Solutions to Extremal Combinatorics

Naman Kumar

March 23, 2024

Abstract

The following document contains solutions to selected problems in the book 'Extremal Combinatorics' by Stasys Jukna. Some problems which are either trivial or require nothing more than brute force/proof mirroring have been omitted.

Contents

1 Counting

1.1 Review

- The binomial theorem and binomial coefficients provide a way to count objects.
- Exact values of binomial coefficients are hard to compute, but they can be roughly estimated. Consider the following inequalities:
  1. $1 + t < e^t$ for $t \neq 0$
  2. $1 - t > e^{-t-t^2/2}$ for $0 < t < 1$

  they can be used to prove
  $$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \qquad \sum_{i=1}^{k} \binom{n}{i} \leq \left(\frac{en}{k}\right)^k.$$
- The Stirling approximation is another approximation which shows
  $$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n} e^{\alpha_n}$$
  with $1/(12n+1) < \alpha_n < 1/12n$. Similarly, an asymptotic approximation for the k-factorial is
  $$(n)_k = n^k e^{-k^2/2n - k^3/6n^2 + o(1)}$$
  which is valid for $k = o(n^{3/4})$.
- The stars and bars technique allows you to partition a set into ordered partitions.
- Double counting can be done using incidence matrices: in this technique you count the rows and the columns of a zero-one matrix, which should yield the same sums.
- The number of vertices in a graph with odd degree is even.
- Jensen's Inequality: a function is convex if
  $$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b)$$
  for $0 \leq \lambda \leq 1$. Suppose that each λ_i is in $[0,1]$, and $\sum \lambda_i = 1$. If f is convex then
  $$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i).$$

- Inclusion-exclusion provides a way to sum up all the elements in some sets that intersect.
- Derangements are permutations which don't fix any points.

1.2 Problems

In-text. Prove:

$$\sum_{x\in Y} d(x) = \sum_{A\in\mathcal{F}} |Y\cap A| \text{ for any } Y\subseteq X. \tag{1.1}$$

$$\sum_{x\in X} d(x)^2 = \sum_{x\in\mathcal{F}}\sum_{x\in A} d(x) = \sum_{A\in\mathcal{F}}\sum_{B\in\mathcal{F}} |A\cap B|. \tag{1.2}$$

Answer. We can proceed by a counting argument. For the first part, consider the incidence matrix $M = (m_{x,a})$ of $\mathcal{F}$. Adding the rows belonging to the elements of $x \in Y$ gives us the left hand side. If we 'slice' the matrix in this way, however (removing all the rows belonging to elements $x \notin Y$) then summing via the columns we get the sum of only those elements which are both in A and in Y, i.e. $|A\cap Y|$.

For the second part, it is enough to notice that if we replace each entry 1 in the incidence matrix with $d(x)$, then adding along the rows gives $\sum_{x\in X} d(x)^2$ while adding along the columns gives the term in the central equality (we add $d(x)$ for every $x \in A$, and then we sum over each of the As). The final equality follows from noticing that

$$\begin{aligned}\sum_{x\in\mathcal{F}}\sum_{x\in A} d(x) &= \sum_{A'\in\mathcal{F}}\left(\sum_{x\in A} |A\cap A'|\right)\\ &= \sum_{A\in\mathcal{F}}\sum_{B\in\mathcal{F}} |A\cap B|.\end{aligned}$$

where the second equality follows from substituting the first part.

Question 1.1 In how many ways can we distribute k balls to n boxes so that each box has at most one ball?

Answer. Depends on whether $k \le n$ or $k > n$. If $k \le n$, we can choose $\binom{n}{k}$ boxes, and the balls can be ordered in $k!$ ways. it follows that the number of ways is $\binom{n}{k}k! = (n)_k$.

If $k > n$, then there are $\binom{k}{n}$ ways we can choose the balls, and we can order them in $n!$ ways. This gives us the answer of $(k)_n$.

Question 1.2 Show that for every k the product of any k consecutive natural numbers is divisible by $k!$.

Answer. Let the k numbers be $n+1, \ldots, n+k$. Consider the number $\binom{n+k}{k}$. This is the number of ways you can choose k numbers from $n+k$ numbers, and is an integer. Expanding, we get

$$\binom{n+k}{k} = \frac{n!}{k!(n-k)!}$$

which clearly shows that $k!$ divides $(n+1)(n+2)\ldots(n+k)$.

Question 1.3 Show that the number of pairs (A, B) of distinct subsets of $\{1, \ldots, n\}$ with $A \subset B$ is $3^n - 2^n$.

Answer. We can proceed as follows. Select a subset S of size k, and then select a subset of S. The former can be done in $\binom{n}{k}$ ways, while the latter can be done in $2^k - 1$ ways. This gives us the sum

$$\sum_{k=0}^{n} \binom{n}{k}(2^k - 1)$$

The binomial theorem tells us that this is equal to $(2+1)^n - (1+1)^n = 3^n - 2^n$.

Question 1.4 Show that

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}.$$

Answer. We will count the number of ways to choose k balls from n balls. Just choosing the k balls is the left hand side. Alternately, we can choose one ball first, and then choose $k-1$ balls from the remaining $n-1$ balls. However, this leads to a k-recounting, since each k size subset is selected k times (once when each element is the ‘fixed’ element). We are done.

Question 1.6 There is a set of $2n$ people: n male and n female. A good party is a set with the same number of male and female. How many possibilities are there to build such a good party?

Answer. For each k there are k ways to choose men and k ways to choose women, so the total number of parties of size k are $\binom{n}{k}^2$. Adding over all the k's we get

$$\sum_{i=0}^{n} \binom{n}{i}^2.$$

Question 1.7 Use Proposition 1.3 to show that

$$\sum_{i=0}^{r} \binom{n+i-1}{i} = \binom{n+r}{r}$$

Answer. We expand the RHS using Pascal's identity, then recursively expand one of the terms.

Question 1.8 Let $0 \leq a \leq m \leq n$ be integers. Show that

$$\sum_{i=m}^{n} \binom{i}{a} = \binom{n+1}{a+1} - \binom{m}{a+1}.$$

Answer. Same trick as the previous question, expand $\binom{n+1}{a+1}$.

Question 1.9 Prove the Cauchy-Vandermonde identity:

$$\binom{p+q}{k} = \sum_{i=0}^{k} \binom{p}{i}\binom{q}{k-i}.$$

Answer. We count twice. The left is selecting k items from $p+q$ items. Another way we can count this is to select i items from p items and $k-i$ items from q items, then add over every i.

Question 1.10 Show that

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

Answer. Start with the previous problem, then set $p = q = k = n$. It follows that

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k} \binom{n}{n-k}.$$

However, we know that to choose k things, we can choose k or choose $n - k$, so the latter terms are equal.

Question 1.11 Prove the following analogy of the binomial theorem for factorials:

$$(x + y)_n = \sum_{k=0}^{n} \binom{n}{k} (x)_k (y)_{n-k}.$$

Answer. Immediately follows from the same consideration as the typical proof of the binomial theorem.

Question 1.28 This took me way too fucking long. Here's how you do it:

First, let $|E|$ be the number of edges in the graph. Note that $|B|D \geq |E| \geq |A|d$ by definition, since the *minimum* number of edges leaving A is $|A|d$, and the *maximum* number of edges leaving B is $|B|D$. Together with the fact given in the question this gives $|A|d = |B|D$. Furthermore, because the number of edges is sandwiched between these two, quantities, we know that $E = |A|d$. This tells us that every vertex in A *must* have degree d. Now take B_0 to be the set such that every neighbor has $\alpha D/2$ vertices to A_0.

It follows that the number of edges leaving A_0 is $|A_0|d = \alpha|A|d = \alpha|B|D$. Of these, some go to B_0, while others go outside of B_0. It follows that $|E_{A_0}| = |E_{A_0 \to B_0}| + |E_{A_0 \to B \setminus B_0}|$. We know that $|E_{A_0 \to B_0}| \leq \alpha|B_0|D/2$. Suppose that $|E_{A_0 \to B \setminus B_0}| \geq |E_{A_0 \to B_0}|$, which implies $|E_{A_0}| = \alpha|B|D \leq \alpha|B_0|D/2 + \alpha|B_0|D/2 = \alpha|B_0|D$, which is preposterous since $|B_0| \leq |B|$. This proves (iii), that $|E_{A_0 \to B \setminus B_0}| \leq |E_{A_0 \to B \setminus B_0}|$, which implies that more than half of the edges leaving A_0 go to B_0.

The first part now easily follows. The number of edges going from A_0 to B_0 is exactly the number of edges going from B_0 to A_0, which we know is greater than

$\alpha|B|D/2$ (which is half the number of edges leaving A_0). The edges leaving B_0 include this number *and more*. This means that $|B_0|D \geq \alpha|B|D/2$, which gives us (i), ie. $|B_0| \geq \alpha|B|/2$.

Very annoying problem. It's not conceptually difficult but there's a number of very shitty moving parts. Happy I solved it, finally.

Question 1.29 This was also mildly annoying, but I got it after a few minutes of serious thought. You apply Jensen's inequality to the parameters $f(x) = x^{t/s}$, $\lambda_i = 1/n$, and $x_i = a_i^s$. The inequality follows from the fact that f is convex iff $t \geq s$.

Question 1.37 This is not interesting, but is used later in the book. Really the only thing here is that this is in some sense 'partial' Inclusion-Exclusion: it's written in a misleading way, but it's actually just inclusion-exclusion with some of the terms chopped off.

In the odd cases, the opposite will happen: it'll be greater for odd k and lesser for even k.

2 Advanced Counting

2.1 Review

Some filling-in-the-details here, in order to make sure that I'm completely understanding what's happening.

Page 29, Prop. 2.8. The final inequality follows from the averaging principle, but it's not as immediate as meets the eye. We basically use the fact that the elements of $\mathcal{F}$ form a partition of the vertices of G, and the fact that every single one of them is a clique. In particular, for any $X, Y \in \mathcal{F}$, the edges from X to Y are not *complete*, and hence $X \cup Y$ is not a clique (also because of the independence of α, none of the subsets are cliques either). This means that this is a complete characterization of *all the cliques* in the graph. We then get that the average size of a clique is

$$\frac{1}{|\mathcal{F}|}\sum X_i \in \mathcal{F}|X_i| = \frac{n}{|\mathcal{F}|}$$

and it follows from the averaging principle there exists some clique with size greater than or equal to the average clique, and hence

$$\omega(G) \geq \frac{n}{|\mathcal{F}|}.$$

2.2 Problems

Question 2.1 Start with the sets A_i and construct new sets $A_{i,j} = A_i \cap A_j$ and $A'_i = A_i \setminus (\cap_{j\neq i} A_j)$. For each i, $A_i = A'_i \cup (\cup_{j\neq i}[A_i \cap A_j])$. Then by the union bound we have

$$|A_i| \leq |A'_i| + \sum_{i\neq j} |A_i \cap A_j|$$

and it follows that

$$\sum_{i=1}^{m} |A_i| \leq \sum_{i=1}^{m} |A'_i| + \sum_i \sum_{j\neq i} |A_i \cap A_j|$$

Since $|A_i \cap A_j| \leq t$, we have the latter half of the RHS to be $\leq t\binom{m}{2}$. For the former, we notice that $A'_i \cap A'_j = \emptyset$ by definition. Then $\sum_{i=1}^{m} |A'_i| \leq n$. We have

$$\sum_{i=1}^{m} |A_i| \leq n + t\binom{m}{2}.$$

Question 2.3 (Did this *without* the hint, very happy.)

Let $d(x)$ be the degree of x in $\mathcal{F}$. Then $p(x) = d(x)^2 + (m - d(x))^2$. This is because the first condition is satisfied by all (A, B) such that x is in both of them, ie. exactly $d(x)$ of them, while the latter is satisfied by (A, B) such that x is in neither of them. We can expand this to write $p(x) = 2d(x)^2 + m^2 - 2md(x) = m^2 - 2d(x)(m - d(x))$. Now $d(x)(m - d(x)) \leq m^2/4$ (which can be easily checked – it reaches its maximum at $d(x) = m/2$) and we have $p(x) \geq m^2 - 2(m^2/4) = m^2/2$.

Question 2.7 We argue using induction. Suppose that every $(r-1)$-partite 2-clique free graph contains $\leq 2m^{r-1-1/2^{r-2}}$ edges. The proposition is trivially true for $r = 1$.

Consider now an r-partite graph with more than $2m^{r-1/2^{r-1}}$ vertices. We can write any r-partite graph G as as a subset of the cartesian product $X \times V = V_1 \times \cdots \times V_{r-1} \times V$. Define the sets $A_v = \{x \in X : (x, v) \in G\}$; we will apply the lemma to them. Clearly the size of X is at most m^{r-1}, and so the size of any A_v is at most m^{r-1}, and since $v \in V$ we have that the number of such sets is at most m. We calculate the average size of the sets as m^{r-1}/m, which is m^{r-2}.

Taking $w = \frac{1}{2}m^{1/2^{r-1}}$, we can apply the lemma to find two sets A_i and A_j such that their intersection is of size more than

$$\frac{n}{2w^2} = \frac{2m^{r-1}}{m^{1/2^{r-2}}} = 2m^{r-1-1/2^{r-2}}.$$

Then $A_i \cap A_j$ must contain a 2-clique. But this 2-clique is connected to i and j, and hence there is an m-partite 2-clique in this graph.

Question 2.8 This is nearly identical to that of the in-text question of chapter 1, but slightly involved. The double counting trick works.

Question 2.10 Let $A_1, \ldots, A_N$ be subsets of some n-element set X, and suppose that these sets have average size at least αn. Show that for every $s \leq (1 - \epsilon)\alpha N$ with $0 < \epsilon < 1$, there are indices $i_1, i_2, \ldots, i_s$ such that

$$|A_{i_1} \cap \cdots \cap A_{i_s}| \geq (\epsilon\alpha)^s n.$$

Answer. (The ideas are correct, but this is unfortunately quite poorly written; I should've been doing this on paper instead of trying to work it out in my head.)

Consider the bipartite graph (X, V, E) where $V = [N]$ and $x \in A_i \iff (x, i) \in E$. Since the size of each A_i is at least αn, the number of edges is at least $\alpha n N$. For any family of s sets $\{A_{i_j}\}_{j\in[s]}$, every element in the intersection forms a $(1, s)$-star. By preceding as in the proof of 2.10, we can bound the number of stars as

$$n\binom{|E|/n}{a} \leq \sum_{x\in V} \binom{d(x)}{a}$$

which, setting $a = s \leq (1-\epsilon)\alpha N$ and $|E| = \alpha n N$ gives us

$$n(\alpha N - (1-\epsilon)\alpha N)^s \leq \sum_{x\in V} \binom{d(x)}{s}$$

which gives

$$n(\epsilon\alpha N)^s \leq \sum_{x\in V} \binom{d(x)}{s}.$$

By expanding the factorial and using $n!/(n-s)! \leq n^s$, we can set

$$n(\epsilon\alpha N)^s \leq \sum_{x\in V} d(x)^s$$

The latter, however, on applying the problem 2.8, gives us

$$n(\epsilon\alpha N)^s \leq \sum_{(i_1,\ldots,i_s)} |A_{i_1} \cap \cdots \cap A_{i_s}| \leq \binom{N}{s}|A_{i_1} \cap \cdots \cap A_{i_s}|$$

where we take the intersection of the largest size to be $(i_1, \ldots, i_s)$. We can again use the factorial inequality to get

$$n(\epsilon\alpha N)^s \leq N^s|A_{i_1} \cap \cdots \cap A_{i_s}|$$
$$n(\epsilon\alpha)^s \leq |A_{i_1} \cap \cdots \cap A_{i_s}|.$$