Olympiad Combinatorics

Pranav A. Sriram

August 2014

Copyright notices

About the Author

Pranav Sriram graduated from high school at The International School Bangalore, India, and is currently pursuing undergraduate studies at Stanford University.

9. THE PROBABILISTIC METHOD

Introduction

Our final chapter will focus on an idea that has had a tremendous impact on combinatorics over the past sixty years, and that is playing a critical role in the "big data" driven applications of today's digitized world. The *probabilistic method* is a technique that broadly refers to using arguments based on probability to prove results in fields not necessarily directly related to probability. Probabilistic combinatorics is an extremely active area of research today, and several, if not most, recent developments in graph theory, extremal combinatorics, computational geometry, and combinatorial algorithms have relied extensively on probabilistic arguments. In addition, the probabilistic method has led to breakthroughs in number theory and additive group theory, and has surprisingly elegant connections to information theory and the theory of error correcting codes. Modern paradigms for efficiently monitoring computer network traffic, finding hidden patterns in large datasets, and solving extremely large scale mathematical problems have leveraged the power of the probabilistic method in advanced ways.

Each section in this chapter illustrates one general method of solving combinatorial problems via probabilistic arguments. We will assume knowledge of basic concepts from high school probability, such as the notion of random variables, expected value, variance, independent events, and mutually exclusive events. No formal background in Lebesgue integration or real analysis is required.

Linearity of Expectation

The linearity of expectation refers to the following (fairly intuitive) principle: if X is the sum of n random variables X_1, X_2, …, X_n, then the expected value of X is equal to the sum of expected values of X_1, X_2, …, X_n. In symbols,

$$E[X] = E[X_1 + X_2 + \dots + X_n] = E[X_1] + E[X_2] \dots + E[X_n].$$

This holds even when the variables X_1, X_2, …, X_n are not independent from each other. Another intuitive principle we will use, sometimes referred to as the "pigeonhole property" (PHP) of expectation, is that a random variable cannot always be less than its average. Similarly, it cannot always be more than its average. For example, if the average age at a party is 24 years, there must be at least one person at most 24 years old, and one person who is at least 24 years old.

How can we use this to solve combinatorial problems?

Example 1 [Szele's Theorem]

For each n, show that there exists a tournament on n vertices having at least $\frac{n!}{2^{n-1}}$ Hamiltonian paths.

Answer:
The basic idea is as follows: instead of explicitly constructing a tournament with the required properties, we use a *randomized procedure* for creating some tournament T. We will show that *on average*, the tournament we (randomly) construct contains $\frac{n!}{2^{n-1}}$ Hamiltonian paths. Then by the pigeonhole principle, *some* tournament must contain at least $\frac{n!}{2^{n-1}}$ Hamiltonian paths.

The randomized procedure is to construct a tournament where each of the matches is decided independently by a flip of a fair coin. Thus for two vertices u and v, the probability that u lost to v is 0.5 and vice versa. To compute the expected total number of Hamiltonian paths, observe that each Hamiltonian path corresponds to a permutation of the vertices of the graph.

Consider a random permutation of the vertices $(v_1, v_2, v_3, \ldots, v_n)$. For this sequence (in order) to form a Hamiltonian path, we need to have v_1 lost to v_2, v_2 lost to v_3, …, v_{n-1} lost to v_n. Each of these happens with probability 0.5, so the probability that they all occur is $(0.5)^{n-1}$. Thus, each permutation of the vertices has probability $(0.5)^{n-1}$ of being a Hamiltonian path, and there are $n!$ permutations that can give rise to Hamiltonian paths. Hence the **expected** total number of Hamilton paths, by the linearity of expectation, is $n!/2^{n-1.}$ It follows that *some* tournament must have at least the expected number of Hamiltonian paths. ■

Example 2 [Generalization of MOP 2010]
Let $\boldsymbol{G}$ be a graph with E edges and n vertices with degrees $d_1, d_2, \ldots, d_n$. Let k be an integer with $k \leq \frac{2E}{n}$. Show that $\boldsymbol{G}$ contains an induced subgraph $\boldsymbol{H}$ such that $\boldsymbol{H}$ contains no $\boldsymbol{K}_{k+1}$ and $\boldsymbol{H}$ has at least $\frac{kn^2}{2E+n}$ vertices. (Recall that a $\boldsymbol{K}_{k+1}$ is a complete graph on k+1 vertices.)

Answer:

The basic idea

Note that it if all the vertices in $\boldsymbol{H}$ have degree at most k-1 (in $\boldsymbol{H}$), then $\boldsymbol{H}$ clearly cannot contain a $\boldsymbol{K}_{k+1}$. This gives us the following idea for a greedy construction.

Take a random permutation of the vertices $v_1v_2\ldots v_n$. Starting from v_1 and continuing, we select vertex v_i if and only if amongst v_i and its neighbors, v_i is one of the first k appearing in the sequence. This ensures that every chosen vertex has at most k-1 chosen neighbors, so the final set of chosen vertices will not contain an induced $\boldsymbol{K}_{k+1}$. Then we just need to show that the expected number of chosen vertices is at least $\frac{kn^2}{2E+n}$.

Computing the expectation

Now, each vertex v_i will be chosen with probability $\frac{k}{d_i+1}$, where d_i is its degree. This is because v_i is selected if and only if amongst v_i and its d_i neighbors, v_i is amongst the first k in the permutation. The expected number of chosen vertices is hence $\sum_{i=1}^{n} \frac{k}{d_i+1}$.

And the rest is just algebra…

The function 1/(x+1) is convex (check this). Hence by Jensen's inequality, $\sum_{i=1}^{n} \frac{1}{d_i+1} \geq n\frac{1}{d+1}$, where d is the average degree ($\frac{\sum_{i=1}^{n} d_i}{n}$). Finally using $d = \frac{2E}{n}$ and multiplying by k on both sides we get the result. ■

Remark 1: Note that "$\frac{kn^2}{2E+n}$" in the problem can be replaced by "$\sum_{i=1}^{n} \frac{k}{d_i+1}$". Also note that $\frac{kn^2}{2E+n}$ is equal to $\frac{kn}{d+1}$, where d is the average degree of a vertex, that is, d = 2E/n. **Both results are extremely useful.** The summation form is obviously more useful in situations when we have (or can find) some information about the degrees of individual vertices in the graph we are analyzing (see example 5), whereas the other two forms are useful when we have some information about the total number of edges in a graph.

Corollary 1: Any graph $\boldsymbol{G}$ with n vertices and E edges contains an independent set containing at least $\frac{n^2}{2E+n}$ vertices. (Recall that an independent set is a set of vertices such that no two vertices in that set have an edge between them.)

Proof: Note that an independent set can be interpreted as a $\boldsymbol{K}_2$-free graph. Apply the above result for k = 1, and the corollary follows. ■

Corollary 2: A tree $\boldsymbol{T}$ on n vertices has an independent set of size greater than $n/3$.

Proof: Apply corollary 1, taking $E = n - 1$. ■
(Exercise: what is the best constant c such that a tree always has an independent set of size at least n/c?)

Remark 2: The above problem is representative of a typical class of problems in extremal graph theory. These problems ask us to show that in graphs under certain conditions there exist sufficiently large or small subgraphs with certain properties.

Example 3 [dominating sets]
Show that if the minimum degree in an n vertex graph $\boldsymbol{G}$ is $\delta > 1$, then $\boldsymbol{G}$ has a *dominating set* containing at most $n\,\frac{1+\ln(\delta+1)}{\delta+1}$ vertices.

Answer:
Form a subset $\boldsymbol{S}$ of the vertices of $\boldsymbol{G}$ by choosing each vertex with probability p, where p is a parameter in (0, 1) we will specify later. Let $\boldsymbol{T}$ be the set of vertices that have no neighbor in $\boldsymbol{S}$. Then $\boldsymbol{S} \cup \boldsymbol{T}$ is a dominating set. We now estimate its expected size.

Clearly $E[|S|]$ = np. The probability that a vertex v is in $\boldsymbol{T}$ is the probability that neither v nor any of its neighbors is in $\boldsymbol{S}$. This is at most $(1-p)^{\delta+1}$, since v has at least d neighbors. Thus $E[|T|] \leq n(1-p)^{\delta+1} \leq ne^{-p(\delta+1)}$.

Finally, $E[|\boldsymbol{S} \cup \boldsymbol{T}|] \leq np + ne^{-p(\delta+1)}$. To minimize this expression, we choose $p = \frac{\ln(d+1)}{d+1}$, and we get $E[|\boldsymbol{S} \cup \boldsymbol{T}|] \leq n\,\frac{1+\ln(\delta+1)}{\delta+1}$. There exists some $\boldsymbol{S}$ and $\boldsymbol{T}$ such that $|\boldsymbol{S} \cup \boldsymbol{T}|$ is at most the expected value, and we are done. ■

Remark 1: The value of p minimizing the expression np + ne$^{-p(d+1)}$ is found by choosing p so that the derivative with respect to p is 0. Remember that p is a parameter we are free to choose, unlike n and d. We could have also directly chosen p to be the value minimizing np + $n(1-\mathrm{p})^{d+1}$, and we could have skipped the step where we estimate $n(1-p)^{d+1} \leq n\mathrm{e}^{-p(d+1).}$ Although this would give a slightly sharper bound, it would have been a lot uglier. Anyway both bounds asymptotically approach $\frac{n\ln d}{d}$.

Remark 2: The basic idea in this proof is quite typical. Pick something randomly, and then make adjustments for whatever got missed.

Example 4 [USAMO 2010-6]

A blackboard contains 68 pairs of nonzero integers. Suppose that for no positive integer k do both the pairs (k, k) and $(-k, -k)$ appear on the blackboard. A student erases some of the 136 integers, subject to the condition that no two erased integers may add to 0. The student then scores one point for each of the 68 pairs in which at least one integer is erased. Determine, with proof, the largest number N of points that the student can guarantee to score regardless of which 68 pairs have been written on the board.

Note: The 68 pairs need not all be distinct; some may be repeated.

Answer:
Note that if (j, j) occurs then $(-j, -j)$ does not, so we can WLOG assume that if (j, j) occurs then $j > 0$ (by replacing j by $(-j)$).

Now for each integer $k > 0$, we can either delete all appearances of k on the board (if any) OR all appearances of $(-k)$ (if any), but not both. So, for each $k > 0$, we delete all k's with probability p, and otherwise delete all $(-k)$'s, where p is a parameter in $(1/2, 1)$ to be specified later. We now consider 3 types of pairs of numbers that can occur, and in each case bound the probability that we score a point.

(i) A pair of the form (k, k)
We score a point for this pair with probability p, since k is deleted with probability p.

(ii) A pair of the form $(k, -k)$
We score a point with probability 1, since we delete either k or $(-k)$

(iii) A pair of the form (a, b), where $b \neq \pm a$
We score a point with probability $(1 - P[a \text{ not deleted}] \times P[b \text{ not deleted}]) \geq (1 - p^2)$. Note that we are using $p > (1-p)$.

In all the above cases, the probability of scoring a point is **at least** $\min\{p, (1-p^2)\}$. Thus the expected number of points we score totally is at least $68 \times \min\{p, (1-p^2)\}$. This quantity is maximized by setting $p = \frac{\sqrt{5}-1}{2}$, and at this point the expectation is at least $68 \times p = 68 \times \frac{\sqrt{5}-1}{2}$, which is greater than 42. Therefore, it is always possible to score **at least 43 points**.

We leave it to the reader to construct an example demonstrating that it is not always possible to score 44 points.

This will show that the bound of 43 is tight, and hence the answer is 43. ■

Example 5 [IMO Shortlist 2012, C7]

There are 2^{500} points on a circle labeled 1, 2, …, 2^{500} in some order. Define the *value* of a chord joining two of these points as the sum of numbers at its ends. Prove that there exist 100 pairwise disjoint (nonintersecting) chords with equal values.

Answer:

Step 1: Basic observations

Let $n = 2^{499}$. There are $\binom{2n}{2}$ chords joining pairs of labeled points, and all chord *values* clearly belong to {3, 4, …, $4n$–1}. Furthermore, note that chords with a common endpoint have different values.

Step 2: Interpretation and reduction to familiar terms

Let $\boldsymbol{G_c}$ denote the graph with vertices representing chords with value c. Two vertices in $\boldsymbol{G_c}$ are neighbors if the corresponding chords intersect. (Note: don't confuse *vertices* in our graphs with *points* on the circle.) So we're basically just looking for a **large independent set** in *some* $\boldsymbol{G_c}$. We already know that by example 2, each $\boldsymbol{G_c}$ has an independent set of size at least $I(\boldsymbol{G_c}) = \sum_{v \in \boldsymbol{G_c}} \frac{1}{d_v+1}$. By the pigeonhole principle, it now suffices to show that the average of $I(\boldsymbol{G_c})$ over all values c is at least 100.

Step 3: From graphs to individual vertices

Note that the average value of $I(\boldsymbol{G_c})$ is

$$\frac{1}{4n-3}\sum_{c=3}^{4n-1} I(\boldsymbol{G_c}) = \frac{1}{4n-3}\sum_{c=3}^{4n-1}\left(\sum_{v \in \boldsymbol{G_c}} \frac{1}{d_v+1}\right)$$

The double summation on the right is nothing but the sum of $\frac{1}{d_v+1}$ over **all 2*n* vertices**, since each vertex belongs to exactly one of the $4n$−3 graphs.

Step 4: Estimating the degrees
For a given chord L, let $m(L)$ denote the number of points contained by its minor arc. $m(L) = 0$ if the chord joins consecutive points and $m(L) = n-1$ if the chord is a diameter. Clearly, a chord L can intersect at most $m(L)$ other chords.

Now for each $i \in \{0, 1, 2, ..., n-1\}$, there are exactly $2n$ chords with $m(L) = i$. Thus for each $i \in \{0, 1, 2, ..., n-1\}$, there are at least $2n$ vertices with degree at most i. Therefore, the sum of $\frac{1}{d_v+1}$ over all vertices is at least $\sum_{1=1}^{n} \frac{2n}{i}$. Finally,

$$\text{Average of } I(\boldsymbol{G}_c) \geq \frac{2n}{4n-3} \sum_{1=1}^{n} \frac{1}{i} > \frac{1}{2} \ln(n+1)$$

$$> \frac{1}{2} \ln(2^{499}) = \frac{499 \ln(2)}{2} > 249.5 \times 0.69 > 100.$$

This proves the desired result. ■

Remark: In fact, $\frac{499 \ln(2)}{2} > 172$. Even with rather loose estimates throughout, this proof improves the bound significantly from 100 to 172. More generally, 2^{500} and 100 can be replaced by $2n$ and $\frac{2n \ln(n+1)}{4n-3}$.

Example 6 [Biclique covering]

In a certain school with n students, amongst any $(\alpha+1)$ students, there exists at least one pair who are friends (friendship is a symmetric relation here). The principal wishes to organize a series of basketball matches $M_1, M_2, ..., M_j$ under certain constraints. Each match M_i is played by two teams A_i and B_i, and each team consists of 1 or more students. Furthermore, each person in A_i must be friends with each person in B_i. Finally, for any pair of friends u and v, they must appear on opposite teams in at least one match. Note that a particular student can participate in any number of matches, and different matches need not be played by a disjoint set of people. Let $|M_i|$ denote the number of students playing match M_i, and let S denote the sum of $|M_i|$ over all matches. Show that $S \geq n\log_2(n/\alpha)$.

Answer:
For each match M_i, randomly fix the winner by flipping a fair coin (note that we are free to do this as the problem mentions nothing about the winners of matches). Hence for each i, A_i wins with probability 0.5 and B_i wins with probability 0.5. Call a student a *champion* if he is on the winning team of each match that he plays.

Key observation: the total number of champions is at most α.
Proof: If there were α+1 champions, some pair must be friends (see first statement of the problem), and hence they must have played against each other in some match. But then one of them would have lost that match and would hence not be a champion; contradiction.

On the other hand, call the students $s_1, s_2, \ldots, s_n$. The probability that a student s_k becomes a champion is 2^{-m_k}, where m_k is the number of matches played by s_k. Hence the expected number of champions is $\sum_{k=1}^{n} 2^{-m_k}$. Applying the AM-GM inequality, $\sum_{k=1}^{n} 2^{-m_k} \geq n \times 2^{-(\sum_{k=1}^{n} m_k)/n} = n \times 2^{-S/n}$. But the number of champions is bounded by α, so the expectation obviously cannot exceed α. Hence $n \times 2^{-S/n} \leq \alpha$. This implies $S/n \geq \log_2(n/\alpha)$. Hence proved. ■

Remark 1: As a corollary, note that the number of matches must be at least $\log_2(n/\alpha)$, since each match contributes at most n people to the sum.

Remark 2: Interpreted graph theoretically, this problem is about covering the edges of a graph by bipartite cliques (complete bipartite graphs). S is the sum of the number of vertices in each bipartite clique and α is the independence number of the graph.

Mutually Exclusive Events

Two events A and B are said to be *mutually exclusive* if they cannot both occur. For instance, if an integer is selected at random, the events A: the integer is divisible by 5, and B: the last digit of the integer is 3 are mutually exclusive. More formally, $P(A|B) = P(B|A) = 0$. If n events are pairwise mutually exclusive, then no two of them can simultaneously occur. A useful property of mutually exclusive events is that if $E_1, E_2, \ldots, E_n$ are mutually exclusive, then the probability that some E_i occurs is equal to the sum of the individual probabilities: $P[\bigcup_{i=1}^{n} E_i] = P[E_1] + P[E_2] + \ldots + P[E_n]$. On the other hand, all probabilities are bounded by 1, so $P[\bigcup_{i=1}^{n} E_i] \leq 1$, which implies $P[E_1] + P[E_2] + \ldots + P[E_n] \leq 1$. Thus we have the following simple lemma:

Lemma: If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then $P[E_1] + P[E_2] + \ldots + P[E_n] \leq 1$.

Like so many other simple facts we have encountered in this book, this lemma can be exploited in non-trivial ways to give elegant proofs for several combinatorial results. The next five examples demonstrate how.

Example 7 [Lubell-Yamamoto-Meshalkin Inequality]
Let $A_1, A_2, \ldots, A_s$ be subsets of $\{1, 2, \ldots, n\}$ such that A_i is not a subset of A_j for any i and j. Let a_i denote $|A_i|$ for each i. Show that

$$\sum_{i=1}^{s} \frac{1}{\binom{n}{a_i}} \leq 1.$$

Answer
Take a random permutation of $\{1, 2, \ldots, n\}$. Let E_i denote the event that A_i appears as an initial segment of the permutation. (For example, if $n = 5$, the permutation is 3, 4, 2, 5, 1 and $A_2 = \{2, 3, 4\}$, then event E_2 occurs since the elements of A_2 match the first three

elements of the permutation.) The key observation is that the events E_i are **mutually exclusive**: if two different sets matched initial segments of the permutation, one set would contain the other. Also note that $P[E_i] = \frac{1}{\binom{n}{a_i}}$, as there are $\binom{n}{a_i}$ different choices for the first a_i elements. Therefore, the probability of *some* event occurring is $P[E_1] + P[E_2] + \ldots + P[E_s] = \sum_{i=1}^{s} \frac{1}{\binom{n}{a_i}}$. But probability is always at most 1, so $\sum_{i=1}^{s} \frac{1}{\binom{n}{a_i}} \leq 1$. ■

Corollary [Sperner's theorem]
Let $A_1, A_2, \ldots, A_s$ be subsets of $\{1, 2, \ldots, n\}$ such that A_i is not a subset of A_j for any i and j. Such a family of sets is known as an *antichain.* Show that $s \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. In other words, an antichain over the power set P^n has cardinality at most $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. Note that equality is achieved by taking all sets of size [n/2], since two sets of the same size cannot contain each other.

Proof:
Since $\binom{n}{\lfloor n/2 \rfloor}$ is the largest binomial coefficient, $\binom{n}{\lfloor n/2 \rfloor} \geq \binom{n}{a_i}$ for each a_i. Therefore $\sum_{i=1}^{s} \frac{1}{\binom{n}{a_i}} \geq \sum_{i=1}^{s} \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} = \frac{s}{\binom{n}{\lfloor n/2 \rfloor}}$. Combining this with the Lubell-Yamamoto-Meshalkin inequality gives $\frac{s}{\binom{n}{\lfloor n/2 \rfloor}} \leq 1$, and the result follows. ■

Example 8 [Bollobas' Theorem]
Let $A_1, A_2, \ldots, A_n$ be sets of cardinality s, and let $B_1, \ldots, B_n$ be sets of cardinality t. Further suppose that $A_i \cap B_j = \emptyset$ if and only if $i = j$. Show that $n \leq \binom{t+s}{s}$.

Answer:
We use a visual trick similar to that used in the previous problem.

Take a random permutation of all elements in $\boldsymbol{S}$, where $\boldsymbol{S}$ is the union of all $2n$ sets. Let E_i be the event that all elements of A_i precede all elements of B_i in the permutation (in other words, the first element of B_i appears only after the last element of A_i). The key observation is that no two events E_i and E_j can occur simultaneously: if E_i and E_j occurred simultaneously, then $A_i \cap B_j$ or $A_j \cap B_i$ would be $\emptyset$, as these sets would be too "far apart" in the permutation. Hence, as in the previous problem, probability of *some* event occurring is $P[E_1] + \ldots + P[E_n] = \frac{n}{\binom{t+s}{s}} \leq 1$, implying the result. ■

Example 9 [Tuza's Theorem (Variant of Bollobas' Theorem)]
Let $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots, B_n$ be sets such that $A_i \cap B_i = 0$ for all i and for $i \neq j$ $(A_i \cap B_j) \cup (A_j \cap B_i) \neq \emptyset$. In other words, either $A_i \cap B_j \neq \emptyset$ or $A_j \cap B_i \neq \emptyset$ (or both). Then show that for any positive real number $p < 1$, $\sum_{i=1}^{n} p^{|A_i|}(1-p)^{|B_i|} \leq 1$.

Answer:
The form of the result we need, namely "$\sum_{i=1}^{n} p^{|A_i|}(1-p)^{|B_i|} \leq 1$", gives us a good clue as to what our random process should be. Let $\boldsymbol{U}$ be the universe of all elements in some A_i or B_i. Randomly select elements from $\boldsymbol{U}$ with probability p. Then $p^{|A_i|}(1-p)^{|B_i|}$ clearly represents the probability that all elements from A_i are selected and no elements from B_i are selected. Call this event E_i. For any pair $i \neq j$, we claim that E_i and E_j are independent. Suppose to the contrary that E_i and E_j both occur for some $i \neq j$. Then that means that all elements of $A_i \cup A_j$ have been selected, but no elements of $B_i \cup B_j$ have been selected. This implies that the sets $(A_i \cup A_j)$ and $(B_i \cup B_j)$ are disjoint, so $A_i \cap B_j = \emptyset$ and $A_j \cap B_i \neq \emptyset$. Contradiction. Thus the n events are pairwise mutually exclusive, so the probability that *some* event occurs is $P[E_1] + \ldots + P[E_n] \leq 1$, which implies $\sum_{i=1}^{n} p^{|A_i|}(1-p)^{|B_i|} \leq 1$. ■

Corollary:
Let $A_1, A_2, \ldots, A_n$ be sets of cardinality a and let $B_1, B_2, \ldots, B_n$ be sets

of cardinality b, such that $A_i \cap B_i = 0$ for all i and for $i \neq j$ $(A_i \cap B_j) \cup (A_j \cap B_i) \neq \emptyset$. Show that $n \leq \frac{(a+b)^{a+b}}{a^a b^b}$.

Proof:
Simply plug in $p = \frac{a}{a+b}$ in Tuza's theorem. ■

Example 10 [Original Problem, Inspired by China TST 1997]
There are n ice cream flavors offered at an ice cream parlor, and k different cone sizes. A group of kids buys ice creams such that no student buys more than one ice cream of a particular flavor. Suppose that for any two flavors, there exists some kid who buys both flavors but in different cone sizes. Show that the total number of ice creams bought by this group is at least $n\log_k n$.

Answer:
How do we go from innocuous ice cream to those nasty logarithms?! Answer: The Probabilistic Method.

Label the k sizes $S_1, S_2, \ldots, S_k$ and the n flavors $F_1, F_2, \ldots, F_n$. Now randomly label each student with one integer in $\{1, 2, \ldots, k\}$. For each flavor F_i, let E_i denote the event that "all students eating an ice cream of flavor F_i are eating the size corresponding to their label". For example, this would mean that a student with label 2 eating flavor F_i must be eating it in cone size S_2.

The probability of E_i is $k^{-|F_i|}$, where $|F_i|$ is the number of kids eating flavor F_i. This is because each of these $|F_i|$ kids has probability $1/k$ of eating flavor F_i from the "correct" size cone, as there are k sizes.

The key observation is that the E_i's are mutually exclusive: this follows from the fact that for any two flavors, there exists some kid who buys both flavors but in different cone sizes, meaning that he cannot be eating from the size corresponding to his label for both flavors.

It follows that $1 \geq P(\bigcup E_i) = \sum_{i=1}^{n} k^{-|F_i|} \geq n \times k^{-\sum_{i=1}^{n}|F_i|/n} =$

$nk^{-T/n}$, where $T = \sum_{i=1}^{n} |F_i|$ is the total number of ice creams bought. Here we have used the AM-GM inequality. Hence $k^{T/n} \geq n$, which implies that $T \geq n\log_k n$, as desired. ■

Example 11 [IMO Shortlist 2009, C4]

Consider a $2^m \times 2^m$ chessboard, where m is a positive integer. This chessboard is partitioned into rectangles along squares of the board. Further, each of the 2^m squares along the main diagonal is covered by a separate unit square. Determine the minimum possible value of the sum of perimeters of all the rectangles in the partition.

Answer:

Squares with position (i, i) have already been covered, and they symmetrically divide the board into two parts. By symmetry, we can restrict our attention to the bottom portion for now.

Suppose the portion of the board below the diagonal has been partitioned. Let R_i denote the set of rectangles covering at least one square in the ith row, and C_i the set of rectangles covering at least one square in the ith column. The total perimeter is $P = 2\sum(|R_i| + |C_i|)$, because each rectangle in R_i traces two vertical segments in row i and each rectangle in C_i traces two horizontal segments in column i.

Now randomly form a set $\boldsymbol{S}$ of rectangles by selecting each rectangle in the partition with probability ½. Let E_i denote the event that **all** rectangles from R_i are chosen and **no** rectangle from C_i is chosen. Clearly, $P[E_i] = 2^{-(|R_i|+|C_i|)}$. But for distinct i and j, E_i implies that the rectangle covering (i, j) **is** in $\boldsymbol{S}$, whereas E_j implies the rectangle covering (i, j) is **not** in $\boldsymbol{S}$. Hence **the $\mathbf{E_i}$'s are independent**.

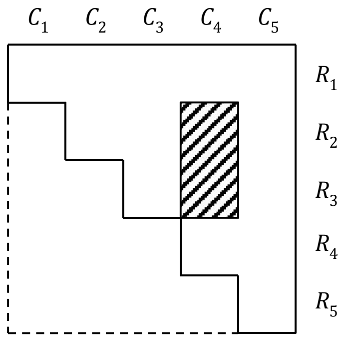

Figure 9.1. The shaded rectangle belongs to R2, R3 and C4. Its perimeter is 6, twice the number of sets it belongs to.

Therefore, $\sum 2^{-(|R_i|+|C_i|)} \leq 1$. But by the AM-GM inequality, $\sum_{i=1}^{2^m} 2^{-(|R_i|+|C_i|)} \geq 2^m \times 2^{-\sum(|R_i|+|C_i|)/2^m} = 2^m \times 2^{-P/2^{m+1}}$. Hence $2^{P/2^{m+1}} \geq 2^m$, so $P \geq m2^{m+1}$.

The same bound holds for the upper portion, so the overall perimeter is at least $2P$ + sum of perimeters of unit squares on diagonal $\geq 2(m2^{m+1}) + 4(2^m) = (m+1)2^{m+2}$. To show that this can be achieved, use a recursive construction: split the board into 4 squares of sizes 2^{m-1} x 2^{m-1}. Two of them have no squares along the main diagonal; leave these squares as they are. Two of them already have unit squares along their diagonals; recursively partition these. ■

Bounding Violated Conditions

In practically all existence problems, we essentially need to prove the existence of an object that satisfies certain conditions. A useful idea is as follows: construct the object randomly, and then show that the expected number of conditions violated is less than 1. In other words, the expected number of "failures" - conditions that fail to hold - should be less than 1. Then *some* object must violate 0 conditions; this is the required object. The next four examples illustrate this idea.

Example 12 [Erdos' theorem on tournaments]
Say that a tournament has the property P_k if for every set of k players there is one who beats them all. If ${}^nC_k \times (1-2^{-k})^{n-k} < 1$, then show that there exists a tournament of n players that has the property P_k.

Answer:
Consider a random tournament of n players, i.e., the outcome of every game is determined by the flip of fair coin. For a set $\boldsymbol{S}$ of k players, let A_S be the event that no y not in $\boldsymbol{S}$ beats all of $\boldsymbol{S}$. Each y not in $\boldsymbol{S}$ has probability 2^{-k} of beating all of $\boldsymbol{S}$ and there are $n-k$ such possible y, all of whose chances are independent of each other. Hence $Pr[A_S] = (1-2^{-k})^{n-k}$. Thus the expected number of sets $\boldsymbol{S}$ such that event A_S occurs is at most ${}^nC_k \times (1-2^{-k})^{n-k}$, which is strictly smaller than 1. Thus, for some tournament $\boldsymbol{T}$ on n vertices, no event A_S occurs and we are done. ■

Example 13 [Taiwan 1997]
For $n \geq k \geq 3$, let $\boldsymbol{X} = \{1, 2, \ldots, n\}$and let $\boldsymbol{F}_k$ be a family of k-element subsets of $\boldsymbol{X}$ such that any two subsets in $\boldsymbol{F}_k$ have at most $k-2$ elements in common. Show that there exists a subset $\boldsymbol{M}$ of $\boldsymbol{X}$ with at least $\lfloor log_2 n \rfloor + 1$ elements and not containing any subset in $\boldsymbol{F}_k$.

Answer:
The basic plan
If $k \geq \log_2 n$ there is nothing to prove, so assume $k < \log_2 n$. Let $m = \lfloor log_2 n \rfloor + 1$. Our idea is to show that the expected number of sets in $\boldsymbol{F}_k$ that a randomly chosen m-element subset would contain is strictly less than 1.

How large is $|F_k|$?

Easy double counting provides this key detail. Since each (k−1) element subset of $\boldsymbol{X}$ lies in at most one set in $\boldsymbol{F}_k$, and each set in $\boldsymbol{F}_k$ contains exactly k (k−1)-element subsets of $\boldsymbol{X}$, $|\boldsymbol{F}_k| \leq \frac{1}{k}\binom{n}{k-1} = \frac{1}{n-k+1}\binom{n}{k}$.

Finding the expectation

Now take a randomly chosen m-element subset $\boldsymbol{S}$ of $\boldsymbol{X}$. $\boldsymbol{S}$ contains $\binom{m}{k}$ k-element subsets of $\boldsymbol{X}$. Thus the expected number of elements of $\boldsymbol{F}_k$ it contains is

$$\frac{\text{number of } k \text{ element subsets of } X \text{ contained}}{\text{total number of } k \text{ element subsets of } X} \times (\text{number of subsets in } \boldsymbol{F}_k)$$

$$= \frac{\binom{m}{k}|F_k|}{\binom{n}{k}} \leq \frac{\binom{m}{k}}{n-k+1}, \text{ using the bound on } |\boldsymbol{F}_k|.$$

The final blow

To prove that this expression is less than 1, it is enough to show that $\binom{m}{k} \leq \frac{3n}{4}$, since $\frac{3n}{4(n-k+1)} < 1$ for $n \geq 3$ (Check this: remember $k < \log_2 n$). Replacing $n \geq 2^{m-1}$, we just need $\binom{m}{k} \leq 3 \times 2^{m-3}$ for all $m \geq k$ and $m \geq 3$, which can be shown by induction on m: The base case holds, and observe that

$$\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k+1} \leq 2 \max\left\{\binom{m}{k}, \binom{m}{k+1}\right\}$$

$$\leq 2 \times 3 \times 2^{m-3} = 2 \times 2^{m-2} \text{ (using the induction hypothesis).}$$

Thus the expectation is less than 1 and we are done. ■

Example 14 [Original problem, inspired by IMO Shortlist '87]
Let n, k, l and j be natural numbers with $k > l$, and $j > l \geq 1$. Let $\boldsymbol{S} = \{1, 2, 3, \ldots, n\}$. Each element of $\boldsymbol{S}$ is colored in one of the colors

c_1, c_2, …, c_j. Call a set of k terms in $\boldsymbol{S}$ in arithmetic progression *boring* if amongst them there are only l or fewer colors. (For instance, if $k = 4$, $l = 2$ and the numbers 1, 4, 7 and 10 are colored c_3, c_5, c_5 and c_3 respectively, then {1, 4, 7, 10} is boring as amongst these 4 numbers in AP there are just 2 colors.) Show that if $n < [(\frac{2k}{e^l})\,(\frac{j}{l})^{(k-l)}]^{1/2}$, there exists a coloring of $\boldsymbol{S}$ such that there are no boring sets. [Note: Here $e \approx 2.718$ is Euler's constant]

Answer:
Looks complicated? The solution is actually surprisingly straightforward. Take a random coloring (each number receives a particular color with probability $1/j$). Our goal is to show that the expected number of boring sets is less than 1. Then *some* outcome of this random coloring produces 0 boring sets. Now note that the expected number of boring sets is (number of k-term APs in $\boldsymbol{S}$) × (probability that a given k-term AP is boring). Denote the later quantity by p and the former by A. We just need to show that $pA < 1$.

Bounding A

Any k-term arithmetic progression $\{x, x+d, \ldots, x+(k-1)d\}$ is uniquely defined by its starting term x and its common difference, d where $x, d > 0$. Since we need all terms to be in $\boldsymbol{S}$, $x+(k-1)d \le n$, or $x \le n - (k-1)d$. Hence, for each $d \le n/(k-1)$, there are exactly $n - (k-1)d$ possible values x can take, and hence there are $(n - (k-1)d)$ k-term arithmetic progressions for each d. Therefore,

$$\mathrm{A} = \sum_{d=1}^{\left\lfloor\frac{n}{k-1}\right\rfloor}[n-(k-1)d]$$

$$= n\left\lfloor\frac{n}{k-1}\right\rfloor - \frac{1}{2}(k-1)\left\lfloor\frac{n}{k-1}\right\rfloor\left(\left\lfloor\frac{n}{k-1}\right\rfloor + 1\right) < \frac{n^2}{2k}$$

where the last step is simple (boring) algebra (do it yourself!). So $\mathrm{A} < \frac{n^2}{2k}$.

Bounding p

For a particular k-term AP, the total number of ways in which it can be colored is j^k. A coloring making it boring can be created by choosing l colors first and then coloring using only these l colors. There are $\binom{j}{l}l^k$ ways of doing this. (In fact, this may be higher than the total number of bad events as there is double counting - irrelevant here but worth noting). Therefore,

$$p \leq \binom{j}{l}l^k / j^k \leq (\frac{ej}{l})^l(\frac{l}{j})^k = e^l(\frac{j}{l})^{l-k}.$$

And finally…

$pA < \frac{n^2}{2k} \times e^l(\frac{j}{l})^{l-k} < 1$, using the bound on n given in the problem. ■

Example 15 [Based on a Result of Kleitman and Spencer]

Let n, m and k be positive integers with $m \geq k\, 2^k \ln(n)$ and $n > k > 3$. Define $\boldsymbol{S} = \{1, 2, \ldots, n\}$. Show that there exist m subsets of $\boldsymbol{S}$, $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_m$, such that for **any** k-element subset $\boldsymbol{T}$ of $\boldsymbol{S}$, the m intersections $\boldsymbol{T} \cap \boldsymbol{A}_i$ for $1 \leq i \leq m$ range over all 2^k possible subsets of $\boldsymbol{T}$.

Answer:

Take a few minutes to completely understand the question. Our basic strategy will be to choose the $\boldsymbol{A}_i$'s randomly and then show that the probability of "failure" is less than one.

Randomly create each set $\boldsymbol{A}_i$ by including each element with probability 0.5. Fix a pair $(\boldsymbol{T}, \boldsymbol{T}_1)$ where $\boldsymbol{T}$ is a k-element subset of $\boldsymbol{S}$ and $\boldsymbol{T}_1$ is a subset of $\boldsymbol{T}$. For each i, the probability that $\boldsymbol{A}_i \cap \boldsymbol{T} = \boldsymbol{T}_1$ is 2^{-k}, because we need each element in $\boldsymbol{T}_1$ to occur in $\boldsymbol{A}_i$ (0.5 probability each) and each element in $\boldsymbol{T} \backslash \boldsymbol{T}_1$ to **not** occur in $\boldsymbol{A}_i$ (0.5 probability each again). Therefore, the probability that **no** set $\boldsymbol{A}_i$ satisfies $\boldsymbol{A}_i \cap \boldsymbol{T} = \boldsymbol{T}_1$ is $(1 - 2^{-k})^m$.

There are $\binom{n}{k}$ choices for $\boldsymbol{T}$ and 2^k choices for $\boldsymbol{T}_1$ once we have chosen $\boldsymbol{T}$. Therefore, the overall probability of failure is at most

$$\binom{n}{k}2^k(1 - 2^{-k})^m \le \frac{n^k}{k!}2^k e^{-m/2^k}$$

$$\le \frac{n^k}{k!}2^k e^{-k2^k \ln(n)/2^k} = \frac{2^k}{k!} < 1, \text{ since } k > 3.$$

Thus, with positive probability, no condition of the problem is violated, and hence there exist m subsets of $\boldsymbol{S}$ satisfying the problem's requirements. ■

Useful Concentration Bounds

In many situations we want the outcome of a random process we design to closely match the expectation of that outcome. Concentration inequalities allow us to achieve this by bounding the probability that a random variable is far from its mean. The following three lemmas represent three of the most basic, widely used concentration inequalities. We omit rigorous measure theoretic proofs, but will provide some intuition as to why they hold.

1. **Markov's Bound**:
   If x is a nonnegative random variable and a is a positive number,
   $$P[x \ge a] \le E[x]/a.$$

   Equivalently written as
   $$P[x \ge aE[x]] \le 1/a.$$

   Intuition: It's not possible for 20% of a country's population to each earn more than 5 times the national average.

2. **Chebychev's Inequality:**
Chebychev's inequality allows us to bound the probability that a random variable is far from its mean in terms of its variance. It is a direct consequence of Markov's bound.

For a random variable x, applying Markov's bound to the positive random variable $(x-E[x])^2$, and applying the definition of variance,

$$P[|x-E[x]| \geq a] = P[(x-E[x])^2 \geq a^2] \leq \frac{\mathrm{E}[(\mathrm{x}-\mathrm{E}[\mathrm{x}])^2]}{a^2} = \frac{Var\{x\}}{a^2}.$$

This result is intuitive: The "spread" of the random variable is directly related to its variance.

3. **Chernoff's Bound**, simple version:
Chernoff's bound provides another bound on the probability that a random variable is far from its mean. It is particularly useful when we want relative guarantees rather than absolute, e.g. when we are not directly concerned about the difference between a random variable and its mean but rather want to estimate the probability that it is ten percent higher than its mean. In addition, we do not need information about the variance, unlike in Chebychev's inequality.

Let X be the sum of n independent 0-1 random variables X_1, X_2, …, X_n, where $E[X_i] = p_i$. Denote $E[X] = \sum_{i=1}^{n} p_i$ by μ. Then for any $0 < \delta < 1$,
(i) $P[X > (1+\delta)\mu] \leq e^{-\delta^2\mu/3}$
(ii) $P[X < (1-\delta)\mu] \leq e^{-\delta^2\mu/2}$

In some cases, it may be clear that you need to use a concentration inequality but you may not be sure which one to use. In such cases, there's no harm in trying all three of these (as long as the variance is not difficult to compute). These three bounds are the only ones you will need for the exercises in this book. However, there exist a wealth of other powerful, more specialized concentration inequalities used in probabilistic combinatorics, and interested readers are encouraged to refer to online resources devoted to these results.

The next example represents a very typical and intuitive use of concentration inequalities. Essentially, we want something to be roughly balanced, so we designed a randomized procedure that would, in expectation, produce something perfectly balanced. Then, we used a concentration inequality to show that with nonzero probability, we would get something "not too unbalanced."

Example 16 [Hypergraph discrepancy]

Let ***F*** be a family of m subsets of {1, 2, …, n} such that each set contains at most s elements. Each element of {1, 2, …, n} is to be colored either red or blue. Define the *discrepancy* of a set ***S*** in ***F***, denoted *disc*(***S***), to be the absolute value of the difference between the number of red elements in ***S*** and blue elements in ***S***. Define the discrepancy of ***F*** to be the maximum discrepancy of any set in ***F***. Show that we can color the elements such that the discrepancy of ***F*** is at most $\left\lceil 2\sqrt{s\ ln(2m)} \right\rceil$.

Answer:

Color each number in {1, 2, …, n} red with probability 0.5 and blue otherwise. We just need to show that for a particular ***S*** in ***F***, the probability that ***S*** has discrepancy higher than our desired threshold is less than $1/m$, for then the result will follow from the union bound.

Let ***S*** be a set with t elements, and denote by b the number of blue elements and by r the number of red elements of ***S***. Note that for a fixed number d < t, *disc*(***S***) > d if and only if either b < $t/2 - d/2$ or r < $t/2 - d/2$. Hence the probability that *disc*(***S***) > d is equal to

$$P[disc(\boldsymbol{S}) > d] = P[b < t/2 - d/2] + P[r < t/2 - d/2]$$

$$= 2P[b < t/2 - d/2] \text{ by symmetry.}$$

Note that b is the sum of t independent random 0-1 variables, and $E[b] = t/2$. Write $\mu = t/2$, $\delta = d/t$. Then applying the Chernoff bound and noting that $t \leq s$,

$$2P[b < t/2 - d/2] = 2P[b < (1 - \delta)\, \mu] \leq 2e^{-\delta^2 \mu/2}$$

$$= 2\exp(-d^2/4t) \leq 2\exp(-d^2/4s).$$

For the last quantity to be less than $1/m$, we just need $d > 2\sqrt{s\, ln(2m)}$, as desired. ■

Remark: Note that using the Chebychev inequality for this problem would prove a weaker bound. It is instructive for readers to try this, as it is a good exercise in computing the variance of a random variable.

The Lovasz Local Lemma

In the section on bounding "violated conditions," we used the fact that if the expected number of violated conditions is less than 1, then there exists an object satisfying all conditions. Unfortunately, we aren't always this lucky: in many cases, the expected number of violated conditions is large. The *Lovasz local lemma* provides a much sharper tool for dealing with these situations, especially when the degree of dependence (defined below) between violated conditions is low.

Definition: Let E_1, E_2, …, E_n be events. Consider a graph $\boldsymbol{G}$ with vertices v_1, …, v_n, such that there is an edge between v_i and v_j if and only if events E_i and E_j are **not** independent of each other. Let d be the maximum degree of any vertex in ***G***. Then d is known as the ***degree of independence*** between the events E_1, E_2, …, E_n.

Theorem [The Lovasz Local Lemma]:
Let E_1, E_2, …, E_n be events with degree of dependence at most d and $P[E_i] \leq p$ for each $1 \leq i \leq n$, where $0 < p < 1$. If $ep(d+1) \leq 1$, then $P[E_1' \cup E_2' \cup \ldots \cup E_n'] > 0$, where E_i' denotes the complement of event E_i. In other words, if $ep(d+1) \leq 1$, then the probability of **none of the *n* events occurring** is strictly greater than 0. (Here e is Euler's constant.)

The Lovasz local lemma hence gives us a new way to solve problems in which we need to show that there exists an object violating 0 conditions. Given n conditions that need to hold, define E_i to be the event that the ith condition is violated. All we need to do is show compute p and d for these events, and show that $ep(d+1) \leq 1$.

Example 17 [Generalization of Russia 2006]
At a certain party each person has at least δ friends and at most Δ friends (amongst the people at the party), where $\Delta > \delta > 1$. Let k be an integer with $k < \delta$. The host wants to show off his extensive wine collection by giving each person one type of wine, such that each person has a set of at least $(k+1)$ friends who all receive different types of wine. The host has W types of wine (and unlimited supplies of each type).

Show that if $W > k\,[e^{k+1}(\Delta^2 - \Delta + 1)]^{1/(\delta-k)}$, the host can achieve this goal.

Answer:
Interpret the problem in terms of a friendship graph (as usual), and interpret "types of wine" as coloring vertices. Note the presence of *local* rather than *global* information- the problem revolves around the neighbors and degrees of individual vertices,

but doesn't even tell us the total number of vertices! This suggests using the Lovasz Local Lemma. We will show that $ep(d+1) \leq 1$, where p is the probability of a *local bad event* E_v (defined below) and d is the degree of dependency between these events.

Defining E_v and bounding $p = P[E_v]$

Randomly color vertices using the W colors (each vertex receives a particular color with probability $1/W$). Let E_v denote the event that vertex v does **not** have $(k+1)$ neighbors each receiving a different color. Hence,

$$P[E_v] \leq \binom{W}{k} \times \left(\frac{k}{W}\right)^{d_v} \leq \binom{W}{k} \times \left(\frac{k}{W}\right)^{\delta} \leq \left(\frac{k}{W}\right)^{\delta} \left(\frac{eW}{k}\right)^{k} = e^k \left(\frac{W}{k}\right)^{k-\delta}.$$

Here we used $d_v \geq \delta$.

Bounding d

E_v and E_u will be non-independent if and only if v and u have a common neighbor - in other words, u must be a neighbor of one of v's neighbors. v has at most Δ neighbors, and each of these has at most $(\Delta-1)$ neighbors apart from v. Hence for fixed v, there are at most $(\Delta^2-\Delta)$ choices for u such that E_v and E_u are dependent. It follows that $d \leq (\Delta^2-\Delta)$.

And finally…

$ep(d+1) \leq e \times e^k\left(\frac{W}{k}\right)^{k-\delta} \times (\Delta^2-\Delta+1)$. Rearranging the condition in the problem (namely $W \geq k\,[e^{k+1}(\Delta^2-\Delta+1)]^{1/(\delta-k)}$) shows that this expression is at most 1. Hence proved. ■

Exercises

1. **[Max cut revisited]**
   Give a probabilistic argument showing that the vertex set $\boldsymbol{V}$ of a graph $\boldsymbol{G}$ with edge set $\boldsymbol{E}$ can be partitioned into two sets $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ such that at least $|\boldsymbol{E}|/2$ edges have one endpoint in $\boldsymbol{V}_1$ and one endpoint in $\boldsymbol{V}_2$.

2. **[Hypergraph coloring]**
   Let $\boldsymbol{F}$ be a family of sets such that each set contains n elements and $|\boldsymbol{F}| < 2^{n-1}$. Show that each element can be assigned either the color red or the color blue, such that no set in $\boldsymbol{F}$ is monochromatic (contains elements of only one color).

3. **[A Bound on Ramsey Numbers]**
   The *Ramsey number* $R(k, j)$ is defined as the least number n such that for any graph $\boldsymbol{G}$ on n vertices, $\boldsymbol{G}$ either contains a clique of size k or an independent set of size j. Show that $R(k, k) \geq 2^{k/2}$ for each k.

4. **[Stronger version of IMO Shortlist 1999, C4]**
   Let $\boldsymbol{A}$ be a set of n different residues mod n^2. Then prove that there exists a set $\boldsymbol{B}$ of n residues mod n^2 such that the set $\boldsymbol{A}+\boldsymbol{B} = \{a+b \mid a \in \boldsymbol{A}, b \in \boldsymbol{B}\}$ contains at least $(e-1)/e$ of the residues mod n^2.

5. **[Another hypergraph coloring problem]**
   Let $\boldsymbol{F}$ be a family of sets such that each set contains n elements. Suppose each set in $\boldsymbol{F}$ intersects at most 2^{n-3} other sets in $\boldsymbol{F}$. Show that each element can be assigned either the color red or the color blue, such that no set in $\boldsymbol{F}$ is monochromatic. Improve this bound (that is, replace 2^{n-3} by a

larger number and prove the result for this number).

6. [Based on Russia 1999]
Let $\boldsymbol{G}$ be a bipartite graph with vertex set $\boldsymbol{V} = \boldsymbol{V}_1 \cup \boldsymbol{V}_2$. Show that $\boldsymbol{G}$ has an induced subgraph $\boldsymbol{H}$ containing at least $|\boldsymbol{V}|/2$ vertices, such that each vertex in $\boldsymbol{V}_1 \cap \boldsymbol{H}$ has odd degree in $\boldsymbol{H}$.

7. [USAMO 2012, Problem 2]
A circle is divided into 432 congruent arcs by 432 points. The points are colored in four colors such that some 108 points are colored Red, some 108 points are colored Green, some 108 points are colored Blue, and the remaining 108 points are colored Yellow. Prove that one can choose three points of each color in such a way that the four triangles formed by the chosen points of the same color are congruent.

8. [List coloring]
Each vertex of an n-vertex bipartite graph $\boldsymbol{G}$ is assigned a list containing more than $\log_2 n$ distinct colors. Prove that $\boldsymbol{G}$ has a proper coloring such that each vertex is colored with a color from its own list.

9. Let $\boldsymbol{A}$ be an $n \times n$ matrix with distinct entries. Prove that there exists a constant $c > 0$, independent of n, with the following property: it is possible to permute the rows of $\boldsymbol{A}$ such that no column in the permuted matrix contains an increasing subsequence of length $c\sqrt{n}$. (Note that consecutive terms in a subsequence need not be adjacent in the column; “subsequence” is different from “substring.”)

10. [IMO Shortlist 2006, C3]
Let $\boldsymbol{S}$ be a set of points in the plane in general position (no three lie on a line). For a convex polygon P whose vertices are in $\boldsymbol{S}$, define $a(P)$ as the number of vertices of P and $b(P)$ as the number of points of $\boldsymbol{S}$ that are outside P. (Note: an empty set , point and line segment are considered convex polygons with 0, 1 and 2 vertices respectively.) Show that for each real

number x,

$$\sum x^{a(P)}(1-x)^{b(P)} = 1,$$

where the sum is taken over all convex polygons P.

11. [Bipartite Expanders]
A bipartite graph $\boldsymbol{G}$ with vertex set $\boldsymbol{V} = \boldsymbol{V}_1 \cup \boldsymbol{V}_2$ is said to be an (n, m, d, β) *bipartite expander* if the following holds:

(i) $|\boldsymbol{V}_1| = n$ and $|\boldsymbol{V}_2| = m$
(ii) Each vertex in $\boldsymbol{V}_1$ has degree d
(iii) For any subset $\boldsymbol{S}$ of $\boldsymbol{V}$ with $|\boldsymbol{S}| \leq n/d$, there are at least $\beta|\boldsymbol{S}|$ vertices in $\boldsymbol{V}_2$ that have a neighbor in $\boldsymbol{S}$

Show that for any integers $n > d \geq 4$, there exists an $(n, n, d, d/4)$ bipartite expander.

12. [Sphere packing]
Let n be a given positive integer. Call a set $\boldsymbol{S}$ of binary strings of length n *α-good* if for each pair of strings in $\boldsymbol{S}$, there exist at least $n\alpha$ positions in which the two differ. For instance, if $n = 100$ and $\alpha = 0.1$, then any pair of strings in $\boldsymbol{S}$ must differ in at least 10 positions. Show that for each integer n and real number $0 < \alpha < 0.5$, there exists an α-good set $\boldsymbol{S}$ of cardinality at least $\lfloor \sqrt{2}e^{n(0.5-\alpha)^2} \rfloor$.

As a (rather remarkable) corollary, deduce that the unit sphere in n dimensions contains a set of $\lfloor \sqrt{2}e^{n/16} \rfloor$ points such that no two of these points are closer than one unit (Euclidean) distance from each other.

Remark: The problem of finding large sets of binary strings of a given length such that any two differ in sufficiently many positions is one of the central problems of coding theory. Coding theory is a remarkable field lying in the intersection of mathematics and computer science, and uses techniques from combinatorics, graph theory, field theory, probability and linear algebra. Its applications range from file compression to mining large scale web data. Coding theory is also the reason your CDs often work even if they get scratched. While this exercise asks for an existence proof, finding constructive solutions to several similar problems remains an active research area with high practical relevance.

13. Let ***F*** be a collection of k-element subsets of {1, 2, …, n} and let $x = |\boldsymbol{F}|/n$. Then there is always a set ***S*** of size at least $\frac{n}{4x^{1/(k-1)}}$ which does not completely contain any member of ***F***.

14. [Another list coloring theorem]
Each vertex of an n-vertex graph ***G*** is assigned a list containing at least k different colors. Furthermore, for any color c appearing on the list of vertex v, c appears on the lists of at most $k/2e$ neighbors of v. Show that there exists a proper coloring of the vertices of ***G*** such that each vertex is colored with a color from its list.

15. [Due to Furedi and Khan]
Let ***F*** be a family of sets such that each set in ***F*** contains at most n elements and each element belongs to at most m sets in ***F***. Show that it is possible to color the elements using at most $1+(n-1)m$ colors such that no set in ***F*** is monochromatic.

16. [Due to Paul Erdos]
A set ***S*** of distinct integers is called *sum-free* if there does not exist a triple $\{x, y, z\}$ of integers in ***S*** such that $x + y = z$. Show that for any set ***X*** of distinct integers, ***X*** has a sum-free subset ***Y*** such that $|\boldsymbol{Y}| > |\boldsymbol{X}|/3$.

17. [IMO 2012, Problem 3]
The *liar's guessing game* is a game played between two players A and B. The rules of the game depend on two positive

integers k and n which are known to both players.

A begins by choosing integers x and N with $1 \le x \le N$. Player A keeps x secret, and truthfully tells N to player B. Player B now tries to obtain information about x by asking player A questions as follows: each question consists of B specifying an arbitrary set $\boldsymbol{S}$ of positive integers (possibly one specified in some previous question), and asking A whether x belongs to $\boldsymbol{S}$. Player B may ask as many questions as he wishes. After each question, player A must immediately answer it with *yes* or *no*, but is allowed to lie as many times as she wants; the only restriction is that, among any $(k+1)$ consecutive answers, at least one answer must be truthful.

After B has asked as many questions as he wants, he must specify a set $\boldsymbol{X}$ of at most n positive integers. If x belongs to $\boldsymbol{X}$, then B wins; otherwise, he loses. Prove that:

1. If $n \ge 2^k$, then B can guarantee a win.
2. For all sufficiently large k, there exists an integer $n \ge 1.99^k$ such that B cannot guarantee a win.

18. [Johnson-Lindenstrauss Lemma]

(i) Let $X_1, X_2, \ldots, X_d$ be d independent Gaussian random variables. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_d)$ be the d-dimensional vector whose coordinates are $X_1, X_2, \ldots, X_d$. Let $k < d$ and let $\boldsymbol{X}'$ be the k-dimensional vector $(X_1, X_2, \ldots, X_k)$. Show that:

a) If α is a constant greater than 1,
$$\text{Prob}[\, ||X_k|| \ge \alpha\, k/d \,] \le \exp\left(\frac{k(1-\alpha+\ln\alpha)}{2}\right)$$

b) If α is a constant smaller than than 1,
$$\text{Prob}[\, ||X_k|| \le \alpha\, k/d \,] \le \exp\left(\frac{k(1-\alpha+\ln\alpha)}{2}\right)$$

(ii) Part (i) says that the norm of a vector of independent randomly distributed Gaussian variables projected onto its first k coordinates is sharply concentrated around its expectation. Equivalently, the norm of a fixed vector projected onto a random k-dimensional subspace is sharply concentrated around its expectation. Use this result to show that given a set $\boldsymbol{S}$ of n points in a d-dimensional space, there is a mapping f to a subspace of dimension $O(n \log n/\varepsilon^2)$ such that for all points u, v in $\boldsymbol{S}$,

$$(1-\varepsilon)\ \mathrm{dist}(u, v) \leq \mathrm{dist}(f(u), f(v)) \leq (1+\varepsilon)\ \mathrm{dist}(u, v).$$