ORIE 5530: Modeling Under Uncertainty

Lecture 2 (Events)

Professor Mark S. Squillante,
Adapted from Professor Itai Gurvich's original notes

We spent most of class today on counting. This is useful and fundamental when outcomes are discrete (such as the number on a die or the number of graphs). Some of this goes back to basics you might be familiar with and my objective is to present these with a bit more elaborate and interesting examples.

When all outcomes on a probability space, $\mathcal{S}$, are equally likely, the probability of an outcome (a single point in the sample space) is

$$\frac{1}{\text{\# of points in the space}}.$$

In the case of a fair die, there are six possible (and equally likely) outcomes and the probability of each is 1/6.

If we throw two dice the number of possible outcomes is 36 and each of the outcomes (i.e., each pair of numbers between 1 and 6) has a likelihood of 1/36. If we want to know the likelihood of a more complex outcomes such as the event $A = \{\text{the sum of the two tosses is 4}\}$ we need to be able to **count** the number of such sample points. There are 3 of these (check that you agree). The likelihood is $|A|$ (the size of A, or the number of points in it) divided by the size, $|\mathcal{S}|$, of the sample space $\mathcal{S}$ (the number of points). So in this case

$$\mathbb{P}\{\text{the sum of the two tosses is 7}\} = \frac{|A|}{|S|} = \frac{3}{36} = \frac{1}{12}.$$

With dice counting is simple. However, it can get fairly complicated; fortunately, there are some basic facts we can use.

Say you have n **different** items.

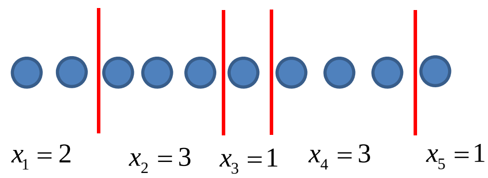
1. There are $n!$ ways of ordering the **n** different items. In other words, the number of ordered sets with n different items is $n!$.
2. There are $P(n,k) = \frac{n!}{(n-k)!}$ ways of obtaining an ordered subset of size $k \leq n$ items from the set of n items. $P(n,k)$ are called the number of k-permutations of n.
3. There are $C(n,k) = \frac{P(n,k)}{P(k,k)} = \frac{n!}{k!(n-k)!} = \binom{n}{k}$ ways of choosing k items out of n if you do not care about the order in which you choose them. $C(n,k) = \binom{n}{k}$ are called the binomial coefficients.
4. There are 2^n ways to choose a subset out of a set of n **different** items.

Two non-trivial examples

1. **Integral equations:** You have an equation

$$x_1 + x_2 + x_3 + x_4 + x_5 = 10.$$

   The question is how many integer (i.e., in whole numbers) different solutions exist for this equation where all variables are **strictly positive**. For example, $x_1 = x_2 = x_3 = x_4 = x_5 = 2$ is a valid solution.



The key is to translate this question into the basic facts we have above. What are we choosing?

We have 10 balls and we are choosing where to place dividers. Everything on the left of the first divider goes to x_1, everything between the 1st and 2nd divider goes to x_2 and everything above the $4th$ divider goes to x_5; see the figure below for an example.

We have 9 spots where we can put the 4 dividers (after the first ball, after the second ball, ..., after the 9th ball). So we need to choose 4 out of 9 spots where to place the divider. The answer is thus $\binom{9}{4}$.

** **Bonus comment:** (I will use ** in my handouts for points/comments that are kind of "bonus" curiosities. You do not have to worry about these)

There is an interesting question as to what happens if we also allow some variables be 0. In that case, $x_1 = x_2 = x_3 = x_4 = 0$ and $x_5 = 10$ is a valid solution.

The simplest way to think about this, building on the previous result, is to have two steps:

- Choose which variables are going to be strictly positive. You are choosing a subset of the variables. There are, for example, 5 ways to choose a single variable to be positive, there are $\binom{5}{2}$ to choose 2, etc.
- For each option resolve the above. For example, if only 1 variables is strictly positive, once we chose this variable, it must equal to 10 for the sum to be 10. If two variables are strictly positive, given the chosen variables, there are (based on part 1) $\binom{9}{1}$ ways to make them strictly positive and sum to 10, etc.

So the solution would be

$$\sum_{i=1}^{5} \binom{5}{i} \binom{9}{i-1}$$

Notice that $\binom{5}{i} = \binom{5}{5-i}$ (the math says this but this is also intuitive – it is the same to choose i variables to be positive as to choose $5 - i$ variables to be 0).

Then, the above is equal to

$$\sum_{i=1}^{5} \binom{5}{i} \binom{9}{i-1} = \sum_{i=1}^{5} \binom{5}{5-i} \binom{9}{i-1} = \binom{14}{4}.$$

If you want to convince yourself that this works, checkout the case where you have only two variables to start with. Then, there are 9 solutions where both variables are strictly positive (you choose the value of x_1 and then x_2 must equal $10 - x_1$) and you have 11 solutions in which one of them could be 0 (once you choose which one is 0, the other one must be 10).

There is an additional way, that requires a bit more creativity, to get to the same result.

2. **Graphs:** How many undirected graphs that have n nodes (individuals) can we build? A graph is defined by the nodes and the edges that connect them. So if the number of nodes is given, the question is only which edges do we include. This means we want to divide our analysis into two steps

   – How many possible edges are there?
   – How many graphs can we construct from these edges?

   Let's start from the second part. If there are m edges (let's use m as a placeholder and we will give it a value shortly), then building a graph is choosing a subset out of the m distinct edges. We know that there are 2^m ways of doing this. So,

   $$\text{Number of possible graphs} = 2^{\text{Number of possible edges}}.$$

   Now, an edge connects to individuals:

   $$\text{number of edges} = \text{number of distinct pairs of individuals.}$$

   How many pairs do we have? Namely, the ways to choose 2 individuals out of n. Hence, the number of edges m equals $\binom{n}{2}$.
   Combining things, we arrive at the conclusion that there are

   $$\text{Number of possible graphs} = 2^{\text{Number of possible edges}} = 2^{\binom{n}{2}}.$$

   The nice thing about this example is that it builds on two of the basic facts listed at the beginning of this note.

Recursions: We also did a couple of things that illustrate the value of counting through recursions. You will do one such question in the homework.

2.1 A refresher of sample spaces and set operations

The first step in analyzing a probabilistic outcome is to have a clear understanding of **all** the possible outcomes even if they have a small probability.

We use $\mathcal{S}$ for the *sample space* – the space of all possible outcomes. An item ω in the sample space is called a *sample point*. So in the toss of a die, the sample space has the 6 sample points $1, 2, \ldots, 6$.

An *event* is a subset of the sample space. In this simple example, $\{1\}, \{1, 2\}, \{2, 4, 6\}$ are all events (the last one is the event that the result of the die is even).

Below are some formal definitions of union and intersection (but really, these are very intuitive notions).

To make definitions easier, let us use the letter ω for a sample point.

If you have two events A and B, the union and intersection are defined by

- $A \cup B = \{w$ in the sample space that are either in A or $B\}$,
- $A \cap B = \{w$ in the sample space that are in **both** A **and** $B\}$

Some basic rules:

- Order of operations does not matter $(A \cup B) \cup C = A \cup (B \cup C)$.
- You can disaggregate operations ("the distributive law"):

  $$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

  and

  $$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$