Tutorial 12 Problems

Problem 1 In lecture, we computed upper and lower bounds on the probability that at least one event out of a set $A_1, \ldots, A_N$ occurs. The upper bound was:

Fact 1. *Let $A_1, A_2, \ldots, A_N$ be events over the same sample space, and let the random variable T be the number of these events that occur. Then*

$$\Pr(T \geq 1) \quad \leq \quad \sum_{i=1}^{N} \Pr(A_i).$$

The lower bound was:

Fact 2. *Let $A_1, A_2, \ldots, A_N$ be events over the same sample space, and let the random variable T be the number of these events that occur. Then*

$$\Pr(T \geq 1) \geq \max_{1 \leq i \leq N} \Pr(A_i).$$

Now assume that the events A_i are mutually independent. In this case, we can obtain a much better upper bound on the probability no events occur:

Theorem 0.1. *Let $A_1, A_2, \ldots A_N$ be independent events, and let T be the number of these events that occur. The probability that none of the events occur is at most $e^{-Ex(T)}$.*

(a) Prove this.

Hint: The proof of the theorem is surprisingly simple. The main trick is to use the fact that $1 - x \leq e^{-x}$ for all x. This inequality follows from the Taylor expansion $e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots$.

Solution:

Proof.

$$\begin{aligned}
\Pr(T = 0) &= \Pr(\bar{A_1} \cap \bar{A_2} \cap \ldots \cap \bar{A_N}) \\
&= \prod_{i=1}^{n} \Pr(\bar{A_i}) \\
&= \prod_{i=1}^{n} 1 - \Pr(A_i) \\
&\leq \prod_{i=1}^{n} e^{-\Pr(A_i)} \\
&= e^{-\sum_{i=1}^{n} \Pr(A_i)} \\
&= e^{-Ex(T)}
\end{aligned}$$

In the first step, we reason that if zero of the events A_i occur, then all of the complementary events $\bar{A}_i$ must occur. Equality holds in the second step because the events are independent. The third step uses the fact that the probability of an event is one minus the probability of the complementary event. The fourth step relies on the inequality $1 - x \leq e^{-x}$ mentioned above. The remainder is simplification. ■

Interestingly, this theorem does not depend on N, the number of events. The theorem gives the same bound whether there are 100 events each with probability 0.1 or 1000 events each with probability 0.01. In both cases, the expected number of events is 10, and so the probability of no event occurring is at most e^{-10} or about 1 in 22,000. (Though the bounds given by the theorem are equal, the actual probabilities are somewhat different in these two cases.)

This theorem can be interpreted as a sort of "Murphy's Law": if we expect some things to go wrong, then something probably will. For example, suppose that we are building a microprocessor, and the fabrication process is such that each transistor is faulty mutually independently with a probability of one in a million. This sounds good. However, microprocessors now contain about ten million transistors, so the expected number of faulty transistors is 10 per chip. Since we expect some things to go wrong, something probably will. In fact, Theorem 0.1 implies that the probability of a defect-free a chip is less than 1 in 22,000!

(b) A Mind-Reading Trick

For this trick, every card has a "value". The value of a number card less than 10 is just the number, and the value of a 10, face card, ace, or joker is 1. For example, $8\clubsuit$ has value 8, $Q\spadesuit$ has value 1, and $10\diamondsuit$ has value 1.

A volunteer shuffles the deck and secretly chooses a number from 1 to 9. She deals this number of cards minus one from the top of the deck and looks at the new top card. For example, if she chooses the number 4, then she deals off 3 cards and looks at the 4th card.

She then takes the value of this top card, deals off that number of cards minus one, and again looks at the new top card. For example if she sees a $6\diamondsuit$, then she deals off 5 more cards and looks at the next one.

She repeats this procedure until she runs out of cards. "Her card" is the last card that she looked at to determine the number to deal off. For example, suppose that the deck is as follows:

$$Q\heartsuit \; 4\spadesuit \; \underline{2\spadesuit} \; J\clubsuit \; \underline{K\spadesuit} \; \underline{5\heartsuit} \; 3\heartsuit \; A\clubsuit \; 7\diamondsuit \; 7\heartsuit \; \underline{3\clubsuit} \; K\clubsuit$$

If the volunteer chooses 3 as her secret number, then she deals off $3 - 1 = 2$ cards and looks at the $2\spadesuit$. Since the value of this card is 2, she deals off $2 - 1 = 1$ more card and looks at the $K\spadesuit$. This card has value 1, so she deals off $1 - 1 = 0$ more cards and looks at the $5\heartsuit$. Since this card has value 5, she deals of $5 - 1 = 4$ more cards and looks at the $3\clubsuit$. This card has value 3, so she tries to deal off $3 - 1 = 2$ more cards. However, she then runs out of cards, so "her card" is $3\clubsuit$.

After the volunteer completes the procedure, the magician is able to name her card!

The secret is that the magician performs the same procedure as the volunteer, starting with the number 1. When he guesses "her card", he is actually naming "his card"! Remarkably, the two are almost certain to be the same! The reason is that if the magician and volunteer ever land on the same card, then they will finish with the same card. In the example, the magician lands on the underlined sequence of cards:

$\underline{Q\heartsuit}$ $\underline{4\spadesuit}$ $2\spadesuit$ $J\clubsuit$ $K\spadesuit$ $\underline{5\heartsuit}$ $3\heartsuit$ $A\clubsuit$ $7\diamondsuit$ $7\heartsuit$ $\underline{3\clubsuit}$ $K\clubsuit$

Even though the volunteer and the magician start with different numbers, they both coincidentally land on the $5\heartsuit$. After this point, they both continue in the same way and both end up with $3\clubsuit$.

All that remains is to prove that the magician and the volunteer are likely to land on the same card at some point. An exact analysis of this probability is difficult, but Theorem 0.1 gives some intuition.

Let A_i be the event that on his i-th jump the magician lands on a card also visited by the volunteer. If the volunteer's next card is 1 away, then the magician has a $\frac{1}{3}$ chance of landing on it (12 face cards, 3 10's, and 2 jokers gives $\frac{18}{54} = \frac{1}{3}$). If the volunteer's next card is more than 1 away, then the magician has a $\frac{2}{27}$ chance of landing on it (4 correct number cards gives $\frac{4}{54} = \frac{2}{27}$).

Since the magician takes 10-15 jumps through the deck, the sum of these probabilities is rather high; that is, $Ex(T)$ is around 3 or so. Consequently, Theorem 0.1 suggests that the magician is likely to eventually jump on a card visited by the volunteer. (As a ball-park estimate, the probability might be around $1 - e^{-3}$.) In this case, the magician and volunteer end up with the same final card and the trick works.

Of course, Theorem 0.1 does not really apply, because the events A_i are not at all independent. For example, event A_i implies event A_{i+1}; once the magician and volunteer match, they always agree. This means that our probability estimates are bogus. However, computer tests suggest that for a full deck with jokers, the trick works more than 90% of the time.

Problem 2 The U.S. Supreme Court has mandated a re-vote in Florida. All 6,000,000 people who voted earlier will get to vote again. They are going to vote randomly, using a strategy that guarantees that the expected value of the number of votes for Bush is exactly 3,000,000. We don't know what the strategy is, though. (For example,
1. Each voter might toss a fair coin independently.
2. The Florida Supreme Court may toss one fair coin, and give all the votes to one candidate or the other based on the result.
3. 1/4 of the voters decide to vote for Bush, 1/4 for Gore, and the other half toss independent fair coins.
Etc.)

(a) What do Markov bounds tell us about the probability that Bush gets at least 6,000 more votes than Gore?

Solution: Let R be the random variable which is equal to the number of votes that Bush gets. Note that for Bush to get at least 6,000 more votes than Gore, Bush only needs to increase his vote total to 3,000 above the expected value. Then, by Markov, $\Pr(R \geq 3,003,000) \leq \frac{Ex[R]}{3,000} = \frac{3,000,000}{3,003,000} = .9990$

(b) Now assume also that the standard deviation of the experiment is 2000. What do Chebyshev bounds tell us about the probability that Bush gets at least 6,000 more votes than Gore?

Solution: According to Mr. Chebyshev, $\Pr(|R - Ex[R]| \geq x) \leq \frac{Var(R)}{x^2}$. Let R be the same random variable as in part a. Plugging in $x = 3,000$, $Ex[R] = 3,000,000$, and $Var[R] = (\text{standard dev})^2 = 4,000,000$ we get:

$$\begin{aligned} Pr(R - 3,000,000 \geq 3,000) &\leq \frac{4,000,000}{3,000^2} \\ &= \frac{4}{9} = .444444 \end{aligned}$$

(c) Now assume that we know the experiment. Specifically, one quarter of the voters decide ahead of time on Gore, one quarter decide on Bush, and the other half flip fair coins, independently. (This doesn't have a standard deviation of 2000.) What do Chernoff bounds tell us about the probability that Bush gets at least 6,000 more votes than Gore?

Solution: This time we'll need to compute the variance on our own before we can apply Mr. Chernoff. Let R be defined as above. Then, let's let R_1 be the random variable corresponding to the number of people who vote for Bush non-randomly, and let R_2 correspond to the number of votes for Bush obtained by the coin-flipping half. These two random variables are independent, so $Var(R) = Var(R_1 + R_2) = Var(R_1) + Var(R_2)$. Additionally, $Var(R_1) = 0$, since $R_1 = Ex(R_1)$ for all possible outcomes of this experiment.
Since R_2 represents a binomial distribution, we know from last tutorial and from Rosen that it has a variance $npq = \frac{n}{4}$. In this case, n is the number of coin-flippers $= 3,000,000$. Putting this all together, we have:

$$\begin{aligned} Pr(R - Ex[R] \geq x) &\leq \frac{Var(R)}{x^2} \\ Pr(R - Ex[R] \geq 3,000) &\leq \frac{Var(R_1)}{3,000^2} + \frac{Var(R_2)}{3,000^2} \\ &= 0 + \frac{3,000,000}{4} * \frac{1}{9,000,000} \\ &= .083333 \end{aligned}$$

Problem 3 *The Probabilistic Method*

A round robin tournament of n contestants is one in which each of the $\binom{n}{2}$ pairs of contestants play each other exactly once, with the outcome of any play being that one of the contestants wins and the other loses. For a fixed integer k, $k < n$, a question of interest is whether it is possible that the tournament outcome is such that for every set of k players there is a player who beat each player of this set. Show that if:

$$\binom{n}{k}\left[1 - \left(\frac{1}{2}\right)^k\right]^{n-k} < 1$$

then such an outcome is possible.

(a) Start by numbering the sets of k contestants. How many such sets are there?

Solution:

$$\binom{n}{k}$$

(b) Let B_i be the event that no contestant beat all the k contestants in set i. Compute $\Pr(B_i)$. (Note that you must choose probabilities for each match in order to compute this).

Solution: Suppose that the results of the game are independent and that each game is equally likely to be won by either contestant. This is an arbitrary choice, but it is easy to work with (and any probability will suffice to prove existance). The probability that a person inside group i beats everyone in i is clearly 0. The probability that a person outside group i beats everyone in i is $\left(\frac{1}{2}\right)^k$, so the probability the person they did not beat everyone in i is $1 - \left(\frac{1}{2}\right)^k$. There are $n - k$ people outside of group i. Thus, B_i has probability

$$\left[1 - \left(\frac{1}{2}\right)^k\right]^{n-k}$$

(c) Give an upper bound on $\Pr(\cup B_i)$.

Solution: Then use Boole's inequality to bound: $\Pr(\cup B_i) \leq \sum \Pr(B_i)$. In other words, $\Pr(\cup B_i)$ can be no greater than if all of the B_i are disjoint. (Overlap will merely reduce the total probability of the union). Since the expression for B_i does not depend on i – the probability is the same for each k-sized group i – the sum of all of the B_i is simply

$$\binom{n}{k}\left[1 - \left(\frac{1}{2}\right)^k\right]^{n-k}$$

(d) Explain why this result can be used to prove the existence of the desired tournament outcome.

Solution: Probabilistic proof. If the overall probability of $\cup B_i$ is less than 1 then there must be an occurrence that is not in $\cup B_i$. It does not matter that we chose the probabilities arbitrarily; the fact that there is *any* probaility at all left over means some outcome that we had not accounted for is possible.

For interest, some numbers that work are:

$k = 1, n = 3; k = 2, n = 21; k = 3, n = 33; k = 4, n = 46;$
$k = 5, n = 59; k = 6, n = 72; k = 7, n = 85; k = 8, n = 98$