_{\mathbb{T}_{(1-\epsilon)n}}\big(|C(\varnothing)| = \infty\big) \tag{3.111}$$

Equation (3.73) in Corollary 3.21 shows that the right-hand side is the largest positive solution to the equation

$$\theta = 1 - \big(1 - \tfrac{\alpha}{n}\theta\big)^{(1-\epsilon)n} \tag{3.112}$$

As

$$\big(1-\tfrac{\alpha}{n}\theta\big)^{(1-\epsilon)n} \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{(1-\epsilon)\alpha\theta} \tag{3.113}$$

this equation and, by convexity of the right-hand side, also its maximal positive solution converge to that of (3.98) in the limits $n \to \infty$ followed by $\epsilon \downarrow 0$. This proves (3.97) with the limits replaced by *limes inferior*.

It remains to prove the corresponding upper bound. Lemma 3.31 gives us

$$\theta_{\epsilon,n} \le \mathbb{P}_{\mathbb{T}_n}\big(|C(\varnothing)| \ge \epsilon n\big). \tag{3.114}$$

We will now use the connection to branching: Let $X_0 = 0, X_1, \dots$ be the number of vertices of $C(\varnothing)$ at integer distances from the root. As we noted before, (X_n) is a Galton-Watson branching process with binomial offspring distribution. On $\{|C(\varnothing)| \ge \epsilon n\}$ the process has survived for at least $1/\epsilon$ generations. This means

$$\mathbb{P}_{\mathbb{T}_n}\big(|C(\varnothing)| \ge \epsilon n\big) \le \mathbb{P}_{\mathbb{T}_n}(X_{1/\epsilon} \ge 1), \tag{3.115}$$

where we pretend, for notational ease, that $1/\epsilon$ is an integer.

First we take the limit $n \to \infty$ of $\mathbb{P}_{\mathbb{T}_n}(X_r \ge 1)$. A direct calculation — or a reference to the convergence of a binomial law to Poisson — shows that

$$\binom{n}{m}\Big(\frac{\alpha}{n}\Big)^m\Big(1-\frac{\alpha}{n}\Big)^{n-m} \underset{n\to\infty}{\longrightarrow} \frac{\alpha^m}{m!}\mathrm{e}^{-\alpha}, \qquad m \ge 0, \tag{3.116}$$

and so, by Lemma 3.32, $\mathbb{P}_{\mathbb{T}_n}(X_r \ge 1)$ converges to the probability that a Poisson branching process with parameter α survived up to time r. As $r \to \infty$ — which corresponds to $\epsilon \downarrow 0$ in the original setting — this probability converges to the maximal positive solution of (3.98). This proves (3.97) with *limes superior* and thus finishes the proof. □

The fact that all vertices of K_n look "the same" suggest that $\theta(\alpha)$ actually represents the fraction of vertices in components of macroscopic size. This is indeed the case, but we will not try to prove it here. In fact, more is known:

Theorem 3.34 *Given a realization of $\mathscr{G}(n, \alpha/n)$, let $\mathcal{C}_1, \mathcal{C}_2, \dots$ be the list of all connected components ranked decreasingly by their size. Then we have:*

(1) If $\alpha < 1$, then $|\mathcal{C}_1| = O(\log n)$, i.e., all components are at most logarithmic.

(2) If $\alpha > 1$, then $|\mathcal{C}_1| = \theta(\alpha)n + o(n)$ and $|\mathcal{C}_2| = O(\log n)$, i.e, there is a unique *giant component and all other components are of at most logarithmic size.*

(3) If $\alpha = 1$, then $|\mathcal{C}_1|, |\mathcal{C}_2|, \dots$ are all of order $n^{2/3}$ with a nontrivial limit distribution of $n^{-2/3}|\mathcal{C}_1|, n^{-2/3}|\mathcal{C}_2|, \dots$.

The case (3) is the one most interesting because corresponds to the critical behavior we saw at $p = p_{\mathrm{c}}$ on the regular tree. The regime actually extends over an entire *critical window*, i.e., for values

$$p = \frac{1}{n} + \frac{\lambda}{n^{4/3}} \tag{3.117}$$

where λ is any fixed real number. One of the key tools to analyze this regime is the tree-search algorithm that we used at the beginning of this section.