

Homework 6

Out: 13 Mar. Due: 20 Mar.

*Instructions: Put your solutions in the homework box on Soda level 2 by 5pm on Tuesday. Take time to write clear and concise answers; confused and long-winded solutions will be penalized. You are encouraged to form small groups (two to four people) to work through the homework, but you **must** write up all your solutions on your own. Depending on grading resources, we reserve the right to grade a random subset of the problems and check off the rest; so you are advised to attempt all problems.*

1. Suppose we throw n balls into n bins, and we are interested in upper bounds for the probability p that the first bin receives at least k balls, where $\log \log n \leq k \leq n$. In this problem, we will compare the bounds we get for p using three different techniques that we have seen in the course so far.

(a) Use the Chernoff bound to show that $p = 2^{-\Omega(k)}$.

(b) Apply a union bound to show that $p \leq \binom{n}{k} \left(\frac{1}{n}\right)^k$. Deduce that $p \leq \left(\frac{e}{k}\right)^k = 2^{-\Omega(k \log k)}$.

[HINT: If the first bin receives at least k balls, then it must be the case that there exists some subset of exactly k balls all of which land in the first bin. You may find the following bound useful: $k! \geq (k/e)^k$.] [NOTE: This bound is rather better than that in part (a).]

(c) Write down the expression for p under the Poisson approximation (this will involve a sum). Then use the inequality from Q2 of HW5 to show that $p \leq \frac{2}{e} \sum_{j=k}^{\infty} \frac{1}{j!}$. [NOTE: This expression is bounded by $\frac{2}{e} \cdot \frac{1}{k^k} \sum_{j=k}^{\infty} \frac{k^j}{j!} \leq \frac{2}{e} \left(\frac{e}{k}\right)^k$, so this bound is essentially the same as that in part (b).]

(d) Now suppose we want to achieve $p = n^{-\log \log n}$. According to each of the bounds in parts (a) and (b) above, how large would k have to be in order to make p this small? Repeat this same comparison for the following two values of p : $p = 1/n^{100}$ and $p = 2^{-0.001n}$. In all cases, you should just give the asymptotic value of k as a function of n (i.e., ignore constants and lower order terms).

2. MU, Exercise 5.16. [NOTE: For each of the three parts, you should start by writing down the expected number of the appropriate objects (cliques, complete bipartite subgraphs, Hamiltonian cycles) as a function of n and p . Then you should figure out what value of p (as a function of n) would make this expectation equal to 1. You should just give the *asymptotic* value of p (i.e., ignore constants and lower order terms).]

3. Recall that a graph (undirected, no self-loops) is *2-colorable* if we can assign colors red and green to each vertex such that the endpoints of every edge are assigned different colors. Suppose we are told that a graph $G = (V, E)$ is “locally 2-colorable”, in the sense that the induced subgraph¹ on every subset of $O(\log n)$ vertices is 2-colorable. Does this imply that G itself is 2-colorable? In this problem we will see that the answer is spectacularly “no”: namely, we will show that there exists a graph that is locally 2-colorable but is “very far away” from being 2-colorable, in the sense that we would have to remove a constant fraction of its edges in order to make it 2-colorable. We will prove the existence of this graph using the probabilistic method.

Throughout, set $p = 16/n$, and let G be a random graph from the model $\mathcal{G}_{n,p}$. The probabilities and expectations refer to the experiment of picking G at random.

(a) Write down the expected number of edges in G .

(b) Apply the Chernoff bound to show that with probability $1 - 2^{-\Omega(n)}$, G has at most $10(n - 1)$ edges.

¹The *induced* subgraph on a subset of vertices $V' \subseteq V$ is the graph with vertex set V' and edge set consisting of all the edges of G both of whose endpoints are in V' .

- (c) Now fix an arbitrary assignment of colors to the vertices. Show that the expected number of violated edges (i.e., edges with endpoints of the same color) in G is at least $4(n-2)$. Deduce by a Chernoff bound that the probability there are more than $n-2$ violated edges is at least $1 - e^{-9(n-2)/8}$. [HINT: For the first part, think of the assignment of colors as being fixed *before* we choose the random edges of G . What is the value for the number of red/green vertices that minimizes the expected number of violated edges? You may assume for simplicity that n is even.]
- (d) Show that for $n \geq 6$, with probability at least $3/4$, G is not 2-colorable even if we delete any $n-2$ of its edges. [HINT: Use the previous part and a union bound over colorings.]
- (e) Show that the expected number of cycles of length exactly k in G is at most 16^k . Deduce that the expected number of cycles of length at most $\frac{1}{8} \log n$ is at most $16\sqrt{n}$.
- (f) Use the previous part to deduce that, with probability at least $3/4$, by deleting only $O(\sqrt{n})$ (suitably chosen) edges of G we can obtain a graph such that the induced subgraph on any subset of $\frac{1}{8} \log n$ vertices is cycle-free (i.e., a forest – a collection of vertex-disjoint trees). (Note that a forest is always 2-colorable.)
- (g) Put parts (b), (d) and (f) together to deduce that, for every sufficiently large n , there exists a graph $G = G_n$ on n vertices such that:
- The induced subgraph on any subset of $\frac{1}{8} \log n$ vertices of G_n is 2-colorable; and
 - G_n is not 2-colorable, and remains not 2-colorable even after deleting any 0.05 fraction of its edges.

[HINT: Remember to take into account the fact that, when we modify G to remove cycles, we may also be deleting violated edges!]