Topics in Theoretical Computer Science: Randomized Algorithms
EPFL, Fall semester, 2025

Homework 1 Solutions

Barras Simon <simon.barras@epfl.ch>
Doll Zachary <zachary.doll@epfl.ch>

---

1 Problem 1 (Lonely Vertices)

Consider the Erdős-Rényi random graph $G(n, p)$ and suppose $\epsilon > 0$ is some constant. Show that:

(a) if $p \geq \frac{(1+\epsilon)\ln n}{n-1}$ then the graph has at no isolated vertices with probability $1 - o(1)$.

(b) if $p \leq \frac{(1-\epsilon)\ln n}{n-1}$ then the graph has at least one isolated vertex with probability $1 - o(1)$.

1.1 Solution

This problem illustrates an application of the first, and later, the second moment method.

1.1.1 Part A

Proof. The probability that a graph build with Erdős-Rényi $G = (n, p)$ contains no isolated vertex can be computed with the first-moment method.

The first step is to define a non-negative integer random variable X that counts the number of isolated vertices (the "bad" events). For each vertex v, set:

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated (degree 0)} \\ 0 & \text{otherwise} \end{cases}$$

Thus,

$$X = \sum_{v=1}^{n} X_v$$

The probability that a fixed vertex v is isolated is the probability that none of its $n-1$ potential incident edges is chosen:

$$\Pr[X_v = 1] = (1-p)^{n-1}$$

Given

$$p \geq \frac{(1+\epsilon)\ln n}{n-1} \quad \epsilon > 0$$

Using the inequality $1 - x \leq e^{-x}$, we can bound:

$$(1-p)^{n-1} \leq e^{-p(n-1)} \leq e^{-(1+\epsilon)\ln n} = n^{-(1+\epsilon)}$$

Therefore, the expected number of isolated vertices satisfies:

$$\mathbb{E}[X] = \sum_{v=1}^{n} \mathbb{E}[X_v] = n * \Pr[X_v = 1] \le n * n^{-(1+\epsilon)} = n^{-\epsilon} \to 0$$

Finality, by Markov's inequality,

$$\Pr[\textit{there exists an isolated vertex}] = \Pr[X \ge 1] \le \mathbb{E}[X] = n^{-\epsilon} = o(1)$$

This concludes that:

$$\Pr[\textit{no isolated vertices}] = 1 - \Pr[X \ge 1] = 1 - o(1)$$

□

1.1.2 Part B

Proof. The probability that an Erdős-Rényi graph $G = (n, p)$ contains at least one isolated vertex can be established using the second moment method.

The first step is also to define a non-negative integer random variable X that counts the number of isolated vertices (the "good" events this time). For each vertex v, set

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated (degree 0)} \\ 0 & \text{otherwise} \end{cases}$$

Thus,

$$X = \sum_{v=1}^{n} X_v$$

The probability that a fixed vertex v is isolated is the probability that none of its $n-1$ potential incident edges is chosen:

$$\Pr[X_v = 1] = (1-p)^{n-1}$$

Given

$$p \le \frac{(1-\epsilon)\ln n}{n-1} \quad \epsilon > 0$$

Using the inequality $\ln(1-x) \ge -\frac{x}{1-x}$ for $x \in (0,1)$, the bound is:

$$q = (1-p)^{n-1} = e^{(n-1)\ln(1-p)} \ge e^{-\frac{p(n-1)}{1-p}}$$

For convenience, we denote this probability by q.

Therefore, as $p \to 0$, the expected number of isolated vertices satisfies

$$\mathbb{E}[X] = \sum_{v=1}^{n} \mathbb{E}[X_v] = n \cdot \Pr[X_v = 1] = n \cdot \mathbb{E}[X_v] = nq \ge n^{\epsilon + o(1)} \xrightarrow{n \to \infty} \infty$$

To apply the Chebyshev's, the next step is to bound the variance:

$$\mathrm{Var}(X) = \sum_{v=1}^{n} \mathrm{Var}(X_v) + \sum_{v \ne u}^{n} \mathrm{Cov}(X_v, X_u)$$

For each v,

$$\mathrm{Var}(X_v) = \mathbb{E}[X_v^2] - \mathbb{E}[X_v]^2 = \mathbb{E}[X_v] - \mathbb{E}[X_v]^2 \leq \mathbb{E}[X_v] = q$$

For $\mathrm{Cov}(X_v, X_u)$ where $v \neq u$,

$$\mathrm{Cov}(X_v, X_u) = \mathbb{E}[X_v X_u] - \mathbb{E}[X_v]^2 = \Pr[X_v = X_u = 1] - q^2$$

The event $\{X_v = X_u = 1\}$ requires that all edges incident to v or u are absent. Since there is a potential edge connection v and v, the number of such edges is $2n - 3$,

$$\mathrm{Cov}(X_v, X_u) = (1-p)^{2n-3} - q^2 = \frac{q^2}{1-p} - q^2 = q^2 \frac{p}{1-p}$$

Substituting these bounds,

$$\mathrm{Var}(X) \leq nq + n(n-1) \cdot q^2 \frac{p}{1-p}$$

Comparing the variance with the squared expectation:

$$\frac{\mathrm{Var}(X)}{\mathbb{E}[X]^2} \leq \frac{nq}{n^2 q^2} + \frac{n(n-1) \cdot q^2 \frac{p}{1-p}}{n^2 q^2} = \frac{1}{nq} + \frac{(n-1)}{n} \frac{p}{1-p}$$

Since $nq = \mathbb{E}[X] \to \infty$ and $p \to 0$, the conclusion is

$$\frac{\mathrm{Var}(X)}{\mathbb{E}[X]^2} \to 0$$

By Chebyshev's inequality,

$$\Pr[X = 0] = \Pr[\textit{no isolated vertices}] \leq \frac{\mathrm{Var}(X)}{\mathbb{E}[X]^2}$$

Therefore,

$$\Pr[\textit{has isolated vertices}] = \Pr[X \geq 1] = 1 - \Pr[X = 0] = 1 - o(1)$$

□

2 Problem 2 (Cutting it Fine)

Given a connected undirected (and unweighted, for simplicity) n-vertex graph G, a k-cut is a paritition of the vertex set into k non-empty parts $S_1, S_2, \ldots, S_k$. The size of the k-cut is the number of edges that cross between distinct parts in this partition. Consider the following algorithm:

- (Phase 1) As long as the number of vertices in G is more than $2k - 2$, pick a random edge and contract it.
- (Phase 2) Now we have a graph on $2k - 2$ vertices. For each of these vertices, choose a random label from $\{L_1, L_2, \ldots, L_k\}$, contract vertices with the same label, and output the resulting cut.

Show the following:

(a) If the min k-cut size is λ, then $\lambda \leq 2(k-1)\frac{m}{n}$.

(b) Any fixed min k-cut survives the first phase with probability at least $1/\binom{n}{2(k-1)}$.

(c) Conditioned on surviving the first phase, it is output in the second phase with probability at least $\frac{k!}{k^k} \cdot \frac{1}{k^{k-2}}$.

Hence, this gives an $\approx n^{2(k-1)}$-time algorithm for the k-cut problem.

2.1 Solution

2.2 Part A

Proof. We want to show the inequality holds for the min k-cut. Thus, it suffices to show that there exists some k-cut C such that $|C| \leq 2(k-1)\frac{m}{n}$, and it will follow that the inequality holds for the min k-cut, by definition.

Consider the family of k-cuts such that $(S_1, ..., S_k)$ contains exactly $(k-1)$ singleton chosen uniformly at random, and a kth set containing the remaining vertices:

$$S_1 = \{v_1\},\ S_2 = \{v_2\},\ ...\ , S_{k-1} = \{v_{k-1}\},\ S_k = V \backslash \{v_1, ..., v_{k-1}\}$$

Consider a particular edge $e \in E$ of this family. Clearly, e is cut if and only if exactly one of its end vertices is chosen as a singleton (i.e. not both of its vertices are in S_k).

For any vertex $v \in V$, the probability that v is among the $(k-1)$ vertices chosen uniformly at random to form singleton sets is:

$$P[\ v \text{ chosen}] = \frac{k-1}{n}$$

Thus, for any edge e_{uv} between two vertices u and v, the probability that e_{uv} is cut is

$$\begin{aligned} P[e_{uv} \text{ cut}] &= P[u \text{ chosen}] + P[v \text{ chosen}] - P[\text{ both } u \text{ and } v \text{ chosen}] \\ &\leq P[u \text{ chosen}] + P[v \text{ chosen}] \\ &= \frac{k-1}{n} + \frac{k-1}{n} = \frac{2(k-1)}{n} \end{aligned}$$

By linearity of expectation, it follows that for the described family of k-cuts, the expected cut size

$$\mathbb{E}[\text{cuts}] \leq 2(k-1)\frac{m}{n}$$

It follows that there exists a cut C such that $|C| \leq 2(k-1)\frac{m}{n}$, and thus by definition

$$\lambda \leq |C| \leq 2(k-1)\frac{m}{n}$$

$\square$

2.3 Part B

Proof. The min k-cut survives phase 1 if and only if we do not contract any of its λ edges throughout the entire process.

Let e be an edge belonging to a fixed min k-cut $\tilde{C}$. Using our result from part a), we can upper bound the probability that e is contracted u.a.r among the m_i possible remaining edges at the ith iteration of phase 1:

$$P[e \text{ contracted}] = \frac{\lambda}{m_i} \leq \frac{2(k-1)}{i}$$

And thus

$$P[e \text{ not contracted}] \geq 1 - \frac{2(k-1)}{i}$$

The probability of $\tilde{C}$ surviving is equal to the probability that none of its edges are contracted:

$$P[\tilde{C} \text{ survives phase 1}] \geq \prod_{i=2k-1}^{n} \left[1 - \frac{2(k-1)}{i}\right] = \prod_{i=2k-1}^{n} \frac{i-2k+2}{i}$$

Where numerator becomes

$$\prod_{i=2k-1}^{n} (i-2k+2) = \prod_{j=1}^{n-2k+2} j = (n-2k+2)!$$

And the denominator yields

$$\prod_{i=2k-1}^{n} i = \frac{n!}{(2k-2)!}$$

Plugging back into our inequality, we finally get:

$$\begin{aligned} P[\tilde{C} \text{ survives phase 1}] &\geq \frac{(n-2k+2)!(2k-2)!}{n!} \\ &= \left(\frac{n!}{(n-2k+2)!(2k-2)!}\right)^{-1} = \left[\binom{n}{2k-2}\right]^{-1} \end{aligned}$$

□

2.4 Part C

Proof. Suppose the fixed min k-cut survives phase 1. In other words, if our fixed min k-cut was $(S_1, ..., S_k)$ at the beginning of the algorithm, then $\{S_1, .., S_k\}$ is still a valid k-cut in the graph obtained after phase 1. This means that each remaining super-node is entirely in exactly one of the parts S_i.

Thus, for the min k-cut to survive phase 2, none of the sets S_i can be merged. This means that each super-node in a same given set S_i must all have the same unique label L_i associated to their set S_i.

Formally, let w_j be the j-th supernode in S_i after phase 1. Then, for the fixed k-min cut to survive, we must have

$$\forall w_j \in S_i : l(w_j) = L_i$$

Where $l(w_j)$ is the label assigned to supernode w_j.

Because we assign each of the k labels uniformly at random, the probability that all $2k-2$ remaining supernodes after phase 1 are assigned their unique label L_i is

$$\left(\frac{1}{k}\right)^{2k-2} = \frac{1}{k^{2k-2}}$$

And since there are $k!$ ways of bijectively assigning every label from $\{L_1, ..., L_k\}$ to each set $\{S_1, ..., S_k\}$, we have that the probability that the fixed min k-cut survives phase 2 is:

$$P = \frac{k!}{k^{2k-2}}$$

□

3 Problem 3 (Cyclic Changes)

An ℓ-cycle in a graph is a cycle with *at most* ℓ nodes. In this problem, we want to show there exist graphs with many edges, and no *short* cycles. It is easy to construct such graphs with $\Omega(n)$ edges—in fact, a tree has $n-1$ edges and no cycles at all! We want slightly denser graphs with no ℓ-cycles for any constant ℓ.

(a) Consider the graph $G(n, p)$ for some $p \in [0, 1]$. Calculate the probability that some sequence of ℓ vertices is a cycle, and hence calculate the expected number of ℓ-cycles.

(b) (Do not submit.) Note that setting $p \approx 1/n$ means the expected number of ℓ-cycles is $o(1)$, but the expected number of edges is $O(n)$—which is not very interesting!

(c) Now consider the following two part algorithm: (i) first pick $G \sim G(n, p)$, and then (ii) for each ℓ-cycle in G, delete an arbitrary edge on it. By construction this graph has no ℓ-cycles. Show that setting $p = n^{\frac{2-\ell}{\ell-1}}$ ensures that the expected number of edges in the resulting graph is $m := \Theta(n^{1+\frac{1}{\ell-1}})$. Hence, infer that there exist graphs with $m = \omega(n)$ edges and no ℓ-cycles.

3.1 Solution

3.1.1 Part A

Proof. Consider an ordered sequence of k distinct vertices. The probability that these vertices form a cycle of length exactly k is the probability that corresponding edges exists, nameley

$$(v_1v_2, v_2v_3, \ldots v_kv_1)$$

Hence,

$$\Pr[\textit{sequence of } k \textit{ vertices forms a cycle}] = p^k$$

In a random graph $G = (n, p)$, each cycle of length k corresponds to $2k$ ordered sequences (k possible starting points and 2 possible directions).

Let X_k denote the number of k exact cycle, The expectation is

$$\mathbb{E}[X_k] = \frac{(n)_k}{2k}p^k = \binom{n}{k}\frac{(k-1)!}{2}p^k$$

Since a cycle of lenth at most l can have any lenth k with $3 \leq k \leq l$, the expected number of such cycles is given by

$$\mathbb{E}[X_{\leq l}] = \sum_{k=3}^{l} \mathbb{E}[X_k]$$

□

3.1.2 Part C

Proof. Let M, $X_{\leq l}$ and Y represent respectively the number of edges, the number of cycles of length at most l and the number of edges in the resulting graph. Since removing one edge from every cycle suffices to eliminate all cycle, we have

$$Y \geq M - X_{\leq l}$$

Taking expectations gives

$$\mathbb{E}[Y] \geq \mathbb{E}[M] - \mathbb{E}[X_{\leq l}]$$

The expected number of edges is

$$\mathbb{E}[M] = \binom{n}{2}p = \left(\frac{1}{2} + o(1)\right)n^2p$$

From Part A, the expected number of short cycles is

$$\mathbb{E}[X_{\leq l}] = \sum_{k=3}^{l}\left(\frac{1}{2k} + o(1)\right)n^kp^k$$

Now set

$$p = n^{\frac{2-l}{l-1}}$$

For each k,

$$n^kp^k = n^{k+k\frac{2-l}{l-1}} = n^{\frac{k}{l-1}}$$

These terms increase with k, so the sum is dominated by $k = l$

$$\mathbb{E}[X_{\leq l}] = \left(\frac{1}{2l} + o(1)\right)n^{1+\frac{1}{l-1}}$$

Meanwhile,

$$\mathbb{E}[M] = \left(\frac{1}{2} + o(1)\right) n^{2+\frac{2-l}{l-1}} = \left(\frac{1}{2} + o(1)\right) n^{1+\frac{1}{l-1}}$$

Therefore,

$$\mathbb{E}[Y] \geq \mathbb{E}[M] - \mathbb{E}[X_{\leq l}] = \left(\frac{1}{2} - \frac{1}{2l} - o(1)\right) n^{1+\frac{1}{l-1}} = \Theta(n^{1+\frac{1}{l-1}})$$

$\square$

4 Problem 4 (Random 3-SAT)

(Taken from Johan Håstad's course "Theoreticians toolkit" at KTH.) Constructing a random 3-SAT formula with n variables and $m = \lceil dn \rceil$ (remember that $\lceil x \rceil$ is the smallest integer larger than x) clauses is done in the following way. Randomly take three different variables (all triples being equally likely). With uniform probability choose one of the eight ways to negate these variables and make them into a clause. Repeat with independent randomness until you have m clauses.

(a) For what value of d is the expected number of satisfying assignments $\Theta(1)$? Call this value d_0.

(b) Prove that the formula is likely (probability $1 - o(1)$) to be unsatisfiable for any constant d such that $d > d_0$.

(c) **(Extra Credit.)** Prove that the formula remains at least somewhat likely to be unsatisfiable also in the case when d is slightly smaller than d_0. The difficulty of this problem is very much dependent on what we mean by "somewhat likely" and "slightly smaller". The exact formulation to prove to get a full score on this problem is that there is some constant $d_1 < d_0$ such that for $d = d_1$ the probability that the corresponding random formula is satisfiable is at most $1/2$. The size of $d_0 - d_1$ does not matter for your score on the problem and the main property of a solution to aim for is a mathematically correct argument. *Hint:* A satisfiable formula that does not depend on all its variables has many satisfying assignments.

4.1 solution

4.2 Part A

Let us fix a truth assignment $\alpha \in \mathcal{A} = \{0,1\}^n$. Any 3-clause is unsatisfied under assingment α if and only if all 3 literals in the clause are false under α.

Since all 3 variables in the clause are distinct and are each negated independently, it follows that for any 3-clause:

$$P[\text{clause unsatisified under } \alpha] = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

And thus

$$P[\text{clause satisfied under } \alpha] = 1 - \frac{1}{8} = \frac{7}{8}$$

Under the assumption of clause independence, the probability that a 3-SAT Ψ of m clauses evaluates to true under α is therefore

$$P[\alpha \text{ satisfies } \Psi] = \left(\frac{7}{8}\right)^m$$

Let X be the number of assignments satisfying the statement Ψ and let $\mathbb{I}_\alpha$ be the indicator variable that the truth assignment α satisfies Ψ:

$$X = \sum_{\alpha \in \mathcal{A}} \mathbb{I}_\alpha \Longrightarrow \mathbb{E}[X] = \sum_{\alpha \in \mathcal{A}} \mathbb{E}[\mathbb{I}_\alpha] = 2^n \left(\frac{7}{8}\right)^m$$

By hypothesis, $m = \lceil dn \rceil \approx dn$. Thus:

$$\mathbb{E}[X] \approx 2^n \left(\frac{7}{8}\right)^{dn} = \exp\left[n(\log 2 + d\log(7/8))\right]$$

And for $\mathbb{E}[X] = \Theta(1)$, we must therefore have

$$\log 2 + d\log(7/8) = 0 \Longleftrightarrow d = -\frac{\log 2}{\log(7/8)} = \frac{\log 2}{\log(8/7)} \stackrel{\triangle}{=} d_0$$

4.3 Part B

Proof. Markov's inequality directly yields:

$$\begin{aligned} P[X \geq 1] \leq \mathbb{E}[X] &\stackrel{\star}{=} \exp\left[n\left(\log 2 + d\log(7/8)\right) + o(n)\right] \\ &= \exp\left(-\Theta(n)\right), \; \forall d < d_0 \end{aligned}$$

Where we used the result from part a in the second equality. It obviously follows that, for all $d < d_0$, we have:

$$\mathbb{E}[X] = \exp\left(-\Theta(n)\right) \xrightarrow{n \to \infty} 0$$

And thus

$$P[\Psi \text{ satisifable}] = P[X \geq 1] = o(1) \Longleftrightarrow P[X < 1] = 1 - o(1)$$

$\square$