6.842 Randomness and Computation	April 14, 2022

Homework 5

Lecturer: Ronitt Rubinfeld — *Due Date: April 27, 2022*

1. **(Noise sensitivity vs. Fourier coefficients** Show that any $f : \{\pm 1\}^n \to \{\pm 1\}$ satisfies

$$NS_\epsilon(f) = \frac{1}{2} - \frac{1}{2}\sum_S (1-2\epsilon)^{|S|}\hat{f}(S)^2$$

2. **(Influence of variables on functions)** For $x = (x_1, \ldots, x_n) \in \{\pm 1\}^n$, let $x^{\oplus i}$ be x with the i-th bit flipped, that is,

$$x^{\oplus i} = (x_1, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n).$$

The *influence of the i-th variable on* $f : \{\pm 1\}^n \to \{\pm 1\}$ is

$$\mathrm{Inf}_i(f) = \Pr_x\left[f(x) \neq f\left(x^{\oplus i}\right)\right].$$

The *total influence* of f is

$$\mathrm{Inf}(f) = \sum_{i=1}^n \mathrm{Inf}_i(f).$$

A function $f : \{\pm 1\}^n \to \{\pm 1\}$ is *monotone* if for all $x, y \in \{\pm 1\}^n$ such that $x_i \leq y_i$ for each i, $f(x) \leq f(y)$.

   (a) Show that for any monotone function $f : \{\pm 1\}^n \to \{\pm 1\}$, the influence of the i^{th} variable is equal to the value of the Fourier coefficient of $\{i\}$, that is $\mathrm{inf}_i(f) = \hat{f}(\{i\})$.

   (b) Show that the majority function $f(x) = \mathrm{sign}(\sum_i x_i)$ maximizes the total influence among n-variable monotone functions mapping $\{\pm 1\}^n$ to $\{\pm 1\}$, for n odd.

3. Consider the sample complexity required to learn the class of monotone functions mapping $\{+1, -1\}^n$ to $\{+1, -1\}$ over the uniform distribution (without queries).

   (a) Show that

$$\sum_{|S| \geq Inf(f)/\epsilon} \hat{f}(S)^2 \leq C \cdot \epsilon$$

   where C is an absolute constant.

   (b) Show that the class of monotone functions can be learned to accuracy ϵ with $n^{\Theta(\sqrt{n}/\epsilon)} = 2^{\tilde{O}(\sqrt{n}/\epsilon)}$ samples under the uniform distribution (where the confidence parameter δ is some small constant).
   Hint: You can use the previous problem.

4. **(Almost k-wise independent random variables)** Let $\epsilon \in (0, 1)$ and $k \in [n]$. A random vector $(X_1, \ldots, X_n) \in \{\pm 1\}^n$ is said to be (ϵ, k)-wise independent if the restriction of $(X_1, \ldots, X_n)$ to any subset of k coordinates in $[n]$ is ϵ-close to the uniform distribution

on $\{0,1\}^k$.[1] Note that $(0,k)$-wise independence coincides with our usual notion of k-wise independence. The goal of this problem is to show that any (ϵ,k)-wise independent distribution is $O(\epsilon n^k)$- close to some k-wise independent distribution.

You may assume that if for all $S \subseteq [n]$ where $|S| \leq k$, we have $\mathbb{P}_{x\sim\mu}[\Pi_{i\in S} x_i = 0] = \frac{1}{2}$, then μ is k-wise independent over $\{\pm 1\}^n$.

(a) Given a random variable $X = \{X_1, X_2, \cdots, X_n\}$ such that X is (ϵ,k)-wise independent, and a subset S with $|S| \leq k$, construct a random variable Y that satisfies
   - $\mathbb{P}[\Pi_{i\in S} Y_i = 0] = \frac{1}{2}$
   - Y is at most ϵ-far from X

(b) Show that there exists a random variable Z which is k-wise independent, such that Z is at most $\epsilon \cdot n^k$ far from X.

[1] By saying that two distributions p and q with support $\{0,1\}^k$ are ϵ-close, we mean that the *statistical distance* $\Delta(p,q) := \frac{1}{2}\sum_{x\in\{\pm1\}^k} |p(x) - q(x)| = \max_{S\subseteq\{0,1\}^k} (p(S) - q(S))$ is at most ϵ.