Undergraduate Seminar: The Probabilistic Method

Organizational Meeting

Topic

- This seminar will closely follow the book *The Probabilistic Method* by Alon and Spencer (which can be found for free online)
- This book covers the probabilistic method, a mathematical method for using probability to solve combinatorics problems
- Required background is a basic understanding of probability, for instance,
    - Random Variables
    - Expectation
    - Independence
    - Conditioning

Logistics

- We will meet for 2 hours each week (either 1 2hr block or 2 1hr blocks depending on scheduling). Our goal will be 1 chapter per week, but we will be flexible with the pace.
- One student will present each week (therefore the number of times each student presents may depend on enrollment, however it should be no more than 3).
- Grading will be based on talk preparation and attendance (see syllabus for more details)
- To determine a meeting time I will send out a when2meet. 15 minute conflicts with other courses can be excused, so please fill out the form with that in mind (scheduling is very difficult!)
- To make scheduling as painless as possible, the class time will be set this weekend and will be final.

Equations in Slides

If creating slides all equations are expected to be LaTeXed. There are several ways to do this, here are a few that I would recommend.

- Install the browser add on called math equations
  - Here is the link to download the extension.
  - Pro: it is probably the most efficient for small things like binomial coefficients since it works directly in Google Slides.
  - Con: Does not have the same versatility you would have in your own TeX document (It doesn't seem to allow you to use external packages such as amsthm, but let me know if you figure out how to do this!)
- Online TeX editor
  - Here is an example.
  - Pro: This has a math keyboard that will help if you are not so comfortable with LaTeX
  - Con: Not as efficient as the browser add on and also lacks the versatility
- Your own TeX editor
  - Pro: All the versatility you could dream of!
  - Con: Also not very efficient

A Sample of the Method! (Section 1.1)

The general idea of the probabilistic method is using probability to prove that certain structures exist. We define probability spaces and show that such a structure occurs on this space with positive probability (and therefore, must be possible). To illustrate this, we will use a simple example, the Ramsey number!

The Ramsey Number

The **Ramsey Number** R(j,k) is the smallest integer n such that in any 2-coloring of a complete graph on n vertices (here on denoted by K(n)) by red or blue there must exist either a completely red K(j) or a completely blue K(k).

Examples:

- R(2,2) = 2
- R(2,3) = 3

We will now use the method to prove the following proposition:

$$\text{If } \binom{n}{k}2^{1-\binom{k}{2}} < 1 \text{ then } R(k,k) > n. \text{ Thus } R(k,k) > \lfloor 2^{k/2} \rfloor \text{ for all } k \geq 3$$

Proof. Consider a random 2-coloring of the edges of K(n) where each edge is colored independently with probability ½ of red and probability ½ of blue. For a set R of k vertices in K(n), we define A(R) to be the event that the subgraph on R is either all red or all blue. Then we have,

$$Pr(A(R)) = 2^{1-\binom{k}{2}}$$

To see this, note that there are $\binom{k}{2}$ edges in the subgraph on R. We can write A(R) as the disjoint union of the event the subgraph on R is all red or all blue, both of which have probability $(\frac{1}{2})^{\binom{k}{2}}$, so we get, $Pr(A(R)) = 2((\frac{1}{2})^{\binom{k}{2}}) = 2^{1-\binom{k}{2}}$. Since there are $\binom{n}{k}$ choices for R, the probability that at least one of these events happens is at most $\binom{n}{k}2^{1-\binom{k}{n}}$

Now we have shown,

$$Pr[\text{K(n) has a complete red or blue subgraph of size k}] \leq \binom{n}{k} 2^{1-\binom{k}{n}} < 1$$

Here we use the key insight of the probabilistic method. Since the probability of this not occuring is positive, there must exist a 2-coloring for which there is no complete red or blue subgraph of size k, which means that n cannot be the Ramsey number R(k,k), and therefore R(k,k) > n.

Furthermore, if we take k ≥ 3 and we set $n = \lfloor 2^{k/2} \rfloor$ then we have,

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{2^{1+k/2}}{k!} \frac{n^k}{2^{k^2/2}} < 1$$

So, in particular this tells us that R(k,k) > $\lfloor 2^{k/2} \rfloor$.

Some Remarks

- Could this have been done using a counting argument?
  - Yes, but it would probably be way more annoying, and later on we will see proofs where the probability is not so easily replaced.

- Can we actually construct a 2-color on K(n) that has no monochromatic K(k)?
  - The proof is non-constructive but does lead us towards a very efficient randomized algorithm!
  - This is much better than an exhaustive search since there are $2^{\binom{n}{2}}$ possible 2-colorings of K(n).