Problem Set 10

Due: Monday, November 24, 7pm

Problem 1. [**20 points**] You are organizing a neighborhood census and instruct your census takers to knock on doors and note the sex of any child that answers the knock. Assume that there are two children in a household, that children are equally likely to be girls and boys, and that girls and boys are equally likely to open the door.

A sample space for this experiment has outcomes that are triples whose first element is either `B` or `G` for the sex of the elder child, likewise for the second element and the sex of the younger child, and whose third coordinate is `E` or `Y` indicating whether the *e*lder child or *y*ounger child opened the door. For example, $(\mathtt{B}, \mathtt{G}, \mathtt{Y})$ is the outcome that the elder child is a boy, the younger child is a girl, and the girl opened the door.

(a) [5 pts] Let T be the event that the household has two girls, and O be the event that a girl opened the door. List the outcomes in T and O.

(b) [5 pts] What is the probability $\Pr(T \mid O)$, that both children are girls, given that a girl opened the door?

(c) [10 pts] Where is the mistake in the following argument for computing $\Pr(T \mid O)$?

> If a girl opens the door, then we know that there is at least one girl in the household. The probability that there is at least one girl is
>
> $$1 - \Pr(\text{both children are boys}) = 1 - (1/2 \times 1/2) = 3/4.$$
>
> So,
>
> $$\begin{aligned} &\Pr(T \mid \text{there is at least one girl in the household}) \\ &\quad = \frac{\Pr(T \cap \text{there is at least one girl in the household})}{\Pr\{\text{there is at least one girl in the household}\}} \\ &\quad = \frac{\Pr(T)}{\Pr\{\text{there is at least one girl in the household}\}} \\ &\quad = (1/4)/(3/4) = 1/3. \end{aligned}$$
>
> Therefore, given that a girl opened the door, the probability that there are two girls in the household is 1/3.

Problem 2. [15 points] In lecture we discussed the Birthday Paradox. Namely, we found that in a group of m people with N possible birthdays, if $m \ll N$, then:

$$\Pr\{\text{all } m \text{ birthdays are different}\} \sim e^{-\frac{m(m-1)}{2N}}$$

To find the number of people, m, necessary for a half chance of a match, we set the probability to 1/2 to get:

$$m \sim \sqrt{(2\ln 2)N} \approx 1.18\sqrt{N}$$

For $N = 365$ days we found m to be 23.

We could also run a different experiment. As we put on the board the birthdays of the people surveyed, we could ask the class if anyone has the same birthday. In this case, before we reached a match amongst the surveyed people, we would already have found other people in the rest of the class who have the same birthday as someone already surveyed. Let's investigate why this is.

(a) [5 pts] Consider a group of m people with N possible birthdays amongst a larger class of k people, such that $m \leq k$. Define $\Pr\{A\}$ to be the probability that m people all have different birthdays *and* none of the other $k - m$ people have the same birthday as one of the m.

Show that, if $m \ll N$, then $\Pr\{A\} \sim e^{\frac{m(m-2k)}{2N}}$. (Notice that the probability of no match is $e^{-\frac{m^2}{2N}}$ when k is m, and it gets smaller as k gets larger.)

Hints: For $m \ll N$: $\frac{N!}{(N-m)!N^m} \sim e^{-\frac{m^2}{2N}}$, and $(1 - \frac{m}{N}) \sim e^{-\frac{m}{N}}$.

(b) [10 pts] Find the approximate number of people in the group, m, necessary for a half chance of a match (your answer will be in the form of a quadratic). Then simplify your answer to show that, as k gets large (such that $\sqrt{N} \ll k$), then $m \sim \frac{N \ln 2}{k}$.

Hint: For $x \ll 1$: $\sqrt{1-x} \sim (1 - \frac{x}{2})$.

Problem 3. [10 points] We're covering probability in 6.042 lecture one day, and you volunteer for one of Professor Leighton's demonstrations. He shows you a coin and says he'll bet you \$1 that the coin will come up heads. Now, you've been to lecture before and therefore suspect the coin is biased, such that the probability of a flip coming up heads, $\Pr\{H\}$, is p for $1/2 < p \leq 1$.

You call him out on this, and Professor Leighton offers you a deal. He'll allow you to come up with an algorithm using the biased coin to *simulate* a fair coin, such that the probability you win and he loses, $\Pr\{W\}$, is equal to the probability that he wins and you lose, $\Pr\{L\}$. You come up with the following algorithm:

1. Flip the coin twice.
2. Based on the results:
   - $TH \Rightarrow$ you win $[W]$, and the game terminates.
   - $HT \Rightarrow$ Professor Leighton wins $[L]$, and the game terminates.
   - $(HH \vee TT) \Rightarrow$ discard the result and flip again.
3. If at the end of N rounds nobody has won, declare a tie.

As an example, for $N = 3$, an outcome of HT would mean the game ends early and you lose, $HHTH$ would mean the game ends early and you win, and $HHTTTT$ would mean you play the full N rounds and result in a tie.

(a) [5 pts] Assume the flips are mutually independent. Show that $\Pr\{W\} = \Pr\{L\}$.

(b) [5 pts] Show that, if $p < 1$, the probability of a tie goes to 0 as N goes to infinity.

Problem 4. [20 points]

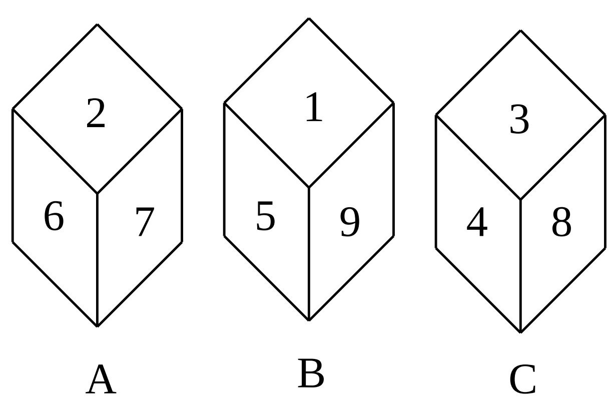
(a) [5 pts] Suppose A and B are *disjoint* events. Prove that A and B are *not independent*, unless $\Pr(A)$ or $\Pr(B)$ is zero.

(b) [5 pts] If A and B are independent, prove that A and $\overline{B}$ are also independent.
Hint: $\Pr\left(A \cap \overline{B}\right) = \Pr(A) - \Pr(A \cap B)$.

(c) [5 pts] Give an example of events A,B, C such that A is independent of B, A is independent of C, but A is not independent of $B \cup C$.

(d) [5 pts] Prove that if C is independent of A, and C is independent of B, and C is independent of $A \cap B$, then C is independent of $A \cup B$.
Hint: Calculate $\Pr(A \cup B \mid C)$.

Problem 5. [20 points] Recall the strange dice from lecture:



In lecture we proved that if we roll each die once, then die A beats B more often, die B beats die C more often, and die C beats die A more often. Thus, contrary to our intuition, the "beats" relation $>$ is not transitive. That is, we have $A > B > C > A$.

We then looked at what happens if we roll each die twice, and add the result. In lecture, we showed that rolling die B twice is more likely to win, i.e., have a larger sum, than rolling die A twice, which is the opposite of what happened if we were to just roll each die once! In fact, we will show that the "beats" relation reverses in this game, that is, $A < B < C < A$, which is very counterintuitive!

(a) [5 pts] Show that rolling die C twice is more likely to win than rolling die B twice.

(b) [5 pts] Show that rolling die A twice is more likely to win that rolling die C twice.

(c) [5 pts] Show that rolling die B twice is more likely to win that rolling die A twice.

Problem 6. [15 points]

(a) [7 pts] Suppose you repeatedly flip a fair coin until you see the sequence `HHT` or the sequence `TTH`. What is the probability you will see `HHT` first?

Hint: Use a bijection argument.

(b) [8 pts] What is the probability you see the sequence `HTT` before you see the sequence `HHT`?

Hint: Try to find the probability that `HHT` comes before `HTT` conditioning on whether you first toss an `H` or a `T`. Somewhat surprisingly, the answer is not $1/2$.