Day 2: Probability

Basics of Probability Review

The probability of an event happening is (desired outcomes)/(total possible outcomes)

AND implies you should multiply, OR implies you should add.

Examples:

The probability I select a pink candy from a set of 5 pink candies and 6 brown candies = 5/(6 + 5) = 5/11.

The probability of selecting pink candy, brown candy, pink candy without replacement in that exact order is 5/11 * 6/10 * 4*9 = 12/99. (pink AND brown AND pink, so multiply)

The probability of selecting a pink or brown candy from a set of 5 pink, 4 brown, and 3 blue is 5/12 + 4/12 = 9/12 = 3/4 (pink OR brown, so add)

Probability Continued, Combinatorics application

Probability of selecting 3 pink or 4 brown candies from a set of 5 pink candies and 6 brown candies?

It is NOT 3/11 + 4/11!

Remember, probability of an event = (desired outcomes)/(total outcomes). The total number of ways I can select 3 pink candies is 5 choose 3 = 5!/(3! * 2!) = 10, and the total number of ways I can select 4 brown candies is 6 choose 4 = 6*5/2 = 15. The total number of ways of selecting 3 candies from a set of 11 is 11 choose 3 = 11!/(3!*8!) = 11*10*9/6 = 165, and the total number of ways of selecting 4 candies from a set of 11 = 11!/(4!*7!) = 330, so the probability is 10/165 + 15/330 = 35/330 = **7/66.**

Probability of selecting a pink candy, a brown candy, and another pink candy, but not necessarily in that exact order?

It is NOT 5/11 * 6/10 * 4*9 = 12/99 like in the previous slide!

We have to multiply this probability by an extra 3!/2! = 3 because there are 3 different orders we can have: {pink, brown, pink}, {brown, pink, pink}, or {pink, pink, brown}. So the correct answer is 12/99 * 3 = **4/11.**

⚠️ Important side note: Sometimes you will be working with much larger numbers, so it's important to understand that the number of permutations of any n distinct objects is n!. Additionally, if some of those objects repeat, you must divide by the number of times they repeat factorial. So for example, if I need the permutations of 13 cars, with 7 of them being blue and 6 of them being red, I would do 13!/(7! * 6!) = 1716 permutations or different ways to rearrange them.

Geometric Probability

We can visualize probabilities using 2d and 3d graphs. For example, the probability two numbers from 0 to 1 sum to 1/2 or more can be modeled by the area above the graph x + y = 1/2 on the unit square.

Example 1: I select two real numbers, both from the interval [0,1]. What is the probability that their sum is less than 1/3?

We can model this with a graph, with the y axis representing the first number chosen and the x axis representing the second. We essentially need the area under x + y = 1/3, or y = 1/3 - x, in a 1x1 unit square. This creates a right triangle with base 1/3 and height 1/3, so the answer is simply 1/3 * 1/3 * 1/2 = **1/18.**

Example 2: I select three real numbers, all from the interval [0,1]. What is the probability that their sum is less than 1/3?

Now, here we need the area of the tetrahedron bounded by x + y + z < 1/3 in a 1x1x1 unit cube. Note that the area of a tetrahedron bounded by the points (0,0,0), (a,0,0), (0,b,0), and (0,0,c) is 1/6 * |abc|. In this case, our tetrahedron is bounded by the points (0,0,0), (1/3,0,0), (0,1/3,0), and (0,0,1/3), so the desired region is 1/6 * 1/3 * 1/3 * 1/3 = 1/6 * 1/27 = **1/162.**

Example 3: (AIME 1998 #9) Two mathematicians take a morning coffee break each day. They arrive at the cafeteria, independently, at random times between 9 a.m. and 10 a.m., and stay for exactly m minutes. The probability that either one arrives while the other is in the cafeteria is 40%, and m = a – b√c, where a, b, and c are positive integers, and c is not divisible by the square of any prime. Find a + b + c.

Again, we can model this problem with a square, only this time it's a 60 x 60 square since there are 60 minutes in an hour. Let the x axis represent the possible times of the first mathematician, and let the y axis represent the possible times of the second mathematician. Essentially, the difference in times of when they arrive must be less than m, because if it were greater, then they would not meet at all. This can be modeled with the equation |x - y| < m, or the area under |x - y| = m in a 60x60 unit square. This is 3600 - (60 - m)^2. Now the problem tells us this probability is equal to 40% or 2/5, so we get (3600 - (60 - m)^2)/3600 = 2/5. ==> m = 60 - 12 √15. So the answer is 60 + 12 + 15 = **87.**

Practice!

1. Johnny randomly selects two numbers, one from the interval [0,2025] and the other from the interval [0,2026]. What is the probability that the sum of the two numbers he chooses is less than 2025?
2. Two mathematicians take a morning coffee break each day. They arrive at the cafeteria independently, at random times between 9 a.m. and 10 a.m., and stay for exactly m minutes. The probability that either one arrives while the other is in the cafeteria is 40%, and $m = a - b\sqrt{c}$, where a, b, and c are positive integers, and c is not divisible by the square of any prime. Find a + b + c.
3. I am rolling 5 standard 6 sided dice. The probability that the faces do not sum to 7 is? (Hint: Stars and Bars!)
4. Two fair dice, each with at least 6 faces, are rolled. On each face of each die is printed a distinct integer from 1 to the number of faces on that die, inclusive. The probability of rolling a sum of 7 is 3/4 of the probability of rolling a sum of 10 and the probability of rolling a sum of 12 is 1/12 . What is the least possible number of faces on the two dice combined?

More difficult problems:

5. There is a 3/4 chance it rains the day after tomorrow and a 1/2 chance it rains tomorrow. However, if it rains tomorrow, then the probability it rains the day after jumps to 15/16. What is the probability it rains at least once in the next two days?
6. A Caltech prefrosh is participating in rotation. There are 8 houses at Caltech: Avery, Blacker, Dabney, Fleming, Lloyd, Page, Ricketts, and Venerable. The prefrosh visits these houses in some order, each of them exactly once. Throughout rotation, the prefrosh maintains a ranking list of all of the houses that the prefrosh has visited. After every visit to a house, the prefrosh updates the ranking list by inserting the most recently visited house to either the top or the bottom of the list, each with probability 1/2, while keeping the order of all previously visited houses the same. Compute the probability that at the end of rotation, third house the prefrosh visited is not ranked fourth or fifth on their list.
7. Jason rolls three fair standard six-sided dice. Then he looks at the rolls and chooses a subset of the dice (possibly empty, possibly all three dice) to reroll. After rerolling, he wins if and only if the sum of the numbers face up on the three dice is exactly 7 . Jason always plays to optimize his chances of winning. What is the probability that he chooses to reroll exactly two of the dice?