The Probabilistic Method
Homework 2

Brett Hemenway

February 29, 2008

Problem 1. *Let X be a random variable taking integral nonnegative values, let $E(X^2)$ denote the expectation of its square, and let* $\mathrm{Var}(X)$ *denote its variance. Prove that*

$$\Pr(X = 0) \leq \frac{\mathrm{Var}(X)}{E(X^2)}.$$

Solution Notice that

$$\Pr(X = 0) = 1 - \Pr(X \geq 1),$$

and

$$\frac{\mathrm{Var}(X)}{E(X^2)} = \frac{E(X^2) - E(X)^2}{E(X^2)} = 1 - \frac{E(X)^2}{E(X^2)},$$

so to prove the result, it suffices to show

$$\Pr(X \geq 1) \geq \frac{E(X)^2}{E(X^2)}.$$

To see this, consider the conditional expectation

$$E(X|X \geq 1) = \sum_{n=0}^{\infty} n \frac{P(X = n \cap X \geq 1)}{P(X \geq 1)} = \frac{1}{P(X \geq 1)} \sum_{n=1}^{\infty} nP(X = n) = \frac{E(X)}{P(X \geq 1)}.$$

Similarly,

$$E(X^2|X \geq 1) = \sum_{n=0}^{\infty} n^2 \frac{P(X = n \cap X \geq 1)}{P(X \geq 1)} = \frac{E(X^2)}{P(X \geq 1)}.$$

Now,

$$E(X|X \geq 1)^2 \leq E(X^2|X \geq 1) \Rightarrow \frac{E(X)^2}{P(X \geq 1)^2} \leq \frac{E(X^2)}{P(X \geq 1)}$$
$$\Rightarrow \frac{E(X)^2}{E(X^2)} \leq P(X \geq 1).$$

Problem 2. *Show that there is a positive constant c such that the following holds. For any n reals $a_1, \ldots, a_n$ satisfying $\sum_{i=1}^n a_i^2 = 1$, if $(\epsilon_1, \ldots, \epsilon_n)$ is a $\{-1,1\}$-random vector obtained by choosing each ϵ_i randomly and independently with uniform distribution to be either -1 or 1, then*

$$\Pr\left(\left|\sum_{i=1}^n \epsilon_i a_i\right| \leq 1\right) \geq c.$$

Solution Define the random variable $X = \sum_{i=1}^n a_i \epsilon_i$. Then

$$E(X) = E\left(\sum_{i=1}^n a_i \epsilon_i\right) = \sum_{i=1}^n a_i E(\epsilon_i) = 0,$$

and because the ϵ_i are independent we have

$$E(X^2) = E\left(\sum_{i=1}^n a_i \epsilon_i\right) = \sum_{i=1}^n a_i^2 E(\epsilon_i^2) = \sum_{i=1}^n a_i^2 = 1,$$

Problem 3. *Let $v_1, v_2, \ldots, v_n$ be n vectors in $\mathbb{R}^n$, each of Euclidean norm at most 1, and let $u = \sum_{i=1}^n p_i v_i$, where $0 \leq p_i \leq 1$ for all i.*

(i) Prove that there are $\epsilon_i \in \{0,1\}$ such that

$$\left\|\sum_{i=1}^n \epsilon_i v_i - u\right\| \leq \sqrt{n}/2.$$

(ii) Prove that the above estimate is tight for all n.

(iii) Prove that even for $m > N$ and for $v_1, \ldots, v_m \in \mathbb{R}^n$, each of norm at most 1, and for $u = \sum_{i=1}^m p_i v_i$ with $0 \leq p_i \leq 1$, there are $\epsilon_i \in \{0,1\}$ such that

$$\left\|\sum_{i=1}^m \epsilon_i v_i - u\right\| \leq \sqrt{n}/2.$$

Solution

(i) If we choose the ϵ_i independently such that $P(\epsilon_i = 1) = p_i$, then we have

$$\begin{aligned}
E\left(\left\|\sum_{i=1}^{n}(\epsilon_i - p_i)v_i\right\|^2\right) &= E\left(\left\langle\sum_{i=1}^{n}(\epsilon_i - p_i)v_i, \sum_{i=1}^{n}(\epsilon_i - p_i)v_i\right\rangle\right) \\
&= E\left(\sum_{i,j=1}^{n}(\epsilon_i - p_i)(\epsilon_j - p_j)\langle v_i, v_j\rangle\right) \\
&= \sum_{i=1}^{n} E\left((\epsilon_i - p_i)^2\right)\|v_i\|^2 + \sum_{i\neq j} E\left((\epsilon_i)(\epsilon_j)\right)\langle v_i, v_j\rangle \\
&= \sum_{i=1}^{n}\left(E(\epsilon_i^2) - 2p_iE(\epsilon_i) + p_i^2\right) + \sum_{i\neq j} E(\epsilon_i - p_i)E(\epsilon_j - p_j)\langle v_i, v_j\rangle \\
&= \sum_{i=1}^{n}(p_i - p_i^2) + \sum_{i\neq j} 0 \\
&= \sum_{i=1}^{n} p_i(1 - p_i) \\
&\leq \frac{n}{4}.
\end{aligned}$$

In particular, there is a choice of the ϵ_i such that

$$\left\|\sum_{i=1}^{n}(\epsilon_i - p_i)v_i\right\|^2 \leq \frac{n}{4}.$$

Taking square roots then gives the result.

(ii) To show that this bound is tight, let $v_i = e_i$, where $\{e_1, \ldots, e_n\}$ is the standard basis for $\mathbb{R}^n$. Then if we take each $p_i = \frac{1}{2}$, we have

$$\left\|\sum_{i=1}^{n}(\epsilon_i - p_i)v_i\right\| = \frac{\sqrt{n}}{2}$$

for all choices of ϵ_i. This is just the geometric fact that the center of the n-cube is a distance $\sqrt{n}/2$ from any of its corners.

(iii)

Problem 4. *Prove that for every set X of at least $4k^2$ distinct residue classes modulo a prime p, there is an integer a such that the set $\{ax \mod p : x \in X\}$ intersects every interval of length at least p/k in $\{0, 1, \ldots, p-1\}$.*
Hint. *Pick random residues a and b and consider $\{ax + b \mod p : x \in X\}$.*

Solution For simplicity, we assume that intervals are allowed to wrap around, so there are exactly p intervals (instead of $p - p/k + 1$).
First, notice that if there exists an $a, b \in \mathbb{Z}/p\mathbb{Z}$ such that $aX + b$ intersects every interval of length p/k, then aX intersects every interval of length p/k as well, since adding b just permutes the intervals. So it is enough to show that there exists and a and a b such that $aX + b$ intersects all intervals of length p/k. Now, divide $\{0, \ldots, p-1\}$ into $2k$ intervals of $B_1, \ldots, B_{2k}$ each of length $p/2k$, specifically

$$B_i = \{(i-1)p/2k + 1, \ldots, ip/2k\}.$$

Now, if we can show that $aX + b$ intersects every interval B_i, then $aX + b$ must intersect every interval of length p/k since each interval of length p/k fully contains at least one B_i. To this end, we choose a, b uniformly from $\mathbb{Z}/p\mathbb{Z}$ and we define the random variables

$$Y_i = |\{aX + b\} \cap B_i|$$

Then by linearity of expectation, we have

$$E[Y_i] = \sum_{x \in X} P(ax + b \in B_i) = \sum_{x \in X} \frac{1}{2k} = 2k.$$

Now, notice that for $x_i, x_j \in X$ the events $ax_i + b \in B_\ell$ and $ax_j + b \in B_\ell$ are pairwise independent since $ax_j + b$ is uniformly distributed in $\mathbb{Z}/p\mathbb{Z}$ even if we condition on fixed value of $ax_i + b$. Thus

$$\begin{aligned}
\mathrm{Var}(Y_i) &= \sum_{x \in X} \mathrm{Var}(|\{ax+b\} \cap B_i|) + 2 \sum_{x \neq y \in X} \mathrm{Cov}(|\{ax+b\} \cap B_i|, |\{ay+b\} \cap B_i|) \\
&= \sum_{x \in X} \mathrm{Var}(|\{ax+b\} \cap B_i|) \quad \text{(pairwise independence)} \\
&= \sum_{x \in X} P(ax+b \in B_i) P(ax+b \notin B_i) \quad \text{(bernoulli r.v.)} \\
&= \sum_{x \in X} \frac{1}{2k}\left(1 - \frac{1}{2k}\right) \\
&= 4k^2 \left(\frac{1}{2k} - \frac{1}{4k^2}\right) \\
&= 2k - 1.
\end{aligned}$$

Thus by problem (1),

$$P(Y_i = 0) \leq \frac{\mathrm{Var}(Y_i)}{E(Y_i^2)} = \frac{\mathrm{Var}(Y_i)}{\mathrm{Var}(Y_i) + E(Y_i)^2} = \frac{2k-1}{2k-1+4k^2} < \frac{1}{2k}.$$

The union bound then gives

$$P(\text{ at least one } Y_i = 0) \leq \sum_{i=1}^{2k} P(Y_i = 0) < 1.$$

In particular, there is some choice of a, b such that $aX + b$ hits every interval B_i.

Problem 5. *Prove that every 3-uniform hypergraph with n vertices and $m \geq n/3$ edges contains an independent set of size at least $\frac{2n^{3/2}}{2\sqrt{3}\sqrt{m}}$.*