Note: Almost all these exercises, as well as almost all the material in the class, is adapted from the phenomenal textbook *The Probabilistic Method* by Alon and Spencer. If you'd like to learn more or see other exercises, this is *the* book to check out. It's pretty rare to find a field of math with *one* canonical textbook, but that's definitely the case here.

1. A *prefix-free code* is a collection $\mathcal{C}$ of finite binary strings with the property that no string in $\mathcal{C}$ is a prefix of another string. For example, $\{0, 11, 101\}$ is prefix-free, whereas $\{0, 11, 110\}$ is not.

   For a binary string x, let $\ell(x)$ denote the length of x. Prove that if $\mathcal{C}$ is a prefix-free code, then
   $$\sum_{x \in \mathcal{C}} \frac{1}{2^{\ell(x)}} \leq 1.$$
   Can you figure out why we care about prefix-free codes, or why this inequality says something meaningful? Come talk to me if not!

2. Let G be a graph on $n \geq 10$ vertices and suppose that if we add to G any edge not in G then the number of copies of a complete graph on 10 vertices in it increases. Show that G has at least $8n - 36$ edges.

   Hint: Use one of the results we proved in class today!

3. Prove the following strengthened version of Bollobás's two families theorem.

   Let $A_1, \ldots, A_m, B_1, \ldots, B_m$ be sets with the property that $A_i \cap B_j = \varnothing$ if and only if $i = j$. Then
   $$\sum_{i=1}^{m} \frac{1}{\binom{|A_i|+|B_i|}{|A_i|}} \leq 1.$$

4. Let $\{1, 2, 3\}^d$ denote the $3 \times 3 \times \cdots \times 3$ grid in d dimensions. We wish to cover this grid by sets of the form $S_1 \times S_2 \times \cdots \times S_d$, where $S_i \subseteq \{1, 2, 3\}$ has size exactly 2. In other words, we wish to cover the grid with side length 3 by (generalized) subgrids of side length 2.

   (a) Prove that we need at least $(3/2)^d$ subgrids of side length 2 for this to be possible.

   (b) Prove that this *is* possible with at most $(3/2)^d \cdot (d \ln 3)$ subgrids.

   ? (c) Can you improve either the lower or the upper bound, and close the $d \ln 3$ gap?

5. Let k, t, n be positive integers.

   (a) Suppose that there exists some $p \in [0, 1]$ so that
   $$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t} (1-p)^{\binom{t}{2}} < 1.$$
   Prove that there exists an n-vertex graph with no clique of size k and no independent set of size t.

$\star$ means that this problem is harder than the other ones.
? means that this is an open problem.

$\star$ (b) Conclude that there exists an n-vertex graph with no K_4 and independence number at most $O(n^{2/3} \log n)$.

$\star\star$ (c) Can you explicitly construct such a graph?

6. Let $\mathcal{F}$ be a collection of k-element subsets of $\{1, 2, \ldots, n\}$. $\mathcal{F}$ is called *intersecting* if for all $A, B \in \mathcal{F}$, we have that $A \cap B \neq \varnothing$.

   (a) If $n < 2k$, find an intersecting family of k-element subset of $\{1, 2, \ldots, n\}$ with $|\mathcal{F}| = \binom{n}{k}$.

   (b) If $n \geq 2k$, find an intersecting family of k-element subsets of $\{1, 2, \ldots, n\}$ with $|\mathcal{F}| = \binom{n-1}{k-1}$.

   (c) For $0 \leq s \leq n-1$, let $A_s = \{s, s+1, \ldots, s+k-1\}$, where addition is modulo n. Prove that if $\mathcal{F}$ is intersecting and $n \geq 2k$, then $\mathcal{F}$ can contain at most k of the sets A_s.

   (d) Let π be a uniformly random permutation of $\{1, 2, \ldots, n\}$, and let $i \in \{1, 2, \ldots, n\}$ be uniformly random as well. Let $A = \{\pi(i), \pi(i+1), \ldots, \pi(i+k-1)\}$. Using part (c), prove that $\Pr(A \in \mathcal{F}) \leq k/n$.

   (e) With the same notation as above, prove that $\Pr(A \in \mathcal{F}) = |\mathcal{F}|/\binom{n}{k}$.

   (f) Using parts (d) and (e), prove that if $n \geq 2k$ and $\mathcal{F}$ is intersecting, then

   $$|\mathcal{F}| \leq \binom{n-1}{k-1},$$

   i.e. that the construction in part (b) is best possible.

1. In class, we proved that for any finite set $A \subseteq \mathbb{N}$, there is a subset $B \subseteq A$ with $|B| \geq |A|/3$ such that B is sum-free.

   Prove the same thing for any finite set $A \subseteq \mathbb{R}$ of *real* numbers.

2. Call a set $B \subseteq \mathbb{N}$ *weirdo-sum-free* if there do not exist $b_1, b_2, b_3, b_4 \in B$ so that

$$b_1 + 2b_2 = 2b_3 + 2b_4. \qquad (*)$$

   Prove that there exists some $c > 0$ so that every $A \subseteq \mathbb{N}$ contains a weirdo-sum-free subset $B \subseteq A$ with $|B| \geq c|A|$.

   For which other equations besides $(*)$ can you prove such a result?

3. Let $v_1, \ldots, v_n \in \mathbb{R}^d$ be vectors in d-dimensional space, with the property that $\|v_i\| \leq 1$ for all i, where $\|\cdot\|$ denotes the usual Euclidean length of a vector. Prove that there exist $\sigma_1, \ldots, \sigma_n \in \{-1, 1\}$ so that

$$\|\sigma_1 v_1 + \cdots + \sigma_n v_n\| \leq \sqrt{n}.$$

4. Let T be an n-vertex tournament. A *Hamiltonian path* in T is some ordering $v_1, \ldots, v_n$ of the vertices so that the arrows go $v_1 \to v_2 \to \cdots \to v_n$.

   (a) Prove that there exists an n-vertex tournament with at least $n!/2^n$ different Hamiltonian paths.

   (b) Prove that every tournament has at least one Hamiltonian path.

5. You may wish to skip this problem if you are unfamiliar with graph theory.

   Let G be an n-vertex graph. Its *independence number* $\alpha(G)$ is the size of the largest independent set in G, i.e. the size of the largest set of vertices with no edges between them. The *chromatic number* $\chi(G)$ is the least number of colors we can use if we want to assign a color to every vertex with the property that adjacent vertices receive different colors.

   (a) Prove that $\chi(G) \geq n/\alpha(G)$.

   (b) Suppose G is *vertex-transitive*, meaning that for all vertices $v, w \in V(G)$, there is some automorphism of G taking v to w. Roughly speaking, this means that G is very symmetric: all vertices "look the same". Prove that in this case,

$$\chi(G) \leq \frac{n}{\alpha(G)} \ln n.$$

   (c) Prove that the vertex-transitive assumption is necessary in part (b), i.e. find non-vertex-transitive graphs for which the lower bound $\chi(G) \geq n/\alpha(G)$ is very far from the truth.

6. Suppose that G is an n-vertex graph with $nd/2$ edges, for some real numer $d \geq 1$. In this problem, you'll show that $\alpha(G) \geq n/(2d)$, which is about a factor of two worse than the result we proved today in class.

(a) Fix some parameter $p \in [0,1]$ that we'll pick later. Let S be a random subset of $V(G)$ obtained by taking each vertex independently with probability p. Let $X = |S|$. Prove that $\mathbb{E}[X] = pn$.

(b) Let Y denote the number of edges in S. Prove that $\mathbb{E}[Y] = p^2 nd/2$.

(c) Pick p to maximize $\mathbb{E}[X - Y]$, and conclude that G has an independent set of size at least $n/(2d)$.

7. Let G be an n-vertex graph with minimum degree δ. A *dominating set* in G is a set $U \subseteq V(G)$ of vertices with the property that every vertex of G is either contained in U or has at least one neighbor in U.

(a) Fix some parameter $p \in [0,1]$ that we'll pick later. Let S be a random subset of $V(G)$ obtained by taking each vertex independently with probability p. Let $X = |S|$. Prove that $\mathbb{E}[X] = pn$.

(b) Let Y denote the number of vertices outside of S that do not have a neighbor in S. Prove that $\mathbb{E}[Y] \leq n(1-p)^{\delta+1}$.

(c) Prove that G has a dominating set of size at most $pn + n(1-p)^{\delta+1}$.

(d) Pick the value of p that minimizes this quantity to conclude that any n-vertex graph with minimum degree δ has a dominating set of size at most

$$n\frac{1+\ln(\delta+1)}{\delta+1}.$$