Combinatorics

Bhairav Chidambaram and Eshaan Nichani

July 10, 2014

1 Topics

1.1 Basics

The number of ways to choose r items from a group of n items is

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

Pascal's Identity:

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$$

Try proving this both combinatorically and algebraically.

Hockey-Stick Identity:

$$\binom{k}{k} + \binom{k+1}{k} + \cdots + \binom{n}{k} = \binom{n+1}{k+1}$$

The number of paths going only up and right from (0,0) to (m,n) is

$$\binom{m+n}{m} = \binom{m+n}{n}$$

(Proof: Consider sequences of m r's and n u's)

1.2 Revisiting Factorials

Let's take a look at how to understand $\binom{n}{r}$. We know that it computes the number of ways to choose r objects from a group of n. But what are the factorials in the formula doing?

Consider randomly ordering the set of n objects, and then analyzing the first r. By dividing by $r!$ and $(n-r)!$, we are essentially unordering the first r objects as well as the remaining $n-r$ objects. This leaves us with just the number of ways to choose those r (or n-r) objects.

$$[1, 2, ...r][r+1, r+2, ...n]$$

Let's apply this idea to a problem to get a better understanding. How many unique ways are there to order the letters in 'MISSISSIPPI'? (1 M, 4 I, 4S, 2P)

If all the letters were unique, the answer would simply be 11!. But this is not the case. We know that the 4 S are not unique, but how can we account for this? In the orginial permutation the possibilities of the S were overcounted by a factor of 4!, because they were treated as unique. Compensating for the 4I, 4S, and 2P, we see that the answer is: $\frac{11!}{4!4!2!} = 34650$
This new understanding of factorials will be helpful in a range of situations, so make sure you see why it works.

1.3 Stars and Bars

Stars and bars is a technique used to solve distribution problems, where some number of items must be distributed between some number of people. For example, how many ways are there to distribute 7 indistinguishable items between 4 people? One way is to list out the all the possibilities and use casework. However, a more elegant solution is to use stars and bars.

To determine the number of ways n items can be distributed between k people, consider n *stars*, which represent the items, and $k-1$ *bars*. Any arrangement of these will result in a distribution, as the first person will get the number of items from the beginning to the first bar, the second person will get the number of items from the first bar to the second bar, and so on. Since there are $\binom{n+k-1}{n}$ ways to arrange these items, that is the number of ways the n items can be distributed. For example, consider the following ordering of 7 stars and 3 bars:

$$*| * * * | * *|*$$

This corresponds to the first person getting 1 item, the second 3, the third 2, and the fourth 1. Stars and bars can also be used to determine the number of non-negative integer solutions to $a_1 + a_2 + \cdots + a_k = n$, since this can be rephrased as distributing n between the k integers.

1.4 Principle of Inclusion-Exclusion

Given two sets A and B, the number of elements in at least one of the sets is the sum of the number of elements in A and B, minus the number of elements in both sets. In other words,

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

For three sets, we have that

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

To see why, consider a venn diagram, and look at which regions each of the terms in the previous expression represent. This generalizes to more than 3 sets. To find the number of elements in the union of n sets just add the number of elements in each set, then subtract the number of elements in at least 2 sets, then add the number of elements in at least 3 sets, and so on.

Example: In a school with 500 students, 276 take Spanish, 54 play baseball, 38 are in the band, 43 play baseball and take Spanish, 20 are in the band and take Spanish, 7 play baseball and are in the band, and 2 do all 3. How many students do none of the activities?

Solution: The number of students that do all three activities is 276 + 54 + 38 - 43 - 20 - 7 + 2 = 300. Thus 200 students do none of the activities.

1.5 Expected Value

The expected value of an event is the "average value" that some event takes. To calculate it, take the sum, over all possible outcomes, of the value an event attains multiplied by the probability that it occurs.

Example: What is the expected value of the number that appears on a fair dice?
Solution: Each number has a probability of $\frac{1}{6}$ of appearing, so the expected value is

$$1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = 3.5$$

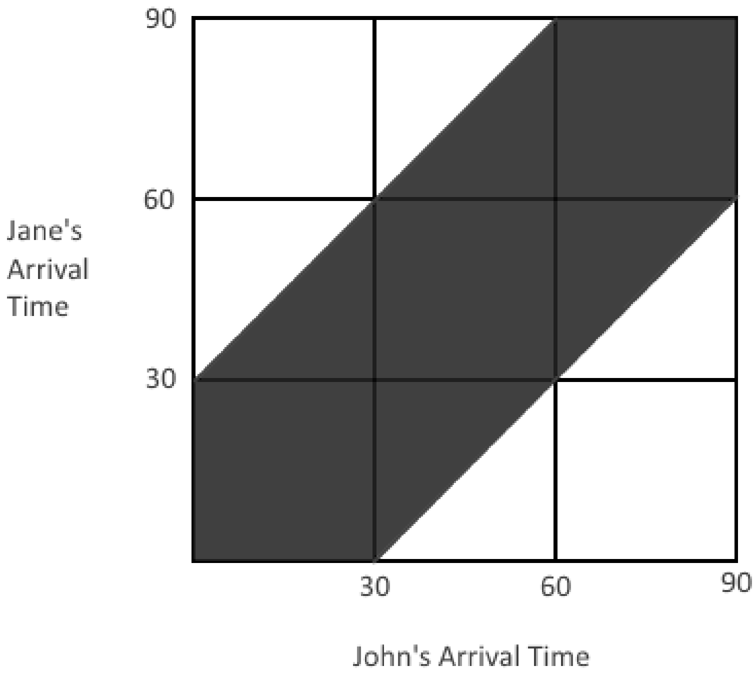
One useful fact about expected value is that for two events X and Y, the expected value of $X + Y$ is the expected value of X plus the expected value of Y. This applies for any events, regardless if they are independent or not.

1.6 Geometric Probability

Geometric probability is useful for problems that involve the overlap of ranges. The basic premise is to set up the probability as an area that represents the set of all instances that satisfy the problem's condition. In particular, this helps find the probability for situations that do not have a finite number of possible events (Ex. set of moments in time, set of all points in a segment, etc).

John and Jane are invited to a 90 minute party. Each will arrive at some time during the party, and stay for 30 minutes or until the party ends. What is the probability they will see eachother?

Set up the graph with two axes as follows. On both the x and y axes is the range 0 to 90. The x axis represents the arrival time of John and the y axis of Jane. Now imagine we shade each segment that represents the times that Jane could arrive to see John given he arrives a particular time x. The probability is simply the area shaded divided by the total area. The graph should look similar to the one on the right. Take some time to understand why this works.



1.7 Recursion

It is often helpful to relate the number of ways to do something with n items to the number of ways to do it with n-1, n-2, etc. items. This method is known as recursion. After finding a recurrence relation, the best way is to calculate each of the terms explicitly up to the number you want to find.

Example: How many ways are there to tile a 1 by 6 rectangle with 1 by 1 and 1 by 2 rectangles?

Solution: Let a_n denote the number of ways to tile a 1 by n rectangle. Consider the first tile. If it is a 1 by 1 tile, then it remains to tile a 1 by n-1 rectangle, which can be done in a_{n-1} ways. If the first tile is a 1 by 2, then you need to tile a 1 by n-2 rectangle, which can be done in a_{n-2} ways. Thus

$$a_n = a_{n-1} + a_{n-2}.$$

Since $a_1 = 1$ and $a_2 = 2$, we can calculate a_6 using the recurrence, yielding $a_6 = 13$. Thus there are 13 ways to tile the rectangle.

2 Tips/Tricks

1. **Look for patterns in smaller cases.** Examining a smaller, easier case can help you identify processes or patterns that may lead you directly to the solution or allow you to make valuable observations. This is a good first step if you are unable to get traction on a problem. Note that not every problem will provide something useful by working with smaller cases.

2. **Identify and exploit symmetry in a problem.** This is difficult but can lead to a very nice solution. Here is a common example: How many two-digit numbers using digits 1-9 have their first digit greater than their second? The solution, using symmetry, is to take half of the numbers that have 2 unique digits, since the second digit has an equal likelihood of being greater than or less than the first.

3. **Organize your casework.** Many problems will require casework. The solution becomes less about ingenuity and more about organization and efficiency. If necessary, start on a new page. They point is to ensure that you never count the same thing twice. Make sure you avoid aimless bashing. If you find yourself wasting time bashing out cases, stop and experiment to see if you can derive a better solution. Briefly evaluate how long the casework will take to see if it is worth your time (which is very precious!!). If you must do the casework, then commit and execute.

4. **Complimentary counting.** Sometimes it may be easier to count the complement of the event you are trying to find.

3 Problems

1. Daphne is visited periodically by her three best friends: Alice, Beatrix, and Claire. Alice visits every third day, Beatrix visits every fourth day, and Claire visits every fifth day. All three friends visited Daphne yesterday. How many days of the next 365-day period will exactly two friends visit her?

2. How many cubic polynomials $f(x)$ with positive integer coefficients are there such that $f(1) = 9$?

3. The numbers 1, 2, 3, 4, 5 are to be arranged in a circle. An arrangement is *bad* if it is not true that for every n from 1 to 15 one can find a subset of the numbers that appear consecutively on the circle that sum to n. Arrangements that differ only by a rotation or a reflection are considered the same. How many different bad arrangements are there?

4. What is the expected number of coin flips to get three heads in a row?

5. Real numbers x, y, and z are chosen independently and at random from the interval $[0, n]$ for some positive integer n. The probability that no two of x, y, and z are within 1 unit of each other is greater than $\frac{1}{2}$. What is the smallest possible value of n?

6. Find the number of five-digit positive integers, n, that satisfy the following conditions:

   (a) the number n is divisible by 5,
   (b) the first and last digits of n are equal, and
   (c) the sum of the digits of n is divisible by 5.

7. Jackie and Phil have two fair coins and a third coin that comes up heads with probability $\frac{4}{7}$. Jackie flips the three coins, and then Phil flips the three coins. Let $\frac{m}{n}$ be the probability that Jackie gets the same number of heads as Phil, where m and n are relatively prime positive integers. Find $m + n$

8. One fair 6 sided die is rolled; let a denote the number that comes up. We then roll a dice; let the sum of the resulting a numbers be b. Finally, we roll b dice, and let c be the sum of the resulting b numbers. Find the expected value of c

9. An integer between 1000 and 9999, inclusive, is called balanced if the sum of its two leftmost digits equals the sum of its two rightmost digits. How many balanced integers are there?

10. In a small pond there are eleven lily pads in a row labeled 0 through 10. A frog is sitting on pad 1. When the frog is on pad N, $0 < N < 10$, it will jump to pad N-1 with probability $\frac{N}{10}$ and to pad N+1 with probability $1 - \frac{N}{10}$. Each jump is independent of the previous jumps. If the frog reaches pad 0 it will be eaten by a patiently waiting snake. If the frog reaches pad 10 it will exit the pond, never to return. What is the probability that the frog will escape without being eaten by the snake?

11. A collection of 8 cubes consists of one cube with edge-length k for each integer $k, 1 \leq k \leq 8$. A tower is to be built using all 8 cubes according to the rules:

    * Any cube may be the bottom cube in the tower.
    * The cube immediately on top of a cube with edge-length k must have edge-length at most $k + 2$.

    Let T be the number of different towers that can be constructed. What is the remainder when T is divided by 1000?

12. Let $A = \{1, 2, 3, 4\}$, and f and g be randomly chosen (not necessarily distinct) functions from A to A. The probability that the range of f and the range of g are disjoint is $\frac{m}{n}$, where m and n are relatively prime positive integers. Find m.