18.218 PROBLEM SET

Instructions:

- All submissions must be **typed in LaTeX** and submitted as PDF on Stellar (try Overleaf if you are looking for an online LaTeX editor without requiring installations). Please name your file `ps#_Lastname_Firstname.pdf` and remember to include your name in each file.
- At the top of each submission, you must **acknowledge all references and people consulted** (other than lectures and the textbook). Failure acknowledge sources will lead to an automatic 10% penalty. Examples include: names of people you discussed homework with, books, and online resources. If you consulted no additional sources, you should write sources consulted: none.
- Please turn in the problems marked $\boxed{\text{ps1}}$ and $\boxed{\text{ps1}\star}$ for problem set 1, etc., by midnight of each due date (see course homepage for due dates). Do not submit the other problems—they are for you to practice. See course homepage for policies (20% per day late penalty; do not look up solutions; collaboration policy).
- **Bonus problems**, marked by $\star$, are more challenging. A grade of A- may be attained by only solving the non-starred problems. To attain a grade of A or A+, you should solve a substantial number of starred problems.
- Please **do not exceed one page** for each unstarred problem (standard 1-inch margins and 11pt font). If you cannot fit your solution within one page, then think about how to better distill your ideas. If necessary, you may skip details of routine calculations (we don't really want to read them either).
- This file will be updated constantly as the term progresses. Please check back regularly. It will be announced when each problem set is complete.
- You are encouraged to include figures whenever they are helpful. Here are some recommended ways to produce figures in decreasing order of learning curve difficulty:
  (1) TikZ or other drawing script
  (2) IPE (which supports LaTeX) or other drawing app
  (3) photo/scan (I recommend the Dropbox app on your phone, which has a nice scanning feature that produces clear monochrome scans)

Problems begin on the next page

Last updated: May 13, 2019

1. 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 $n = \left(\frac{1}{e\sqrt{2}} + o(1)\right) k2^{k/2}$.
   (b) For each k, the maximum value of $n - \binom{n}{k}2^{1-\binom{k}{2}}$ as n ranges over positive integers is $\left(\frac{1}{e} + o(1)\right) k2^{k/2}$.
   (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 $n = \left(\frac{\sqrt{2}}{e} + o(1)\right) k2^{k/2}$.
2. 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 $c > 0$.
3. (Extension 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.
4. ps1 Let $A_1, \ldots, A_m$ be r-element sets and $B_1, \ldots, B_m$ 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 $m \leq (r+s)^{r+s}/(r^r s^s)$.
5. ps1 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 $S = S_1 \cup \cdots \cup S_r$ 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 $1 \leq i \leq r$.
6. ps1 Prove that every set of 10 points in the plane can be covered by a union of disjoint unit disks.
7. ps1⋆ Prove that $[n]^d$ cannot be partitioned into fewer than 2^d sets each of the form $A_1 \times \cdots \times A_d$ where $A_i \subsetneq [n]$.
8. 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.
9. ps1 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:)
10. 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 $(1, 5, 2, 4, 3)$.)
11. Let G be a graph with n vertices and m edges. Prove that K_n can be written as a union of $O(n^2(\log n)/m)$ copies of G (not necessarily edge-disjoint).
12. ps1 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 $|\mathcal{F}| = O(k2^k \log n)$ such that for every k-element subset A of $[n]$, $\mathcal{F}|_A$ contains all 2^k subsets of A.
13. Let A be a 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).

14. Let $\boldsymbol{r} = (r_1, \dots, r_k)$ be a vector of nonzero integers whose sum is nonzero. Prove that there exists a real $c > 0$ (depending on $\boldsymbol{r}$ 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 $r_1 b_1 + \cdots + r_k b_k = 0$.

ps1 15. 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 $|B_1| + |B_2| > 2n/3$. Can you do it if A is a set of nonzero reals?

ps1⋆ 16. Prove that every graph with n vertices and $m \geq n^{3/2}$ edges contains a pair of vertex-disjoint and isomorphic subgraphs (not necessarily induced) each with at least $cm^{2/3}$ edges, where $c > 0$ is a constant.

17. Let $M(n)$ denote the maximum number of edges in a 3-uniform hypergraph on n vertices without a clique on 4 vertices.
    (a) Prove that $M(n+1)/\binom{n+1}{3} \leq M(n)/\binom{n}{3}$ for all n, and conclude that $M(n)/\binom{n}{3}$ approaches some limit α as $n \to \infty$.
    (This limit is called the *Turán density* of the hypergraph $K_4^{(3)}$, and its exact value is currently unknown and is a major open problem.)
    (b) Prove that for every $\delta > 0$, there exists $\epsilon > 0$ and n_0 so that every 3-uniform hypergraph with $n \geq n_0$ vertices and at least $(\alpha + \delta)\binom{n}{3}$ edges must contain at least $\epsilon\binom{n}{4}$ copies of the clique on 4 vertices.

18. Using the alteration method, prove that the Ramsey number $R(4, k)$ satisfies $R(4, k) \geq c(k/\log k)^2$ for some constant $c > 0$.

19. 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.

20. (Zarankiewicz problem) Prove that for every positive integer $k \geq 2$, there exists a constant $c > 0$ such that for every n, there exists an $n \times n$ matrix with $\{0, 1\}$ entries, with at least $cn^{2-2/(k+1)}$ 1's, such that there is no $k \times k$ submatrix consisting of all 1's.

ps2 21. Fix k. Prove that there exists a constant $c_k > 1$ so that for every sufficiently large n, 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 $[n] \setminus F_i$.

22. *Acute sets in* $\mathbb{R}^n$
    (a) Prove that there exists a family of $\Omega((2/\sqrt{3})^n)$ subsets of $[n]$ containing no three distinct members A, B, C satisfying $A \cap B \subseteq C \subseteq A \cup B$.
    (b) Prove that there exists a set of $\Omega((2/\sqrt{3})^n)$ 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 (c) 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 the angle spanned by every three distinct points is at most $61°$.

ps2⋆ 23. *Covering complements of sparse graphs by cliques*

(a) 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.

(b) 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⋆ 24. Let $G = (V, E)$ be a graph with n vertices and minimum degree $\delta \geq 2$. Prove that there exists $A \subseteq V$ with $|A| \leq Cn(\log \delta)/\delta$, where $C > 0$ is a constant, so that every vertex in $V \setminus A$ contains at least one neighbor in A and at least one neighbor not in A.

25. Let X be a nonnegative real-valued random variable. Suppose $\mathbb{P}(X = 0) < 1$. Prove that

$$\mathbb{P}(X = 0) \leq \frac{\operatorname{Var} X}{\mathbb{E}[X^2]}.$$

ps2 26. Let X be a random variable with mean μ and variance σ^2. Prove that for all $\lambda > 0$,

$$\mathbb{P}(X \geq \mu + \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

27. What is the threshold function for $G(n, p)$ to contain a cycle?

ps2 28. Show that, for each fixed k, there is a sequence p_n such that

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

29. Let $p = (\log n + f(n))/n$. Show that, as $n \to \infty$,

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

ps2 30. Let $v_1 = (x_1, y_1), \ldots, v_n = (x_n, y_n) \in \mathbb{Z}^2$ 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 $\sum_{i \in I} v_i = \sum_{j \in J} v_j$.

ps2⋆ 31. 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 $\mathbb{Z}/p\mathbb{Z}$.

ps2⋆ 32. Let $S_1, \ldots, S_k$ be subsets of $[n]$. Prove that if $k \leq 1.99n/\log_2 n$ and n is sufficiently large, then there are two distinct subsets $X, Y \subseteq [n]$ such that $|X \cap S_i| = |Y \cap S_i|$ for all $i \in [k]$.

In addition, show that there is some constant C such that the claim is false for $k \geq Cn/\log_2 n$. What is the best constant C?

ps2⋆ 33. Let X be a collection of pairwise orthogonal unit vectors in $\mathbb{R}^n$ and suppose that the projection of each of these vectors on the first k coordinates has norm at least ϵ. Show that $|X| \leq k/\epsilon^2$, and show that this is tight if $\epsilon^2 = k/2^r < 1$ for some integer r.

ps2⋆ 34. 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 $\{-1, 1\}^n$.
(Give your best c. Can you get $c \geq 3/8$? It is conjectured that $c = 1/2$ works.)

ps2 35. Prove that, with probability approaching 1 as $n \to \infty$, $G(n, n^{-1/2})$ has at least $cn^{3/2}$ edge-disjoint triangles, where $c > 0$ is some constant.
(Hint in white:)

ps2 36. *Simple nibble.* Prove that for some constant C, with probability approaching 1 as $n \to \infty$,
(a) $G(n, Cn^{-2/3})$ has at least $n/100$ vertex-disjoint triangles.
(b) $G(n, Cn^{-2/3})$ has at least $0.33n$ vertex-disjoint triangles
(Hint in white:)
(You are asked to solve the above problem using the second moment method. Later in the course we will learn a different method to solve this problem.)

37. Let $X \sim \text{Binomial}(n, p)$. Prove that for $0 < q \le p < 1$,

$$\mathbb{P}(X \le nq) \le e^{-nH(q\|p)} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X \le nq) = -H(\|p)$$

and for $0 < p \le q < 1$,

$$\mathbb{P}(X \ge nq) \le e^{-nH(q\|p)} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X \ge nq) = -H(\|p),$$

where

$$H(q\|p) := q \log \frac{q}{p} + (1-q) \log \frac{1-q}{1-p}.$$

is known as the *relative entropy* or *Kullback-Leibler divergence*, in this case, between two Bernoulli distributions.

38. Prove that there is a constant $C > 0$ so that, with probability $1 - o(1)$ as $n \to \infty$, the maximum number of edges in a bipartite subgraph of $G(n, 1/2)$ is at most $n^2/8 + Cn^{3/2}$.

39. (a) Prove that there is some constant $c > 1$ so that there exists $S \subset \{0,1\}^n$ with $|S| \ge c^n$ so that every pair of points in S differ in at least $n/4$ coordinates.
(b) Prove that there is some constant $c > 1$ so that the the unit sphere in $\mathbb{R}^n$ contains at least c^n points, where each pair of points is at distance at least 1 apart.

ps3 40. *Planted clique.* Give a deterministic polynomial-time algorithm solving the following problem so that it succeeds over the random input with probability approaching 1 as $n \to \infty$:
Input: an n-vertex unlabeled graph G created as the union of $G(n, 1/2)$ and a clique on vertex subset of size $t = \lfloor 100\sqrt{n \log n} \rfloor$
Output: a clique in G of size t

ps3 41. 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.

42. Prove that there is some constant C so 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 edge has at most $C \log k$ vertices of each color.

ps3 43. 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.

ps3⋆ 44. 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.

ps3 45. Prove that, for every $\epsilon > 0$, there exists 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.

ps3 46. A *periodic path* in a graph G with respect to a vertex coloring $f\colon V(G) \to [k]$ is a path $v_1v_2\dots v_{2\ell}$ for some positive integer ℓ with $f(v_i) = f(v_{i+\ell})$ for each $i \in [\ell]$ (reminder: no repeated vertices allowed in a path).

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 with no periodic paths.

ps3 47. 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.

(A *proper edge-coloring* is one where no two adjacent edges receive the same color.)

ps3⋆ 48. 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.

ps3⋆ 49. 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.

50. Let $H = (V, E)$ be a hypergraph satisfying, for some $\lambda > 1/2$,

$$\sum_{f \in E : v \in f} \lambda^{|f|} \le \frac{1}{2} - \frac{1}{4\lambda} \quad \text{for every } v \in V$$

(here $|f|$ is then number of vertices in the edge f). Prove that H is 2-colorable.

51. Prove that there exists k_0 and a red/blue coloring of $\mathbb{Z}$ without any monochromatic k-term arithmetic progressions with $k \ge k_0$ and common difference less than 1.99^k.

52. *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.)

ps3 (a) Prove that every k-regular directed graph has at least $ck/\log k$ vertex-disjoint directed cycles, where $c > 0$ is some constant.

ps3⋆ (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:)

ps3⋆ 53. 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.

(Hint in white:)

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

$$x_1x_2\cdots x_n(x_1 + \cdots + x_n)^{n-2} = \sum_T x_1^{d_T(1)} x_2^{d_T(2)} \cdots x_n^{d_T(n)}$$

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) *Independence property for uniform spanning tree of K_n.* 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 $n^{n-2} \prod_{i=1}^{s} (f_i / n^{f_i - 1})$. Deduce 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.

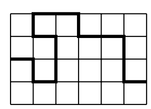
ps3⋆ (c) *Packing rainbow spanning trees.* Prove that there is a constant $c > 0$ so that for every edge-coloring of K_n where each color appears at most cn times, there exist at least cn edge-disjoint spanning trees, where each spanning tree has all its edges colored differently.

ps4 55. Let $G = (V, E)$ be a graph. Color every edge with red or blue independently and uniformly at random. Let E_0 be the set of red edges and E_1 the set of blue edges. Let $G_i = (V, E_i)$ for each $i = 0, 1$. Prove or disprove:

$$\mathbb{P}(G_0 \text{ and } G_1 \text{ are both connected}) \le \mathbb{P}(G_0 \text{ is connected})^2.$$

ps4 56. A set family $\mathcal{F}$ is *intersecting* if $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{F}$. Let $\mathcal{F}_1, \dots, \mathcal{F}_k$ each be a collection of subsets of $[n]$ and suppose that each $\mathcal{F}_i$ is intersecting. Prove that $\left|\bigcup_{i=1}^{k} \mathcal{F}_i\right| \le 2^n - 2^{n-k}$.

57. Let $G_{m,n}$ be the grid graph on vertex set $[m] \times [n]$ (m vertices wide and n vertices tall). A *horizontal crossing* is a path that connects some left-most vertex to some right-most vertex. See below for an example of a horizontal crossing in $G_{7,5}$.



Let $H_{m,n}$ denote the random subgraph of $G_{m,n}$ obtained by keeping every edge with probability $1/2$ independently.

Let $\mathsf{RSW}(k)$ denote the following statement: there exists a constant $c_k > 0$ such that for all positive integers n, $\mathbb{P}(H_{kn,n} \text{ has a horizontal crossing}) \ge c_k$.

ps4 (a) Prove that $\mathsf{RSW}(2)$ implies $\mathsf{RSW}(100)$.

ps4⋆ (b) Prove $\mathsf{RSW}(1)$.

(c) (Challenging. Not to be turned in) Prove $\mathsf{RSW}(2)$.

ps4 58. Let U_1 and U_2 be increasing events and D a decreasing event of independent boolean random variables. Suppose U_1 and U_2 are independent. Prove that $\mathbb{P}(U_1 | U_2 \cap D) \le \mathbb{P}(U_1 | U_2)$.

ps4 59. *Coupon collector.* Let $s_1, \dots, s_m$ be independent random elements in $[n]$ (not necessarily uniform or identically distributed; chosen with replacement) and $S = \{s_1, \dots, s_m\}$. Let I and J be disjoint subsets of $[n]$. Prove that $\mathbb{P}(I \cup J \subseteq S) \le \mathbb{P}(I \subseteq S)\mathbb{P}(J \subseteq S)$.
(Hint in white:)

ps4⋆ 60. Prove that there exist $c, \epsilon > 0$ such that if $A_1, \dots, A_k$ are increasing events of independent boolean random variables with $\mathbb{P}(A_i) \le \epsilon$ for all i, then, letting X denote the number of events A_i that occur, one has $\mathbb{P}(X = 1) \le 1 - c$. (Give your largest c.)

ps4 61. Prove that with probability $1 - o(1)$, the size of the largest subset of vertices of $G(n, 1/2)$ inducing a triangle-free subgraph is $\Theta(\log n)$.

62. *Lower tails of small subgraph counts.* Fix graph H and $\epsilon \in (0,1]$. Let X_H denote the number of copies of H in $G(n,p)$. Prove that for all n and $0 < p < 1/2$,

$$\mathbb{P}(X_H \le (1-\epsilon)\mathbb{E}X_H) = e^{-\Theta_{H,\epsilon}(\Phi_H)} \quad \text{where } \Phi_H := \min_{H' \subseteq H : e(H')>0} n^{v(H')} p^{e(H')}.$$

Here the hidden constants in $\Theta_{H,\epsilon}$ may depend on H and ϵ (but not on n and p).

63. *Vertex-disjoint triangles in $G(n,p)$ again.* Using Janson inequalities this time, give another solution to Problem 36 in the following generality.

ps4 (a) Prove that for every $\epsilon > 0$, there exists $C_\epsilon > 0$ such that such that with probability $1-o(1)$, $G(n, C_\epsilon n^{-2/3})$ contains at least $(1/3-\epsilon)n$ vertex-disjoint triangles.

ps4⋆ (b) Compare the the dependence of the optimal C_ϵ on ϵ you obtain using the method in Problem 36 versus this problem (don't worry about leading constant factors).

ps4⋆ 64. Show that $\operatorname{ch}(G(n,1/2)) = (1+o(1))\frac{n}{2\log_2 n}$ with probability $1-o(1)$.
Here $\operatorname{ch}(G)$ is the *list-chromatic number* (also called *choosability*) of a graph G and it is defined to the minimum k such that if every vertex of G is assigned a list of k acceptable colors, then there exists a proper coloring of G where every vertex is colored by one of its acceptable colors.

ps4 65. For each part, prove that there is some constant $c > 0$ so that, for all $\lambda > 0$,

$$\mathbb{P}(|X - \mathbb{E}X| \ge \lambda\sqrt{\operatorname{Var} X}) \le 2e^{-c\lambda^2}.$$

(Such families of random variables are called *sub-Gaussian.*)

(a) X is the number of triangles in $G(n,1/2)$.

(b) X is the number of inversions of a uniform random permutation of $[n]$ (an *inversion* of $\sigma \in S_n$ is a pair (i,j) with $i<j$ and $\sigma(i) > \sigma(j)$).

ps4⋆ 66. Let $k \le n/2$ be positive integers and G an n-vertex graph with average degree at most n/k. Prove that a uniform random k-element subset of the vertices of G contains an independent set of size at least ck with probability at least $1-e^{-ck}$, where $c > 0$ is a constant.

ps5 67. True or False: In the definition of a martingale, the condition $\mathbb{E}[X_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0] = x_{n-1}$ may be replaced by simply $\mathbb{E}[X_n | X_{n-1} = x_{n-1}] = x_{n-1}$.

ps5 68. Prove that for every $\epsilon > 0$ there exists $\delta > 0$ and n_0 such that for all $n \ge n_0$ and $S_1, \dots, S_m \subset [2n]$ with $m \le 2^{\delta n}$ and $|S_i| = n$ for all $i \in [m]$, there exists a function $f \colon [2n] \to [n]$ so that $(1-e^{-1}-\epsilon)n \le |f(S_i)| \le (1-e^{-1}+\epsilon)n$ for all $i \in [m]$.

ps5 69. *Simultaneous bisections.* Fix Δ. Let $G_1, \dots, G_m$ with $m = 2^{o(n)}$ be connected graphs of maximum degree at most Δ on the same vertex set V with $|V| = n$. Prove that there exists a partition $V = A \cup B$ so that every G_i has $(1+o(1))e(G_i)/2$ edges between A and B.

ps5 70. Show that for every $\epsilon > 0$ there exists $C > 0$ so that every $S \subset [4]^n$ with $|S| \ge \epsilon 4^n$ contains four elements whose pairwise Hamming distance at least $n - C\sqrt{n}$.

ps5⋆ 71. *Tighter concentration of chromatic number*

(a) Prove that with probability $1-o(1)$, every vertex subset of $G(n,1/2)$ with at least $n^{1/3}$ vertices contains an independent set of size at least $c\log n$, where $c > 0$ is some constant.

(b) Prove that there exists some function $f(n)$ and constant C such that for all $n \ge 2$,

$$\mathbb{P}(f(n) \le \chi(G(n,1/2)) \le f(n) + C\sqrt{n}/\log n) \ge 0.99.$$

ps5⋆ 72. Let $G = (V, E)$ with chromatic number $\chi(G) = k$ and S a uniform random subset of V. Prove that for every $t \geq 0$,

$$\mathbb{P}(\chi(G[S]) \leq k/2 - t) \leq e^{-ct^2/k},$$

where $c > 0$ is a constant and $G[S]$ is the subgraph induced by S.

ps5⋆ 73. Prove that for all n there exists some $k \sim 2\log_2 n$ and some n-vertex graph that contains every graph on k vertices as an induced subgraph.

ps5⋆ 74. Prove that there exists a constant $c > 0$ so that the following holds. Let G be a d-regular graph and $v_0 \in V(G)$. Let $m \in \mathbb{N}$ and consider a simple random walk $v_0, v_1, \dots, v_m$ where each v_{i+1} is a uniform random value of v_i. For each $v \in V(G)$, let X_v be the number times that v appears among $v_0, \dots, v_m$. For that for every $v \in V(G)$ and $\lambda > 0$

$$\mathbb{P}\left(\left|X_v - \frac{1}{d}\sum_{w \in N(v)} X_w\right| \geq \lambda + 1\right) \leq 2e^{-c\lambda^2/m}$$

Here $N(v)$ is the neighborhood of v.

ps5⋆ 75. Let $\mathsf{maxcut}(G)$ denote the maximum number of edges in a bipartite subgraph of G. Prove there is a constant $c > 0$ so that $\mathsf{maxcut}(G(n, 1/2)) > n^2/8 + cn^{3/2}$ with probability $1 - o(1)$.

For the next three exercises, use Talagrand's inequality

ps5 76. Let Q be a subset of the unit sphere in $\mathbb{R}^n$. Let $\boldsymbol{x} \in [0,1]^n$ be a random vector with independent coordinates. Let $X = \sup_{\boldsymbol{q} \in Q} \langle \boldsymbol{x}, \boldsymbol{q} \rangle$ and m a median of X. Let $t > 0$. Prove

$$\mathbb{P}(|X - m| \geq t) \leq 4e^{-t^2/4}.$$

ps5⋆ 77. Prove that there are constants $c, C > 0$ such that if A is a symmetric $n \times n$ matrix with independent entries in $[-1, 1]$, then the second largest eigenvalue $\lambda_2(A)$ satisfies

$$\mathbb{P}(|\lambda_2(A) - \mathbb{E}\lambda_2(A)| > t) \leq Ce^{-ct^2}.$$

(Hint: use this Courant–Fischer characterization of $\lambda_2(X)$: for every pair of unit vectors $u, v \in \mathbb{R}^n$, there exist $a, b \in \mathbb{R}$ with $a^2 + b^2 = 1$ and $w = au + bv$ satisfying $w^t X w \leq \lambda_2(X)$.)

78. Let $q = q_n \gg n$. Let $\boldsymbol{x} = (x_1, \dots, x_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$ be two random sequences whose entries are chosen independently and uniformly at random from $[q]$. Let X be the length of the longest common subsequence between $\boldsymbol{x}$ and $\boldsymbol{y}$ (i.e., X is the largest k such that there exist $i_1 < \cdots < i_k$ and $j_1 < \cdots < j_k$ with $x_{i_1} = y_{j_1}, \dots, x_{i_k} = y_{j_k}$). Show that with probability $1 - o(1)$, X lies within $\sqrt{n}$ of its median.

Entropy methods (You are encouraged to find solutions using entropy)

79. (Submodularity) Prove that $H(X, Y, Z) + H(X) \leq H(X, Y) + H(X, Z)$.

ps5⋆ 80. (Uniquely decodable codes) Let $[r]^*$ denote the set of all finite strings of elements in $[r]$. Let A be a finite subset of $[r]^*$ and suppose no two distinct concatenations of sequences in A can produce the same string. Prove that $\sum_{a \in A} r^{-|a|} \leq 1$ where $|a|$ is the length of $a \in A$.

ps5 81. Let $\mathcal{G}$ be a family of graphs on vertices labeled by $[2n]$ such that the intersection of every pair of graphs in $\mathcal{G}$ contains a perfect matching. Prove that $|\mathcal{G}| \leq 2^{\binom{2n}{2} - n}$.

$\boxed{\text{ps5}\star}$ 82. Let X, Y, Z be independent $\mathbb{Z}$-valued random variables. Prove that

$$2H(X+Y+Z) \leq H(X+Y) + H(X+Z) + H(Y+Z).$$

$\boxed{\text{ps5}\star}$ 83. *Triangles versus vees in a directed graph.* Let V be a finite set, $E \subseteq V \times V$, and

$$\triangle = \left|\left\{(x,y,z) \in V^3 : (x,y),(y,z),(z,x) \in E\right\}\right|$$

(i.e., cyclic triangles; note the direction of edges) and

$$\wedge = \left|\left\{(x,y,z) \in V^3 : (x,y),(x,z) \in E\right\}\right|.$$

Prove that $\triangle \leq \wedge$.