

PROBABILISTIC METHODS IN COMBINATORICS
MIT 18.226 (FALL 2024)
PROBLEM SET

<https://sammy-luo.github.io/18-226/>

A. INTRODUCTION AND LINEARITY OF EXPECTATIONS

A1. Verify the following asymptotic calculations used in Ramsey number lower bounds:

- (a) For each k , the largest n satisfying $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$ has $n = \left(\frac{1}{e\sqrt{2}} + o(1)\right) k 2^{k/2}$.
- (b) For each k , the maximum value of $n - \binom{n}{k} 2^{1-\binom{k}{2}}$ as n ranges over positive integers is $\left(\frac{1}{e} + o(1)\right) k 2^{k/2}$.
- (c) For each k , the largest n satisfying $e \left(\binom{k}{2} \binom{n}{k-2} + 1\right) 2^{1-\binom{k}{2}} < 1$ satisfies $n = \left(\frac{\sqrt{2}}{e} + o(1)\right) k 2^{k/2}$.

A2. Prove that, if there is a real $p \in [0, 1]$ such that

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t} (1-p)^{\binom{t}{2}} < 1$$

then the Ramsey number $R(k, t)$ satisfies $R(k, t) > n$. Using this show that

$$R(4, t) \geq c \left(\frac{t}{\log t}\right)^{3/2}$$

for some constant $c > 0$.

ps1

- A3. Let G be a graph with n vertices and m edges. Prove that K_n can be written as a union of $O(n^2(\log n)/m)$ isomorphic copies of G (not necessarily edge-disjoint).
- A4. Prove that there is an absolute constant $C > 0$ so that for every $n \times n$ matrix with distinct real entries, one can permute its rows so that no column in the permuted matrix contains an increasing subsequence of length at least $C\sqrt{n}$. (A subsequence does not need to be selected from consecutive terms. For example, $(1, 2, 3)$ is an increasing subsequence of $(1, 5, 2, 4, 3)$.)
- A5. *Generalization of Sperner's theorem.* Let \mathcal{F} be a collection of subset of $[n]$ that does not contain $k+1$ elements forming a chain: $A_1 \subsetneq \cdots \subsetneq A_{k+1}$. Prove that \mathcal{F} is no larger than taking the union of the k levels of the Boolean lattice closest to the middle layer.
- A6. Let G be a graph on $n \geq 10$ vertices. Suppose that adding any new edge to G would create a new clique on 10 vertices. Prove that G has at least $8n - 36$ edges.

Hint in white:

A7. Let $k \geq 4$ and H a k -uniform hypergraph with at most $4^{k-1}/3^k$ edges. Prove that there is a coloring of the vertices of H by four colors so that in every edge all four colors are represented.

ps1

A8. Given a set \mathcal{F} of subsets of $[n]$ and $A \subseteq [n]$, write $\mathcal{F}|_A := \{S \cap A : S \in \mathcal{F}\}$ (its *projection* onto A). Prove that for every n and k , there exists a set \mathcal{F} of subsets of $[n]$ with $|\mathcal{F}| = O(k 2^k \log n)$ such that for every k -element subset A of $[n]$, $\mathcal{F}|_A$ contains all 2^k subsets of A .

ps1

A9. Let A_1, \dots, A_m be r -element sets and B_1, \dots, B_m be s -element sets. Suppose $A_i \cap B_i = \emptyset$ for each i , and for each $i \neq j$, either $A_i \cap B_j \neq \emptyset$ or $A_j \cap B_i \neq \emptyset$. Prove that $m \leq (r+s)^{r+s}/(r^r s^s)$.

- ps1★** A10. Show that in every non-2-colorable n -uniform hypergraph, one can find at least $\frac{n}{2} \binom{2n-1}{n}$ unordered pairs of edges with each pair intersecting in exactly one vertex.
- A11. Let A be a measurable subset of the unit sphere in \mathbb{R}^3 (centered at the origin) containing no pair of orthogonal points.
- ps1** (a) Prove that A occupies at most $1/3$ of the sphere in terms of surface area.
- ps1★** (b) Prove an upper bound smaller than $1/3$ (give your best bound).
- ps1★** A12. Prove that every set of 10 points in the plane can be covered by a union of disjoint unit disks.
- A13. Let $\mathbf{r} = (r_1, \dots, r_k)$ be a vector of nonzero integers whose sum is nonzero. Prove that there exists a real $c > 0$ (depending on \mathbf{r} only) such that the following holds: for every finite set A of nonzero *reals*, there exists a subset $B \subseteq A$ with $|B| \geq c|A|$ such that there do not exist $b_1, \dots, b_k \in B$ with $r_1 b_1 + \dots + r_k b_k = 0$.
- ps1** A14. Prove that every set A of n nonzero integers contains two disjoint subsets B_1 and B_2 , such that both B_1 and B_2 are sum-free, and $|B_1| + |B_2| > 2n/3$.
- ps1** A15. Let G be an n -vertex graph with pn^2 edges, with $n \geq 10$ and $p \geq 10/n$. Prove that G contains a pair of vertex-disjoint and isomorphic subgraphs (not necessarily induced) each with at least $cp^2 n^2$ edges, where $c > 0$ is a constant.
- ps1★** A16. Prove that for every positive integer r , there exists an integer K such that the following holds. Let S be a set of rk points evenly spaced on a circle. If we partition $S = S_1 \cup \dots \cup S_r$ so that $|S_i| = k$ for each i , then, provided $k \geq K$, there exist r congruent triangles where the vertices of the i -th triangle lie in S_i , for each $1 \leq i \leq r$.
- ps1★** A17. Prove that $[n]^d$ cannot be partitioned into fewer than 2^d sets each of the form $A_1 \times \dots \times A_d$ where $A_i \subsetneq [n]$.

B. ALTERATION METHOD

- B1. Using the alteration method, prove the Ramsey number bound

$$R(4, k) \geq c(k/\log k)^2$$

for some constant $c > 0$.

- B2. Prove that every 3-uniform hypergraph with n vertices and $m \geq n$ edges contains an independent set (i.e., a set of vertices containing no edges) of size at least $cn^{3/2}/\sqrt{m}$, where $c > 0$ is a constant.
- B3. Prove that every k -uniform hypergraph with n vertices and m edges has a transversal (i.e., a set of vertices intersecting every edge) of size at most $n(\log k)/k + m/k$.
- ps2** B4. *Zarankiewicz problem*. Prove that for every positive integers $n \geq k \geq 2$, there exists an $n \times n$ matrix with $\{0, 1\}$ entries, with at least $\frac{1}{2}n^{2-2/(k+1)}$ 1's, such that there is no $k \times k$ submatrix consisting of all 1's.
- ps2** B5. Fix k . Prove that there exists a constant $c_k > 1$ so that for every sufficiently large $n > n_0(k)$, there exists a collection \mathcal{F} of at least c_k^n subsets of $[n]$ such that for every k distinct $F_1, \dots, F_k \in \mathcal{F}$, all 2^k intersections $\bigcap_{i=1}^k G_i$ are nonempty, where each G_i is either F_i or $[n] \setminus F_i$.

- B6. *Acute sets in \mathbb{R}^n .* Prove that, for some constant $c > 0$, for every n , there exists a family of at least $c(2/\sqrt{3})^n$ subsets of $[n]$ containing no three distinct members A, B, C satisfying $A \cap B \subseteq C \subseteq A \cup B$.

Deduce that there exists a set of at least $c(2/\sqrt{3})^n$ points in \mathbb{R}^n so that all angles determined by three points from the set are acute.

Remark. The current best lower and upper bounds for the maximum size of an “acute set” in \mathbb{R}^n (i.e., spanning only acute angles) are $2^{n-1} + 1$ and $2^n - 1$ respectively.

ps2★

- B7. *Covering complements of sparse graphs by cliques*

- Prove that every graph with n vertices and minimum degree $n - d$ can be written as a union of $O(d^2 \log n)$ cliques.
- Prove that every bipartite graph with n vertices on each side of the vertex bipartition and minimum degree $n - d$ can be written as a union of $O(d \log n)$ complete bipartite graphs (assume $d \geq 1$).

ps2★

- B8. Let $G = (V, E)$ be a graph with n vertices and minimum degree $\delta \geq 2$. Prove that there exists $A \subseteq V$ with $|A| = O(n(\log \delta)/\delta)$ so that every vertex in $V \setminus A$ contains at least one neighbor in A and at least one neighbor not in A .

ps2★

- B9. Prove that every graph G without isolated vertices has an induced subgraph H on at least $\alpha(G)/2$ vertices such that all vertices of H have odd degree. Here $\alpha(G)$ is the size of the largest independent set in G .

C. SECOND MOMENT METHOD

ps2

- C1. *Threshold for k -APs.* Let $[n]_p$ denote the random subset of $\{1, \dots, n\}$ where every element is included with probability p independently. For each fixed integer $k \geq 3$, determine the threshold for $[n]_p$ to contain a k -term arithmetic progression.

- C2. Show that, for each fixed positive integer k , there is a sequence p_n such that

$$\mathbb{P}(G(n, p_n) \text{ has a connected component with exactly } k \text{ vertices}) \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Hint in white:

ps2

- C3. *Poisson limit.* Let X be the number of triangles in $G(n, c/n)$ for some fixed $c > 0$.

- For every nonnegative integer k , determine the limit of $\mathbb{E}\binom{X}{k}$ as $n \rightarrow \infty$.
- Let $Y \sim \text{Binomial}(n, \lambda/n)$ for some fixed $\lambda > 0$. For every nonnegative integer k , determine the limit of $\mathbb{E}\binom{Y}{k}$ as $n \rightarrow \infty$, and show that it agrees with the limit in (a) for some $\lambda = \lambda(c)$.

We know that Y converges to the Poisson distribution with mean λ . Also, the Poisson distribution is determined by its moments.

- Compute, for fixed every nonnegative integer t , the limit of $\mathbb{P}(X = t)$ as $n \rightarrow \infty$.

(In particular, this gives the limit probability that $G(n, c/n)$ contains a triangle, i.e., $\lim_{n \rightarrow \infty} \mathbb{P}(X > 0)$. This limit increases from 0 to 1 continuously when c ranges from 0 to $+\infty$, thereby showing that the property of containing a triangle has a coarse threshold.)

ps2

- C4. *Central limit theorem for triangle counts.* Find a real (non-random) sequence a_n so that, letting X be the number of triangles and Y be the number of edges in the random graph

$G(n, 1/2)$, one has

$$\text{Var}(X - a_n Y) = o(\text{Var } X).$$

Deduce that X is asymptotically normal, that is, $(X - \mathbb{E}X)/\sqrt{\text{Var } X}$ converges to the normal distribution.

(You can solve for the minimizing a_n directly similar to ordinary least squares linear regression, or first write the edge indicator variables as $X_{ij} = (1 + Y_{ij})/2$ and then expand. The latter approach likely yields a cleaner computation.)

C5. *Isolated vertices.* Let $p_n = (\log n + c_n)/n$.

(a) Show that, as $n \rightarrow \infty$,

$$\mathbb{P}(G(n, p_n) \text{ has no isolated vertices}) \rightarrow \begin{cases} 0 & \text{if } c_n \rightarrow -\infty, \\ 1 & \text{if } c_n \rightarrow \infty. \end{cases}$$

(b) Suppose $c_n \rightarrow c \in \mathbb{R}$, compute, with proof, the limit of LHS above as $n \rightarrow \infty$, by following the approach in [C3](#).

ps2* C6. Is the threshold for the bipartiteness of a random graph coarse or sharp?

(You are not allowed to use any theorems that we did not prove in class/notes.)

ps2 C7. *Triangle packing.* Prove that, with probability approaching 1 as $n \rightarrow \infty$, $G(n, n^{-1/2})$ has at least $cn^{3/2}$ edge-disjoint triangles, where $c > 0$ is some constant.

Hint in white:

ps3 C8. *Nearly perfect triangle factor.* Prove that, with probability approaching 1 as $n \rightarrow \infty$,

(a) $G(n, n^{-2/3})$ has at least $n/100$ vertex-disjoint triangles.

(b) *Simple nibble.* $G(n, Cn^{-2/3})$ has at least $0.33n$ vertex-disjoint triangles, for some constant C .

Hint in white:

C9. *Permuted correlation.* Recall that the *correlation* of two non-constant random variables X and Y is defined to be $\text{corr}(X, Y) := \text{Cov}[X, Y]/\sqrt{(\text{Var } X)(\text{Var } Y)}$.

Let $f, g \in [n] \rightarrow \mathbb{R}$ be two non-constant functions. Prove that there exist permutations π and τ of $[n]$ such that, with Z being a uniform random element of $[n]$,

$$\text{corr}(f(\pi(Z)), g(Z)) - \text{corr}(f(\tau(Z)), g(Z)) \geq \frac{2}{\sqrt{n-1}}.$$

Furthermore, show that equality can be achieved for even n .

Hint in white:

ps3 C10. Let $v_1 = (x_1, y_1), \dots, v_n = (x_n, y_n) \in \mathbb{Z}^2$ with $|x_i|, |y_i| \leq 2^{n/2}/(100\sqrt{n})$ for all $i \in [n]$. Show that there are two disjoint sets $I, J \subseteq [n]$, not both empty, such that $\sum_{i \in I} v_i = \sum_{j \in J} v_j$.

ps3* C11. Prove that there is an absolute constant $C > 0$ so that the following holds. For every prime p and every $A \subseteq \mathbb{Z}/p\mathbb{Z}$ with $|A| = k$, there exists an integer x so that $\{xa : a \in A\}$ intersects every interval of length at least Cp/\sqrt{k} in $\mathbb{Z}/p\mathbb{Z}$.

ps3* C12. Prove that there is a constant $c > 0$ so that every hyperplane containing the origin in \mathbb{R}^n intersects at least c -fraction of the 2^n closed unit balls centered at $\{-1, 1\}^n$.

D. CHERNOFF BOUND

D1. Prove that with probability $1 - o(1)$ as $n \rightarrow \infty$, every bipartite subgraph of $G(n, 1/2)$ has at most $n^2/8 + 10n^{3/2}$ edges.

ps3

D2. *Unbalancing lights*. Prove that there is a constant C so that for every positive integer n , one can find an $n \times n$ matrix A with $\{-1, 1\}$ entries, so that for all vectors $x, y \in \{-1, 1\}^n$, $|y^T Ax| \leq Cn^{3/2}$.

ps3

D3. Prove that there exists a constant $c > 1$ such that for every n , there are at least c^n points in \mathbb{R}^n so that every triple of points form a triangle whose angles are all less than 61° .

ps3

D4. *Planted clique*. Give a deterministic polynomial-time algorithm for the following task so that it succeeds over the random input with probability approaching 1 as $n \rightarrow \infty$.

Input: some unlabeled n -vertex G created as the union of $G(n, 1/2)$ and a clique on $t = \lfloor 100\sqrt{n \log n} \rfloor$ vertices.

Output: a clique in G of size t .

D5. *Weighing coins*. You are given n coins, each with one of two known weights, but otherwise indistinguishable. You can use a scale that outputs the combined weight of any subset of the coins. You must decide in advance which subsets $S_1, \dots, S_k \subseteq [n]$ of the coins to weigh. We wish to determine the minimum number of weighings needed to identify the weight of every coin. (Below, X and Y represent two possibilities for which coins are of the first weight.)

ps3*

(a) Prove that if $k \leq 1.99n/\log_2 n$ and n is sufficiently large, then for every $S_1, \dots, S_k \subseteq [n]$, there are two distinct subsets $X, Y \subseteq [n]$ such that $|X \cap S_i| = |Y \cap S_i|$ for all $i \in [k]$.

(There is a neat solution to part (a) using information theory, though here you are explicitly asked to solve it using the Chernoff bound.)

ps3*

(b) Show that there is some constant C such that (a) is false if 1.99 is replaced by C . (What is the best C you can get?)

E. LOVÁSZ LOCAL LEMMA

ps3

E1. Show that it is possible to color the edges of K_n with at most $3\sqrt{n}$ colors so that there are no monochromatic triangles.

E2. Prove that it is possible to color the vertices of every k -uniform k -regular hypergraph using at most $k/\log k$ colors so that every color appears at most $O(\log k)$ times on each edge.

ps3*

E3. *Hitting thin rectangles*. Prove that there is a constant $C > 0$ so that for every sufficiently small $\epsilon > 0$, one can choose exactly one point inside each grid square $[n, n+1) \times [m, m+1) \subset \mathbb{R}^2$, $m, n \in \mathbb{Z}$, so that every rectangle of dimensions ϵ by $(C/\epsilon) \log(1/\epsilon)$ in the plane (not necessarily axis-aligned) contains at least one chosen point.

ps4

E4. *List coloring*. Prove that there is some constant $c > 0$ so that given a graph and a set of k acceptable colors for each vertex such that every color is acceptable for at most ck neighbors of each vertex, there is always a proper coloring where every vertex is assigned one of its acceptable colors.

E5. Prove that, for every $\epsilon > 0$, there exist ℓ_0 and some $(a_1, a_2, \dots) \in \{0, 1\}^{\mathbb{N}}$ such that for every $\ell > \ell_0$ and every $i > 1$, the vectors $(a_i, a_{i+1}, \dots, a_{i+\ell-1})$ and $(a_{i+\ell}, a_{i+\ell+1}, \dots, a_{i+2\ell-1})$ differ in at least $(\frac{1}{2} - \epsilon)\ell$ coordinates.

ps4 E6. *Avoiding periodically colored paths.* Prove that for every Δ , there exists k so that every graph with maximum degree at most Δ has a vertex-coloring using k colors so that there is no path of the form $v_1 v_2 \dots v_{2\ell}$ (for any positive integer ℓ) where v_i has the same color as $v_{i+\ell}$ for each $i \in [\ell]$. (Note that vertices on a path must be distinct.)

ps4 E7. Prove that every graph with maximum degree Δ can be properly edge-colored using $O(\Delta)$ colors so that every cycle contains at least three colors.

(An edge-coloring is *proper* if it never assigns the same color to two edges sharing a vertex.)

ps4★ E8. Prove that for every Δ , there exists g so that every bipartite graph with maximum degree Δ and girth at least g can be properly edge-colored using $\Delta + 1$ colors so that every cycle contains at least three colors.

ps4★ E9. Prove that for every positive integer r , there exists C_r so that every graph with maximum degree Δ has a *proper* vertex coloring using at most $C_r \Delta^{1+1/r}$ colors so that every vertex has at most r neighbors of each color.

E10. *Vertex-disjoint cycles in digraphs.* (Recall that a directed graph is k -regular if all vertices have in-degree and out-degree both equal to k . Also, cycles cannot repeat vertices.)

ps4 (a) Prove that every k -regular directed graph has at least $ck/\log k$ vertex-disjoint directed cycles, where $c > 0$ is some constant.

ps4★ (b) Prove that every k -regular directed graph has at least ck vertex-disjoint directed cycles, where $c > 0$ is some constant.

Hint in white:

E11. (a) *Generalization of Cayley's formula.* Using Prüfer codes, prove the identity

$$x_1 x_2 \cdots x_n (x_1 + \cdots + x_n)^{n-2} = \sum_T x_1^{d_T(1)} x_2^{d_T(2)} \cdots x_n^{d_T(n)}$$

where the sum is over all trees T on n vertices labeled by $[n]$ and $d_T(i)$ is the degree of vertex i in T .

(b) Let F be a forest with vertex set $[n]$, with components having f_1, \dots, f_s vertices so that $f_1 + \cdots + f_s = n$. Prove that the number of trees on the vertex set $[n]$ that contain F is exactly $n^{n-2} \prod_{i=1}^s (f_i/n^{f_i-1})$.

(c) *Independence property for uniform spanning tree of K_n .* Show that if H_1 and H_2 are vertex-disjoint subgraphs of K_n , then for a uniformly random spanning tree T of K_n , the events $H_1 \subseteq T$ and $H_2 \subseteq T$ are independent.

ps4★ (d) *Packing rainbow spanning trees.* Prove that there is a constant $c > 0$ so that for every edge-coloring of K_n where each color appears at most cn times, there exist at least cn edge-disjoint spanning trees, where each spanning tree has all its edges colored differently.

(In your submission, you may assume previous parts without proof.)

The next two problems use the lopsided local lemma.

ps4 E12. *Packing two copies of a graph.* Prove that there is a constant $c > 0$ so that if H is an n -vertex m -edge graph with maximum degree at most cn^2/m , then one can find two edge-disjoint copies of H in the complete graph K_n .

- ps4★ E13. *Packing Latin transversals.* Prove that there is a constant $c > 0$ so that every $n \times n$ matrix where no entry appears more than cn times contains cn disjoint Latin transversals.