DIMACS: Probability Crash Course — Homework

Instructor: Pat Devlin — prd41@math.rutgers.edu

Day 4, Summer 2015

(1) (a) [**Markov's inequality**] Suppose that X is a random variable and that X is never negative. Prove that $\mathbb{P}(X \geq a) \leq \dfrac{\mathbb{E}[X]}{a}$ for all $a > 0$. (**Hint:** Consider the r.v. Y:

$$Y = \begin{cases} 0, & \text{if } X < a \\ a, & \text{if } X \geq a. \end{cases}$$

What is $\mathbb{E}[Y]$? How does this compare to $\mathbb{E}[X]$? Why?)

(b) Use Markov's inequality to show that if X is a random variable that only takes non-negative integer values then $\mathbb{P}(X \neq 0) \leq \mathbb{E}[X]$.

(c) Show by example that Markov's inequality need not apply if the r.v. can be negative.

(2) (**Pat's absolute favorite math problem**) Let S be the r.v. corresponding to the sum of two standard dice. The goal is try to find two weird dice (dice not necessarily numbered 1 to 6) with the property that if T is the r.v. for the sum of these weird dice, then $\mathbb{P}(S = n) = \mathbb{P}(T = n)$ for all n.

(a) Find $\mathbb{P}(S = n)$ for $0 \leq n \leq 20$.

(b) If the two weird dice are numbered $1, 1, 3, 3, 7, 9$ and $1, 2, 3, 4, 5, 6$, then what would be $\mathbb{P}(T = n)$ for $1 \leq n \leq 4$?

(c) Try to find two dice with the desired property. Assume that the smallest number on either of these dice is 1 (otherwise, you could do something "silly" like $0, 1, 2, 3, 4, 5$ and $2, 3, 4, 5, 6, 7$).

(3) (**This is *exactly* the type of question that Pat studies for a living**) Suppose we start with the complete graph K_n. Then for each of these $\binom{n}{2}$ edges, we independently flip a coin, and we decide to keep that edge with probability $p = 1/2$ (otherwise we throw it away). We then only consider the n vertices and the random edges that we decided to keep. This object is called a *random graph*, and we denote it by $G_{n,1/2}$.

(a) Suppose we make a random graph, $G_{n,1/2}$, in this way, where each possible edge is present independently with probability $1/2$. What is the probability that this random graph will have all $\binom{n}{2}$ edges? What is the probability that it won't have any edges?

(b) What is the probability that the random graph will be a perfect matching with no other edges at all?

(c) What is the probability that the random graph will be one big cycle with no other edges at all?

(d) (Optional) Let F be the event that the random graph is a forest (i.e., has no cycles). Show that $\mathbb{P}(F)$ is very close to 0 if n is large. [**Hint:** recall that a forest on n vertices has at most $n - 1$ edges.]

(4) Bob tosses a fair coin N times in a row, and the outcomes of these flips are $f_1, f_2, \ldots, f_N$. Let Y_k be the number of indices j such that $f_j = f_{j+1} = \cdots = f_{j+k-1} = H$. (So Y_k is counting the number of times that you see k consecutive H in a row.)
For example: Suppose $N = 12$ and the flips were $HHTHHHTTHHHH$. Then $Y_1 = 9$, and $Y_2 = 6$, and $Y_3 = 3$, and $Y_4 = 1$, and $Y_k = 0$ for $k \geq 5$. Note for instance that $Y_2 = 6$ since you can find two H in a row by starting at any of the indices $1, 4, 5, 9, 10, 11$.

(a) For $N = 4$, compute $\mathbb{E}[Y_k]$ for $1 \leq k \leq 4$.

(b) For $1 \le j \le N-k+1$, let $A_k^{(j)}$ denote the event that $f_j = f_{j+1} = \cdots = f_{j+k-1} = H$. Find $\mathbb{P}(A_k^{(j)})$.

(c) Show that $\mathbb{E}[Y_k] = (N-k+1)/2^k$. (Hint: you could use indicator r.v.s for $A_k^{(j)}$.)

(d) Someone hands you a sequence of 10000 coin flips. You notice that $Y_6 > 0$ and $Y_7 = 0$ (i.e., the longest consecutive string of H had length 6). Using your answer to (c), would you find this surprising?

(e) (Optional) How do you think the number 6 from part (d) should be changed to make this more believable?

(5) (**Hat matching problem**) There are N people who all toss their hats into a big pile. At the end of the night, nobody remembers which hat is which, and each person picks one uniformly at random. Let X denote the number of people who happen to get their own hat back. Find $\mathbb{E}[X]$. How does this change when N is very small or very large? (Optional challenge: Compute $\mathbb{P}(X = 0)$.)

(6) Suppose there are N people who all have different heights. These people randomly arrange themselves in a line from left to right. At the front of the line is a chalkboard. The very first person in line can see the board since nobody is in front of him. But everyone else in line can only see the chalkboard if *nobody taller is in front of him.* Let Z be the random variable for the number of people who can see the board, and let $f(N) = \mathbb{E}[Z]$.

(a) Compute $f(N)$ when $1 \le N \le 4$.

(b) Show that

$$f(N) = \mathbb{E}[Z] = \sum_{i=1}^{N} \frac{1}{i}.$$

Hint: You can do this in a few ways. For instance, you could use indicator r.v.s for the events that the i^{th} person can see and then use linearity of expectation. Or you could prove $f(N) = \dfrac{1}{N} + f(N-1)$. There are other ways too.

(c) (Optional) Argue that we expect the tallest person to be approximately in the middle, and since nobody behind him can see, we have $f(N) \approx f(N/2) + 1$. [This is actually not a great approximation, but it's decent.]

(d) (Optional) Use part (c) to conclude that

$$\sum_{i=1}^{N} \frac{1}{i} \approx \log_2(N).$$

(It turns out that approximating this sum by $\ln(N)$ gives a *very* close answer.)

(7) See next page.

(7) One of the most popular US lottery games is called **Powerball**. The following description of the game is taken from their webpage:

> Every Wednesday and Saturday night at 10:59 p.m. Eastern Time, we draw five [different] white balls [without replacement] out of a drum with 59 balls and [then we draw] one red ball out of a drum with 35 red balls. ...Players win by matching one of the *nine ways to win*. ...Each ticket costs $2.

When you buy a ticket, you specify 5 different 'white' numbers from 1 to 59 (order does not matter), and then you specify one "powerball number" from 1 to 35 [for predicting what the red ball will be]. Here are the *nine ways to win* and the amount paid for each. **Note:** the order that the white balls are drawn does not matter for determining the pay-outs.

Prize	**Exact number of white balls matched**	**Is red ball matched?**	**Amount paid**
A	5	Yes	*Jackpot*
B	5	No	$\$1,000,000$
C	4	Yes	$\$10,000$
D	4	No	$100
E	3	Yes	$100
F	3	No	$7
G	2	Yes	$7
H	1	Yes	$4
I	0	Yes	$4

(a) Find the probability of winning each prize. For instance, the probability of winning prize F is

$$\mathbb{P}(F) = \frac{\binom{5}{3}\binom{54}{2}}{\binom{59}{5}} \times \frac{34}{35}.$$

[Hint: $\mathbb{P}(I)$ is *not* equal to $1/35$.]

(b) What is the probability that a ticket won't win *any* of the prizes?

(c) *Assume for a moment that lottery winnings are not taxed.* How high would the jackpot prize need to be in order to make the expected winnings from a ticket more than $2 (its cost)? Assume that multiple winners would not affect how much money each winning ticket is given.

(d) Now answer part (c) again but with the (correct) assumption that if you win at least a million dollars, then you owe 35% of that money to the government in taxes. Assume that smaller winnings are taxed at 15%.

(e) Suppose now that you work at the lottery, and you are thinking about removing one of the prizes. You decide you will not remove the jackpot prize. Which of the other prizes should you remove in order to maximize the expected amount of profit the lottery makes per ticket sale?