CS59200-PRS: Pseudorandomness

Assignment 1 *Due: Oct 6*

Problem 1 (k-wise uniformity vs almost k-wise uniformity) [5 pts]. In this problem we will derive a general (but often not optimal) method to transfer every result about k-wise uniformity into one about ε-almost k-wise uniformity, by bounding their statistical distance.

In the questions below, let $p\colon \{\pm 1\}^n \to \mathbb{R}$ be a ε-almost k-wise uniform distribution for some $\varepsilon \le 1/2$. We will construct another distribution q such that q is k-wise uniform and $d_{\mathrm{TV}}(p, q)$ is small.

1. **[1 pt].** Let $S \subseteq [n]$ be a set of indices such that $1 \le |S| \le k$. Assume that $\widehat{p}(S) \ge 0$, show that for some properly chosen $\alpha \in [0, 2\varepsilon]$, the following function $p_{\alpha,S}\colon \{\pm 1\}^n \to \mathbb{R}$,

$$p_{\alpha,S}(x) = (1 - \alpha) \cdot p(x) + \alpha \cdot 2^{-n}(1 - \chi_S(x))$$

   is a distribution over $\{\pm 1\}^n$ and $\widehat{p_{\alpha,S}}(S) = 0$.

   Hint. Use the fact that the Fourier transformation is linear, and $|\widehat{p}(S)| \le 2^{1-n}\varepsilon$ as we showed in class.

2. **[1 pt].** Prove that for the function $p_{\alpha,S}$ that we chose above, it holds that

$$\left|\widehat{p_{\alpha,S}}(T)\right| \le |\widehat{p}(T)|$$

   for all $T \subseteq [n]$.

3. **[2 pt].** Use the claim in the above two questions to conclude that there exists a k-wise uniform distribution q such that $d_{\mathrm{TV}}(p, q) \le 2n^k\varepsilon$.

4. **[1 pt].** In the $2k$-wise independent Chebyshev's inequality, if instead we have $X_1, \ldots, X_n \in [0, 1]$ being ε-almost $2k$-wise independent, show that for $X = \frac{1}{n}\sum_{i=1}^{n} X_i$,

$$\Pr\left[|X - \mathbb{E}[X]| \ge \delta\right] \le \left(\frac{k^2}{n\delta^2}\right)^k + \frac{\varepsilon}{\delta^{2k}}.$$

   Hint. Try to directly follow the proof for $2k$-wise independent Chebyshev from class. In fact, in the regime where the inequality is meaningful ($\delta > 1/\sqrt{n}$), even for $2k$-wise uniform random variables, this is strictly better than the bound we would have obtained using the result from question 3.

Problem 2 (the longest distance in the world) [5 pts]. Suppose that Alice and Bob live on the unit sphere of an n-dimensional Euclidean space. Their locations are represented by unit vectors $a, b \in \mathbb{R}^n$ respectively. They want to know if they are really close, say $\|a-b\|_2 \le 0.1$, or really far away, say $\|a-b\|_2 \ge 1$, by using as little communication as possible.

Their plan is as follows: Alice first draws a uniformly random vector $x \in \{\pm 1\}^n$, computes $\langle a, x\rangle = \sum_{i=1}^n a_i x_i$, and send both x and $\langle a, x\rangle$ (ignoring the accuracy issue here) to Bob. Bob then computes $\langle b, x\rangle$ and checks how close it is to $\langle a, x\rangle$.

1. **[1 pt].** Show that $\mathop{\mathbb{E}}_x\left[(\langle a, x\rangle - \langle b, x\rangle)^2\right] = \|a-b\|_2^2$.

2. **[1 pt].** Show that $\mathbf{Var}\left[(\langle a, x\rangle - \langle b, x\rangle)^2\right] \le 2\,\|a-b\|_2^4$.

   Hint. Write the variance as expectation and expand it. Which summands survive under expectation?

3. **[1 pt].** Use Chebyshev's inequality to conclude that, by repeating the plan constantly many times (drawing independent x each time), Bob can correctly tell if $\|a-b\|_2 \le 0.1$ or $\|a-b\|_2 \ge 1$ with at least 99% success probability.

4. **[2 pt].** In the above plan, Alice has to send at least n bits to communicate the random x with Bob. Show how they can modify the plan so that the number of bits communicated is $O(\log n)$ while not changing the success probability.

   Hint. What type of pseudorandom $x \in \{\pm 1\}^n$ could make the claims in all previous questions still hold?