EECS 70 Discrete Mathematics and Probability Theory
Fall 2015 Jean Walrand, Satish Rao

HW 10

Due Wednesday Nov 4 at 10PM

1. Card Game

A game is played with six double-sided cards. One card has "1" on one side and "2" on the other. Two cards have "2" on one side and "3" on the other. And the last three cards have "3" on one side and "4" on the other. A random card is then drawn and held in a random orientation between two players, each of whom sees only one side of the card. The winner is the one seeing the smaller number. If the card that was drawn was a "2/3" card, compute the probabilities each player thinks he/she has for winning.

Solution: For the player that sees the "2" side:

$$\Pr[\text{Win} \mid \text{Player sees } 2] = \frac{\Pr[\text{Win} \cap \text{Player sees } 2]}{\Pr[\text{Player sees } 2]} = \frac{\Pr[\text{Card is "2/3"} \cap \text{Player sees } 2]}{\Pr[\text{Player sees } 2]}$$
$$= \frac{\Pr[\text{Card is "2/3"}] \cdot \Pr[\text{Player sees 2} \mid \text{Card is "2/3"}]}{\Pr[\text{Card has a "2"}] \cdot \Pr[\text{The side with "2" is chosen} \mid \text{Card has a "2"}]}$$
$$= \frac{\frac{2}{6} \times \frac{1}{2}}{\frac{3}{6} \times \frac{1}{2}} = \frac{2}{3}$$

For the player that sees the "3" side:

$$\Pr[\text{Win} \mid \text{Player sees } 3] = \frac{\Pr[\text{Win} \cap \text{Player sees } 3]}{\Pr[\text{Player sees } 3]} = \frac{\Pr[\text{Card is "3/4"} \cap \text{Player sees } 3]}{\Pr[\text{Player sees } 3]}$$
$$= \frac{\Pr[\text{Card is "3/4"}] \cdot \Pr[\text{Player sees 3} \mid \text{Card is "3/4"}]}{\Pr[\text{Card has a "3"}] \cdot \Pr[\text{The side with "3" is chosen} \mid \text{Card has a "3"}]}$$
$$= \frac{\frac{3}{6} \times \frac{1}{2}}{\frac{5}{6} \times \frac{1}{2}} = \frac{3}{5}$$

2. Boys and Girls

There are three children in a family. A friend is told that at least two of them are boys. What is the probability that all three are boys? The friend is then told that the two are the oldest two children. Now what is the probability that all three are boys? Use Bayes' Law to explain this. Assume throughout that each child is independently either a boy or a girl with equal probability.

Solution: Let A be the information that you are told and B the event that all three are boys. By Bayes' rule, we have $\Pr[B \mid A] = \frac{\Pr[B] \cdot \Pr[A|B]}{\Pr[A]}$. However, note that whether $A =$ "at least two boys" or $A =$ "oldest two are boys", $\Pr[A \mid B]$ is simply 1. So in either case we have $\Pr[B \mid A] = \frac{\Pr[B]}{\Pr[A]}$.

Hence,

$$\Pr[\text{All 3 boys} \mid \text{At least 2 boys}] = \frac{\Pr[\text{All 3 boys}]}{\Pr[\text{At least 2 boys}]} = \frac{\frac{1}{8}}{\frac{4}{8}} = \frac{1}{4}$$

$$\Pr[\text{All 3 boys} \mid \text{Oldest 2 are boys}] = \frac{\Pr[\text{All 3 boys}]}{\Pr[\text{Oldest 2 are boys}]} = \frac{\frac{1}{8}}{\frac{2}{8}} = \frac{1}{2}.$$

In this case we saw Bayes' rule simplify to $\Pr[B \mid A] = \frac{\Pr[B]}{\Pr[A]}$. Since B is a subset of A, the formula directly shows that this conditional probability depends only on the number of possibilities contained in A. When we are told that at least two children are boys, any of the three children could be a girl. In contrast, if we are told that the oldest two children are boys, then only the youngest child has the possibility of being a girl. Therefore the latter case has fewer possibilities and therefore larger conditional probability.

3. Lie Detector

A lie detector is known to be 80% reliable when the person is guilty and 95% reliable when the person is innocent. If a suspect is chosen from a group of suspects of which only 1% have ever committed a crime, and the test indicates that the person is guilty, what is the probability that he is innocent?

Solution: Let A denote the event that the test indicates that the person is guilty, and B the event that the person is innocent. Note that

$$\Pr[B] = 0.99, \quad \Pr[\overline{B}] = 0.01, \quad \Pr[A \mid B] = 0.05, \quad \Pr[A \mid \overline{B}] = 0.8$$

Using the Bayesian Inference Rule, we can compute the desired probability as follows:

$$\Pr[B \mid A] = \frac{\Pr[B]\Pr[A \mid B]}{\Pr[B]\Pr[A \mid B] + \Pr[\overline{B}]\Pr[A \mid \overline{B}]} = \frac{0.99 \cdot 0.05}{0.99 \cdot 0.05 + 0.01 \cdot 0.8} \approx 0.86$$

4. Cliques in random graphs

Consider a graph $G(V,E)$ on n vertices which is generated by the following random process: for each pair of vertices u and v, we flip a fair coin and place an (undirected) edge between u and v if and only if the coin comes up heads. So for example if $n = 2$, then with probability $1/2$, $G(V,E)$ is the graph consisting of two vertices connected by an edge, and with probability $1/2$ it is the graph consisting of two isolated vertices.

1. What is the size of the sample space?
2. A k-clique in graph is a set of k vertices which are pairwise adjacent (every pair of vertices is connected by an edge). For example a 3-clique is a triangle. What is the probability that a particular set of k vertices forms a k-clique?
3. Prove that the probability that the graph contains a k-clique for $k = 4\lceil \log n \rceil + 1$ is at most $1/n$.

Solution:

1. There are two choices for each of the $\binom{n}{2}$ pairs of vertices, so the size of the sample space is $2^{\binom{n}{2}}$.
2. For a fixed set of k vertices to be a k-clique, all of the $\binom{k}{2}$ pairs of those vertices have to be connected by an edge. The probability of this event is $1/2^{\binom{k}{2}}$.
3. Let A_S denote the event that S is a k-clique, where $S \subseteq V$ is of size k. Then, the event that the graph contains a k-clique can be described as the union of A_S's over all $S \subseteq V$ of size k. Using the union bound,

$$\Pr\left[\bigcup_{S \subseteq V, |S|=k} A_S\right] \le \sum_{S \subseteq V, |S|=k} \Pr[A_S] = \sum_{S \subseteq V, |S|=k} \frac{1}{2^{\binom{k}{2}}}.$$

Now, since there are $\binom{n}{k}$ ways of choosing a subset $S \subseteq V$ of size k, the right-hand side of the above equality is

$$\frac{\binom{n}{k}}{2^{\binom{k}{2}}} = \frac{\binom{n}{k}}{2^{\frac{k(k-1)}{2}}} \le \frac{n^k}{\left(2^{\frac{(k-1)}{2}}\right)^k} \le \frac{n^k}{\left(2^{\frac{(4\log n+1-1)}{2}}\right)^k} = \frac{n^k}{(2^{2\log n})^k} = \frac{n^k}{n^{2k}} = \frac{1}{n^k} \le \frac{1}{n}.$$

5. College applications

There are n students applying to n colleges. Each college has a ranking over all students (i.e. a permutation) which, for all we know, is completely random and independent of other colleges.

College number i will admit the first k_i students in its ranking. If a student is not admitted to any college, he or she might file a complaint against the board of colleges, and colleges want to avoid that as much as possible.

1. If for all i, $k_i = 1$, i.e. if every college only admits the top student on its list, what is the chance that all students will be admitted to at least one college?

   Solution: If we consider the first choices of all colleges, there are n^n different possibilities, all of which are equally likely because colleges are indepently sorting students in a random manner. Out of these we want the possibilities that have all students covered, which is the same as those that have no repeated student (because the number of colleges is the same as the number of students). So we are counting permutations, and we know that there are $n!$ of them. So the probability is $\frac{n!}{n^n}$.

2. What is the chance that a particular student, Alice, does not get admitted to any college? Prove that if the average of all k_i's is $2\ln n$, then this probability is at most $1/n^2$. (Hint: use the inequality $1-x < e^{-x}$)

   Solution: The chance that Alice does not get admitted to college i is $1-\frac{k_i}{n}$. This is because out of all the $n!$ permutations that college i can have on students $k_i \times (n-1)!$ of them result in Alice being one of the top k_i (we first choose Alice's place and then randomly permute the remaining students). So the probability that Alice ends up in the top k_i is k_i/n and the probability that she does not is $1-\frac{k_i}{n}$.
   The probability that she does not get admitted to any college is just

$$\prod_{i=1}^{n}(1-\frac{k_i}{n})$$

   Now using the inequality $1-x \le e^{-x}$, we get $1-\frac{k_i}{n} \le e^{-k_i/n}$. Multiplying over all i we get

$$\prod_{i=1}^{n}(1-\frac{k_i}{n}) < \prod_{i=1}^{n} e^{-k_i/n} = e^{-\sum_{i=1}^{n} k_i/n}$$

   But $\sum_{i=1}^{n} k_i/n$ is simply the average of all k_i. If this average is $2\ln n$, the last expression simply reduces to $e^{-2\ln n}$ which is just $1/n^2$.

3. Prove that when the average k_i is $2\ln n$, then the probability that at least one student does not get admitted to any college is at most $1/n$. (Hint: use the union bound)

   Solution: If A_i is the event that student i does not get admitted to any college is at most $1/n^2$ by the previous part. $\cup_{i=1}^{n} A_i$ is the event that at least one of the students does not get admitted to any college. By using the union bound we get

$$\Pr[\cup_{i=1}^{n} A_i] \le \sum_{i=1}^{n} \Pr[A_i] \le \sum_{i=1}^{n} \frac{1}{n^2} = \frac{1}{n}.$$

6. Balls and bins, again

Suppose that I throw n balls into n bins.

1. What is the probability that I throw the i^{th} ball into the same bin as the j^{th} ball?
2. Suppose that every time I throw a ball into a bin, I lose \$1 for every ball that was already in the bin. How much money should I expect to lose?
   Hint: Note that when you throw ball i in the bin, you can only pay because of the presence of balls numbered $j < i$. Define an indicator random variable $X_{i,j}$ for all $j < i$ and express the total money you lose in terms of the $X_{i,j}$.

Solution:

1. Assume without loss of generality that $j < i$ (the case where $i < j$ is similar). Once the j^{th} ball has been thrown into some bin, the i^{th} ball lands in that bin with probability $\frac{1}{n}$.
2. Let $X_{i,j} = 1$ if balls i, j land in the same bin, and 0 otherwise. Observe that the amount of money that I lose by throwing ball i is $\sum_{j<i} X_{i,j}$. It follows that the total money I lose is $X = \sum_{i=1}^{n} \sum_{j=1}^{i-1} X_{i,j}$. By linearity of expectation,

$$\mathbf{E}[X] = \sum_{i=1}^{n} \sum_{j=1}^{i-1} \mathbf{E}[X_{i,j}] = \binom{n}{2} \frac{1}{n} = \frac{n-1}{2}$$

7. Runs

Suppose I have a biased coin which comes up heads with probability p, and I flip it n times. A "run" is a sequence of coin flips all of the same type, which is not contained in any longer sequence of coin flips all of the same type. For example, the sequence "HHHTHH" has three runs: "HHH," "T," and "HH."

Compute the expected number of runs in a sequence of n flips.

Solution: Let X_i be the indicator variable for the event that position i is the beginning of a run. Then the total number of runs is simply $X = \sum_{i=1}^{n} X_i$. Now, obviously X_1 is always 1, whereas for $i \geq 2$, $X_i = 1$ if and only if the outcome of the i-th flip is different from that of the $(i-1)$-th flip. This happens with probability $p(1-p) + (1-p)p = 2p(1-p)$. By linearity of expectation,

$$\mathbf{E}[X] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \mathbf{E}[X_3] + \cdots + \mathbf{E}[X_n] = 1 + (n-1) \cdot 2p(1-p) = 1 + 2(n-1)p(1-p).$$