How to Make Math Make Choices for You

Discrete Probabilistic Methods

Robin Truax

April 2021

Contents

1 Motivation

Given enough time and a typewriting, a monkey will eventually type any string of letters. It will type the complete works of Shakespeare, every paper of Erdös, and some *very* clever combinatorial proofs.

Almost all of math can be reframed in terms of "questions of existence"; that is, questions about if an object with a particular property exists or doesn't exist. The most famous open questions ($P = NP$, the Riemann Hypothesis, the Collatz Conjecture, etc.) ask questions about the existence of counterexamples. These questions of existence are often solved with a clever algorithm, a neat perspective, a new insight. But making these "smart choices" is hard. In this paper, we'll cover some examples of a method, called the probabilistic method, which can allow you to avoid making these smart choices. With the probabilistic method, randomness and probability do the heavy lifting for us.

The probabilistic method is, in essence, the process of making a totally random choice according to some distribution, and then showing that *some choice* has the desired outcome. The reason why the probabilistic method is so useful is because mathematicians are often unconcerned with how to actually *find* the objects they prove the existence of. In other words, the probabilistic method is highly nonconstructive (although it has often motivated and led to constructive proofs and algorithms).

Most of this post relies on nothing more complicated than basic probability theory and combinatorics. However, the language of abelian groups is utilized for brevity and rigor in the final example.

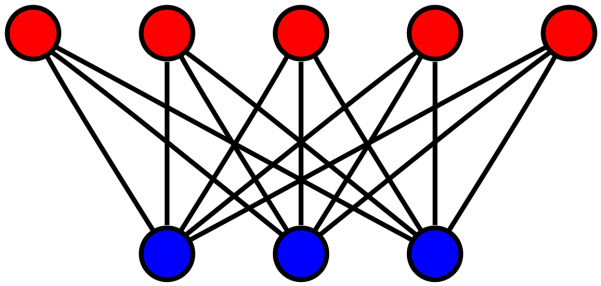
Finally, credit for these examples goes to Matthew Kwan and his course Math 159 (Discrete Probabilistic Methods) at Stanford. All of these were covered in lecture or assigned as problems in the first week of class. As a side note, I would recommend Matthew's classes (Math 159 and everything else he teaches) to anyone.

2 Bipartite Subgraphs

Our first question of existence will be about "bipartite graphs".

Definition 1 (Bipartite Graph)**.** A graph G with vertex set V is called *bipartite* if there exists a partition V_1, V_2 of V such that every edge of E is between a vertex of V_1 and V_2.

Following is an example of a bipartite graph (image source):



Understanding bipartite graphs can help us understand where *divisions* exist in many real-world situations. For example, if we consider a graph of social relationships, where an edge between two people signifies mutual dislike, then if this graph is bipartite, it could signify that the group can be divided into two tight-knit groups which dislike one another. The itch to cite American politics as an example is overwhelming.

Yet the real world is messy. Often our graphs are "almost" bipartite. A natural, if informal, question might to ask "how bipartite is my graph G?" A more formal way to phrase this question might be "what is the largest subgraph of G which is bipartite?" As we'll see, every graph is at least "half bipartite", in some sense.

Theorem 1. *If G is a graph with m edges, there exists a bipartite subgraph of G with at least $m/2$ edges.*

Proof. Suppose we have colored the vertices of G red or blue. Then, by keeping all the edges between vertices of different colors, we get a bipartite subgraph of G. Yet choosing the right coloring of vertices seems like it might be difficult – at the very least, it is nontrivial. Luckily, we can apply the probabilistic method.

Let E be the set of edges of G, and V be the set of edges. Instead of coloring each vertex intelligently, flip a coin to decide whether to color each vertex red or blue. Then, let E' be the subset of edges whose vertices are different colors, so that $G = (V, E')$ is a bipartite subgraph.

Now, notice that for any edge $e \in E$, the probability that the vertices of e are different colors is just $\frac{1}{2}$. Symbolically, we can write that $\mathbb{P}(e \in E')$ is $\frac{1}{2}$. Now, let 1_e be the following random variable:

$$1_e = \begin{cases} 1 & e \in E' \\ 0 & e \notin E' \end{cases}$$

Then, $\mathbb{E}[1_e] = \mathbb{P}(e \in E') = \frac{1}{2}$. Why might this be useful? Well, the number of edges in E' is simply the sum of these "indicator random variables"; that is, $|E'| = \sum_{e \in E} 1_e$. Therefore, via the linearity of expectation,

$$\mathbb{E}|E'| = \sum_{e \in E} \mathbb{E}[1_e] = \sum_{e \in E} \frac{1}{2} = \frac{|E|}{2} = \frac{m}{2}.$$

Yet consider what this means. This means that, on average, our coloring gives us a bipartite subgraph with $\frac{m}{2}$ edges. Therefore, clearly *some* coloring has to give us a bipartite subgraph with at least $\frac{m}{2}$ edges. □

Corollary 1.1. *If G has $m > 0$ edges, there exists a bipartite subgraph of G with* more than *$m/2$ edges.*

Proof. Since there is a positive-probability way to color the edges such that $|E'| = 0$ (ex. if all vertices are colored the same color), there *must* be some coloring with $E' > \frac{m}{2}$ in order to have $\mathbb{E}|E'| = \frac{m}{2}$. □

3 Coloring Graphs

A related question has to do with "coloring graphs".

Definition 2 (Proper Coloring)**.** A *proper coloring* of a graph G is an assignment of colors to its vertices such that no edge has two vertices of the same color.

Proper colorings of graphs are related to bipartite graphs in the sense that a graph has a proper coloring with 2 colors if and only if it is bipartite. However, most questions about graph coloring presume that all vertices can be colored the same. What if we can only color vertices with a small list of legal colors, where each vertex gets its own (possibly overlapping) list? Here, we give an answer for bipartite graphs.

Theorem 2. *Let $G = (V, E)$ be a bipartite graph with n vertices. For each vertex v, give a list $L(v)$ containing more than $\log_2 n$ colors. Then there is a proper coloring of G where each vertex v is colored with a color from its list $L(v)$.*

Proof. Let's let C be the set of all colors in any of the lists $L(v)$. Formally, C is the union $\bigcup_{v \in V} L(v)$. Also let V_1 and V_2 be the two parts of G. Then, if we can partition C into two sets C_1 and C_2 such that every vertex in V_1 can be "legally" colored with a color from C_1 and every vertex in V_2 can be "legally" colored with a color from C_2, then we are assured to get a proper coloring of G. This is because, if G is bipartite, any edge goes between V_1 and V_2, so the vertices of said edge will use different colors.

Therefore, our "smart choice" is the right partition of C. There are many options; in fact, there are $2^{|C|}$ partitions $\{C_1, C_2\}$ of C. Can you see why? In any case, choose one of these partitions uniformly at random.

Now consider a vertex v in part V_i. Let E_v be the event that there is no color in C_i contained in $L(v)$; more simply, E_v is the event that we cannot color v according to the rules using a color in C_i. Notice that for any given color c, $P(c \in C_i) = \frac{1}{2}$. Also notice that this probability is independent from the probability that any *other* color is contained in C_i. Therefore, we may conclude that

$$\mathbb{P}(E_v) = \bigcap_{c \in L(v)} \mathbb{P}(c \in C_i) = \left(\frac{1}{2}\right)^{|L(v)|} < \left(\frac{1}{2}\right)^{\log_2(n)} = \frac{1}{2^{\log_2(n)}} = \frac{1}{n}.$$

Now let E be the event that there exists a vertex v in either part V_i that cannot be legally colored with a color in C_i. Since $E = \bigcup_{v \in V} E_v$ – that is, E is the event that at least one of E_v happens – we see that

$$\mathbb{P}(E) \leq \sum_{v \in V} \mathbb{P}(E_v) < \sum_{v \in V} \frac{1}{n} = 1.$$

Therefore, E is not certain – in other words, failure is *not* inevitable. Formally, this implies that there exists some "smart choice" of a partition $\{C_1, C_2\}$ of C satisfying the following:

For every vertex v in either part V_i, there exists at least one color in C_i and $L(v)$.

Therefore, we can color each vertex v in part V_i with a color in C_i. Recall, as we discussed in the beginning, that this is a proper coloring; since G is bipartite, any edge is between a vertex in V_1 and a vertex in V_2, and since $\{C_1, C_2\}$ is a partition, these vertices will be colored differently. □

4 Families of Sets

Next, we'll study two totally different types of families of sets. Since sets are the most basic objects in modern math, the applicability of these results is widespread.

4.1 Antichains

Definition 3. Let $\mathcal{F}$ be a collection of subsets of $[n] = \{1, \ldots, n\}$. We say $\mathcal{F}$ is an *antichain* if no two distinct subsets $A, B \in \mathcal{F}$ satisfy $A \subseteq B$.

For example, let $\mathcal{F}$ be the set of all subsets of size $\lfloor n/2 \rfloor$ of $[n]$. Then, $|\mathcal{F}| = \binom{n}{\lfloor n/2 \rfloor}$. Notice that any element of $\mathcal{F}$ cannot contain another element of $\mathcal{F}$, since all elements of $\mathcal{F}$ are the same size. Therefore, $\mathcal{F}$ is an antichain. Can you find a bigger one? I wouldn't waste your time: this is as big as antichains get.

Theorem 3 (Sperner's Theorem)**.** *Every antichain of subsets of $[n] = \{1, \ldots, n\}$ has size at most $\binom{n}{\lfloor n/2 \rfloor}$.*

Proof. Let $\mathcal{F}$ be a family of subsets of $[n]$. Consider a chain $\varnothing = A_0 \subsetneq A_1 \subsetneq \cdots \subsetneq A_n$, where $|A_i| = i$. Now, these chains are in one-to-one correspondence with permutations σ of $\{1, \ldots, n\}$ by the correspondence

$$\sigma \Leftrightarrow \varnothing, \{\sigma(1)\}, \{\sigma(1), \sigma(2)\}, \ldots, \{\sigma(1), \ldots, \sigma(n)\}.$$

Therefore, choose one of the $n!$ random permutations σ of $\{1, \ldots, n\}$, and let X be the number of sets of $\mathcal{F}$ which appear in the corresponding chain. For each $A \subseteq \{1, \ldots, n\}$, let E_A be the event that A appears in the chain. As it turns out, the probability of E_A is dependent only on the size of A. Can you see why? In fact,

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

Now, by the linearity of expectation, $\mathbb{E}[X] = \sum_{A \in \mathcal{F}} \mathbb{P}(E_A)$. Yet notice that the binomial coefficient $\binom{n}{k}$ is maximized when k is "in the middle"; precisely, when $k = \lfloor n/2 \rfloor$ or $\lceil n/2 \rceil$. Therefore,

$$\binom{n}{|A|} \leq \binom{n}{\lfloor n/2 \rfloor} \Rightarrow \frac{1}{\binom{n}{|A|}} \geq \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \Rightarrow \mathbb{E}[X] \geq \sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{\lfloor n/2 \rfloor}}$$

In particular, if $\mathcal{F}$ has more than $\binom{n}{\lfloor n/2 \rfloor}$ elements, then $\mathbb{E}[X] > 1$. But this implies that there is a chain such that $\mathcal{F}$ contains at least two elements of said chain. But then one of these elements must contain the other, so $\mathcal{F}$ cannot be an antichain. Contrapositively, if $\mathcal{F}$ is an antichain, then it has at most $\binom{n}{\lfloor n/2 \rfloor}$ elements. □

4.2 Intersecting Families

Definition 4. Let $\mathcal{F}$ be a collection of subsets of $[n] = \{1, \ldots, n\}$. We say $\mathcal{F}$ is *intersecting* if any sets $A, B \in \mathcal{F}$ have nonempty intersection.

Now, finding the maximum size of an intersecting family of subsets is easy:

Theorem 4. *Any intersecting family of subsets of $[n]$ has cardinality at most 2^{n-1}. This bound is tight.*

Proof. Observe that for any set $A \subseteq [n]$, at most of one A and $[n] \setminus A$ can be contained in the family. Therefore, the family has maximum size 2^{n-1}, and this bound is maximized precisely when one of each pair is taken to form the family. Another way to show this bound is tight is by considering, for example, all the subsets which contain 1. There are 2^{n-1} such subsets. □

Yet, if we are only allowed to have subsets of the same size, the problem becomes slightly more interesting.

Theorem 5. *Any intersecting family $\mathcal{F}$ of k-element subsets of $[n] = \{1, \ldots, n\}$ has cardinality at most $\binom{n}{k}$ if $n < 2k$ and $\binom{n-1}{k-1}$ if $n \geq 2k$. This bound is tight.*

Proof. If $n < 2k$, any two k-element subsets of $[n]$ are intersecting, so we can choose up to $\binom{n}{k}$ such subsets.

On the other hand, assume that $n \geq 2k$. Now, we could get $\binom{n-1}{k-1}$ k-element subsets by, for example, taking all k-element subsets which contain 1. Indeed, this is best possible, which we show via the following argument:

Consider the integers $\{1, \ldots, n\}$ arranged in a circle. Say a subset of the integers is *contiguous* if it forms an unbroken segment in the circle. Observe that if an intersecting family $\mathcal{F}$ consists only of contiguous subsets

of $\{1, \ldots, n\}$ of size k, then $|\mathcal{F}| \leq k$ (if this is unclear, draw a picture). This relies on the assumption $n \geq 2k$, so that no contiguous subset can "wrap around".

Now, consider a uniformly random circle ordering of $1, \ldots, n$. For any $A \in \mathcal{F}$, the probability it is contiguous is $\frac{n}{\binom{n}{k}}$, as n is the number of contiguous k-element subsets (can you see why?) and $\binom{n}{k}$ is the number of all k-element subsets. Let $\mathcal{G}$ is the subfamily of $\mathcal{F}$ with only contiguous sets. We know from our above work that $|\mathcal{G}| \leq k$. Yet, if we take the expected value of the number of contiguous sets, we find that

$$\mathbb{E}[\mathcal{G}] = \frac{n}{\binom{n}{k}} |\mathcal{F}| \leq k$$

But with a little rearranging, we get the desired result:

$$|\mathcal{F}| \leq \frac{k\binom{n}{k}}{n} = \binom{n-1}{k-1}$$

□

This might feel like cheating, since the choice of randomness is itself a "smart choice". But I assure you that it's much easier than trying to make a direct smart choice which proves this bound.

5 Sum-Free Sets

Finally, we'll conclude with my favorite example, which almost serves as an opposite of Schur's Theorem (see §4.3 of Graph Theory to learn more).

Definition 5. A subset A of an abelian is *sum-free* if there do not exist $x, y, z \in A$ with $x + y = z$.

Theorem 6. *Every set A of n nonzero integers contains a sum-free subset of at least size $n/3$.*

Proof. For a real number x, denote the fractional part of x by $x \bmod 1$. Then $\mathbb{R}/\mathbb{Z}$ is an abelian group on the underlying set $(0, 1]$ where addition happens mod 1. Yet notice that $\left(\frac{1}{3}, \frac{2}{3}\right)$ is a sum-free subset of $\mathbb{R}/\mathbb{Z}$. Now let $A_x = \{a \in A \mid ax \bmod 1 \in \left(\frac{1}{3}, \frac{2}{3}\right)\}$. Then, for any $x \in \mathbb{R}$, A_x is a sum-free subset of A. This follows because if $a + b = c$, then $ax + bx \equiv cx \bmod 1$, which is impossible since $\left(\frac{1}{3}, \frac{2}{3}\right)$ is sum-free in $\mathbb{R}/\mathbb{Z}$.

Choose x randomly (and uniformly) in $(0, 1]$, and note that if a is an integer then $ax \bmod 1$ is also uniform in $\mathbb{R}/\mathbb{Z}$. Therefore, via the linearity of expectation,

$$\mathbb{E}|A_x| = \sum_{a \in A} \mathbb{P}(a \in A_x) = \sum_{a \in A} \mathbb{P}\left(ax \bmod 1 \in \left(\frac{1}{3}, \frac{2}{3}\right)\right) = \sum_{a \in A} \frac{1}{3} = \frac{n}{3}$$

But that means there must be a choice of x for which $|A_x| \geq \frac{n}{3}$. Then, A_x is sum-free, so we are done. □

Unfortunately, this proof required another smart choice, but I hope the amazing brevity and power of this result makes up for it. Luckily, this smart choice generalizes to other sums (say $x + y = z + w$ or $x + 2y = 3w + z$) quite easily. Try to prove similar results about "strange-sum-free" sets!

6 Conclusion

As we learned near the end, even with the probabilistic method, smart choices can be helpful. Indeed, maximizing the power of the technique and finding tighter bounds can be extremely difficult, and the field has many open questions. Yet, like a game with cheats, it would've been a lot less fun otherwise, wouldn't it? As disappointing as it may seem, the fact that math still requires creativity even with these clever techniques is reassuring. Perhaps it'll help us stave off the AI revolution for just a bit longer.