Problem Set 11

Due: Tuesday, November 29, 7:30pm

Problem 1. **[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 2. **[20 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) [7 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 $1/2$.

(c) [6 pts] Suppose you flip three fair, mutually independent coins. Define the following events:

- Let A be the event that *the first* coin is heads.
- Let B be the event that *the second* coin is heads.
- Let C be the event that *the third* coin is heads.
- Let D be the event that *an even number of* coins are heads.

Use the four step method to determine the probability of each of A, B, C, D.

Problem 3. [20 points] Suppose you have seven standard dice with faces numbered 1 to 6. Each die has a label corresponding to a letter of the alphabet (A through G). A *roll* is a sequence specifying a value for each die in alphabet order. For example, one roll is $(6, 1, 4, 1, 3, 5, 2)$ indicating that die A showed a 6, die B showed 1, die C showed 4, ...

(a) [5 pts] What is the probability of a roll where *exactly* two dice have the value 3 and the remaining five dice all have different values?

Example: $(3, 2, 3, 1, 6, 4, 5)$ is a roll of this type, but $(1, 1, 2, 6, 3, 4, 5)$ and $(3, 3, 1, 2, 4, 6, 4)$ are not.

(b) [5 pts] What is the probability of a roll where two dice have an even value and the remaining five dice all have different values?

Example: $(4, 2, 4, 1, 3, 6, 5)$ is a roll of this type, but $(1, 1, 2, 6, 1, 4, 5)$ and $(6, 6, 1, 2, 4, 3, 4)$ are not.

(c) [10 pts] What is the probability of a roll where two dice have one value, two different dice have a second value, and the remaining three dice a third value?

Example: $(6, 1, 2, 1, 2, 6, 6)$ is a roll of this type, but $(4, 4, 4, 4, 1, 3, 5)$ and $(5, 5, 5, 6, 6, 1, 2)$ are not.