Poisson branching process

We now consider branching process whose offspring distribution is Poisson with parameter $\lambda > 0$. We write P_λ for its law. The generating function of this distribution is

(2.39) $$G_\lambda(s) = E_\lambda[s^X] = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} s^i = \exp(\lambda(s-1)).$$

Hence, the extinction probability $\eta = \eta_\lambda$ solves

(2.40) $$\eta_\lambda = \exp(\lambda(\eta_\lambda - 1))$$

As $E_\lambda[Z_1] = \lambda$, the branching process is supercritical, i.e. $\eta_\lambda < 1$, iff $\lambda > 1$. The duality theorem (2.25) then becomes

(2.41) Theorem (Poisson duality) Let $\lambda > 1$ and $\mu < 1$ such that

(2.42) $$\mu e^{-\mu} = \lambda e^{-\lambda} \qquad \text{(conjugate pair)}$$

Then Pois(λ)-branching process conditioned on $T < \infty$ has the same distribution as Pois(μ)-branching process

Proof: We only need to compute p_i' of (2.24)

$$p_i' = \eta_\lambda^{i-1} p_i \overset{(2.40)}{=} \frac{(\eta_\lambda \lambda)^i}{i!} e^{-\lambda \eta_\lambda}.$$

so $(p_i')_{i \geq 0}$ is Poisson with parameter $\lambda \eta_\lambda = \mu$.

But (2.40) then implies that μ solves (2.42). $\square$

Similarly, for the total progeny, (2.34)–(2.35) yields

(2.43) Proposition: $(\lambda > 0)$.

(2.44) $$P_\lambda(T = n) = \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}$$

Hence, as $n \to \infty$, setting $I_\lambda = \lambda - 1 - \log \lambda$

(2.45) $$P_\lambda(T = n) = \frac{1}{\lambda \sqrt{2\pi n^3}} e^{-I_\lambda n} \left(1 + O\left(\tfrac{1}{n}\right)\right)$$

and, for $\lambda = 1$,

(2.46) $$P_1(T = n) = \frac{1}{\sqrt{2\pi n^3}} \left(1 + O\left(\tfrac{1}{n}\right)\right)$$

Proof: (2.44) follows directly from (2.34), (2.35), since $\sum_{i=1}^{n} X_i \sim \text{Pois}(n\lambda)$. (2.45) then follows from Stirling formula

$$n! = \sqrt{2\pi n}\, e^{-n} n^n \left(1 + O\left(\tfrac{1}{n}\right)\right).$$

Binomial branching process.

This process will play substantial role in investigations of random graphs. We write $P_{n,p}$ for the law of branching process whose offspring distribution is binomial w.p. n,p.

(2.47) Lemma (Comparison of Poisson/Binomial branching processes)

Let $\lambda = n \cdot p$. Then, for every $k \ge 1$,

(2.48) $$P_{n,p}(T \ge k) = P_\lambda(T \ge k) + e_n(k)$$

where

(2.49) $$|e_n(k)| \le \frac{\lambda^2}{n} \sum_{s=1}^{k-1} P_\lambda(T \ge s)$$

In particular

(2.50) $$|e_n(k)| \le k \frac{\lambda^2}{n}.$$

Proof: We use a coupling argument based on the following

(2.51) Exercise: Let $X \sim \mathrm{Bin}(n, \frac{\lambda}{n})$ and $X^* \sim \mathrm{Pois}(\lambda)$. Then there exists coupling $(\hat{X}, \hat{X}^*)$ of X, X^* such that $P(\hat{X} \ne \hat{X}^*) \le \frac{\lambda^2}{n}$.

We use this exercise to construct two iid sequences, $(X_i)_{i \ge 1}$, $(X_i^*)_{i \ge 1}$ on a common probability space $(\Omega, \mathcal{A}, P)$ such that $X_i \sim \mathrm{Bin}(n,p)$, $X_i^* \sim \mathrm{Pois}(\lambda)$ and $P(X_i \ne X_i^*) \le \frac{\lambda^2}{n}$.

Hence, writing $H_0 = \inf\{k \ge 0 : 1 + \sum_{i=1}^{k} (X_i - 1) = 0\}$, $H_0^* = \inf\{k \ge 0 : 1 + \sum_{i=1}^{k} (X_i^* - 1) = 0\}$,

$$|P_{n,p}(T \ge k) - P_\lambda(T \ge k)| = |P(H_0 \ge k) - P(H_0^* \ge k)|$$

(2.52) $$= |P(H_0 \ge k, H_0^* < k) - P(H_0^* \ge k, H_0 < k)|$$

$$\le \max\{P(H_0 \ge k, H_0^* < k), P(H_0 < k, H_0^* \ge k)\}.$$

(2.53) Events $\{H_0 < k\}$, $\{H_0^* < k\}$ are $\sigma(X_i, X_i^*; i < k)$-measurable. Hence, if $H_0 \ge k$ and $H_0^* < k$, there must be $i \le k-1$ with $X_i \ne X_i^*$, and similarly if $H_0 < k$ and $H_0^* \geq k$. Hence

$$P[H_0 \geq k, H_0^* < k] \leq \sum_{j=1}^{k-1} P[X_i = X_i^* \, \forall i < j, \; X_j \neq X_j^*, \; H_0 \geq k]$$

On $\{X_i = X_i^* \, \forall i < j, \; H_0 \geq k\}$ we have $H_0^* \geq j$. Therefore

$$\leq \sum_{j=1}^{k-1} P[X_j \neq X_j^*, \; H_0^* \geq j]$$

$$\overset{(2.53)}{=} \sum_{j=1}^{k-1} P[H_0^* \geq j] \cdot P[X_j \neq X_j^*],$$

$$\leq \frac{\lambda^2}{n} \sum_{j=1}^{k-1} P[H_0^* \geq j] = \frac{\lambda^2}{n} \sum_{j=1}^{k-1} P_\lambda[T \geq j].$$

Similarly one obtains

$$P[H_0^* \geq k, H_0 < k] \leq \frac{\lambda^2}{n} \sum_{j=1}^{k-1} P_\lambda[T \geq j].$$

Inserting this into (2.52) yields (2.48), (2.49).
(2.50) is then obvious. $\square$.

III. ERDŐS-RÉNYI RANDOM GRAPH

In this chapter we study the connected components of the Erdős-Rényi random graph. We will show that there is a phase transition in the behaviour of these components corresponding to the phase transition in branching processes.

The vertex set of ER graph is $[n] := \{1, \dots, n\}$. Every edge $\{x,y\}$, $x \neq y \in [n]$, is present in the graph with probability p, independently of remaining edges, where $p \in [0,1]$ is a fixed parameter. We use $ER(n,p)$ to denote such random graph and write $P = P^{ER}_{n,p}$ for its distribution.

For $x, y \in [n]$, we say that x is connected to y, and write $x \leftrightarrow y$, if there is a path $x = x_0, \dots, x_n = y$ joining x, y, that is all edges $(x_{i-1}, x_i)_{i \le n}$ are present in $ER(n,p)$. We set

(3.1) $$C(x) = \{y \in [n] : x \leftrightarrow y\}$$

to be the connected component containing x, and write C_{max} for the largest connected component (e.g. containing the vertex with the smallest label, to break possible ties).

(3.2) Then $$|C_{max}| = \max_{x \in [n]} |C(x)|.$$

Procedure to find $C(x)$

Similarly as in the tree case (see pp. 10-11), it is useful to explore $C(x)$ step by step. The algorithm is as follows: In its course, every vertex is in one of three states active, neutral, inactive. At time 0 only x is active all other vertices are neutral, we set $S_0 = 1$.

At each time $k \geq 1$, we pick one active vertex y, by an arbitrary rule, and explore the state of all edges $\{y, z\}$ where z runs over neutral vertices. If $\{y, z\}$ is an edge in the graph, then z becomes active, otherwise it remains neutral. After searching all neutral vertices, we set y inactive and set S_k equal the new number of active vertices at time k.

At the moment $T = \inf\{k \geq 0 : S_k = 0\}$ the algorithm stops, and by construction $C(x)$ is exactly the set of

(3.3) inactive vertices, that is $|C(x)| = T$.

As in the tree case $S_0 = 1$ and

(3.4) $$S_k = S_{k-1} + X_k - 1, \qquad k \geq 1$$

where X_k is the number of vertices becoming active in the k^{th} step.

(3.5) Observe that in the algorithm, the state of every edge is explored at most once.

In the Erdős–Rényi case, every edge is present w.p. p, independently of each others. Due to (3.5) it means that X_k has a binomial distribution with parameters N_{k-1}, p, where N_{k-1} is the number of neutral vertices at the end of $(k-1)^{th}$ step. Since in every step exactly one vertex becomes inactive,

$$N_{k-1} = n - (k-1) - S_{k-1} \qquad \text{and thus}$$

(3.6) $$X_k \sim \mathrm{Bin}\,(n - (k-1) - S_{k-1}, p)$$

Note that (3.4) is the same as (2.19). The new feature is that the law of X_k depends on S_{k-1}. However, if k and S_k are small w.r.t. n, then this dependence can "almost" be ignored, as we will see.

From Theorem (2.4), we should expect that the behaviour of $|C(x)|$ changes when $p \simeq \frac{1}{n}$, since then $EX_i \simeq 1$, at least for i small. We will thus be mainly interested in the regime when

(3.7) $$p = \frac{\lambda}{n}, \quad \lambda \geq 0$$

Since $\mathrm{Bin}(n, \frac{\lambda}{n}) \xrightarrow{n\to\infty} \mathrm{Pois}(\lambda)$, it explains why we studied Poisson branching process before.

(3.8) Lemma (domination by branching process) For every $n, k \geq 1$, $p \in [0,1]$

$$P^{ER}_{n,p}(|C(1)| \geq k) \leq P^{BP}_{n,p}(T \geq k)$$

where $P^{BP}_{n,p}$ is the law of Binomial(n,p) branching process and T is its total progeny.

(3.9) Remark: (3.8) implies that T stochastically dominates $|C(1)|$, in particular there is a coupling of T and $|C(1)|$ s.t. $|C(1)| \leq T$.

Proof of (3.8): Recall (3.4)-(3.6) and (2.19). Let $X_i \sim \mathrm{Bin}(N_{i-1}, p)$ and $Y_i \sim \mathrm{Bin}(n - N_{i-1}, p)$, be conditionally independent given N_{i-1}, and set $\tilde{X}_i = X_i + Y_i$. Since X_i, Y_i are conditionally independent, $\tilde{X}_i \sim \mathrm{Bin}(n,p)$ and since this distribution does not depend on $(\tilde{X}_j)_{j<i}$, $\tilde{X}_i$'s are i.i.d. Also $\tilde{X}_i \geq X_i$, a.s., since $Y_i \geq 0$. Set $\tilde{S}_k = \sum_{i=1}^{k} \tilde{X}_i - (k-1)$. Then

$$P^{ER}_{n,p}(|C(1)| \geq k) \overset{(3.3)}{=} P(S_i > 0 \; \forall i \leq k-1) \overset{\tilde{X}_i \geq X_i}{\leq} P(\tilde{S}_i > 0 \; \forall i \leq k-1)$$
$$= P^{BP}_{n,p}(T \geq k) \qquad \square.$$

We now obtain a corresponding lower bound.

(3.10) <u>Lemma</u>: For every $n \geq 1$, $k \in [n]$, $p \in [0,1]$,
$$P^{ER}_{n,p}\left(|C(1)| \geq k\right) \geq P^{BP}_{n-k,p}\left(T \geq k\right).$$

<u>Proof</u>: Let
$$T_k = \inf\{j \geq 0 : N_j \leq n-k\}.$$
As all active vertices eventually become inactive, $T_k < \infty$ iff $|C(1)| \geq k$. Moreover, trivially $T_k \leq k-1$ on $\{|C(1)| \geq k\}$, since $N_{k-1} \leq n-(k-1)-1 = n-k$. Hence

(3.11) $$P^{ER}_{n,p}\left(|C(1)| \geq k\right) = P\left(S_j > 0 \ \forall j \leq T_k\right).$$

We now take $(\hat{X}_i)_{i \geq 1}$ iid $\mathrm{Bin}(n-k, p)$ and for $i \leq T_k$, $Y_i \sim \mathrm{Bin}(N_{n-1} - (n-k), p)$, independently of anything else. Then $X_i = \hat{X}_i + Y_i \sim \mathrm{Bin}(N_{n-1}, p)$, i.e. $X_i \geq \hat{X}_i \ \forall i \leq T_k$, a.s. Setting $\hat{S}_i = \sum_{j=1}^{i} \hat{X}_j - (i-1)$, we get
$$\{S_j > 0 \ \forall j \leq T_k\} \supset \{\hat{S}_j > 0 \ \forall j \leq T_k\} \supset \{\hat{S}_j > 0 \ \forall j \leq k-1\}$$
and thus, by (3.11)
$$P^{ER}_{n,p}\left(|C(1) \geq k|\right) \geq P\left(\hat{S}_j > 0 \ \forall j \leq k-1\right) = P^{BP}_{n-k,p}\left(T \geq k\right) \quad \square.$$

<u>Subcritical regime of ER graph.</u>

We now study the size $|C_{max}|$ of the maximal connected component in the case $p \overset{(3.7)}{=} \frac{\lambda}{n}$, with $\lambda < 1$.

(3.12) <u>Exercise</u> (Large deviations of Poisson distribution).
Let X be a Poisson random variable with parameter λ.
Show that for every $a > \lambda$

(3.13) $$P(X \geq a) \leq e^{-I_\lambda(a)}$$

and for every $0 \leq a < \lambda$

(3.14) $$P(X \leq a) \leq e^{-I_\lambda(a)}$$

with $$I_\lambda(a) = a\left(\log \frac{a}{\lambda} - 1\right) + \lambda > 0 \quad \text{if } a \neq \lambda.$$
$$= \sup_{t \in \mathbb{R}} \left(at - \log E[e^{tX}]\right)$$

(3.15) <u>Theorem</u> (the largest subcritical cluster). For $\lambda<1$, $p=\frac{\lambda}{n}$.

$$P^{ER}_{n,p}\left(\left|\frac{|C_{max}|}{\log n}-\mu\right|\geq\varepsilon\right)\leq n^{-\delta}$$

for some $\delta(\varepsilon,\lambda)>0$, where $\mu=\frac{1}{I_\lambda(1)}$. In particular, $\frac{|C_{max}|}{\log n}\xrightarrow{n\to\infty}\mu$ in probability.

(3.16) <u>Remark</u>: To understand the value μ, recall that $|\mathcal{C}(1)|$ can be approximated by the total progeny of $\mathrm{Pois}(\lambda)$ branching process. By (2.45), $P^{BP}_\lambda[T\geq a\mu\log n]\approx e^{-I_\lambda(1)\,a\mu\log n}=n^{-a}$, $a>0$. If $|C(x)|$, $x\in[n]$, were independent, this would yield $|C_{max}|=\max_{x\in[n]}|\mathcal{C}(x)|\simeq\mu\log n$.

<u>Proof of (3.15)</u>: We write I_λ for $I_\lambda(1)$.

<u>Upper bound</u>: We show: For every $a>\mu$

(3.17) $$P^{ER}_{n,p}\left(|C_{max}|\geq a\log n\right)\leq n^{-\delta(a,\lambda)}.$$

Indeed

(3.18) $$P\left(|C_{max}|\geq a\log n\right)=P\left(\exists x\leq n: |C(x)|\geq a\log n\right)\leq n\cdot P\left(|C(1)|\geq a\log n\right).$$

By Lemma (3.8)

$$P\left(|C(1)|\geq a\log n\right)\leq P^{BP}_{n,p}\left(T\geq a\log n\right)\overset{(2.47)-(2.50)}{\leq}P^{BP}_\lambda\left(T\geq a\log n\right)+\underbrace{a\log n\,\frac{\lambda^2}{n}}_{\leq c n^{-1/2}}$$

By (2.43),

$$P_\lambda\left(T\geq a\log n\right)\leq c\sum_{j=a\log n}^{\infty}e^{-jI_\lambda}=c\,\frac{e^{-I_\lambda a\log n}}{1-e^{-I_\lambda}}\leq c\,n^{-a\cdot I_\lambda}\overset{(a>\mu)}{\leq}n^{-1-\delta(a,\lambda)}$$

Coming back to (3.18), this implies (3.17)

Lower bound: We use 2nd moment method. Let

(3.19) $$Z_k = \#\{x \in [n] : |\mathcal{C}(x)| \geq k\}$$

Obviously $|\mathcal{C}_{max}| = \max\{k : Z_k \geq k\} = \max\{k : Z_k > 0\}$.

Hence, by Chebyshev inequality,

(3.20) $$P[|\mathcal{C}_{max}| < k] = P[Z_k = 0] \leq \frac{Var(Z_k)}{(E Z_k)^2}.$$

We now set $k_n := a \log n$ with $a < I_\lambda^{-1}$ and estimate the variance and the expectation on the RHS of (3.20).

(3.21) Lemma: If $a < I_\lambda^{-1}$, then, for n large,

$$E^{ER}_{n,\frac{\lambda}{n}}[Z_{k_n}] \geq n^{(1 - I_\lambda a)(1+o(1))}$$

Proof:

$$E^{ER}_{n,\frac{\lambda}{n}}[Z_{k_n}] \overset{(3.19)}{=} n \cdot P[|\mathcal{C}(1)| \geq k_n]$$

$$\overset{(3.10)}{\geq} n \cdot P^{BP}_{n-k_n,\frac{\lambda}{n}}(T \geq k_n)$$

(3.22)

$$\overset{(2.47)}{\geq} n \cdot P^{BP}_{\lambda_n}(T \geq k_n) - \underbrace{k_n \frac{\lambda_n^2}{n}}_{\leq c n^{-1/2}}$$

with $\lambda_n = \frac{\lambda(n-k_n)}{n} \xrightarrow{n\to\infty} \lambda$. Using (2.43)

$$P^{BP}_{\lambda_n}(T \geq k_n) \geq P^{BP}_{\lambda_n}(T = \lceil k_n \rceil) =$$

(3.23) $$= \frac{(\lambda_n k_n)^{k_n}}{k_n!} e^{-\lambda_n k_n} \overset{\text{Stirling}}{=} \frac{(\lambda_n k_n)^{k_n}}{k_n^{k_n} e^{-k_n} \sqrt{2\pi k_n}} e^{-\lambda_n k_n} (1+o(1))$$

$$= \exp\{-k_n I_{\lambda_n}(1+o(1))\} \overset{k_n = a\log n}{=} n^{-I_\lambda a(1+o(1))},$$

by continuity of $\lambda \mapsto I_\lambda$. Putting (3.22), (3.23) together implies (3.21) $\square$

(3.24) Lemma: (Variance of Z_k). For every $\lambda > 0$, $n \geq 1$, $k \in [n]$,

$$Var^{ER}_{n,\frac{\lambda}{n}}[Z_k] \leq n\, E^{ER}_{n,\frac{\lambda}{n}}[|\mathcal{C}(1)| \mathbb{1}_{|\mathcal{C}(1)| \geq k}].$$

Proof: Since $Z_k = \sum_{x \in [n]} \mathbb{1}_{|\mathcal{C}(x)| \geq k}$, by easy computation

(3.25) $$Var(Z_k) = \sum_{x,y \in [n]} \Big\{ P(|\mathcal{C}(x)| \geq k, |\mathcal{C}(y)| \geq k) - \underbrace{P(|\mathcal{C}(x)| \geq k) P(|\mathcal{C}(y)| \geq k)}_{= P(|\mathcal{C}(x)| \geq k)^2} \Big\}$$

We now split the first probability, depending on $\{x \leftrightarrow y\}$.

(3.26)
$$P(|\mathcal{C}(x)| \geq k, |\mathcal{C}(y)| \geq k, x \not\leftrightarrow y) = \sum_{\ell \geq k} P(|\mathcal{C}(x)| = \ell, |\mathcal{C}(y)| \geq k, x \not\leftrightarrow y) =$$

$$\overset{\leq}{=} \sum_{\ell \geq k} P(|\mathcal{C}(y)| \geq k \mid |\mathcal{C}(x)| = \ell, x \not\leftrightarrow y) \cdot P(|\mathcal{C}(x)| = \ell).$$

On $\{|\mathcal{C}(x)| = \ell, x \not\leftrightarrow y\}$, by independence of edges occurrences in ER graph, $|\mathcal{C}(y)|$ has the same law as $|\mathcal{C}(1)|$ under $P^{ER}_{n-\ell,p}$.

$$\leq \sum_{\ell \geq k} P_{n-\ell,p}(|\mathcal{C}(1)| \geq k) \cdot P(|\mathcal{C}(x)| = \ell)$$

$$\overset{\leq}{=} P_{n,p}(|\mathcal{C}(1)| \geq k)^2.$$

where in the last inequality we used obvious monotonicity

$$P_{n-\ell,p}(|\mathcal{C}(1)| \geq k) \leq P_{n,p}(|\mathcal{C}(1)| \geq k).$$

Hence, by (3.25), (3.26)

$$\mathrm{Var}(Z_k) \leq \sum_{x,y \in [n]} P(|\mathcal{C}(x)| \geq k, |\mathcal{C}(y)| \geq k, x \leftrightarrow y)$$

(3.27)
$$= \sum_{x,y} P(|\mathcal{C}(x)| \geq k, x \leftrightarrow y)$$

$$= n \cdot E\Big[\mathbb{1}_{|\mathcal{C}(1)| \geq k} \cdot \sum_{y \in [n]} \mathbb{1}_{y \in \mathcal{C}(1)}\Big]$$

$$= n \cdot E\big[|\mathcal{C}(1)| \, \mathbb{1}_{|\mathcal{C}(1)| \geq k}\big]$$

as claimed $\square$.

(3.28) <u>Lemma</u>: If $k_n = a \log n$ then $E\big[|\mathcal{C}(1)| \mathbb{1}_{|\mathcal{C}(1)| \geq k}\big] \leq C k_n n^{-aI_\lambda}$.

<u>Proof</u>: For any $\mathbb{N}$-valued r.v. (<u>Exercise</u>)

$$E[X \mathbb{1}_{X \geq k}] = k \cdot P[X \geq k] + \sum_{\ell=k+1}^{\infty} P[X \geq \ell]$$

Hence,

$$E\big[|\mathcal{C}(1)| \mathbb{1}_{|\mathcal{C}(1)| \geq k_n}\big] \overset{\leq}{=} k_n P[|\mathcal{C}(1)| \geq k_n] + \sum_{\ell \geq k_n} P[|\mathcal{C}(1)| \geq \ell]$$

$$\underset{\substack{\text{as on page 21} \\ \text{below (3.18)}}}{\overset{(3.8),(2.47)(2.43)}{\leq}} k_n \cdot e^{-k_n I_\lambda} + \sum_{\ell \geq k_n} e^{-\ell I_\lambda} \leq C n^{-aI_\lambda} \quad \square.$$

We may now conclude the proof of (3.15). Combining (3.20) with Lemmas (3.21), (3.24), (3.28), for $k_n = a\log n$, $a < \mu$

$$\mathbb{P}\left[\,|C_{max}| \le k_n\right] \le \frac{C n \cdot k_n\, n^{-aI_\lambda}}{n^{2(1-I_\lambda a)(1+o(1))}} = n^{-(1-I_\lambda a)(1+o(1))} \le n^{-\delta(\lambda, a)}$$

This completes the proof. $\square$

Supercritical regime of ER graph.

We show that if $\lambda > 1$, then ER graph has a "giant" component. We use ζ_λ to denote the survival probability of Poisson(λ) branching process

(3.29) $$\zeta_\lambda = 1 - \eta_\lambda = \mathbb{P}_\lambda^{BP}(T = \infty)$$

(3.30) Theorem: (LLN for giant component). When $p = \frac{\lambda}{n}$, $\lambda > 1$, then for every $\varepsilon > 0$

$$\mathbb{P}^{ER}_{n,p}\left(\left| \frac{|C_{max}|}{n} - \zeta_\lambda \right| > \varepsilon \right) \xrightarrow{n\to\infty} 0$$

(3.31) Remark: There is a CLT corresponding to this LLN

$$\frac{1}{\sqrt{n}}\left(|C_{max}| - n\zeta_\lambda\right) \xrightarrow[n\to\infty]{\mathbb{P}^{ER}_{n,p}} \mathcal{N}(0, \sigma_\lambda^2)$$

See [vdH] Section 4.5.

Proof: The proof has four steps:

Step 1: We show first that considerable proportion of vertices is in "rather large" components, on average.

(3.32) Lemma: ($\lambda > 1$). Let $a > \frac{1}{I_\lambda}$ and $k_n \ge a \log n$, $k_n \ll n$. Then

$$E^{ER}_{n,p}\left[Z_{\ge k_n} \right] = n\zeta_\lambda + O(k_n) \quad \text{as } n \to \infty.$$

Proof: As $E[Z_{\ge k_n}] = n \cdot \mathbb{P}[|C(1)| \ge k_n]$ we need to show

(3.33) $$\mathbb{P}(|C(1)| \ge k_n) = \zeta_\lambda + O\left(\frac{k_n}{n}\right)$$