Homework 3 – Due Friday, February 7, 2020 at noon

Submit solutions to all problems on separate sheets. They will be graded by different people.

Page limit You can submit **at most** 1 sheet of paper per problem, even if the problem has multiple parts. If you submit a longer solution for some problem, only the first sheet of paper will be graded.

Reminder Collaboration is permitted, but you must write the solutions *by yourself without assistance*, and be ready to explain them orally to the instructor if asked. You must also identify your collaborators and whether you gave help, received help, or worked something out together. Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden.

Exercises Please practice on exercises in Chapter 2 of Mitzenmacher-Upfal and the following exercise.

1. Let R_1 and R_2 be independent random variables and let f be a function from real numbers to an arbitrary range.
   (a) Prove that for every subset S of the range of f, the events $f(R_1) \in S$ and $f(R_2) \in S$ are independent.
   (b) Use part (a) to justify our analysis of amplification, where we repeat the same basic algorithm multiple times with independent random coins to decrease error probability.

Problems

1. (**Detecting defects**) You are in charge of inspecting cookies baked in your company. A worker is *unreliable* if the proportion of defective cookies he bakes is α or higher. (For example, if he bakes 1000 cookies in a day, $\alpha 1000$ of them will be defective.)
   (a) You would like to make sure that if a worker is unreliable, you will find at least one defective cookie with probability 99% and fire him. You do not have time to inspect every single cookie. Instead you decide to choose k cookies uniformly at random. For simplicity, let's say that you will choose cookies with replacement. That is, on each try, you find a defective cookie with probability p, where p is the proportion of defective cookies the worker baked. Prove that if
   $$k \geq \frac{\ln 100}{\alpha}$$
   and $p \geq \alpha$ then you will find at least one defective cookie with probability 99%.
   (b) You got a raise, and now you are inspecting cookies made by n workers. You would like to ensure that with probability at least 99%, all unreliable workers will be fired, that is, you will detect a defective cookie for every one of them. You use the same random strategy as in part (a) for each worker, but now you have to inspect more cookies per worker because you want to achieve overall success probability of 99%, not just per worker. What should you choose k to be as a function of α and n? Find the smallest k you can.
   Hint: Use the union bound.

2. (**Improved Randomized Min-Cut Algorithm**) Exercise 1.25 in the book.

3. (**Jensen's Inequality**)

   (a) Suppose f is concave. How does Jensen's inequality change for this case? (Briefly justify.)

   (b) The geometric mean of a collection of n positive real numbers is the nth root of the product of the numbers, and the arithmetic mean is just the average. Use part (a) to prove that, for any set of n numbers, the arithmetic mean is greater or equal to the geometric mean.

   (c) Use part (a) to prove that for any $\triangle ABC$, we have $\sin A + \sin B + \sin C \leq \frac{3\sqrt{3}}{2}$.

4. (**Random children**) After getting married, Cinderella and Prince decide to have children. They want to have at least one daughter, so they keep having children until the first girl is born.

   (a) Assuming that each child is a boy or a girl independently with equal probability, what is the expected number of girls and the expected number of boys the couple will have?

   (b) How does the answer change if the probability of having a girl is only 0.4 instead of 0.5?

   After thinking about feeding an army of children, Cinderella and Prince decide to change the strategy. They want to have children until either they have their first girl or they have k children, where $k \geq 1$. Assume there are no multiple births.

   (c) The same question as in part (a).

   (d) The same question as in part (b).