Lecture 15: Random walk representation of a branching process

Last lecture we related the size of the component containing a vertex v in the Erdős-Rényi random graph to the total population of a Poisson BP. Here we study the distribution of the latter quantity.

Random walk representation of BP: Recall the BP recursion (12.1) from Lecture 12:

$$A_t = A_{t-1} + \mathbb{I}[A_{t-1} > 0](-1 + Z_t), \tag{15.1}$$

where Z_t are i.i.d. with the offspring distribution, and recall that

$$\tau_0 := \inf\{t > 0 : A_t = 0\}$$

is the total size of the population. Thus, if $(Y_i)_{i\geq 1}$ are i.i.d. with distribution $\mathscr{L}(Z_1 - 1)$ and we define the *random walk* $(S_j)_{j\geq 0}$ by

$$S_j = \sum_{i=1}^{j} Y_i,$$

and $T_0 := \inf\{j \geq 1 : S_j = -1\}$, then

$$\mathscr{L}\big((1 + S_j)_{0\leq j\leq T_0}\big) = \mathscr{L}\big((A_t)_{0\leq t\leq \tau_0}\big).$$

This is because up to time τ_0, $1 + S_j$ satisfies the same recursion (15.1). (If you are having trouble understanding this, a small example will clarify.) By Lemma 12.4 of Lecture 12, τ_0 equals the total population size of the BP, which we denote by W. Thus we have shown the following lemma.

Lemma 15.1. *If W is the total population size of a BP with offspring distribution $\mathscr{L}(Z)$, and T_0 is the first time the random walk started from one with increments $\mathscr{L}(Z-1)$ reaches 0, then*

$$\mathscr{L}(W) = \mathscr{L}(T_0).$$

Note that it is possible that W or T_0 is infinity. The next lemma gives a useful expression for $\mathscr{L}(T_0)$, and thus $\mathscr{L}(W)$.

Lemma 15.2. *Let $(1 + S_j)_{j\geq 0}$ be a random walk started from one, with increment supported on $\{-1, 0, 1, \ldots\}$, and $T_0 = \inf\{j \geq 1 : 1 + S_j = 0\} = \inf\{j \geq 1 : S_j = -1\}$. Then*

$$\mathbb{P}(T_0 = k) = \frac{1}{k}\mathbb{P}(S_k = -1).$$

The lemma is useful (and remarkable) because the event on the left hand side depends on the entire path $(S_0, \ldots, S_k)$, whereas the event of the right hand side only depends on the final value S_k. Before the proof, we state an easy but key consequence of the last two lemmas.

Theorem 15.3. *If W is the total population size of a BP with offspring distribution $\mathscr{L}(Z)$, then for $(Z_j)_{j=1}^k$ i.i.d. with distribution $\mathscr{L}(Z)$,*

$$\mathbb{P}(W = k) = \frac{1}{k}\mathbb{P}\left(\sum_{j=1}^{k} Z_j = k - 1\right).$$

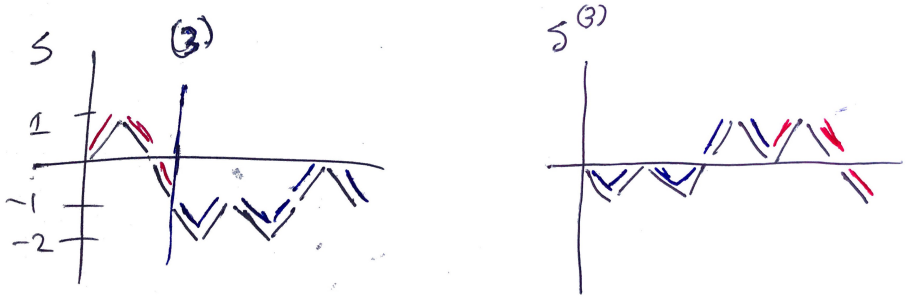
Proof of Lemma 15.2. The first part of the proof is the following *cycle lemma*. Let $(y_1, \ldots, y_k)$ be numbers in $\{-1, 0, 1, \ldots\}$, with $\sum_{j=1}^k y_j = -1$. For the "walk"

$$(s_i)_{i=0}^k := \big(\sum_{j=1}^{i} y_j\big)_{i=0}^k$$

and $\ell \in \{1, \dots, k\}$, we define the ℓ-cyclic rearrangement $(s_i)_{i=0}^k$ by

$$\big(s_i^{(\ell)}\big)_{i=0}^k = \Big(\sum_{j=1}^{i} y_{\ell+j}\Big)_{i=0}^k,$$

where the indices on the right hand side are mod k. The graph of the walk $(s_i^{(\ell)})_{i=0}^k$, is obtained from that of $(s_i)_{i=0}^k$ by moving the origin to (ℓ, s_ℓ), and attaching the fragment of the path before step ℓ to $(k, s_k) = (k, -1)$. The following picture illustrates this cyclic rearrangement.

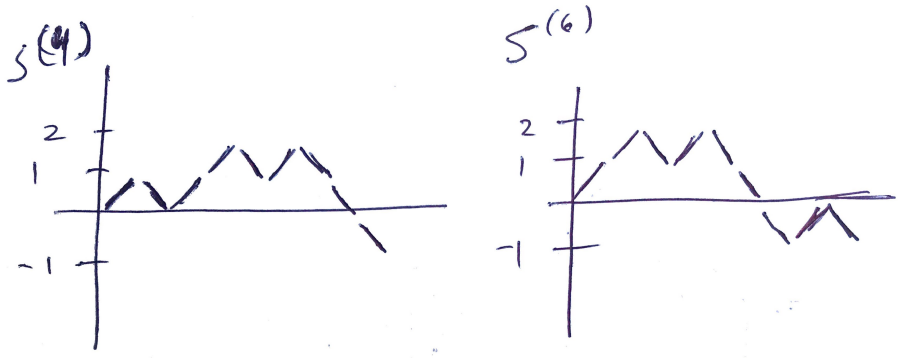


Graphs of $(s_i)_{i=0}^k$ and $(s_i^{(3)})_{i=0}^k$ for $(y_1, \dots, y_9) = (1, -1, -1, -1, 1, -1, 1, 1, -1)$ and $k = 9$. The 3-cyclic rearrangement is obtained by moving the red path to the end of the blue path, and adjusting the origin.

The key fact is that there is exactly one $\ell \in \{1, \dots, k\}$ such that the walk $\big(s_i^{(\ell)}\big)_{i=0}^k$ hits -1 for the first time at step k, and ℓ is the index of the first time of the minimum of the walk:

$$s_i^{(\ell)} > -1 \text{ all } i = 0, \dots, k-1 \iff \ell = \min\{i : s_i = \min_j s_j\}. \tag{15.2}$$

The fact follows (roughly) since such a path clearly has to start at a minimum, and if it starts at a minimum that is not the first, then since the last step is down, the shift will end at -1, which is the same value reached by the first minimum (here is where it is crucial that all down steps are of size 1). The following picture illustrates.



Two cyclic rearrangements starting from the minimums of the path in the previous picture. The path $s^{(4)}$ is the unique path reaching -1 for the first time at the final step of the walk, since step 4 is the first occurrence of the minimum of the walk s.

From here, the result follows by considering $\big(S_i^{(U_k)}\big)_{i=0}^k$, where $U_k \sim \mathrm{U}\big(\{1, \dots, k\}\big)$ is independent of the walk. Since the increments are i.i.d. and U_k is independent of them, we have

$$\big(S_i^{(U_k)}\big)_{i=0}^k \stackrel{d}{=} \big(S_i\big)_{i=0}^k.$$

Combining this with the fact that $S_k^{(U_k)} = S_k$, we have

$$\begin{aligned}\mathbb{P}\big(\cap_{j=1}^{k-1}\{S_j > -1\}, S_k = -1\big) &= \mathbb{P}\big(\cap_{j=1}^{k-1}\{S_j^{(U_k)} > -1\}, S_k^{(U_k)} = -1\big) \\ &= \mathbb{P}\big(\cap_{j=1}^{k-1}\{S_j^{(U_k)} > -1\} | S_k = -1\big)\mathbb{P}(S_k = -1) \\ &= \frac{1}{k}\mathbb{P}(S_k = -1),\end{aligned}$$

where the last equality is from the cycle lemma (15.2), since we need U_k to equal the unique index that keeps the path above -1 until the end. □

For the Erdős-Rényi exploration process, Lemma 13.1 says that the component size can be related to a BP with Poisson offspring distribution. For this case, Theorem 15.3 gives the following result. The simple proof is omitted.

Lemma 15.4. *If W_λ is the total size of a BP with Poisson mean λ offspring distribution and, for $k \geq 1$, $Y_{\lambda k} \sim \mathrm{Po}(\lambda k)$, then*

$$\mathbb{P}(W_\lambda = k) = \frac{1}{k}\mathbb{P}(Y_{\lambda k} = k-1) = e^{-\lambda k}\frac{(\lambda k)^{k-1}}{k!}.$$

Recall from the high-level overview in Lecture 11 that we are trying to show that with high probability, the size of the largest component $C_{\max}$, is $\Theta(\log(n))$ for $\lambda < 1$ and $\Theta(n)$ for $\lambda > 1$, and we will do this by analysing C_v, the size of a component containing the vertex v; for example, a union bound implies

$$\mathbb{P}(C_{\max} \geq k) \leq \sum_{v \in \{1,\ldots,n\}} \mathbb{P}(C_v \geq k).$$

For this, Lemma 13.1 and Lemma 15.4 imply that we need to understand

$$\mathbb{P}(W_\lambda \geq k) = \sum_{j \geq k} \frac{1}{j}\mathbb{P}(Y_{\lambda j} = j-1).$$

In particular, we need to understand the point probabilities $\mathbb{P}(Y_{\lambda j} = j-1)$ for j large, which is the topic of the next lecture.