Lecture 20: The giant component in supercritical Erdős-Rényi random graphs

Our goal now is to prove the supercritical $\lambda > 1$ part of Theorem 17.2 from Lecture 17. Let ζ_λ denote the probability that a branching process with Poisson(λ) offspring distribution does not go extinct. Recall from Example 12.3 of Lecture 12 that ζ_λ satisfies

$$\zeta_\lambda = 1 - e^{-\lambda\zeta_\lambda}.$$

Theorem 17.2 (Supercritical case)**.** *Let $C_{\max}$ be the size of the largest component in $G = G(n, \lambda/n) \sim \mathrm{ER}(n, \lambda/n)$ with $\lambda > 1$. Then for any $\alpha \in (0, 1/2)$ and $\beta \in (0, 1 - 2\alpha)$,*

$$\mathbb{P}\left(\left|\frac{C_{\max}}{n} - \zeta_\lambda\right| > n^{-\alpha}\right) = \mathrm{O}(n^{-\beta}).$$

Moreover, with probability tending to one, there is a unique largest component, and the second largest component is $\Theta(\log(n))$.

The strategy of proof is as follows. We first show that with good probability, the number of vertices in components with at most $(1+\kappa)I_\lambda^{-1}\log(n)$ vertices is suitably close to $n(1-\zeta_\lambda)$. We then show that the number of vertices in components having between $(1+\kappa)I_\lambda^{-1}\log(n)$ and (just smaller than) $\zeta_\lambda n$ vertices is small. These two facts imply that $\zeta_\lambda n$ vertices must be in components with at least $\zeta_\lambda n$ vertices, but then they must all be in the same component, and also give the basic picture (one giant component, rest of size at most $\log(n)$). The following two lemmas more precisely state these ideas.

Lemma 20.2. *For $v \in \{1, \dots, n\}$, let C_v denote the size of the component containing vertex v in $G \sim \mathrm{ER}(n, \lambda/n)$ with $\lambda > 1$, and let $Y_k = \sum_{v=1}^n \mathbb{I}[C_v \le k]$, where $k = \lceil (1+\kappa)I_\lambda^{-1}\log(n)\rceil$ and $\kappa > 0$. Then for any $\alpha \in (0, 1/2)$, $\beta \in (0, 1-2\alpha)$ and fixed constant $c > 0$,*

$$\mathbb{P}\big(|Y_k - (1-\zeta_\lambda)n| > c\, n^{1-\alpha}\big) = \mathrm{O}(n^{-\beta}). \tag{20.1}$$

Lemma 20.3. *For $v \in \{1, \dots, n\}$, let C_v denote the size of the component containing vertex v in $G \sim \mathrm{ER}(n, \lambda/n)$ with $\lambda > 1$, let $0 < \gamma < \zeta_\lambda$, and let $V_k = \sum_{v=1}^n \mathbb{I}[k < C_v \le \gamma n]$. Then for any $\alpha \in (0, 1/2)$ and $\beta \in (0, 1-2\alpha)$, and fixed constant $c > 0$, there is a $\kappa > 0$ such that for $k = \lceil (1+\kappa)I_\lambda^{-1}\log(n)\rceil$,*

$$\mathbb{P}(V_k > c\, n^{1-\alpha}) = \mathrm{O}(n^{-\beta}). \tag{20.2}$$

Our goal over the next two lectures will be to prove these lemmas, but first we show precisely how the $\lambda > 1$ part of Theorem 17.2 follows from them.

Proof of Theorem 17.2 ($\lambda > 1$). First note that

$$\mathbb{P}\left(\left|\frac{C_{\max}}{n} - \zeta_\lambda\right| > n^{-\alpha}\right) = \mathbb{P}(C_{\max} > \zeta_\lambda n + n^{1-\alpha}) + \mathbb{P}(C_{\max} < \zeta_\lambda n - n^{1-\alpha}).$$

For the first term on the right hand side, note that if $C_{\max} > \zeta_\lambda n + n^{1-\alpha}$, then for Y_k and k defined in Lemma 20.2, we must have $Y_k < (1-\zeta_\lambda)n - n^{1-\alpha}$, since there are n vertices total. Thus, by Lemma 20.2, for any $\beta \in (0, 1-2\alpha)$,

$$\mathbb{P}(C_{\max} > \zeta_\lambda n + n^{1-\alpha}) \le \mathbb{P}\big(|Y_k - (1-\zeta_\lambda)n| > n^{1-\alpha}\big) = \mathrm{O}(n^{-\beta}).$$

For the other term, let $\kappa > 0$ and $k = \lceil (1+\kappa)I_\lambda \log(n)\rceil$ such that (20.2) holds for $c = 1/2$. Using Lemmas 20.2 and 20.3 with $\gamma = (3/4)\zeta_\lambda$ (and the notation there), we have

$$\mathbb{P}(C_{\max} < \zeta_\lambda n - n^{1-\alpha})$$

$$\leq \mathbb{P}\big(C_{\max} < \zeta_\lambda n - n^{1-\alpha}, |Y_k - (1-\zeta_\lambda)n| \leq (1/2)n^{1-\alpha}, V_k \leq (1/2)n^{1-\alpha}\big) + \mathrm{O}(n^{-\beta}).$$

Now notice that if $|Y_k - (1-\zeta_\lambda)n| \leq (1/2)n^{1-\alpha}$ and $V_k \leq (1/2)n^{1-\alpha}$, then the number of vertices in components with at least $\gamma n = (3/4)\zeta_\lambda n$ vertices must be in the interval $[n\zeta_\lambda - n^{1-\alpha}, n\zeta_\lambda + (1/2)n^{1-\alpha}]$. But then they must all be in the same component; this is because two components of sizes at least $(3/4)\zeta_\lambda n$ would consist of greater than $n\zeta_\lambda + (1/2)n^{1-\alpha}$ vertices. To summarise: If $|Y_k - (1-\zeta_\lambda)n| \leq (1/2)n^{1-\alpha}$ and $V_k \leq (1/2)n^{1-\alpha}$, then there is at most one component of size at least $(3/4)\zeta_\lambda n$, and if there is one, it must be of size at least $n\zeta_\lambda - n^{1-\alpha}$, and so in particular, $C_{\max} \geq n\zeta_\lambda - n^{1-\alpha}$. Thus

$$\mathbb{P}\big(C_{\max} < \zeta_\lambda n - n^{1-\alpha}, |Y_k - (1-\zeta_\lambda)| \leq (1/2)n^{1-\alpha}, V_k \leq (1/2)n^{1-\alpha}\big) = 0,$$

and so the first assertion of the theorem follows. For the second assertion, the argument above shows that with probability tending to one, the giant component is unique, and the other components are at most order $\log(n)$. That there must be a component of strict order $\log(n)$ follows from the second moment method for $\sum_{v=1}^n \mathbb{I}[C_v \in (1\pm\kappa)I_\lambda^{-1}\log(n)]$, similar to the proofs of Lemma 18.1 from Lecture 18 and Lemma 20.5 below, and we leave this as an exercise. □

We now work on proving Lemma 20.2, which follows by showing that $\mathbb{E}[Y_k]$ is close to $(1-\zeta_\lambda)n$, and that that $\mathrm{Var}(Y_k)$ is small.

Lemma 20.4. *For $v \in \{1,\dots,n\}$, let C_v denote the size of the component containing vertex v in $G \sim \mathrm{ER}(n, \lambda/n)$ with $\lambda > 1$, and let $Y_k = \sum_{v=1}^n \mathbb{I}[C_v \leq k]$, where $k = \lceil (1+\kappa)I_\lambda \log(n)\rceil$.*

$$\mathbb{E}[Y_k] = n(1-\zeta_\lambda) + \mathrm{O}\big(\log(n)^3\big).$$

Proof. First we compute

$$\mathbb{E}[Y_k] = n\mathbb{P}(C_v \leq k) = n(1 - \mathbb{P}(C_v > k)).$$

To estimate $\mathbb{P}(C_v > k)$, we again use Lemmas 13.1, 15.4, and 16.5, which imply there is a constant c' depending only on λ

$$\begin{aligned}\mathbb{P}(C_v > k) &= \zeta_\lambda + \mathrm{O}\big(e^{-kI_\lambda}\big) + \mathrm{O}(k^3/n)\\ &= \zeta_\lambda + \mathrm{O}\big(n^{-(1+\kappa)}\big) + \mathrm{O}\big(\log(n)^3/n\big)\\ &= \zeta_\lambda + \mathrm{O}\big(\log(n)^3/n\big).\end{aligned}$$

□

Next time we prove the following lemma bounding the variance of Y_k, and whose proof is similar to Lemma 18.1 from Lecture 18, and is deferred to next lecture.

Lemma 20.5. *For $v \in \{1,\dots,n\}$, let C_v denote the size of the component containing vertex v in $G \sim \mathrm{ER}(n, \lambda/n)$ and $Y_k =: \sum_{v=1}^n \mathbb{I}[C_v \leq k]$, where $k < n$. Then*

$$\mathbb{E}[Y_k^2] \leq \mathbb{E}[Y_k]^2 + \mathbb{E}[Y_k](\lambda k^3 + k + 1)$$

With these lemmas, we can now prove Lemma 20.2.

Proof of Lemma 20.2. For $\alpha \in (0, 1/2)$, we use the triangle inequality, Lemma 20.4, Chebychev's inequality, and Lemma 20.5 to find

$$\begin{aligned}&\mathbb{P}\big(|Y_k - (1-\zeta_\lambda)n| > c\,n^{1-\alpha}\big)\\ &\quad\leq \mathbb{P}\Big(\big|Y_k - \mathbb{E}[Y_k]\big| > \big(\tfrac{c}{2}\big)n^{1-\alpha}\Big) + \mathbb{I}\Big(\big|\mathbb{E}[Y_k] - (1-\zeta_\lambda)n\big| > \big(\tfrac{c}{2}\big)n^{1-\alpha}\Big)\\ &\quad\leq 4c^{-2}\frac{\mathrm{Var}(Y_k)}{n^{2(1-\alpha)}} + \mathbb{I}[\mathrm{O}(\log(n)^3) > n^{1-\alpha}]\\ &\quad= \mathrm{O}\big(n^{1-2(1-\alpha)}\log(n)^3\big)\\ &\quad= \mathrm{O}\big(n^{-(1-2\alpha)}\log(n)^3\big) = \mathrm{O}\big(n^{-\beta}\big),\end{aligned}$$

for any $\beta \in (0, 1-2\alpha)$, as desired. □