Axioms and assumptions

18.600 Problem Set 2, due February 24

Welcome to your second 18.600 problem set! We will continue to explore some combinatorics along with probability and the axioms of probability.

Before we get to work, let's indulge in just a bit of reflection. When we say "The probability that A will happen is p" where does p come from? Sometimes the evidence convinces pretty much everyone that A will or will not happen. Informally, the probability that a predicted lunar eclipse will happen on schedule is pretty much 1, and the probability that Mars and Jupiter will collide this month is pretty much 0. In other simple situations (die rolls, coin tosses, etc.) experience may lead us to agree on probabilities that aren't 0 or 1. The assumption that all outcomes are equally likely (for random permuations or die rolls or coin tosses) is sometimes a natural starting point. This assumption is implicitly made in a few of the problems here.

In more complicated real world settings, one can sometimes define the *risk neutral* probability, a probability measure derived from the market prices of contracts whose values depend on future events. If we want to know the risk neutral probability that a given candidate will win an election, or that an athletic team will win a game, we can look at betting markets. (Check out predictwise.com, electionbettingodds.com, oddschecker.com, predictit.com, and similar websites.) As we will see later in the course, if we want to know the risk neutral probability that the price of a share of Apple stock will exceed some value by the end of the year, we can work this out by looking at current prices of *derivatives* (contracts whose future value depends on future share prices). The total amount of money at stake in derivative markets is estimated at over a quadrillion dollars per year (try googling derivatives quadrillion).

Some argue that betting markets set up perverse incentives. If I buy a contract that gives me $\$500,000$ if my house burns down, that's useful insurance. But if I buy a contract that gives me $\$500,000$ if *your* house burns down, that gives me an unhealthy incentive to burn your house down. People similarly worry about a world in which hedge funds can bet that a company will collapse and then actively cause it to collapse. Rules are required to prevent such things, but foolproof (and evil-genius-proof) rules are hard to design and enforce.

On the other hand, one might argue that the absence of betting markets is part of the reason that some questions in politics and law are divisive. It is hard to place a bet on the proposition that "my candidate would do more to advance long term happiness and prosperity than yours" or "my client is innocent," so there is no market mechanism for producing a commonly accepted probability. Different groups can *claim* to have different probability estimates, the expression of which may advance their own agendas, but without a market we cannot tell which parties would actually be *willing to bet money* at the corresponding rates. Some studies claim that people answering questions about the economy are both more accurate and less partisan when they are paid (even a very small amount) for correct answers. Maybe there is something to be said for having money on the line.

Please stop by my weekly office hours (2-249, Wednesday 3 to 5) for discussion.

A. FROM ROSS 8TH EDITION CHAPTER TWO:

1. **Problem 25:** A pair of dice is rolled until a sum of either 5 or 7 appears. Find the probability that a 5 occurs first. *Hint.* Let E_n denote the event that a 5 occurs on the nth roll and no 5 or 7 occurs on the first $(n-1)$ rolls. Compute $P(E_n)$ and argue that $\sum_{n=1}^{\infty} P(E_n)$ is the desired probability.

2. **Theoretical Exercise 10:** Prove that
$P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(E^cFG) - P(EF^cG) - P(EFG^c) - 2P(EFG)$.

3. **Theoretical Exercise 20:** Consider an experiment whose sample space consists of a countably infinite number of points. Show that not all points can be equally likely. Can all points have a positive probability of occurring?

B. Suppose that there are 8 job candidates and 8 companies with job openings. Each candidate (independently, uniformly at random) develops a serious interest in one of 8 companies and each company (independently, uniformly at random) develops a serious interest in one of the 8 candidates. What is the probability that there is at least one company-candidate pair that are seriously interested in each other? (Hint: let E_j be the event that the jth applicant's interest is requited. Use inclusion-exclusion on these events. You can write the probability as a sum. Don't worry about simplifying further.)

C. Alice and Bob are playing a game of tennis and have reached the game state called "deuce." From here the players keep playing points until one player's total exceeds the other player's total by 2, at which point the player ahead by 2 points is declared winner of the game. Suppose that Alice wins each point with probability p (independently of all previous points) and Bob wins each point with probability $q = (1-p)$ (independently of all previous points). Find the probability that Alice wins the game, as a function of p. (Hint: consider what happens over the course of the next *two* points. Either Alice wins both and the game is over, or Bob wins both and the game is over, or each player wins a point and the players are back where they started. Compute the probabilities of these three outcomes. Then apply the ideas from the first problem on this problem set.) Based on your answer, do you agree or disagree with the following statement? *If Alice is k times as likely as Bob to win a point, then Alice is k^2 times as likely as Bob to win the game if the current score is deuce.*

D. The online comic strip xkcd.com has a "random" button one can click to choose one of the previous $n \approx 1800$ strips. Assume that there are exactly 1800 numbered strips and that each time one clicks the "random" button, one gets the kth strip where k is chosen uniformly from $\{1, 2, \ldots, n\}$. If one observes m strips in this way, what is the probability that one sees at least one strip more than once? (This is a variant of the birthday problem.) Give approximate numerical values for $m = 30$ and $m = 50$ and $m = 70$. (Hint: try going to wolframalpha.com and entering something like Prod[(1-k/1800), {k,0,29}].) Based on your answer, do you agree or disagree with the following statement? *The number of clicks required before you see the same strip twice is a random quantity whose median is about* 50*, and which lies between* 30 *and* 70 *about half the time*

E. 20 athletes show up to play a game of kickball, including you and your two good friends. It is determined that a subset of 10 of the people (chosen uniformly at random from the set of all such subsets) will be chosen to be the first team, and the remaining 10 people will form the second team. What is the probability that you and your two friends end up on the same team? Is it more or less than .25?

F. The following is a popular and rather instructive puzzle. A standard deck of 52 cards (26 red and 26 black) is shuffled so that all orderings are equally likely. We then play the following game: I place the deck face down and begin turning over the cards from the top of the deck one at a time so that you can see them. At some point (before I have turned over all 52 cards) you say "now." At this point I turn over the next card and if the card is red, you receive one dollar; otherwise you receive nothing. You would like to design a strategy to maximize the probability that you will receive the dollar. How should you decide when to say "now"?

Your first observation is that a good time to say "now" is when you know that a high fraction of the cards remaining in the deck are red. On the other hand, if you wait for this fraction to increase, there's always a chance you'll see more red cards while you wait, so that the fraction actually *decreases*. What's the right way to balance these concerns, i.e., what is the *optimal* strategy for deciding when to say "now"? A hint for this puzzle appears on the next page, but don't look at it until you have to.

HINT: Imagine a variant of the game in which, after you say "now," I turn over the *bottom* card on the deck. Observe that in this variant, it makes no difference when you say "now" (since you're going to see the same bottom card of the deck regardless). Now try to argue that your probability of winning with a given strategy in the modified game is the same as your probability of winning with that strategy in the original game. Conclude that in the original game, it also makes no difference which strategy you choose. Your odds of winning the dollar are .5 regardless.

REMARK: The fraction of red cards in the deck turns out to be something called a *martingale* and the fact that it makes no difference when you bet can be derived from something called the *optional stopping theorem*. We'll have more on this at the end of the course. But I like this puzzle because it's something you can appreciate right now. This is a strategy game that Warren Buffett and a chimpanzee would play equally well. Unlike the game shown here `https://www.youtube.com/watch?v=JkNV0rSndJ0` and playable here `http://www.novelgames.com/en/ayumu/`