Probabilistic method

Last week, we used the probabilistic method to find upper bounds on Ramsey numbers. We'll do another small example. The following proposition can be proved constructively, but we can also do it probabilistically. The latter allows for a slight improvement that is harder to achieve constructively.

Proposition 1. *For any graph G, there exists a subset $S \subseteq V(G)$ such that*

$$|[S, \overline{S}]| \geq \frac{|E|}{2}.$$

Exercise 1. *Prove Proposition 1.* 1

Exercise 2. *Can you improve on Proposition 1 by sampling S differently?* 2

A dominating set in a graph $G = (V, E)$ is a set S such that every vertex is in S or has a neighbor in S. Dominating sets have many applications, for example in communication routing in networks. Often, the goal is to find a smallest possible dominating set. This problem is NP-complete in general.
Let $d(v)$ be the degree of v in G, $\delta(G)$ be the minimum degree in G (and $\Delta(G)$ be the maximum degree in G).

Lemma 2. *Every graph G has a dominating set of at most size $\frac{\log(1+\delta)+1}{1+\delta}n$.*

Suppose $\delta \geq 1$. Create the set X by including each $v \in V$ with probability p i.i.d. Then create $Y = \{v \in V \backslash X : v \text{ has no neighbor in } X\}$. By definition, $X \cup Y$ is a dominating set in G. Since each v is included in X with probability p, the expected size of X is $\mathbb{E}(|X|) = np$. The expected size of Y is

$$\begin{aligned}
\mathbb{E}(|Y|) &= \sum_{v \in V} \mathbb{E}(1_{v \in Y}) = \sum_{v \in V} \mathbb{P}(v \in Y) = \sum_{v \in V} \mathbb{P}(v \notin X, v\text{'s neighbors} \notin X) \\
&= \sum_{v \in V} (1-p)^{d(v)+1} \leq \sum_{v \in V} (1-p)^{\delta+1} = n(1-p)^{\delta+1}
\end{aligned}$$

Since X and Y are disjoint we have

$$\mathbb{E}(|X \cup Y|) \leq np + n(1-p)^{\delta+1}.$$

Exercise 3. *Show that* 1

$$\mathbb{E}(|X \cup Y|) \leq np + ne^{-p(\delta+1)}.$$

The following fact is helpful:

$$(1+a)^n \leq e^{an}, \quad \textit{for } a > -1,\ n \geq 0.$$

Then, find the value of p that minimizes this function, and finish the proof of Lemma 2.

Random graphs

The two models $G(n,m)$ and $G(n,p)$ are the two simplest models for generating a random graph on n vertices. The model $G(n,m)$ was introduced by Erdős and Rényi in a series of papers that form the foundation of random graph theory. The model $G(n,p)$ was introduced by Gilbert, although it is commonly referred to as the Erdős-Rényi model. The model $G(n,m)$ generates a graph uniformly at random over all simple (no multi-edges or self-loops) graphs on n vertices and m edges. There are $\binom{\binom{n}{2}}{m}$ such graphs. Therefore, for any graph G on n vertices and m edges, we have

$$\mathbb{P}(G(n,m) = G) = \binom{\binom{n}{2}}{m}^{-1}.$$

The model $G(n,p)$ is a graph on n vertices, where each of the $\binom{n}{2}$ possible edges is present with probability p, independently of other edges. This implies that for any graph G on n vertices and m edges, we have

$$\mathbb{P}(G(n,p) = G) = p^m(1-p)^{\binom{n}{2}-m}.$$

The model $G(n,p)$ is a lot easier to work with and is therefore much more common. When $m = p\binom{n}{2}$, the models behave "almost" the same.

Exercise 4. *An* isolated vertex *is a vertex that has no neighbors (i.e. is not an endpoint of any edge in E). Let A_n be the event that a graph generated by $G(n,p)$ has an isolated vertex. Show that* 1

$$\mathbb{P}(A_n) \leq n(1-p)^{n-1}.$$

Exercise 5. *Show that for any $\epsilon > 0$, if $p = (1+\epsilon)\frac{\log n}{n}$, then* 1

$$\lim_{n\to\infty} \mathbb{P}(A_n) = 0.$$

Let's look for the appearance of certain fixed subgraphs in $G(n,p)$. For example, at what values of p do we expect to start seeing copies of K_3 in $G(n,p)$?

Exercise 6. *Show that when $p \ll n^{-1}$, we have that $\mathbb{P}(X_3 > 0) \to 0$, where X_3 is the number of triangles in $G(n,p)$.* 1

Exercise 7. *Do a similar analysis for the graph $K_3 + K_2$. Do you notice anything strange?* 1