18.310 Homework 2

Due September 18th at 6PM

Instructions: Remember to submit a separate PDF for each question. Do not forget to include a list of your collaborators or to state that you worked on your own.

1. The following theorem and its proof are mathematically incorrect.

   > Let n be any positive integer and recall that $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$. For a set $X \subseteq \mathbb{Z}_n$, let $2X = \{2x \pmod{n} : x \in X\}$.
   >
   > **Theorem 1.** *Suppose that X and Y are drawn independently and uniformly at random among all 2^n subsets of $\mathbb{Z}_n$. Then $\mathbb{P}(2X \cap Y = \emptyset) = \frac{3^n}{4^n}$.*
   >
   > *Proof.* We know that a uniformly random set X can be generated by independently deciding to include i in X with probability $\frac{1}{2}$, for each $i \in \mathbb{Z}_n$. Thus, we obtain that
   >
   > $$\mathbb{P}(2X \cap Y = \emptyset) = \prod_{i=0}^{n-1} \mathbb{P}(\text{if } i \in X \text{ then } (2i \pmod{n}) \notin Y) = \left(\frac{3}{4}\right)^n.$$
   >
   > $\square$

   - **To be handed in in recitation on 9/12/2013.** Show that the theorem is false by explicitly calculating $\mathbb{P}(2X \cap Y = \emptyset)$ for $n = 2$. Does your counterexample generalize to $n = 3$? (Also, think about what step of the current proof is incorrect prior to recitation.)

   - **Writing assignment.** (To be submitted with the rest of the problem set on 9/18/2013.) Correct the statement of the theorem above so that it is true for every n. Provide a well-written proof. Pay attention to notation, and also to the issues that made the "proof" above wrong.

2. The classroom that we are in has six blackboard frames. In some of the lectures, the instructor enjoys showing his (lack of) drawing skills and draws a pigeon on one or several board frames. Show that over the course of a semester with 36 lectures, there exist two lectures and three board frames such that these three frames either all had no pigeons drawn on them in both lectures, or all had at least one pigeon drawn on them in both lectures.

3. A random variable $Y : \Omega \to \mathbb{Z}$ is distributed according to the Poisson distribution with parameter $\lambda \geq 0$ if for all $i \geq 0$:

   $$\mathbb{P}(Y = i) = \mathrm{e}^{-\lambda}\frac{\lambda^i}{i!}.$$

   - Verify that $\sum_{i=0}^{\infty} \mathbb{P}(Y = i) = 1$.

- Show that $\mathbb{E}[Y] = \text{Var}(Y) = \lambda$.

- Suppose that each random variable $X_1, X_2, \ldots, X_n$ follows the Poisson distribution with parameter λ_i. Assume that all X_i are independent and let $X = \sum_{i=1}^{n} X_i$. Show that, for $\mu \geq \mathbb{E}[X]$ and for all $\delta > 0$:

$$\mathbb{P}(X > (1+\delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu.$$

Hint: Compare this theorem with part (i) of Theorem 1 in the lecture notes on Chernoff Bound. Try to follow the proof of Theorem 1 closely.

4. You may have heard recently some story about former MIT students (and other groups) winning a fair amount of money at the Massachusetts lottery game Cash WinFall (if not, just google 'Cash WinFall MIT students'). Let's analyze the game (or some simplification of it).

In Cash WinFall, a customer can buy a ticket for $2 which let him/her choose 6 numbers between 1 and 46 hoping to match the 6 (distinct) numbers being randomly selected at the next drawing. If the 6 numbers on the ticket match the 6 numbers that are drawn, he/she wins the jackpot, which is at least $500,000. The customer also wins prize money if 5, 4 or 3 of the numbers are matched, see the second column in the table below for the prize money in each case.

Match	Prize money	Example (from 2/8/2010) of prize money when rolldown
6 out of 6	jackpot	-
5 out of 6	$4,000	$22,096
4 out of 6	$150	$807
3 out of 6	$5	$26

The MIT students and the other groups exploited the fact that if the jackpot reaches $2,000,000 and the jackpot is not won then part of the jackpot money is used to considerably increase the prize money for matching 5, 4 or 3 of the numbers; see the third column in the table above. Notice that the increase is more than 5-fold. Such a drawing is known as a *rolldown drawing*. The precise increase for a rolldown drawing is based on formulas that are not (quite) revealed to the public (and depends on the amount of the jackpot, etc.), but the increase is always very significant and of the order of magnitude shown in the 3rd column above.

(a) For $i = 6, 5, 4, 3$, what is the probability p_i that one ticket matches precisely i of the 6 numbers that are randomly drawn? Give a formula and also numerically compute these probabilities.

(b) Let A be the event that one wins any amount of prize money when buying a single ticket. What is $\mathbb{P}(A)$?

(c) Let the random variable X be the prize money for a single ticket, assuming (i) that the jackpot amount is $1,900,000 and (ii) that the drawing is not a rolldown drawing. What is $\mathbb{E}(X)$? Compute its numerical value. (Should you play?)

(d) Assume that we have a rolldown drawing (i.e. no one wins the jackpot which happens to be over \$2,000,000). Suppose furthermore that the prize money for matching 5, 4 or 3 numbers are as in the 3rd column in the table. Let Y be the prize money for a single ticket under these assumptions.

What is $\mathbb{E}(Y)$ and $\text{Var}(Y)$? Compute their values.

(e) If you purchase only one ticket, you have a large probability of not recovering your bet. Now suppose you purchase 1,000,000 tickets[1], each randomly drawn. Let Z be the total prize money received.

What is $\mathbb{E}(Z)$? What is $\text{Var}(Z)$? Use Chebyshev's inequality to compute an upper bound on the probability that $Z < 2,000,000$ (i.e. that you are losing money).

(f) Now use the Chernoff-Hoeffding bound to compute a better upper bound on the probability that $Z < 2,000,000$. How much better is your result?

(g) If the jackpot goes over \$2,000,000, a rolldown might not happen since some ticket might win the jackpot. Suppose that, for a given drawing, the total number of (distinct) tickets sold[2] is $1,000,000$. Let B be the event that someone wins the jackpot. What is $\mathbb{P}(B)$? To evaluate this numerically, it is convenient to use the approximation[3] $1 - x \sim e^{-x}$.

The following exercises should **not** be handed in, but we nevertheless encourage you to do them.

1. The union bound states that, for any set of n events, the probability that at least one of the events happens is no greater than the sum of the probabilities of the individual events, i.e.

$$\mathbb{P}\left(\bigvee_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}\left(A_i\right),$$

where $\bigvee_{i=1}^{n} A_i = A_1 \vee A_2 \vee \cdots \vee A_n$. Prove the union bound using linearity of expectation. Recall that the indicator variable for an event A, $\mathbb{1}_A$ is a random variable that's 1 when A holds and 0 otherwise.

2. Informally, this question asks you to show that any real number (even irrational) can be approximated by a fraction with small denominator without incurring too much error. More formally, prove that for any $x \in \mathbb{R}$ and for any $n \in \mathbb{N}$, there exist integers p and q with $1 \leq q \leq n$ satisfying

$$\left|x - \frac{p}{q}\right| < \frac{1}{qn}.$$

Hint: Prove that there exist p and q with $1 \leq q \leq n$ such that $|qx - p| \leq \frac{1}{n}$.

---

[1]The MIT students purchased up to 700,000 tickets for one drawing...

[2]In 2004-2005, the number of tickets sold in anticipation of a rolldown drawing was never more than 950,000 and typically less than 600,000 while in 2007, the number of tickets sold in a rolldown drawing was typically between 1,200,000 and 1,400,000.

[3]One has $1 - x \leq e^{-x}$ for *all* x, and the approximation $1 - x \sim e^{-x}$ is very good for x close to 0.

18.310 Principles of Discrete Applied Mathematics
Fall 2013