CS 127/227 Fall 2021 - Homework Zero

Due on **Thursday September 9, 2021** (see gradescope) This homework is meant to ensure you brush up on the mathematical background needed for this course, see the mathematical background lecture notes on the course notes website

Some policies: (See the course syllabus for the full policies.)

- You can collaborate with other students that are currently enrolled in this course in brainstorming and thinking through approaches to solutions but you should write the solutions on your own and cannot share them with other students.
- Sharing questions or solutions with anyone outside this course, including posting on outside websites, is a violation of the honor code policy. Collaborating with anyone except students currently taking this course or using material from past years from this or other courses is a violation of the honor code policy.
- The submitted PDF should be typed and in the same format and pagination as ours. Please include the text of the problems and write **Solution X:** before your solution. (This is easy to do if you use our markdown template.) Please mark in gradescope the pages where the solution to each question appears. Points will be deducted if you submit in a different format or do not mark the pages.

By writing my name here I affirm that I am aware of all policies and abided by them while working on this problem set:

Your name: (Write name and HUID here)

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

Number of late days used so far: (not including this pset; it is your responsibility to make sure you do not go over the late days budget)

Number of late days used for this pset:

Probability questions

As we'll see in the first lecture **there is no secrecy without randomness** and so much of cryptography relies on probability theory. 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 up 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. Otherwise, feel free to ask questions on Ed 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 1 (30 points): 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 (10 points): If X and Y are two distributions over the same set S, we define the *statistical distance* of X and Y, denoted as $\Delta(X, Y)$ (also known as *total variation* distance of X and Y) to be $\frac{1}{2}\sum_{x \in S} |\Pr[X = x] - \Pr[Y = x]|$. Prove that for every function $f : S \to [0, 1]$, $|\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]| \le \Delta(X, Y)$.

Solution 1.1:

Question 1.2 (10 points): Prove that the statistical distance satisfies the *triangle inequality*: For every three distributions X, Y, Z over the same set S, $\Delta(X, Z) \le \Delta(X, Y) + \Delta(Y, Z)$.

Solution 1.2:

Question 1.3 (10 points): Let X, Y, Z be three distributions over some set S. Denote by $X \circ Y$ the distribution over S^2 obtained by sampling independently x from X and y from Y and outputting the pair (x, y), and similarly $X \circ Z$ denotes the distribution obtained by sampling independently x from X and z from Z and outputting the pair (x, z). Prove that

$$\Delta(X \circ Y, X \circ Z) = \Delta(Y, Z)$$

In other words, prove that concatenating an independent sample from some distribution X does not change the statistical distance of Y and Z.

Solution 1.3:

Question 2 (30 points): 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 2.1 (5 points): Suppose that $X_1, \dots, X_k$ are independent 0/1 valued random variables such that $\Pr[X_i = 1] = 1/2 + \epsilon$ for every $i \in \{1, \dots, k\}$. Prove that the standard deviation of $Y = \sum_{i=1}^{k} X_i$ is at most $\sqrt{k}$.

Solution 2.1:

Question 2.2 (10 points): Use Chebychev's inequality (feel free to look it up) to prove the following:

If $k > 100/\epsilon^2$, then there exists some function $f : \{0,1\}^k \to \{0,1\}$ such that if $X_1, \dots, X_k$ are as above then $\Pr[f(X_1, \dots, X_k) = 1] \geq 0.9$. However, if $Z_1, \dots, Z_k$ are independent unbiased coins (i.e. $\Pr[Z_i = 1] = 1/2$ for every i) then $\Pr[f(Z_1, \dots, Z_k) = 1] \leq 0.1$. Hence f can distinguish between unbiased coins and coins that are ϵ biased using $O(1/\epsilon^2)$ samples.

Solution 2.2:

Question 2.3 (15 points): (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 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 2.3:

Question 3: We will now consider the converse problem: showing a *lower bound* on the number of coin tosses needed. To do so, we will use the Hybrid Argument which is one of the central tools in cryptography.

Question 3.1 (hybrid argument - 10 points): Suppose that $X_1, \dots, X_k$ and $Y_1, \dots, Y_k$ are some distributions over $\{0,1\}^n$, prove that

$$\Delta(X_1 \circ \cdots \circ X_k, Y_1 \circ \cdots \circ Y_k) \leq \sum_{i=1}^{k} \Delta(X_1 \circ \cdots \circ X_i \circ Y_{i+1} \circ \cdots \circ Y_k, X_1 \circ \cdots \circ X_{i-1} \circ Y_i \circ \cdots \circ Y_k)$$

where $\circ$ denotes concatenation. **Hint:** The notation in this question might be a bit confusing (ask the TFs if it's unclear!) but it's not actually hard to solve using the triangle inequality.

Solution 3.1:

Question 3.2 (10 points): Let $X_1, \dots, X_k$ be binary random variables where $\Pr[X_i = 1] = 1/2$, and let $Y_1, \dots, Y_k$ be binary random variables where $\Pr[Y = 1] = 1/2 + \epsilon$. Next, let $X = X_1 \circ \cdots \circ X_k$ and $Y = Y_1 \circ \cdots \circ Y_k$. Using results proven in previous questions, prove that $\Delta(X, Y) \leq 10k\epsilon$ and show that if $k < 1/(100\epsilon)$ 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 3.2:

Question 3.3 (bonus problem: optional and more challenging - 10 points bonus): Let $X_1, \dots, X_k$ and $Y_1, \dots, Y_k$ be as defined in Question 3.2. In that question you proved that $\Delta(X_1 \circ \cdots \circ X_k, Y_1 \circ \cdots \circ Y_k) \leq O(k\epsilon)$. Prove the better bound

$$\Delta(X_1 \circ \cdots \circ X_k, Y_1 \circ \cdots \circ Y_k) \leq O(\sqrt{k}\epsilon)$$

Note that this implies that distinguishing between an unbiased coin and an ϵ-biased one requires a constant times $1/\epsilon^2$ samples. See footnote for hint.[1]

Solution 3.3:

And now for a little crypto

Question 4 (very important! points on perusall): Read the lecture notes for lecture 1: introduction, and annotate them on the Perusall website.

Question 5.1 (20 points): Prove that an encryption scheme (E, D) with messages of length ℓ and keys of length n is *perfectly secret* if and only if 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 5.1:

Question 5.2 (bonus but please do attempt it - 20 points): Let $\epsilon > 0$. Define an encryption scheme (E, D) with messages of length ℓ and keys of length n to be ϵ-secret if for every $x, x' \in \{0,1\}^\ell$, $\Delta(Y^x, Y^{x'}) < \epsilon$. Suppose that (E, D) is ϵ secret. Prove that for every adversary Eve, which we model as function

[1] **Hint:** One way to prove this is to use the notion of KL divergence which is another notion of distance between the distributions. You can show that the KL divergence of X and Y is $O(\epsilon^2)$ and use known relations between the KL divergence and the statistical distance (also known as total variation distance) and the Hellinger distance, and specifically 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 KL divergence, the Pinsker Inequality, and the Hellinger distance in Wikipedia as well in several other sources.

$Eve : \{0,1\}^m \to \{0,1\}$ where m is the length of the ciphertexts, and a pair of messages $x_0, x_1 \in \{0,1\}^\ell$, the probability that Eve wins in the game described below is less than $1/2 + 10\epsilon$.

The game is defined as follows:

1. We pick k at random in $\{0,1\}^n$.
2. We pick b at random in $\{0,1\}$.
3. We let $y = E_k(x_b)$.
4. We let $b' = Eve(y)$.
5. Eve *wins* iff $b' = b$.

Solution 5.2: