Homework 3
Due 2006/6/14

Amnon Ta-Shma
Dept. of Computer Science
Tel Aviv University

Topic today: Tail inequalities and ε-bias.

Notation: For $X_1, \ldots, X_n$, $X = \sum_{i=1}^{n} X_i$, $\mu_i = \mathbb{E}(X_i)$ and $\mu = \mathbb{E}(X)$.

Fact (Bellare, Rompel: Randomness efficient oblivious sampling): If $X_1, \ldots, X_n$ are *independent* and take values in $[0, 1]$, then: $\mathbb{E}[(X - \mu)^t] \leq c^t(t\mu + t^2)^{t/2}$ for some (small) constant c.

1. (Tail bound for k-wise independence)
   - Let $X_1, \ldots, X_n$ be $k-wise$ independent random variables taking values in $[0, 1]$. Prove that $\Pr[|X - \mu| \geq A] \leq c_k(\frac{k\mu+k^2}{A^2})^{k/2}$, for some constant c_k that depends only on k.
   - Let $X_1, \ldots, X_n$ be $k - wise$ independent random variables taking values in $[-1, 1]$. Prove that $\Pr[|X - \mu| \geq A] \leq B_{k,\mu,A} = c_k(\frac{k(n+\mu)+2k^2}{A^2})^{k/2}$, for some constant c_k that depends only on k.
   - Let $X_1, \ldots, X_n$ be $k - wise$ ε-biased random variables taking values in $\{-1, 1\}$, for some even k. Prove that $\Pr[|X - \mu| \geq A] \leq B_{k,\mu,A} + \frac{n^k\varepsilon}{A^k}$, for some constant c_k that depends only on k.
2. $X_1, \ldots, X_n$ are binary random variables with $Prob(X_i = -1) = Prob(X_i = 1) = 1/2$. Find a bound on $Prob(|\Sigma_{i=1}^{n} X_i| \geq t)$ when:
   - Nothing else is known.
   - The variables are pair-wise independent.
   - The variables are $2k$-wise ε–biased.
   - The variables are $2k$-wise independent.
   - The variables are independent.
3. This is your lucky day! n coins are laid covered on a table, $k < \frac{n}{3}$ of which are pure gold and the rest copper, and you are told to uncover and take $\frac{2n}{3}$ coins. You are allowed to use any algorithm, no matter what its complexity is. The only catch is that the adversary knows your algorithm and places the gold coins based on your algorithm.
   - Show that if you are deterministic, you can guarantee no gold coin.
   - Show that if you use n random coins you can almost certainly get $\Omega(k)$ gold coins.
   - Show that with $O(\log n)$ random coins, you can guarantee $\Omega(k)$ gold coins with probability at least $1 - O(\frac{1}{k})$.
   - Show that for $\varepsilon \geq \frac{1}{k}$, with $O(\log\log n + \log\frac{1}{\varepsilon})$ coins, you can guarantee $\Omega(k)$ gold coins with probability at least $1 - \varepsilon$.

- Formulate the problem as an extractor-like question. What are the differences between your problem and the extractor problem?

Hint: Use pair-wise independence, and pair–wise ε–bias independence.

4. Let Π be an ε-biased distribution over $\{0,1\}^n$ with support size S.

   - Show how to build a linear $[S, n, \frac{1}{2} - \varepsilon]_2$ code, where each non-zero code-word has weight in $(\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)$.
   - Show how to construct 2^n unit vectors in $\mathbb{R}^S$ such that the inner product between any two is at most $O(\varepsilon)$.
   - How many vectors do you get that way as a function of the dimension and the error.

5. (The [NN] construction)

   - Prove a converse to Question (4.1).
   - Let C be an asymptotically good $[n, k, 1/3]_2$ linear code. Use a disperser $G = (V = [m], W = [n], \varepsilon)$ of degree D to get a $[m, k, (1-\varepsilon)m]_{2^D}$ linear code. What is m? Can you use this code to get an ε–biased distribution?
   - Use concatenation with a Hadamrd code to get a $[m2^D, k, \frac{1}{2} + \delta]_2$ binary linear code. Can you use this code to get an ε–biased distribution ?
   - Choose the parameters for the disperser to be those of the best non-explicit disperser. What is the support size of the ε–biased distribution?

6. Let $M_{n\times k}$ be a generating matrix of a Boolean linear code.

   - Show that there exists a code-word of weight at least $\frac{n}{2}$.
   - Show how to find such a heavy codeword fast in parallel.

     hint: use epsilon-bias.