Discrete Math and The Probabilistic Method

Mark Gillespie

October 26, 2017

1 Discrete Math

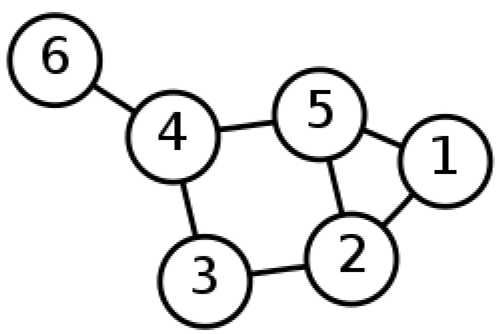
The word *discrete* means separate or distinct. Discrete math is a general term for the study of things that are composed of separate chunks. Most of the things you've studied so far are 'continuous'.

For instance, if you have a function f which takes in real numbers, its input can vary smoothly. You can plug in all of the numbers between 0 and 1. By contrast, someone studying discrete mathematics might look at a function which takes in integers. Then, you can't smoothly change the input. The function can only take in *discrete* values.

I don't have a great definition of what discrete math is in general. But I'll be talking about some ideas from discrete math today, and hopefully you can see how they have a different flavor from calculus and continuous math.

1.1 Graphs

One very important idea in discrete math is the idea of a *graph*. A graph is a composed of a set of vertices, and a set of edges that go between the vertices. You can draw a graph like this.



The edges can have a direction associated with them if you want

Often, you use graphs to represent connections between things.

- Roads between cities
- Friendships on Facebook
- Atoms with bonds between them
- Particles that feel forces from each other
- Websites that link to each other
- Food webs

Can you think of other examples?

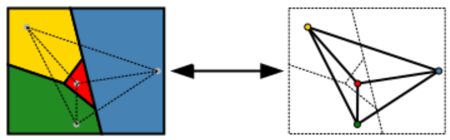
1.2 Graph Coloring

1.2.1 Vertex Coloring

One classic thing that people do with graphs is color them! We say that a vertex coloring of a graph is an assignment of a color to each vertex, where no two vertices connected by an edge have the same color.

The motivation for this problem comes from map-making.

Imagine you have a map of several countries. You want to color the countries different colors, and you want to make sure that no two countries that share a border are the same color. This would make it hard to read your map. We can turn this problem into a problem about graphs by making each country a vertex, and connecting two countries by an edge if they share a border.



Then, a graph coloring gives us a way of coloring our map in the way that we wanted!

Theorem 1 (Four Color Theorem, Appel and Haken 1976)**.** *Any map can be colored with four colors.*

The proof is hard.

1.2.2 Edge Coloring and Ramsey Theory

We can also color the edges of a graph instead of the vertices. An edge coloring of a graph is an assignment of a color to every edge in the graph. This time, we won't enforce any constraints on how the colors are assigned.

Here's an example. Consider the complete graph on n vertices, which is the graph with n vertices and an edge between every pair of vertices. Suppose we color the edges of the graph with two colors, red and blue. We can give this sort of coloring a fun interpretation. We can say that the vertices represent people, and two people have a red edge between them if they know each other and a blue edge otherwise.

> Here's a fun problem: What's the smallest number of people you need in order to guarantee that either there are 3 who all know each other, or there are 3 none of whom know each other. In fancier math words, what is the smallest n such that every 2-coloring of the complete graph on n vertices has a monochromatic triangle. This is a fun problem, and has a really cool solution. I don't have time to get into it at the moment, but I encourage you to think about it at some point.

The study of this sort of question (questions about how big a structure needs to be before you can guarantee that it has some property) is known as Ramsey Theory.

2 Combinatorics

Combinatorics is the math of counting. That makes it sound easy, but it is deceptively hard. There is a neat application of combinatorics to the *Silver Bar Problem.* Fiona told me that you were trying to prove that you can't solve the problem with fewer than 5 pieces. You can prove this really nicely just by counting!

Proof. We want to pay the landlady 1 unit of silver every day for 31 days. In particular, this means that the landlady should have a different amount of silver every day for 31 days. So you need to be able to giver her at least 31 distinct amounts of silver.

Suppose the bar is cut into 4 or fewer pieces. With 4 pieces, you can give her at most $2^4 = 16$ combinations of the pieces. So you can't possibly give her 31 distinct amounts of silver. The same is true if you cut the bar in fewer than 4 pieces. Therefore, you have to cut your bar into at least 5 pieces in order to pay through the whole month! □

This sort of argument is what combinatorics is all about - counting things in clever ways.

There are a lot of combinatorial questions in graph theory. For example, Ramsey theory falls under the domain of combinatorics - answering the question of whether a certain type of graph exists or not is the simplest type of counting of all! Pretty much everything I'll said today can be called combinatorics.

3 The Probabilistic Method

Reference: *The Probabilistic Method* by Alon and Spencer

In a nutshell, the probabilistic method is about using probability to prove definite results (i.e. results that don't mention probability).

A lot of questions in combinatorics are of the form "Does a structure with property X exist"? Ramsey theory is an example of this. We can phrase the question I asked earlier about monochromatic triangles as: what is the largest n for which there exists a graph with n vertices and no monochromatic triangles?

The probabilistic method is a powerful, and really cool, method for answering this sort of question. The general strategy is as follows: We consider picking a random structure somehow. Then, we show that the probability that it has property X is nonzero. Thus, it must be possible to pick a structure with property X. So such a structure must exist!

This method can seem very roundabout and overly convoluted, but it can actually lead to some beautiful proofs of facts that are hard to prove otherwise.

3.1 Ramsey Theory Example (Will be included if I have time)

Here, we'll look at a slight generalization of the Ramsey Theory problem I mentioned earlier. Before, we looked at how big a graph has to be before there's a set of 3 people who are all friends or who are all strangers to each other. Now, let's ask how big a graph has to be before there's guaranteed to be a set of k people who are all friends or who are all strangers. This turns out to be a very hard question. But we can find some big graphs that don't have this property using the probabilistic method.

Theorem 2. *If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, then it's possible to have a set of n people such that no subset of k people are all friends or all strangers.*

Proof. Consider a random assignment of friendships, where every pair of people is said to be friends with probability $\frac{1}{2}$. Given any set R of k people, we let A_R denote the event that they are all friends or all strangers.

$\Pr[A_R] = 2^{1-\binom{k}{2}}$ (Why?).

Since there are $\binom{n}{k}$ choices for R, the probability that at least one of the A_R happens is $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$. Therefore, the probability that no A_R happens is greater than 0. So it must be possible to have a set of n people where no set of k are all friends or all strangers! $\square$

After some algebra[1] you can show that this means that there's always a set of $\lfloor 2^{k/2} \rfloor$ people with no set of k friends or k strangers, as long as $k \geq 3$.

3.2 Useful Tool: Linearity of Expectation

A very useful tool in probability is the expected value. Suppose X is a random variable that has outcomes $x_1, \ldots, x_k$ with probabilities $p_1, \ldots, p_k$. (E.g. a die is a random variable D that has outcomes $1, \ldots, 6$ with

[1] Let $k \geq 3$ and $n = \lfloor 2^{k/2} \rfloor$

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

probabilities $\frac{1}{6}, \ldots, \frac{1}{6}$). Then the expected value of X is

$$\mathbb{E}[x] = \sum_{i=1}^{k} x_i p_i$$

In our die example, this gives us that $\mathbb{E}[D] = \frac{1}{6}(1+2+3+4+5+6) = \frac{1}{6}(21) = 3.5$

An important, and sometimes counter-intuitive, fact about expected values is that they are linear. Suppose we want to know the expected value of the sum of two die rolls. Linearity of expectation tells us that we can compute this by adding together the expected values of the individual die rolls. That tells us that $\mathbb{E}[D+D] = 7$.

There's a useful trick that we can use to convert from probabilities to expected values. Suppose we're flipping a biased coin that lands on heads with probability p and tails with probability $1-p$. We can express this as an expected value as follows: let $\mathbb{1}_p$ be a function which is 1 if the coin lands on heads and 0 if the coin lands on tails. $\mathbb{1}_p$ is called an *indicator function*. Then $\mathbb{E}[\mathbb{1}_p] = 1 \cdot p + 0 \cdot (1-p) = p$. So we've converted a probability into an expected value! This might not seem very exciting, but it is, because once you have an expected value you can use linearity of expectation, which is very powerful.

Here's another fun riddle that I don't have time to answer.
N gentlemen check their hats in the lobby of the opera, but after the performance the hats are handed back at random. How many men, on average, get their own hat back?

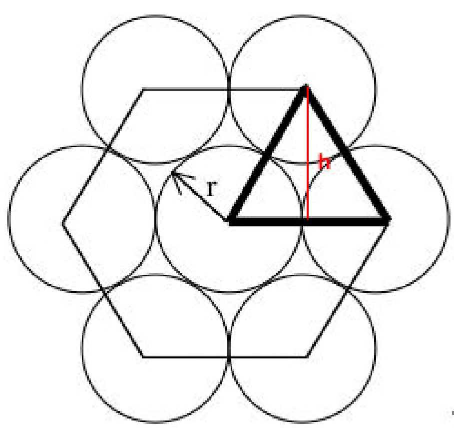
Opaque hint: use linearity of expectation

3.3 Coins on Dots Example

Here's a fun problem: On the table before us are 10 dots, and in our pocket are 10 nickels. Prove the coins can be placed on the table (no two overlapping) in such a way that all the dots are covered.

Proof. Fix any set of dots on the table. We want to use the probabilistic method, so we want to place our coins randomly on the table. But placing each coin randomly on the table while enforcing the constraint that no two overlap sounds really complicated. So we'll do something else instead.

Imagine tiling the whole table with hexagonally-packed circles.



Then, whenever a circle contains a dot, we'll place a coin on that circle. Because we're using hexagonally-packed circles, the coins we place will never overlap. Since there are 10 dots and 10 coins, we can cover all of the dots as long as each dot is in a circle.

Now, we want to show that if we position the lattice randomly, we have a nonzero probability of getting all of the dots inside of the circles.

First, let's work out the probability of 1 dot winding up inside of some circle. Since we place the lattice randomly, this is just the fraction of the plane that the circles cover. This is the same as the fraction of the area inside of that triangle that the circles cover. Suppose the triangle has side-length 2. Then the triangle's area is

$$A_\Delta = \sqrt{3}$$

And the area covered by the circles is

$$A_\circ = 3 \cdot \frac{1}{6} \cdot \pi \cdot 1^2 = \frac{\pi}{2}$$

Therefore, the proportion of area covered by the circles is $\frac{\pi}{2\sqrt{3}} \sim 0.906$.

So the probability that we cover any given dot with a circle is $p = \frac{\pi}{2\sqrt{3}} \sim 0.906$. Now, we want to think about covering multiple circles. To combine probabilities about different events, we want to use linearity of expectation! So let's apply our trick with indicator variables. Let $d_1, \dots, d_{10}$ be the 10 dots. Let $\mathbb{1}_{d_i}$ be the function that is 1 when dot i is covered by a circle and 0 otherwise. Then $\mathbb{E}[\mathbb{1}_{d_i}] = p$.

Now, what does $\mathbb{1}_{d_1} + \cdots + \mathbb{1}_{d_{10}}$ mean? It's the number of dots that are covered by circles.

$$\mathbb{E}[\mathbb{1}_{d_1} + \cdots + \mathbb{1}_{d_{10}}] = \mathbb{E}[\mathbb{1}_{d_1} + \cdots + \mathbb{E}[\mathbb{1}_{d_{10}}] = 10p \sim 9.06 > 9$$

So the expected number of dots covered by circles is greater than 9. This means that there must be some placement of the lattice such that all 10 dots are inside circles! Therefore, we can cover all 10 dots with non-overlapping coins. $\square$