Lecture 26: Probability theory

A *very* important application of counting arises in the context of computing probabilities in various situations. We won't have time to go very deep into probability theory (and this is not how you would formally learn it in a statistics class), but the basics of probability can be directly understood by invoking the standard counting riles.

Assume that we perform some type of procedure or experiment. The set of all possible outcomes of the experiment is called the *sample space*. An *event* is a specific subset of the sample space.

For instance, if we toss a coin three times, the set of possible outcomes (i.e., the sample space) is given by:

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

The event "All Heads" would correspond to the (singleton) subset:

$$A = \{HHH\}.$$

The event "At least two heads in a row" would correspond to the subset:

$$A = \{HHH, HHT, THH\}.$$

If all possible outcomes of an experiment have the same likelihood of occurrence, then the probability of an event $A \subseteq S$ is given by:

$$P(A) = \frac{|A|}{|S|}.$$

Applying the above formula, the probability of getting at least two heads in a row in the above experiment is 3/8, whereas the probability of getting "All heads" is 1/8.

In general, computing the probability of an event will involve (carefully) enumerating the number of elements in the subset that defines the event, and dividing by all possible outcomes.

A somewhat different, and more illustrative, approach to arrive at the same answer is to draw the *tree* of possible outcomes. In the above example, it will be a binary tree with 3 levels, with each edge indicating either a head or a tail, and each edge probability being equal to 1/2. There will be eight leaves (corresponding to the 8 outcomes listed in the sample space above), and each leaf has a probability that is the product of the edges leading up to it.

Another example: if we were to roll a (regular) dice twice, what is the probability that the *sum* of the numbers that came up equals 7?

Here again, we can draw a tree with 2 levels; the root, the 6 children in the 1st level (corresponding to the first roll), and 36 leaves (6 children for each node in the first level). Each leaf is one outcome, and can be labeled by the pair of numbers (i, j) where i is the first number and j is the second number.

We need to enumerate the set of outcomes (leaves) in which $i + j = 7$. Observe that there are 6 such outcomes:

- (1,6)
- (2,5)
- (3,4)
- (4,3)
- (5,2)
- (6,1)

The total number of outcomes is 36, so the probability of getting a seven is $6/36 = 1/6$.

Properties of probabilities

Let us now list some properties of probabilities. The justification for each of these is evident from the definition.

- By definition, since A is a subset of S, we always have $0 \leq P(A) \leq 1$. A probability value that is negative, or bigger than 1, is meaningless.
- For every event A, the probability of $\bar{A}$ (its complement) is given by:
$$P(\bar{A}) = 1 - P(A).$$
- By the addition rule of counting, if A_1 and A_2 are mutually exclusive events, then we know that $|A_1| + |A_2| = |A_1 \cup A_2|$. Therefore,
$$P(A_1) + P(A_2) = P(A_1 \cup A_2).$$
- By the Principle of Inclusion-Exclusion, for general A_1 and A_2, we know that $|A_1| + |A_2| - |A_1 \cap A_2| = |A_1 \cup A_2|$. Therefore,
$$P(A_1) + P(A_2) - P(A_1 \cap A_2) = P(A_1 \cup A_2).$$

Funny Dice

Probability calculations can often be misleading and counter-intuitive. This is in fact why people trip up all the time while discussing probabilities. Our brains are not wired to think of probabilities the same way we think of ordinary quantities. Here is a somewhat strange example of when this happens.

Imagine that you have 3 specially constructed dice cubes.

- Die A has 3 numbers – (1,5,9) – listed twice each, so that the probability of rolling 1, or 5, or 9, is 1/3.
- Die B has 3 other numbers – (2,6,7).
- Die C has the numbers (3,4,8).

You and a friend play a betting game. You let your friend pick a die, and then you pick a die. You mutually agree to roll the dice a zillion times, and whoever wins a larger number of times wins the bet. The question is: among the three, what is the best die to choose?

(Of course, no one has time to actually sit and roll dice a zillion times, so you agree that you will simply simulate this game by computing the likelihoods of various outcomes, and whoever is more likely to win the game gets the prize.)

At first glance, notice that the numbers of all the three dice add up to the same amount (15), so the *average* value of each roll, irrespective of the choice of die, is 15/3 = 5. So it seems that there is no comparative advantage.

However, let's actually play the game. Suppose your friend picks die A. Then, you can pick die B. Let's roll!

There are 9 outcomes, listed as (their roll, your roll). Assuming perfectly fair dice, each happens with a probability $1/9$.

- (1,2) - *you win*
- (1,6) - *you win*
- (1,7) - *you win*
- (5,2) - *they win*
- (5,6) - *you win*
- (5,7) - *you win*
- (9,2) - *they win*
- (9,6) - *they win*
- (9,7) - *they win*

Therefore, the event that you win has a probability of 5/9. Great! You win!

OK, so maybe they picked wrong, and you give them a chance to play again. This time, they pick die B. But you counter by picking die C. Let's roll again! The tuples below, again, are (their roll, your roll):

- (2,3) - *you win*
- (2,4) - *you win*
- (2,8) - *you win*
- (6,3) - *they win*
- (6,4) - *they win*
- (6,8) - *you win*
- (7,3) - *they win*
- (7,4) - *they win*
- (7,8) - *you win*

Again, the event that you win has a probability of 5/9. So clearly, die C is the best...right?

You are feeling generous, so you give them a chance to play one more time. This time, they choose C. But your response is to pick A. Let's roll!

- (3,1) - *they win*
- (3,5) - *you win*
- (3,9) - *you win*
- (4,1) - *they win*
- (4,5) - *you win*
- (4,9) - *you win*
- (8,1) - *they win*
- (8,5) - *they win*

- (8,9) - *you win*

Again, the event that you win has a probability of 5/9 !!?

So we are in this strange situation where die B beats die A, die C beats die B, and die A beats die C. Weird.

This example goes to show that the "probably beats" relation is *not transitive*.

Intuition often fails when discussing probability. (This is also why polls routinely are misinterpreted.) The safest option, whenever possible, is to enumerate the sample space, calculate the probabilities of the outcomes and events, and confirm your answer via other means. Of course, in complicated problems this is not always easy to do.