Problem Set 2
Due: Noon, Monday September 26, email the pdf to toolkit2016homework@gmail.com

Homework policy: Exactly the same as last time.

Notational conventions: The notation \( [n] = \{1, 2, \ldots, n\} \) is very standard in theoretical computer science. Some people like to use boldface to denote random variables; you might like to do this too.

1. (a) Let \( X \) be a random variable which is 1 with probability \( p \) and 0 with probability \( 1 - p \). We “empirically estimate the mean of \( X \)”, by defining \( \overline{X} = \frac{1}{n}(X_1 + \cdots + X_n) \), where \( X_1, \ldots, X_n \) are independent copies of \( X \). We want to choose \( n = n(\varepsilon, \delta) \) sufficiently large so that “\( \overline{X} \) is \( \varepsilon \)-accurate with \( \delta \)-confidence”, meaning \( \Pr[|\overline{X} - p| > \varepsilon] \leq \delta \). Show that \( n = O\left(\frac{1}{\varepsilon^2} \log(1/\delta)\right) \) is sufficient (as \( \varepsilon, \delta \to 0^+ \)).

(b) Let \( Y \) be a random variable with a continuous probability distribution. We estimate the median of \( Y \) by defining \( m = \text{median}(Y_1, \ldots, Y_n) \), where \( Y_1, \ldots, Y_n \) are independent copies of \( Y \). We wish to have

\[
\frac{1}{2} - \varepsilon \leq \Pr[Y \leq m] \leq \frac{1}{2} + \varepsilon,
\]

except with probability at most \( \delta \). Again, show that \( n = O\left(\frac{1}{\varepsilon^2} \log(1/\delta)\right) \) is sufficient.

2. Prove the following “one-sided” version of Chebyshev’s Inequality: If \( X \) is a random variable with \( \mathbb{E}[X] = \mu \) and \( \text{stddev}[X] = \sigma > 0 \), then for every \( t > 0 \),

\[
\Pr[X \geq \mu + t\sigma] \leq \frac{1}{t^2 + 1}.
\]

(Hint: Mimic the proof of Chebyshev’s Inequality. “Standardize” \( X \), then prove and use the fact that \( \frac{(x+1/t)^2}{(t+1/t)^2} \geq 1_{\{x \geq t\}} \).)

3. It is a basic fact of linear algebra that if we have \( m \) orthogonal unit-length vectors \( \vec{u}_1, \ldots, \vec{u}_m \) in \( \mathbb{R}^n \), then \( m \leq n \). (Recall that

“orthogonal” \( \iff \angle(\vec{u}_i, \vec{u}_j) = \pi/2 = 90^\circ \iff \vec{u}_i \cdot \vec{u}_j = 0 \),

where \( \vec{u}_i \cdot \vec{u}_j = ||\vec{u}_i|| ||\vec{u}_j|| \cos(\angle(\vec{u}_i, \vec{u}_j)) \) is the dot-product.)

However, in this problem you will show the rather surprising fact that if we are willing for the unit vectors to only be \( \text{almost} \) orthogonal, we can have \( \text{exponentially} \) many such vectors.

Suppose we define random vectors \( \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m \in \mathbb{R}^n \) by choosing every coordinate of each of the vectors to be \( \pm 1 \) with probability \( 1/2 \) each. Then we put \( \vec{w}_i = \vec{v}_i / \sqrt{n} \) for each \( i \in [m] \) so as to get unit vectors (meaning \( ||\vec{w}_i|| = 1 \) for all \( i \in [m] \)).
(a) Suppose \( i \neq j \). Let \( \theta_{ij} = \angle(\mathbf{w}_i, \mathbf{w}_j) \). Show that
\[
\Pr[|\cos \theta_{ij}| \geq \delta] \leq \exp(-\Omega(\delta^2n)).
\]
Deduce
\[
\Pr[|\pi/2 - \theta_{ij}| \geq \delta] \leq \exp(-\Omega(\delta^2n)).
\]
(b) Show that even for some \( m = \exp(\Omega(\delta^2n)) \) we will have
\[
\Pr[|\pi/2 - \delta \leq \theta_{ij} \leq \pi/2 + \delta| \geq 0.99.
\]

4. Let \( X \) be a random variable that is always nonnegative. Assume also that \( X \) only takes on finitely many different values.\(^1\)

(a) Prove
\[
\mathbb{E}[X] = \int_0^\infty \Pr[X \geq t] \, dt.
\]
(b) Prove
\[
\mathbb{E}[X^2] = 2 \int_0^\infty t \Pr[X \geq t] \, dt.
\]

5. Let \( X \) be a nonnegative random variable.

(a) Prove that \( \Pr[X = 0] \leq \frac{\text{Var}[X]}{\mathbb{E}[X]^2} \).
(b) Prove that \( \Pr[X > 0] \geq \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^2} \).
(c) In the Erdős–Rényi random graph model, we start with \( n \) vertices, and then each of the \( \binom{n}{2} \) potential edges is included independently with probability \( p \) (where \( p \) may be a function of \( n \)). This is denoted \( G \sim G(n, p) \). Suppose that \( p = o(n^{-2/3}) \). Show that
\[
\Pr[G \text{ contains a 4-clique}] = o(1) \quad (n \to \infty).
\]
(Hint: Let \( X \) be the number of 4-cliques in \( G \). Compute \( \mathbb{E}[X] \) exactly as a function of \( n \) and \( p \); then use Markov.)
(d) On the other hand, show that if \( p = \omega(n^{-2/3}) \) then
\[
\Pr[G \text{ doesn’t contain a 4-clique}] = o(1) \quad (n \to \infty).
\]
(Hint: use part (a) or (b). You’ll have to carefully calculate the probability of 4-cliques occurring simultaneously on vertex sets \( A \) and \( B \) when \( |A \cap B| \geq 2 \).)

6. In this problem, let \( Z \sim N(0,1) \) denote a standard Gaussian random variable, with probability density function \( \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \). Before solving part (a) below, you might try differentiating \( \varphi(x) \), just for fun.

(a) Compute \( \int_0^\infty x \varphi(x) \, dx \). Deduce \( \mathbb{E}[|Z^2|] = \sqrt{2/\pi} \).

---

\(^1\)This isn’t really necessary, but it keeps things simple.
(b) Let $a_1, \ldots, a_n$ be real numbers satisfying $\sum_{i} a_i^2 = 1$ and write $\varepsilon = \max\{|a_i| : i \in [n]\}$.
Let $x_1, \ldots, x_n$ be i.i.d. random variables, each being $\pm 1$ with equal probability.
Let $S = \sum_i a_i x_i$. Show that

$$\left| \mathbb{E}[|S|] - \sqrt{2/\pi} \right| = o(1) \quad (\text{as } \varepsilon \to 0^+).$$

Here the $o(1)$ function may not depend on $n$ or the $a_i$’s; it must be a function of $\varepsilon$ only.

For full credit, you should achieve a bound of $O(\varepsilon \sqrt{\log(1/\varepsilon)})$.

Hint: for this problem you will need the Berry–Esseen Theorem, which will be covered on Wednesday:

**Berry–Esseen Theorem.** There is a universal constant $c$ (e.g., $c = .56$ suffices) such that the following holds: Let $X_1, \ldots, X_n$ be independent random variables with $\mathbb{E}[X_i] = 0$ for all $i \in [n]$, $\text{Var}[X_i] = \sigma_i^2$, $\sum_{i=1}^n \sigma_i^2 = 1$, and $\sum_{i=1}^n \mathbb{E}[|X_i|^3] = \beta$. Write $S = X_1 + \cdots + X_n$. Then for all $u \in \mathbb{R}$,

$$\left| \Pr[S \leq u] - \Pr[Z \leq u] \right| \leq c\beta.$$