Random Sums and Graphs

Kent Quanrud

April 29, 2021

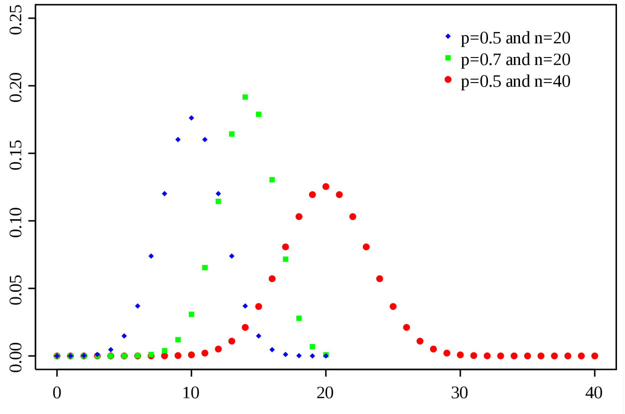
1 Random sums

If, out of 100 coin tosses, you were told that 50 of them were heads, would you be surprised? Actually, you should be a little surprised. The odds of getting exactly 50 heads is about 8%. But if you were told that the number was in the range, say, 45 to 55, you probably wouldn't think much of it.

If you were told that all 100 coin tosses came up heads, you wouldn't believe it. The odds of that, we know, is $1/2^{100}$. If you bet money and lost on this event, you would be outraged (and, at even odds, certainly broke for the rest of eternity).

Suppose you were told that at most 25 coin tosses came up heads. Should you be surprised? On one hand, 25 is half of the expected amount. On the other hand, the claim is not that there was exactly 25 heads, but *at most* 25 heads. There could be 25, 24, 23, etc., down to 0. Even though the event of getting any one of these counts should be low, being far from average, there are also 26 of these events. Do the probabilities add up to very much? It turns out that the probability of getting 25 or fewer heads is tiny: about 2.818×10^{-7}.

The *scale* is very important in this discussion. If, out of 10 coin tosses, you got 4 or fewer heads, you shouldn't be too surprised. There is a (roughly) a 37.7% chance of getting at most 4 heads. But if at most 40% of 1000 coin tosses came up heads, you should be *very* surprised. The odds of this occurring is occurring is roughly 1.364×10^{-10}.



We generalize the discussion to coins with any fixed probability of heads, $p \in [0,1]$. The **binomial distribution**, denoted $\mathcal{B}(n,p)$, is the distribution of the number of heads over n independent coin tosses that each flip heads with probability p. The probabilities of different binomial distributions is plotted above. (See also [8].)

We write $B \sim \mathcal{B}(n,p)$ to denote a random variable $B \in \{0,\dots,n\}$ drawn from the binomial distribution $\mathcal{B}(n,p)$. The expected value of B is $\mathbf{E}[B] = pn$. The following lemma bounds the probability of B being a multiplicative factor smaller than its mean, pn. Note that the probability decays *exponential fast in the mean*.

Lemma 1.1. *Let $B \sim \mathcal{B}(n,p)$ and $\epsilon \in (0,1)$. Then*

$$\mathbf{P}[B \leq (1-\epsilon)pn] \leq e^{-\epsilon^2 pn/2}.$$

Proof. We have

$$\mathbf{P}[B \leq (1-\epsilon)pn] = \mathbf{P}\left[e^{-\epsilon B} \geq e^{-\epsilon(1-\epsilon)pn}\right] \overset{\text{(a)}}{\leq} \frac{\mathbf{E}\left[e^{-\epsilon B}\right]}{e^{-\epsilon(1-\epsilon)pn}} = e^{\epsilon(1-\epsilon)pn}\,\mathbf{E}\left[e^{-\epsilon B}\right] \tag{1}$$

by (a) Markov's inequality. It remains to analyze $\mathbf{E}\left[e^{\epsilon B}\right]$. Write B as the sum $B = X_1 + \cdots + X_n$, where each X_i is an independent $\{0,1\}$-random variable with $\mathbf{P}[X_i = 1] = p$. Now we have

$$\mathbf{E}\left[e^{-\epsilon B}\right] = \mathbf{E}\left[e^{-\epsilon X_1 - \epsilon x_2 - \cdots - \epsilon X_n}\right] \overset{\text{(b)}}{=} \mathbf{E}\left[e^{-\epsilon X_1}\right]\mathbf{E}\left[e^{-\epsilon X_2}\right]\cdots\mathbf{E}\left[e^{-\epsilon X_n}\right],$$

where (b) is by independent of the X_i's. For each X_i, we have

$$\begin{aligned} \mathbf{E}\left[e^{-\epsilon X_i}\right] &= pe^{-\epsilon} + (1-p) = 1 + p(e^{-\epsilon-1}) \\ &\overset{\text{(c)}}{\leq} p\left(1 - \epsilon + \epsilon^2/2\right) + (1-p) = 1 - (\epsilon - \epsilon^2/2)p \\ &\overset{\text{(d)}}{\leq} e^{-(\epsilon-\epsilon^2/2)p}. \end{aligned}$$

Here (c) uses the inequality $e^{-x} \leq 1 - x + x^2/2$ for all $x > 0$[1]. (d) is by the inequality $1 + x \leq e^x$ for all x. Thus,

$$\mathbf{E}\left[e^{-\epsilon B}\right] = \mathbf{E}\left[e^{-\epsilon X_1}\right]\mathbf{E}\left[e^{-\epsilon X_2}\right]\cdots\mathbf{E}\left[e^{-\epsilon X_n}\right] \leq e^{-(\epsilon-\epsilon^2/2)pn},. \tag{2}$$

Putting everything together, we have

$$\mathbf{P}[B \leq (1-\epsilon)pn] \overset{\text{(e)}}{\leq} e^{\epsilon(1-\epsilon)pn}\,\mathbf{E}\left[e^{-\epsilon B}\right] \overset{\text{(f)}}{\leq} e^{\epsilon(1-\epsilon)pn - (\epsilon-\epsilon^2/2)pn} = e^{-\epsilon^2 pn/2},$$

by (e) inequality (1) and (f) inequality (2), as desired. ■

One can prove a similar inequality bounding the probability that B exceeds its mean by a multiplicative factor. The proof is similar to Lemma 1.1 and left as Exercise 1.

Lemma 1.2. *Let $B \sim \mathcal{B}(n,p)$ and $\epsilon \in (0,1)$. Then*

$$\mathbf{P}[B \geq (1+\epsilon)pn] \leq e^{-\epsilon^2 pn/3}.$$

[1] To see that

$$1 - x + x^2/2 \geq e^{-x}$$

for all $x \geq 0$, observe first that both sides equal 1 at $x = 0$. The derivative of the LHS, $-(1-x)$, is always at least the RHS of the derivative of the RHS, $-e^{-x}$, by the inequality $1 + y \leq e^y$ for all y.

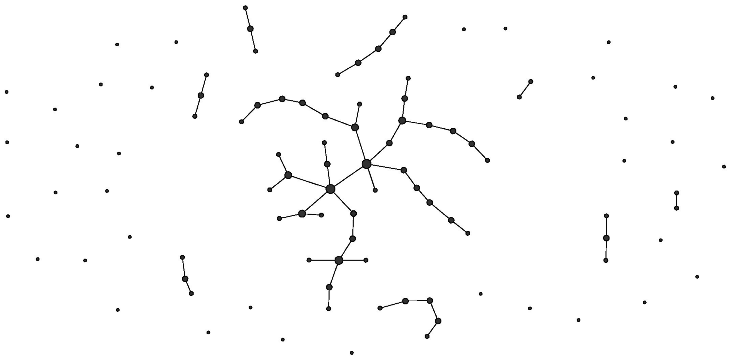


Figure 1: A random graph sampled from $\mathcal{G}(n, .01)$ [7].

2 Random graphs

Paul Erdös, inspired by Ramsey [5] before him, had a series of work analyzing *random graphs*, producing a large body of results that can mostly be grouped into two broad categories. First, he designed elebarate randomized constructions of graphs and showed that with nonzero probability, they can possess certain counterintuitive, seemingly impossible properties. This general approach is now called *Ramsey theory*. Second, he showed that for natural random graph models, these graphs – however random – tend to be extremely consistent about certain properties. Today we will study the $\mathcal{G}(n, p)$ random graph, sometimes called Erdös-Rényi graphs based on work by Erdös and Rényi [2, 3]. A random graph from $\mathcal{G}(n, p)$ is an undirected graph over n vertices, where every edge is sampled independently with probability p. By now there is a large catalog of nontrivial and useful properties that, depending on p, are almost certain to appear or not appear in such a graph (for sufficiently large n). Moreover, Erdös and Rényi showed that these properties can vary dramatically with very small changes in p. Consider the following theorem.

Theorem 2.1. *Consider a random graph $G \sim \mathcal{G}(n, p)$ for $p = c/n$, where c is a constant.*

1. *If $c > 1$, then with high probability, there is exactly one connected component of G with $\Omega(n)$ vertices, and all other components have size $\leq O\big(\log n\big)$.*
2. *For $c < 1$, then with high probability, all connected components of G has size $< O\big(\log n\big)$.*

The parameter c above models the average degree (in expectation). The drama lies in the fact that a tiny change in the average degree c – from .999 to 1.0001 – flips the qualitative nature of a typical random graph from one of many tiny components to essentially one giant component. This is an example of a *threshold phenomena*; alternatively, a *nonlinear dynamic*. Such phenomena is not rare: it occurs in many situations in physics, as well as in models for epidimiology and social networks. Let us briefly mention - without claiming to be very precise - that the sensitivity to c gives some motivation for controlling the "reproductive number" when analyzing and preventing the spread of infectious diseases. The reproductive number is the expected number of healthy individuals that a sick individual effects.

We note that a line of research has obtained a much more refined and detailed understanding than stated in Theorem 2.1. We refer the reader to [1, Chapter 7] for further details and other results in this area.

2.1 Overview of the proof

We will prove part 1 of Theorem 2.1 in roughly three parts.

Part 1: the gap theorem. Observe that in Theorem 2.1 above, regardless of the value of p, there are simply no "medium"-size components, like a component of size $\sqrt{n}$ or of size $n/\log(n)$. The intermediate sizes are ruled out by the following "gap theorem".

Lemma 2.2. *There is a universal constant $C > 0$, such that for all $\epsilon \in (0,1)$, and for all $n > 0$ sufficiently large, and $p = (1+\epsilon)/n$, we have the following. For a random graph $\mathcal{G}(n,p)$, with probability of error $\leq 1/n^2$, no component has k vertices for any value k in the interval*

$$\frac{C\log(n)}{\epsilon^2} \leq k \leq \frac{\epsilon n}{C}.$$

We analyze Lemma 2.2 theorem in Section 3. The proof makes a surprising connection to our discussion on random sums in Section 1.

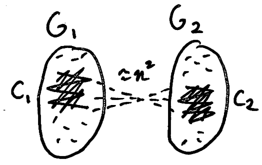
Part 2: Existence of a large component. Lemma 2.2 establishes that all components are either very small or very big. However it does not assert that there are any big components. The next theorem, proven in Section 4 and based on analyzing a *Galton-Watson branching process*, shows that any single vertex has a reasonable chance of being in a component that is not small.

Lemma 2.3. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \leq h \leq \epsilon n$, with probability at least $1/h$, v has at least $1+h$ vertices in its connected component.*

Lemma 2.3 implies that there is almost certainly at least one giant component as follows. Let $h = c\log(n)/\epsilon^2$ for a sufficiently large constant c, and let $q = 1/h = \Omega\big(\epsilon^2/\log(n)\big)$. Call a component "small" if it has at most h vertices. We want to argue that, for $p > 1/n$, there is at least one component that is not small. In conjunction with the gap theorem (Lemma 2.2), which rules out all intermediate sizes, this implies that there is at least one giant component of size $\Omega\big(\epsilon^2 n\big)$.

By Lemma 2.3, any vertex v has at least a probability q of not being in a small component. Now imagine a process where we first randomly select a vertex v and inspect its component. If it is not small, then we have obtained the non-small component we seek. Otherwise, if the component is small, then we throw out v and its component, and randomly select another vertex as v, and repeat. Each vertex we inspect has probability q of not being in a small component. We would have to fail on the order of n/h consecutive samples to conclude there is no small component - which happens with diminishingly small probability. Thus with very high probability, there is at least one component that is not small.

Part 3: Uniqueness of the giant component. Can there be two giant components? The answer is no (with high probability) and here is a quick explanation. Instead of sampling from $\mathcal{G}(n, p)$ directly, we can first sample two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ from $G(n/2, p)$. In the second stage we can sample each cross-edge (v_1, v_2), where $v_1 \in V_1$ and $v_2 \in V_2$, independently with probability p. Now, by applying the theory we have already developed to G_1 and G_2, G_1 and G_2 will have some giant components, each of size $\Omega(\epsilon^2 n)$. Note that each graph can only have $O(1/\epsilon^2)$ of them.



Let C_1 be a giant component in G_1 and let C_2 be a giant component in G_2. We can sample up to $|C_1||C_2| \geq \Omega(\epsilon^4 n^2)$ edges between C_1 and C_2. Recalling that p is greater than $1/n$, the odds that all $\Omega(\epsilon^4 n^2)$ edges fail to be sampled is vanishingly small. That is, we almost certainly connect C_1 and C_2. Since there is a limited number of giant components we will almost certainly connect all of them together. Thus, for $p > (1 + \epsilon)/n$ for $\epsilon > 0$, we get a *unique* giant component. This establishes Theorem 2.1 for $c > 1$.

$\boldsymbol{c < 1.}$ The proof for $c < 1$ is simpler and only requires the ideas underlying Lemma 2.2. See Section 3.1.

2.2 Directed graphs

One could naturally ask the same questions for directed graphs. Let $D(n, p)$ denote the distribution over *directed graphs* where every directed edge appears independently with probability p. We might similarly ask for the maximum number of vertices reachable from any component, or the size of the maximum strongly component.

It turns out that the analysis of directed graphs can be largely reduced to undirected graphs, as shown by Karp [4] in the following delightfully simple way.

Theorem 2.4. *Let $G \sim G(n, p)$ and $D \sim G(n, p)$, and fix a vertex v. Then the size of the connected component of v in G, and the number of vertices reachable from v in D, are identically distributed.*

Proof. Let us introduce a second distribution of directed random graphs. Let $B(n, p)$ be the distribution of directed graphs where we sample each *undirected* edge $\{u, v\}$ independently with probability p, and for each sampled edge, add both directions (u, v) and (v, u) to the graph. Clearly for a fixed vertex v, the size of the v's (undirected) component in $G(n, p)$ is distributed identically to the number of vertices reachable from v in $B(n, p)$. We claim that the number of vertices reachable from v in $B(n, p)$ is identically distributed as in $D(n, p)$. At this point let us simply quote Karp [4, Lemma 1] (with minor changes in notation) whose proof is very elegant.

> ...To see that the last two random variables are identically distributed, note that the probability spaces $B(n, p)$ and $D(n, p)$ differ in only one respect: a digraph

> G drawn from $B(n,p)$, are (u,v) is present if and only if arc (v,u) is present, while, in a digraph D drawn from $D(n,p)$, then event that (v,u) is present is independent of the event that (u,v) is present. Thus no experiment based on checking for the presence or absence of arcs can distinguish between the two probability spaces unless it checks both an arc and its reversal. But any standard sequential algorithm, such as breadth-first search or depth-first search, for building a search tree containing exactly the vertices reachable from vertex 1, checks for the presence of arc (u,v) only if vertex u is in the search tree and v is not; thus it never checks both an arc and its reversal, and accordingly cannot distinguish $B(n,p)$ from $D(n,p)$.

To summarize the excerpt, standard search algorithms for reachability do not distinguish $B(n,p)$ and $D(n,p)$ anyway, so the number of reachable vertices is identically distributed. ■

3 A gap in component size

In this section we prove Lemma 2.2, which asserts that when $p = (1+\epsilon)/n$ for a constant $\epsilon > 0$, then with high probability, all components are either very small or very large. Our analysis follows an approach due to Karp [4]. His proof is also described in [1]. We will also reuse some of the ideas in the proof to analyze the $p < 1/n$ in Section 3.1. We first restate Lemma 2.2 for the reader's convenience.

Lemma 2.2. *There is a universal constant $C > 0$, such that for all $\epsilon \in (0,1)$, and for all $n > 0$ sufficiently large, and $p = (1+\epsilon)/n$, we have the following. For a random graph $\mathcal{G}(n,p)$, with probability of error $\leq 1/n^2$, no component has k vertices for any value k in the interval*

$$\frac{C\log(n)}{\epsilon^2} \leq k \leq \frac{\epsilon n}{C}.$$

For a vertex $v \in V$, let $C(v) \subset V$ be the (randomized) component of v. To analyze $C(v)$, we imagine revealing $C(v)$ by a search algorithm. We maintain a collection of vertices known to be connected to v; initially just $\{v\}$. Each iteration i, starting from v, select a vertex v_i that is known to be in $C(v)$, but has not been explored. Then "explore" v by inspecting all of the edges incident to v_i, possibly adding to the collection of vertices known to be connected to v (but not yet explored).

We annotate this process as follows. For $i \in \mathbb{N}$, let

- v_i be the vertex that is explored in the ith iteration (or `nil` if all of $C(v)$ has already been explored).

For each $i \in \mathbb{Z}_{\geq 0}$, let

- A_i be the set of vertices known to be in $C(v)$ after i iterations, and let
- B_i be the set of vertices that have been explored.

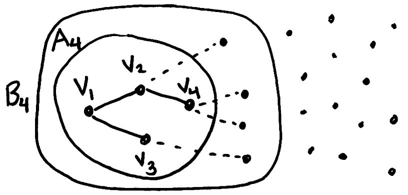
For the sake of concreteness, one can imagine processing the v_i's in BFS order. Recall that BFS marks each vertex when the vertex first encountered, and if the vertex was unmarked, it is added to a queue. The next vertex visited is drawn from the queue. In terms of BFS, then, A_i is the set of vertices marked after i iterations, and B_i is the set of vertices that have left the queue and been fully processed.

Ultimately, B_i, A_i, v_i are built up incrementally as follows.

1. Initially, we have $B_0 = \emptyset$ and $A_0 = \{v\}$.

2. In the first iteration, set $v_1 = v$, set $B_1 = \{v_1\}$, and set $A_1 = A_0 \cup N(v_1)$, where $N(v_1)$ is the (randomized) neighborhood of v_1.

3. In the ith iteration, is $B_{i-1} \neq A_{i-1}$, then select (any) $v_i \in A_{i-1} \setminus B_{i-1}$. Set $B_i = B_{i-1} \cup \{v_i\}$ and $A_i = A_{i-1} \cup N(v_i)$. Otherwise we terminate with $C(v) = B_{i-1} = A_{i-1}$.



The process terminates when $B_i = A_i$. But since $B_i \subseteq A_i$ and $|B_i| = i$, this is precisely when $|A_i| = i$. As long as $|A_i| \neq i$, A_{i+1} is generated by taking the union of A_i and a random sample of $V - A_i$ where each vertex is included with probability p. Thus we could have generated the sequence of A_i's instead by the following equivalent process, which omits any mention of B_i or v_i.

1. Initially set $A_0 = \{v\}$.

2. For each $i \in \mathbb{N}$, let S sample each vertex in $V \setminus A_{i-1}$ independently with probability p and set $A_i = A_{i-1} \cup S$.

3. Let i be the first index such that $|A_i| = i$, and return $C(v) = A_i$.

Fix an iteration i. The alternative (but equivalent) process described above exposes a simple distribution for A_i. For any vertex $x \neq v$, we have $x \notin A_i$ iff x failed to be added in each of the first i rounds, which occurs with probability exactly $(1-p)^i$. Moreover this event is independent across vertices. Thus $|A_i|$ is distributed exactly as the binomial distribution with $n-1$ coins and probability $1-(1-p)^i$; i.e., $|A_i| \sim \mathcal{B}\big(n-1, 1-(1-p)^i\big)$.

Lemma 3.1. *Let* $i \leq \epsilon n/2((1+\epsilon))$. *Then* $\mathbf{E}[|A_i|] \geq (1+\epsilon/2)i$.

Proof. We have

$$(1-p)^i \leq e^{-ip} \leq 1 - ip + (ip)^2 \leq 1 - ip + \epsilon ip/2 = 1 - (1-\epsilon/4)ip.$$

where (a) is because $ip = (1+\epsilon)i/n \leq \epsilon/4$. Thus

$$\mathbf{E}[|A_i|] = 1 + \Big(1 - (1-p)^i\Big)(n-1) \geq (1-\epsilon/4)ipn \geq (1+\epsilon/2)i.$$

■

Lemma 3.2. *Let $i \leq \epsilon n/2(1+\epsilon)$. Then* $\mathbf{P}[|A_i| \leq i] \leq e^{-\epsilon^2 i/8}$.

Proof. We have

$$\mathbf{P}[|A_i| \leq i] \overset{\text{(a)}}{\leq} \mathbf{P}[|A_i| \leq (1-\epsilon/2)\,\mathbf{E}[|A_i|]] \overset{\text{(b)}}{\leq} e^{-\epsilon^2/8}.$$

Here (a) is by Lemma 3.1. (b) is by the tail inequality on binomial distributions, Lemma 1.1. ■

Let $I = \left\{i \in \mathbb{N} : 32\ln(n)/\epsilon^2 \leq i \leq \epsilon n/2((1+\epsilon))\right\}$. For all $i \in I$, we have

$$\mathbf{P}[|A_i| \leq i] \leq 1/n^4.$$

By the union bound, we have

$$\mathbf{P}[|A_i| > i \text{ for all } i \in I] \geq 1 - \sum_{i \in I} \mathbf{P}[|A_i| \leq i] \geq 1 - 1/n^3.$$

Thus with probability $\geq 1 - 1/n$, the number of vertices in the connected component of v, $|C(v)|$, does not lie in the range I. Taking the union bound over all $v \in V$ establishes part 1 of Lemma 2.2.

3.1 Probabilities < 1

Suppose instead that $p = (1-\epsilon)/n$. Then we have $\mathbf{E}[A_i] \leq (1-\epsilon/2)i$ unless i is very close to n. In particular, for $i = O\big(\log(n)/\epsilon^2\big)$, the probability of $A_i > i$ is $1/\text{poly}(n)$. Thus we see that all components will have size $O\big(\log(n)/\epsilon^2\big)$ with high probabilities, establishing part 2 of Lemma 2.2.

4 Galton-Watson branching processes

We now move onto the second part of the analysis. By now we have established that there are (with high probability) no "medium" components – all component sizes have either at most $O\big(\log(n)/\epsilon^2\big)$ vertices, or at least $\Omega\big(\epsilon^2 n\big)$ vertices. Now we want to prove Lemma 2.3, which we first restate for the reader's convenience.

Lemma 2.3. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \leq h \leq \epsilon n$, with probability at least $1/h$, v has at least $1+h$ vertices in its connected component.*

The proof is by relation to the so-called *Galton-Watson process* that arises in the study of reproducing populations. In the simplest case, imagine a population of size 1. Each generation, each member of the current generation flips k coins, each heads with probability $1/k$. For each heads, we generate another member of the next generation. The probabilities and number of coins are configured so that each member expects to have one child.

What is the probability that the population survives for h iterations, for a given parameter h? This is answered by the following.

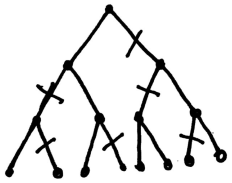


Figure 2: A complete binary tree of height 3, where each edge was deleted with probability 1/2.

Theorem 4.1. *Let T be a complete k-ary tree of height h, and suppose every edge is deleted independently with probability at most $1 - 1/k$. Then the probability that there is a leaf connected to the root is $\geq 1/h$ for $h \geq 3$, and $\geq (1 - e^{-1})^h$ for $h \leq 2$.*

An example of the case $k = 2$ is drawn in Figure 2.

Proof. For $i \in \mathbb{N}$, let p_i be the probability that a particular node at height i is connected to a subleaf. We have $p_0 = 1$. For a node at height $i + 1$, the probability that there is no path to a leaf via a particular child is

$$1 - \frac{1}{k} + \frac{1}{k}(1 - p_i) = 1 - \frac{p_i}{k}.$$

By independence, we have

$$p_{i+1} = 1 - \left(1 - \frac{p_i}{k}\right)^k.$$

Observe that the RHS is increasing in p_i; thus to lower bound p_{i+1}, we can substitute any lower bound for p_i. We have

$$\begin{aligned}
p_0 &= 1, \\
p_1 &= 1 - (1 - 1/k)^k \geq 1 - e^{-1} \geq .63, \\
p_2 &= 1 - (1 - .63/k)^k \geq 1 - e^{-.63} \geq .467, \\
p_3 &\geq 1 - (1 - .467/k)^k \geq 1 - e^{-.467} \geq .373 \geq 1/3.
\end{aligned}$$

We claim by induction on i that $p_i \geq 1/i$ for all $i \geq 3$. The base case $i = 3$ was just proven. For the general case,

$$p_{i+1} \overset{(a)}{\geq} 1 - (1 - 1/ik)^k \geq 1 - e^{-1/i} \overset{(b)}{\geq} \frac{1}{i} - \frac{1}{2i^2} \geq \frac{1}{i+1}$$

Here (a) is by induction. (b) applies the inequality $e^x \leq 1 + x + \frac{1}{2}x^2$ for $x \leq 0$. ■

4.1 Likelihood of small components

We can use the above branching process to analyze the probability that a given vertex v is in a component of size $\leq h$, for any $h \leq \epsilon n/(1 + \epsilon)$. Recall the sets $B_0, B_1, B_2, \ldots$ from Section

3. Given that $|B_i| \leq h$, we can think of $B_{i+1} - B_i$ as adding (at least) $n/(1+\epsilon)$ children each with probability $p = (1+\epsilon)/n$. Either we find new elements for all h rounds - which forces $|B_h| \geq h$ - or we hit $|B_i| = h$ at some point $i < h$. Thus the odds of v acquiring h vertices in its connected component is at least the odds produced by Theorem 4.1 for this value of h; namely, $1/h$. This gives us Lemma 2.3.

4.2 A more general analysis

One can consider a more general model as follows. Let $X \in \mathbb{Z}_{\geq 0}$ be a random variable taking nonnegative integer values. For each node at a generation i, we sample an independent copy of X, and generate this many children in the next generation. Here we have the following remarkably precise theorem.

Theorem 4.2. *Let n_i denote the number of children in the ith generation in the process described above.*

1. *If $\mathbf{E}[X] < 1$, then $\lim_{i\to\infty} \mathbf{P}[n_i = 0] = 1$.*
2. *If $\mathbf{E}[X] > 1$, then $\lim_{i\to\infty} \mathbf{P}[n_i = 0] = q$, where q is the unique solution to $x = \sum_{i\geq 0} \mathbf{P}[X = 0]x^i$.*

See the lecture notes by Sinclair [6] for a proof.

5 Exercises

Exercise 1. Prove Lemma 1.2. (Here the important part is not the constant, 1/3 – any constant $c > 0$ is already interesting.)

References

[1] Béla Bollobás. *Modern Graph Theory*. 1st ed. Graduate Texts in Mathematics 184. Springer-Verlag New York, 1998.

[2] Paul Erdös and Alfred Rényi. “On Random Graphs I”. In: *Publicationes Mathematicae Debrecen* 6 (1959), p. 290.

[3] Paul Erdös and Alfred Rényi. “On the evolution of random graphs”. In: *Publ. Math. Inst. Hungary. Acad. Sci.* 5 (1960), pp. 17–61.

[4] Richard M. Karp. “The Transitive Closure of a Random Digraph”. In: *Random Struct. Algorithms* 1.1 (1990), pp. 73–94.

[5] Frank Plumpton Ramsey. “On a Problem of Formal Logic”. In: *Proceedings of the London Mathematical Society* s2-30.1 (Jan. 1930), pp. 264–286.

[6] Alistair Sinclair. “Lecture 16 of CS271: Randomness & Computation”. Available at `https://people.eecs.berkeley.edu/~sinclair/cs271/n16.pdf`. 2020.

[7] Vonfrisch. *Erdos generated network-p0.01.jpg*. URL: `https://commons.wikimedia.org/wiki/File:Erdos_generated_network-p0.01.jpg`.

[8] Wikipedia contributors. *Binomial distribution — Wikipedia, The Free Encyclopedia*. URL: https://en.wikipedia.org/wiki/Binomial_distribution.