The Probabilistic Method in Combinatorics

Jacob Green
Department of Mathematical Sciences, The University of Bath

December 7, 2024

Abstract

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

Keywords: Combinatorics, Probability, Lovasź Local Lemma, Alteration

Preface

Notation

- $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$
- $[n] = \{1, \dots, n\}$
- if a is defined to be b we may say $a := b$ (but I will often forget)
- $A^k = \overbrace{A \times \cdots \times A}^{k \text{ times}}$ for a set A, $\times$ being the Cartesian product.

Contents

1 Introduction

Typically, the probabilistic method is used to prove the existence of a construction with desirable properties without actually constructing it. We do this by showing it exists with some positive probability.

A simple tool for doing so is the following.

Lemma 0 (a random variable must take its mean value)

Let X be an integer-valued random variable. Then

$$\mathbb{P}(X \geq \mathbb{E}[X]) > 0$$

Proof. Suppose $\mathbb{P}(X \geq \mathbb{E}[X]) = 0$. Then,

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{x=-\infty}^{\mathbb{E}[X]-1} x\mathbb{P}(X = x) + \sum_{x=\mathbb{E}[X]}^{\infty} x\mathbb{P}(X = x) \\
&\leq (\mathbb{E}[X] - 1) \sum_{x=-\infty}^{\mathbb{E}[X]-1} \mathbb{P}(X = x) + \sum_{x=\mathbb{E}[X]}^{\infty} x\mathbb{P}(X \geq x) \\
&= (\mathbb{E}[X] - 1)\mathbb{P}(X \leq \mathbb{E}[X] - 1) = \mathbb{E}[X] - 1
\end{aligned}$$

Absurd! $\square$

Clearly our argument works for any X taking values on a countable set. A continuous analogue also exists.

We can use this to prove any graph has a large bipartite subgraph. The idea will be to 2-colour its vertices uniformly, using the fact a bipartite subgraph is a graph where we can two colour the vertices such that every edge has endpoints with distinct colours.

Proposition 1 (large bipartite subgraph)

Every graph with m edges has a bipartite subgraph with at least $m/2$ edges.

Proof. Let our graph be $G = G(V, E)$ and assign each vertex a colour from $\{1, 2\}$. Let $E' \subseteq E$ be the set of edges with distinctly coloured endpoints. Then, enumerating $E = \{e_1, \ldots, e_m\}$ and letting A_i be the event edge e_i has distinctly coloured endpoints,

$$\mathbb{E}[|E'|] = \mathbb{E}\left[\sum_{i=1}^{m} \mathbf{1}(A_i)\right] = \sum_{i=1}^{m} \mathbb{P}(A_i) = \frac{1}{2}m$$

via the linearity of expectations and the fact $\mathbb{E}[\mathbf{1}(A)] = \mathbb{P}(A)$ for any event A, where $\mathbf{1}(\cdot)$ is the indicator function[1]. Hence, by our previous lemma

$$\mathbb{P}(|E'| \geq m/2) > 0$$

and we have our large bipartite subgraph. $\square$

Linearity of expectations as a tool is extremely prevelant in probabilistic combinatorics, so much so we dedicate an entire section to its enumerable applications.

[1]That is, $\mathbf{1}(A) = 1$ where A is true and 0 elsewhere

1.0.1 Lower Ramsey Bounds

Definition 2 (Ramsey number)

Let $k, \ell \in \mathbb{N}$. Then the Ramsey number $R(k,\ell)$ is the smallest $n \in \mathbb{N}$ such that in every 2-colouring of K_n we have a monochromatic K_k of colour or a monochromatic K_ℓ of colour 2.

For example, one can prove $R(3,3) = 6$ with the Pidgeonhole principle. WLOG we can choose a vertex such that 3 of the edges from this vertex have colour 1. Now consider the triangle formed by the edges connecting these 3 endpoints. If any one of the edges in this triangle have colour 1 then we have a monochromatic triangle, if not then we also have a monochromatic triangle, giving $R(3,3) \leq 6$. Then colour K_5 nicely to prove $R(3,3) > 5$.

Remark 3 (a nice interpretation)

Proving $R(3,3) \geq 6$ is a classical olympiad problem, often stated as "At a party with 6 people, there are 3 people that all know eachother or 3 people that do not know one another."

Ramsey first proved that these numbers exist.

Theorem 4 (Ramsey's Theorem - 1929)

For all $k, \ell \in \mathbb{N}$, $R(k,\ell)$ exists and is finite.

Proof. Omitted. □

Estimating the Ramsey number is a fundamental problem in Ramsey theory, a vast subfield of graph theory. We will prove 3 quantitative lower bounds for the diagonal Ramsey number $R(k,k)$, each serving as an exhibition of an important technique in probabilistic combinatorics.

The first technique we use will be the so called union bound, often referred to by probability theorists as Boole's inequality.

Lemma 5 (union bound / Boole's inequality)

Let $A_1, \ldots, A_n$ be events on a probability space $(\omega, \mathcal{F}, \mathbb{P})$. Then,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$$

We'll prove the $n = 2$ case, which gives the $n \in \mathbb{N}$ case immediately by induction.

Proof. Write $A_1 \cup A_2 = (A_1 \setminus A_2) \cup (A_2 \setminus A_1)$ as a disjoint union, noting $A_i \setminus A_j \subseteq A_i$ for $i \neq j$. Then,

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1 \setminus A_2) + \mathbb{P}(A_2 \setminus A_1) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2)$$

as required. □

Alternatively, the 2 variable inclusion-exclusion principle gives this immediately.

We may use this to bound unions from above, giving us a way to bound the probability of some "bad" event A_i from a set of "bad" events $A_1, \ldots, A_n$ occuring. Erdős used this methodology to obtain the following diagonal Ramsey lower bound, one of the first known applications of the probabilistic method in combinatorics.

Proposition 6 (Ramsey LB via union bound - Erdős 1947)

If $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$ then $R(k,k) > n$.

We will consider random 2-colourings of the edges of K_n, showing with positive probability if k satisfies $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ then K_n has no monochromatic K_k.

Proof. Independently and uniformly 2-colour the edges of K_n at random. There are $\binom{n}{k}$ subgraphs isomorphic to K_k, and these are exactly the graphs obtained by taking the subgraph induced by keeping k edges from K_n. Let $m = \binom{n}{k}$ and $A_1, \ldots, A_m$ be the events these graphs are monochromatic. Then,

$$\mathbb{P}\left(\bigcup_{i=1}^{m} A_i\right) \leq \sum_{i=1}^{m} \mathbb{P}(A_i) = \binom{n}{k} \times (2 \times 2^{-\binom{n}{k}})$$

so that, via DeMorgan's laws,

$$\binom{n}{k}2^{1-\binom{k}{2}} < 1 \implies \mathbb{P}\left(\bigcap_{i=1}^{m} A_i^c\right) > 0$$

giving us a 2-colouring of the edges of K_n with no monochromatic K_n. □

This bound doesn't, at first, tell us anything about the size of k relative to n. To deduce this fact we'll need to do some work with asymptotics.

Proposition 7 (quantitative Erdős lower bound)

For $k \in \mathbb{N}$,

$$R(k,k) > \left(\frac{1}{\sqrt{e}} + o(1)\right) k2^{k/2}$$

We will work with Stirling's approximation.

Proof. I can't get it to bloody work :/ □

The next approach we'll take is via the so called method of *alteration*. In this approach, generally, we start with a random construction, fix the undesirable features of said construction and prove our construction still has an interesting size. This approach is so important we'll dedicate a whole section to it.

In this application, as before, we'll take a 2-colouring of the edges of K_n independently and uniformly. This time though, instead of just analysing these colourings, we'll turn them into graphs K_X that certainly contain no monochromatic K_k and analyse the size of X.

Proposition 8 (Ramsey LB via alteration - unknown)

For all $k, n \in \mathbb{N}$, $R(k,k) > n - \binom{n}{k}2^{1-\binom{k}{2}}$

Proof. 2-colour the edges of the clique K_n independently and uniformly at random. Iteratively delete a vertex from each monochromatic K_k. Let $m = \binom{n}{k}$ and let $A_1, \ldots, A_m$ be the events the i^{th} k-clique is monochromatic after the initial colouring (before the deletion of any vertices). Let X be the number of vertices we delete during the deletion process. Then,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{m} \mathbf{1}(A_i)\right] = m\mathbb{P}(A_i) = \binom{n}{k}2^{1-\binom{k}{2}}$$

so, as we have $n - X$ vertices remaining after the alteration, we have

$$\mathbb{P}\left(n - X \geq n - \binom{n}{k}2^{1-\binom{k}{2}}\right) > 0$$

i.e. a 2-colouring of the edges of the clique size $n - \binom{n}{k}2^{1-\binom{k}{2}}$ containing no monochromatic K_k. □

Notice, again we used linearity of expectations here, as well as lemma 2.1. This will often be the case in alteration proofs.

As before, we can optimise n for fixed k to gain a quantitative bound.

Proposition 9 (quantitative Ramsey LB via alteration)

$$R(k,k) > \left(\frac{1}{e} + o(1)\right) k2^{k/2}$$

Proof. $\square$

Note this improves the Erdős bound by a factor of $\frac{1}{\sqrt{e}}$.

Now we introduce our third method, again, warranting its own section. The idea is that the so called "bad" events $A_1, \ldots, A_n$ we want to avoid are often only weakly dependent, and we have a great bound for the probability that none of the "bad" events occur in such a setup (courtesy of Lovasź).

Lemma 10 (Lovasź Local Lemma - random variable model)

Let $X_1, \ldots, X_n$ be independent random variables. Let $B_1, \ldots, B_m \subset [n]$ and, for each $1 \le i \le m$, let A_i be an event depending only on the random variables $\{X_j, j \in B_i\}$.

If for each i, $B_i \cap B_j \neq \emptyset$ for at most d other B_j's and $\mathbb{P}(A_i) \le 1/e(d+1)$ then with positive probability none of the A_i occur.

We will save the proof of this for our section dedicated to this lemma.

Now lets use this to obtain another improvement on our Ramsey lower bound. As before we'll take random 2-colourings and consider the events $A_1, \ldots, A_m$ where clique i is monochromatic, but, observing the weak dependence between these events (and letting the colour of the edges be our random variables) this makes a perfect candidate for the Lovasź Local Lemma.

Proposition 11 (Ramsey LB via LLL - Spencer 1977)

If $\left(\binom{k}{2}\binom{n}{k-2} + 1\right) 2^{1-\binom{k}{2}} < \frac{1}{e}$ then $R(k,k) > n$.

Proof. 2-Colour the edges of K_n independently and uniformly at random, letting the colour of $1 \le i \le \binom{n}{2} =: N$ be the random variable X_i. Enumerate the k-cliques of K_n as $G_1, \ldots, G_M$ where $M := \binom{n}{k}$ and let $A_j : 1 \le j \le M$ be the event G_i is monochromatic. Let $B_j \subset [N] : 1 \le j \le M$ be the edges of G_i. Clearly A_i depends only on the random variables $\{X_j : j \in B_i\}$.

Now $B_i \cap B_j \neq \emptyset$ exactly when G_i and G_j share at least one edge. This is equivalent (edges have two endpoints and two vertices make an edge!) to G_i and G_j sharing at least two vertices. Fixing G_i, we may count the number of G_j satisfying this condition by counting the number of ways to choose 2 vertices from G_i (our shared vertices) and multiplying by the number of ways to choose the remaining $k-2$ vertices from the vertices of K_n (noting we may choose from all n, as B_i can share any number of vertices with B_j) giving us a candidate d of

$$d = \binom{k}{2}\binom{n}{k-2}$$

Now, as we've already seen, $\mathbb{P}(A_i) = 2^{1-\binom{k}{2}}$, so to apply the Lovasź local lemma we need

$$\binom{k}{2}\binom{n}{k-2} 2^{1-\binom{k}{2}} < \frac{1}{e}$$

If this inequality in n and k holds, we have with positive probability a colouring of the edges of K_n with no monochromatic k-clique, and thus $R(k,k) > n$. $\square$

We may optimise one final time for n in k to get a quantitative bound.

Proposition 12 (quantitative Ramsey LB via LLL)

$$R(k,k) > \left(\frac{\sqrt{2}}{e} + o(1)\right) k2^{k/2}$$

Proof. □

This is improves our alteration bound by a factor of $\sqrt{2}$, and is currently the best known lower bound for the diagonal Ramsey number. Philosophically, it is interesting to note that all three of these different techniques only improve our lower bound by a constant factor. With these methods, we always obtain a $O(k2^{k/2})$ lower bound.

We are very far from knowing the exact diagonal ramsey number. Recently... (Campos upper bound exponential dialogue via a kind of alteration).

1.0.2 Set Systems

A *set system* is simply a family $\mathcal{F}$ of subsets of some space X. Often we take $X = [n]$. In *extremal set theory* we are usually concerned with finding the maximal/minimal size of a set system that has some desirable property.

The first type of set system we study are the so called *antichains*. These are families that are pairwise incomparable by containment, that is $\forall A \in \mathcal{F} \nexists B \in \mathcal{F}$ with $A \subseteq B$.

We will write $\binom{[n]}{k}$ for the family of k-element subsets of $[n]$.

Claim: The family $\binom{[n]}{k}$ is an antichain for any $0 \le k \le n$

Proof. Suppose there are k element subsets of $A, B \subseteq [n]$ with $A \subseteq B$. Then, as $|A| = k = |B|$ we have $A = B$ forced but, as our family is a set, its elements must be unique. Contradiction! □

The size of $\binom{[n]}{k}$ is $\binom{n}{k}$, which is maximised by taking $k = \lfloor n/2 \rfloor$. It turns out the family $\binom{[n]}{\lfloor n/2 \rfloor}$ is our biggest antichain in $[n]$.

Theorem 13 (Sperner's theorem - 1928)

If $\mathcal{F}$ is an antichain on $[n]$, then $|\mathcal{F}| \le \binom{n}{\lfloor n/2 \rfloor}$

Rather than proving this we will prove a stronger result independently discovered by several mathematicians, often referred to as the LYM inequality.

Theorem 14 (LYM inequality - 1954/1963/1965/1966)

If $\mathcal{F}$ is an antichain on $[n]$ then

$$\sum_{A \in \mathcal{F}} \binom{n}{|A|}^{-1} \le 1$$

Note this immediately uncovers Sperner's theorem as $\binom{n}{|A|} \le \binom{n}{\lfloor n/2 \rfloor}$.

The idea will be to consider random permutations of n and the canonical chain generated from this permutation. If we have two events in said canonical chain we no longer have an anti-chain, so an easy disjoint probabilities sum arises.

Proof. Choose a permutation σ or $[n]$ uniformly at random and consider the canonical chain

$$\mathcal{S}(\sigma) = \{\emptyset, \{\sigma(1)\}, \{\sigma(1), \sigma(2)\}, \ldots, \{\sigma(1), \ldots, \sigma(n)\}\}$$

For subsets $A \subset [n]$ let E_A be the event $A \in \mathcal{S}(\sigma)$. Then we have $A \in \mathcal{S}(\sigma)$ if and only if the elements of A are exactly the first $|A|$ elements in the listed permutation $(\sigma(1), \ldots, \sigma(|A|))$ up to reordering, leaving the remaining $n-|A|$ elements of $[n]$ to be the remaining elements $(\sigma(|A|+1), \ldots, \sigma(n))$ up to reordering. Thus, as we have $|A|!(n - |A|!)$ such permutations, we have

$$\mathbb{P}(E_A) = \frac{|A|!(n - |A|)!}{n!} = \binom{n}{|A|}^{-1}$$

Now suppose $\mathcal{F}$ is an antichain. Then clearly we have at most one $A \in \mathcal{F}$ with $A \in \mathcal{S}(\sigma)$, hence $\{E_A : A \in \mathcal{F}\}$ are disjoint events and we have

$$1 \geq \mathbb{P}\left(\bigcup_{A \in \mathcal{F}} E_A\right) = \sum_{A \in \mathcal{F}} \mathbb{P}(E_A) = \sum_{A \in \mathcal{F}} \binom{n}{|A|}^{-1}$$

$\square$

Here our probabilistic tool was actually just an axiom of measure, namely the sum of the measure of disjoint subsets is equal the measure of the union of said subsets. One could quite easily rephrase this as a counting argument by using disjointness and the fact we have at most $n!$ permutations, then dividing through by $n!$ at the very end.

Exercises

1.1 Linearity of Expectation

Exercises

1.2 Alteration

Exercises

1.3 Lovasź Local Lemma

Exercises

2 Solutions