Homework 0

Homework is due by **11pm of Sep 9**. Send by email to both "regev" (under the cs.nyu.edu domain) and "des480" (under the nyu.edu domain) with subject line "CSCI-GA 3210 Homework 0" and name the attachment "YOUR NAME HERE HW0.tex/pdf". There is no need to print it. Start early!

1. Send a short email to Oded (`regev` at `cims`) with subject `CSCI-GA 3210 student` containing (1) a few words about yourself and your background (including your department, graduate program, how long in program), and (2) your comfort level with the following: mathematical proofs, elementary probability theory, big-O notation and analysis of algorithms. Please also mention any courses you've taken covering these topics.

2. *(Working with negligible functions.*[1]*)* Recall that a non-negative function $\nu : \mathbb{N} \to \mathbb{R}$ is *negligible* if it decreases faster than the inverse of any polynomial (otherwise, we say that ν is *non-negligible*). More precisely, $\nu(n) = o(n^{-c})$ for every fixed constant $c > 0$, or equivalently, $\lim_{n\to\infty} \nu(n) \cdot n^c = 0$.

   State whether each of the following functions is negligible or non-negligible, and give a brief justification. In the following, $\mathrm{negl}(n)$ denotes some arbitrary negligible function, and $\mathrm{poly}(n)$ denotes some arbitrary polynomial in n. (If you are not comfortable with these notion, read Section 4.2 of Lecture 2 in Peikert's notes)

   (a) (1 point) $\nu(n) = 1/2^{100 \log n}$.

   (b) (1 point) $\nu(n) = n^{-\log\log\log n}$. (Compare with the previous item for "reasonable" values of n.)

   (c) (1 point) $\nu(n) = \mathrm{poly}(n) \cdot \mathrm{negl}(n)$. (State whether ν is *always* negligible, or not necessarily.)

   (d) (1 point) $\nu(n) = (\mathrm{negl}(n))^{1/\mathrm{poly}(n)}$. (Same instructions as previous item.)

   (e) (1 point)
   $$\nu(n) = \begin{cases} 2^{-n} & \text{if } n \text{ is composite} \\ 100^{-100} & \text{if } n \text{ is prime.} \end{cases}$$

3. *(Statistical distance.)* Recall that given two distributions over a (finite) set Ω, their statistical distance (also known as variational or L_1 distance) is defined as
   $$\Delta(X,Y) := \frac{1}{2}\sum_{\omega\in\Omega} |X(\omega) - Y(\omega)| \ .$$

   (a) (3 points) Show that Δ defines a metric (see here for the definition).

   (b) (3 points) Show that the following is an equivalent definition:
   $$\Delta(X,Y) := \sup_{A\subseteq\Omega} |X(A) - Y(A)| \ ,$$
   where $X(A)$ denotes the probability of X to be in A, and similarly for $Y(A)$. Give an "operational" interpretation to this definition (i.e., in terms of an algorithm trying to distinguish X and Y).

   (c) (3 points) Let D_0 and D_1 be two distributions over the same support Ω. Suppose that we play the following game with an algorithm $\mathcal{A}$. First, we pick at random a bit $b \leftarrow \{0,1\}$ and then we pick $x \leftarrow D_b$ and we give x to $\mathcal{A}$. Finally, $\mathcal{A}$ returns a bit $\mathcal{A}(x)$. It wins if the bit returned is equal to b. Show that the highest success probability in this game is exactly $\frac{1}{2} + \frac{1}{2}\Delta(D_0, D_1)$.

[1] Based on a question from Peikert's class

4. (Pairwise independence)
   (a) (4 points) Assume that $r_1, \dots, r_t$ are independent uniform strings in $\{0,1\}^n$. Show that the collection of all $2^t - 1$ nontrivial XORs, $\{\bigoplus_{i \in S} r_i\}_{\emptyset \neq S \subset [t]}$ is pairwise independent, i.e., any two of them are jointly distributed like an independent uniform pair of strings in $\{0,1\}^n$.
   (b) (4 points) Let p be a prime number. Let Y and Z be uniform and independent random variables in $\mathbb{Z}_p$. For $k = 0, \dots, p-1$ define the random variables $X_k = Yk + Z \bmod p$. Show that $X_0, \dots, X_{p-1}$ are pairwise independent, i.e., that for any $k \neq j$, X_k and X_j are jointly distributed like an independent uniform pair of elements in $\mathbb{Z}_p$.

5. *(Large deviation bounds.)* Assume that $X_1, \dots, X_n$ are independent identically distributed (i.i.d.) random variables, each taking 1 with probability p and 0 with probability $1-p$. Recall that Chernoff's bound says that for all $\varepsilon > 0$,

$$\Pr\left[\left|\frac{1}{n}\sum_i X_i - p\right| > \varepsilon\right] \leq 2e^{-2n\varepsilon^2} .$$

   If you are rusty on Chernoff's bound, read about it, e.g., here or search Google; there are lots of forms of the bound, the above being the most convenient for our applications.
   (a) (2 points) How large should n be if we want the average of the X_i to be within $\pm\varepsilon$ of p with probability at least $1-\delta$? (asymptotic expression for n is enough)
   (b) (3 points) Imagine we used Chebyshev's bound instead of Chernoff's, and if you wish, assume for simplicity that $p = 1/2$. What bound on n would you get then? Do you see any advantage of Chebyshev's bound over Chernoff's?

6. *(Error-correcting codes (optional, no credit).)* This is a bit off topic, but will give you an idea of the kind of math we use in this course. It will also give you a glimpse to an immensely important topic that also dates back to Shannon's seminal work. These ideas are used in pretty much all digital communication protocols: cell phones, Internet, satellites, etc.
   (a) Assume we choose $2^{n/20}$ strings from the set $\{0,1\}^n$ uniformly at random. Show that with positive probability (in fact, high probability) the Hamming distance (i.e., number of different coordinates) between *any* two strings in the set is more than $n/4$. I need a hint! (ID 84542)
   (b) Show how Alice can communicate to Bob a message of k bits by sending only $n = 20k$ bits in such a way that Bob can recover the message even if an adversary flips up to $n/8$ bits of the communication. Would simply repeating the message 20 times be good enough?