Math 96:
Combinatorics and Linearity of Expectation

November 22nd, 2021

1 Pigeonhole Principle

Combinatorics is often about counting things. One of the most important applications for counting is that you can often obtain results just from knowing that the set of possibilities is too big or too small. One common application of this is the Pigeonhole Principle:

Lemma. *If n pigeons are placed in $m < n$ holes, then there must be some hole with at least 2 pigeons in it.*

This hold because if there were at most one pigeon per hole, we could fit at most m pigeons. We can generalize this to say that if $n > km$, there must be a hole with at least $k+1$ pigeons.

1956 A2: Prove that every positive integer has a multiple whose decimal representation involves all ten digits.

2 Binomial Coefficients

One of the most basic counting tools are binomial coefficients. The number $\binom{n}{k}$ counts the number of ways to select k items from a set of size n. One way to count this is to note that if you want to pick an *ordered* set of k different items, there are n ways to select the first, $n-1$ ways to select the second and so on through $(n-k+1)$ ways to select the last. However, this overcounts what we want since each set has $k!$ different orders. However, $\binom{n}{k}$ can be computed as the ratio,

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

Binomial coefficients also satisfy a number of useful identities like the Binomial Theorem:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k},$$

and the recurrence

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

1956 A5: Given n objects arranged in a row, a subset of these objects is called unfriendly if no two its elements are consecutive. Show that the number of unfriendly subsets having k elements is

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

3 Basic Counting Techniques

A lot of counting problems can be solved by cleverly splitting the problem up into simpler pieces.

If the objects you are trying to count each belong to one of a few types, the total number of objects is equal to the sum over types of the number of objects of that type. Note that this only works when each object belongs to exactly one type. If an object could have more than one type, you need to instead use something called Inclusion-Exclusion.

If you construct the objects you are trying to count by making some series of choices and there are n_1 ways to make the first choice, n_2 ways to make the second choice and so on. Then the total number of objects will be given by the product of these n_i.

A lot of counting problems can be solved by cleverly breaking it down into simpler parts and using these techniques on each of them.

1996 B1 (modified): Define a **selfish** set to be one which has its own cardinality (number of elements) as an element. Find, with proof, the number of subsets of $\{1, 2, \ldots, n\}$ selfish.

4 Linearity of Expectation

Another surprisingly versatile tool is known as *linearity of expectation*. It says that given two sequences $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ that the average value of $a_i + b_i$ is the average value of a_i plus the average value of b_i. Formally, this can be worded as

$$\frac{1}{n}\sum_{k=1}^{n}(a_k + b_k) = \frac{1}{n}\sum_{k=1}^{n} a_k + \frac{1}{n}\sum_{k=1}^{n} b_k.$$

This result also holds if the sum is replaced by an integral, or an expectation if one is working with random variables. Although this result seems simple, it is surprisingly powerful that this holds no matter how the sequences a and b might be correlated with each other.

This idea is particularly useful when you want to find the average value of some quantity X that can be written as a sum of relatively simple terms. One common way to do this if X counts the number of objects with a given property is to write X as a sum of *indicator functions*. For each object i that *might* be counted by X, let f_i be 1 if i is included and 0 if it is not. Then $X = \sum_i f_i$ and linearity of expectation often allows you to compute X.

For example, suppose that S is a random subset of $\{1, 2, \ldots, n\}$ and we want to know the average size of $|S|^2$. This is the number of ordered pairs (i, j) where both i and j are in S. Thus, letting $f_{i,j} = 1$ if i and j are in S and $f_{i,j} = 0$ otherwise,

$$|S|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{i,j}.$$

Now the average value of $f_{i,j}$ can be seen to be $1/2$ if $i = j$ and $1/4$ if $i \neq j$, so this sum is $\frac{n^2+n}{4}$.

2016 B4: Let A be a $2n \times 2n$ matrix with entries chosen independently at random. Every entry is chosen to be either 0 or 1, each with probability $1/2$. Find the expected value of $\det(A - A^t)$ (as a function of n), where A^t is the transpose of A.