CS 127 Spring 2018 - Homework Zero

Due on **Thursday January 25 at midnight** but I recommend you do this homework even before the first lecture.

Please write the names of yourself and your collaborators below: (this page does not contain any solutions, and so is not viewed by the people grading your problems on gradescope.)

Collaboration policy for homework zero: You can work on this problem set and submit it in **groups of one or two students**. In addition you can discuss these problems with other students that are considering taking this course. However, each pair needs to write up their solution on their own, and you need to write down the names of your collaborators below.

Submitters: (List your names and HUIDs here)

Collaborators: (List here anyone you discussed problems or ideas for solutions with)

If submitting late, for every pair member add a sentence of the form **"Submitting X days late, Y has used (including this time) Z days out of the 6 total late days"**

For any questions or clarifications, please see the Piazza board. See the course syllabus for policies.

This homework needs to be submitted as a PDF document. The PDF can be generated using either markdown or LaTeX. The Markdown source for this homework is posted online.

Probability questions

As we'll see in the first lecture, much of cryptography relies on probability theory, and so basic knowledge of probability will be essential. The CS 121 probability review is one source for the probability theory we will need. During the course I will assume you are familiar with (or can pick on your own) all notions presented there. However, if you find these questions unfamiliar or difficult you should not despair! There are plenty of sources on probability on the web, and in particular Harvard STAT 110 and its textbook are of course wonderful resources. If any of the notation is unfamiliar, looking at the lecture notes might help, and otherwise feel free to ask questions on Piazza, even before the semester starts!

Notation: While often in probability theory people use the name "random variable" for a distribution over the set $\mathbb{R}$ of real numbers, it will be convenient for us to generalize this to arbitrary sets, and hence we will use the following notation. We define a *random variable* or *distribution* X over a finite set S to correspond to the probabilistic experiment where we draw an element x from S with some probability, which we denote by $\Pr[X = x]$. All we need from these probabilities is that they are non-negative and sum up to one. (One can also consider distributions over infinite sets, though almost always in this course we will restrict ourselves to the finite case.) We use $x \leftarrow_R X$ as shorthand for saying that x is drawn accoring to the distribution X. If $f : S \to T$ is a function, then the random variable $f(X)$ corresponds to the probabilistic experiment where we draw $x \leftarrow_R X$ and output $f(x)$.

Question 0. Join the course Piazza board at http://piazza.com/harvard/spring2018/cs127.

In this question we will study the notion known as Total Variation or statistical distance. It is a basic notion of distance between probability distribution, and its computational analog is fundamental for cryptography.

Question 1.1: If X and Y are two distributions over the same set S, we define $\Delta(X,Y)$ (also known as the *statistical* or *total variation* distance of X and Y) to be $\sum_{x \in S} |\Pr[X = x] - \Pr[Y = x]|$. Prove that for every function $f : S \to [0,1]$, $|\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]| \leq \Delta(X,Y)$.

Solution 1.1:

Question 1.2: Prove that the statistical distance satisfies the *triangle inequality*: For every three distributions X, Y, Z over the same set S, $\Delta(X,Z) \leq \Delta(X,Y) + \Delta(Y,Z)$.

Solution 1.2:

We will now use the notion of statistical distance to study one of the most basic questions in probability theory: if we are given a coin that is either completely unbiased, or has bias $\epsilon > 0$ towards "heads", how many tosses will it take for us to distinguish between the two cases.

Question 2: Prove that this can be done in at most $O(1/\epsilon^2)$ coin tosses. Specifically prove that if $k > 100/\epsilon^2$, then there exists some function $f : \{0,1\}^k \to \{0,1\}$ such that if X is the uniform distribution over $\{0,1\}^k$ and Y is the distribution obtained by tossing k independent coins, each equalling 1 with probability $1/2 + \epsilon$ and equalling 0 with probability $1/2 - \epsilon$, then $\Pr[f(X) = 0] > 0.9$ and $\Pr[f(Y) = 1] > 0.9$, and hence f can distinguish between X and Y.[1]

Solution 3:

[1] **Hint:** Use the Chernoff bound.

We now study the converse problem: showing a *lower bound* on the number of coin tosses needed.

Question 3.1: (This problem involves some notation, so take your time reading it carefully and parsing what it means.) For every $k \in \mathbb{N}$, $0 \leq i \leq k$, and $\epsilon > 0$, let $X_i^{k,\epsilon}$ be the following distribution over $\{0,1\}^k$: the first i bits of $X_i^{k,\epsilon}$ are chosen independently and uniformly at random, and the last $k-i$ bits are chosen independently at random with but each is equal to 1 with probability $1/2 + \epsilon$ and equal to 0 with probability $1/2 - \epsilon$. Prove that for very k, $i < k$ and ϵ, $\Delta(X_i^{k,\epsilon}, X_{i+1}^{k,\epsilon}) \leq 10\epsilon$.

Solution 3.1:

Question 3.2: Prove that, in the notation of Question 2, $\Delta(X,Y) \leq 10k\epsilon$. Show that this implies if $k < 1/(100\epsilon)$, there does not exists a function $f : \{0,1\}^k \to \{0,1\}$ such that $\Pr[f(X) = 0] > 0.9$ and $\Pr[f(Y) = 1] > 0.9$. Use your solution for Question 3.1 (this is a special case of the Hybrid Argument which is used time and again in cryptography).

Solution 3.2:

Question 4 (bonus problem: optional and more challenging): Prove that in fact $\Delta(X,Y) \leq O(\sqrt{k}\epsilon)$.[2] Conclude that the method of Question 2 is essentially *optimal* in the sense that there exist some absolute constant δ (independent of ϵ) such that for every ϵ and distribution X, Y as in Question 2, if $k < \delta/\epsilon^2$ then there does not exists a function $f : \{0,1\}^k \to \{0,1\}$ such that $\Pr[f(X) = 0] > 0.9$ and $\Pr[f(Y) = 1] > 0.9$.

Solution 4:

[2] **Hint:** One way to prove this is to use the notion of KL divergence which is another notion of distance between the distributions that satisfies the triangle inequality. You can show that the KL divergence of $X_i^{k,\epsilon}$ and $X_{i+1}^{k,\epsilon}$ is $O(\epsilon^2)$ and then use the Pinsker Inequality that shows that the statistical distance between two distributions is at most the square root of their KL divergence. You can read about both KL divergence and the Pinsker Inequality in Wikipedia as well in several other sources.

And now for a little crypto

Question 5 (very important! but no grades :)): Read the lecture notes for lecture 1: introduction.

Solution 5: Write here "I solemnly swear that I read every word of the lecture notes". (Just kidding, I trust you, and in any case I expect you to read every word *twice*.)

Question 6: Prove that an encryption scheme (E, D) with messages of length ℓ and keys of length n is *perfectly secret* if and only for every $x, x' \in \{0,1\}^\ell$, $\Delta(Y^x, Y^{x'}) = 0$, where Y^x is the distribution obtained by choosing a random $k \leftarrow_R \{0,1\}^n$ and outputting $E_k(x)$.

Solution 6: