CMPT 881: Pseudorandomness Prof. Valentine Kabanets

Lecture 3: Power of Randomness

September 14, 2004 *Scribe: Ladan A. Mahabadi*

1 Markov Inequality

For a non-negative random variable X where $\mathrm{E}[X]$ represents the expectation of X, the following holds:

$$\Pr[X \geq t] \leq \frac{E[X]}{t} \tag{1}$$

The Markov Inequality holds for any *non-negative* random variable, and since it doesn't require restrictions such as independence for the random variables, it results in a laxer bound.

2 Chebyshev Inequality

The Chebyshev inequality, can be easily driven from the Markov inequality, indicates that for a random variable X,

$$\Pr[|X - E[X]| \geq \lambda] \leq \frac{Var[X]}{\lambda^2} \tag{2}$$

(where $\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}(X))^2] = \mathrm{E}[X^2] - E^2[X]$).

Note: If $\mathrm{Var}[X]$ is $\mathrm{o}(E^2[X])$, then $X = (1 \pm \epsilon)\ \mathrm{E}[X]$ with probability close to 1 (i.e $1 - o(1)$).

Before commencing to a more powerful bound, it is useful to recall a few properties:

- It is easy to show that expectation is linear:

  $E[X + Y] = E[X] + E[Y]$ for all random variables X and Y, and more generally:

  $E[\Sigma_i X_i] = \Sigma_i\ E[X_i]$, $\forall$ random variables X_i

- $\mathrm{Var}[cX] = c^2\mathrm{Var}[X]$

- For *pairwise independent* random variables $X_1 ... X_n$, $\mathrm{Var}\ [\Sigma_{i=1}^n\ x_i] = \Sigma_{i=1}^n\ \mathrm{Var}[X_i]$

3 Chernoff Inequality

Let $X_1, ..., X_n$ be *independent, identically distributed* random variables where $X_i \in \{0,1\}$.
Let $\mu_i = \mathrm{E}[X_i] = 1.\Pr[X_i = 1] + 0.\Pr[X_i = 0] = \Pr[X_i = 1], X = \Sigma_i X_i$, and
$\mu = \Sigma_i \mu_i = \Sigma_i E[X_i] = E[\Sigma_i X]$ by linearity of the expectation. Then, for $0 < \epsilon \leq 1$:

$$\Pr[|X - \mu| \geq \epsilon\mu] \leq 2e^{-\frac{\epsilon^2\mu}{4}} \tag{3}$$

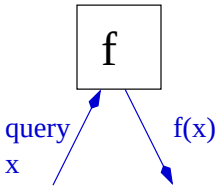
Exercises 1 *Show how to reduce the error probability in any BPP-style algorithm using Chernoff bounds.*

Chernoff 's inequality is the most restrictive of all in a sense that it requires independence amongst the random variables, and yet it exhibits a very strong result. Namely, if the chosen random variables are independent, the error probability of the deviation decreases exponentially. The following sampling problem is a vivid example of the application of the Chernoff bound. For detailed proof of these bounds see [1] and [2].

4 Sampling Problem

This problem which will be analyzed in what follows, is an example of a problem that in the chosen model of computation (namely in the Black box or Oracle Model) demonstrates an advantage for using randomness. In this model, as demonstrated below, this problem can not be solved deterministically, and yet it has an efficient randomized algorithm.

Given:
Oracle Access to some Boolean function $f : \{0,1\}^n \to \{0,1\}$. Oracle access can be thought of as "black-box access" where no extra information is given about f or its complexity. However, through querying the oracle, one can determine the value of the function at any x:



Compute:
$(1 \pm \epsilon)\mu$ where μ is the expected number of 1's of this function with respect to the uniform distribution:

$\mu = \frac{1}{2^n}\Sigma_x f(x) = \frac{|\{x_i | f(x_i)=1\}|}{2^n}$

Randomized Algorithm:
Choose t random points x_1, ... , $x_t \in \{0,1\}^n$ for $t = O(\frac{1}{\epsilon^2}\log\frac{1}{\delta})$, and output:

$\frac{|\{x_i | f(x_i)=1\}|}{t}$

Claim 2 *This gives us $(1 \pm \epsilon)\mu$ with probability $\geq 1 - \delta$*

Proof:

Define random variables $X_i = f(x_i)$ where $i = 1..t$, each x_i is a random point, and $\mu_i = E[X_i] = E[f(x_i)]$.
Let $X = \Sigma_i X_i$ = number ofx_i's where $f(x_i)$ evaluates to 1. We want to argue that $\frac{1}{t}\Sigma X_i$ will be close to $E[f]$. In other words, we'd like to show that $\frac{1}{t}\Sigma X_i = (1 \pm \epsilon)\mu$ with high probability.
Now, by Chernoff, it follows that:
$\Pr\left[|\frac{1}{t}\Sigma X_i - \mu| > \epsilon\right] = \Pr\left[|\Sigma X_i - \mu t| > \epsilon t\right] = \Pr\left[|\Sigma X_i - \mu t| > \frac{\epsilon}{\mu}t\mu\right] < 2e^{-(\frac{\epsilon}{\mu})^2\frac{1}{4}t\mu} = 2e^{\frac{-\epsilon^2 t}{4\mu}}$
Note that $\mu \leq 1$ since $\mu = \Pr$ [f evaluating to 1].
Thus, for t = $O(\frac{1}{\epsilon^2}\log\frac{1}{\delta})$,
$\Pr\left[|\frac{1}{t}\Sigma X_i - \mu| > \epsilon\right] < 2e^{\frac{-\epsilon^2 t}{4\mu}} < \delta$.

It is crucial to observe that the above calculation and the number of samples taken are *independent* from n. Furthermore, the running time is linear in n (the size of the input).

■

Proving a Lower Bound:

Claim 3 *No deterministic algorithm with less than $\frac{2^n}{2}$ queries to the black box can solve the sampling problem for the error term $\epsilon \leq \frac{1}{4}$*

Proof:

The following will provide a function f that will cause the failure of any deterministic algorithm.
Run the deterministic algorithm for the sampling problem for $< \frac{2^n}{n}$ steps, answering 0 to all the queries. This fixes some partial Boolean function f, where $f = 0$ $\forall$ query points. Extend the function in two possible ways:

extend to $f_0 \equiv 0$ (identically zero function)
extend to $f_1(x) = \begin{cases} 0 & \text{if } x \text{ was queried} \\ 1 & \text{otherwise} \end{cases}$

Note that the deterministic algorithm must answer with the same value a on both functions f_0, f_1. However, f_0 and f_1 differ from each other on at least 2^{n-1} values:

$E[f_0] = 0$
$E[f_1] \geq (1 \ . \ \frac{2^{n-1}}{2^n} + 0) = \frac{1}{2}$

Thus, $\frac{1}{2} \leq |E[f_0] - E[f_1]| = |\Big(E[f_0] - a\Big) + \Big(a - E[f_1]\Big)| \leq |E[f_0] - a| + |E[f_1] - a|$
Consequently, one of the averages is farther than $\frac{1}{4}$ from a.

■

BPP $\neq$ *P* **??**

No! The above analysis does not prove $BPP \neq P$, since it is restricted to the black-box model(i.e. the Oracle model). Generally in BPP and P languages, function f must be given explicitly as opposed implicitly through an oracle. Furthermore, the above randomized algorithm provides well bounded approximation as opposed to a definite solution for the decision problem.

5 Graph Reachability

Undirected *st*-Connectivity (USTCON)

Given:
An undirected graph $G = (V, E)$ on n vertices. Given two particular vertices s, $t \in V$, decide whether or not there is a path from s to t in G.
This decision problem can be solved deterministically, using breath-first-search in polynomial time, and linear space [1]. However, further reduction of the used space is desirable.

Open Problem 4 *Can USTCON be solved in LOGSPACE?*

By Savitch's Theorem [3], there is a $O(\log^2 n)$ space deterministic algorithm for USTCON (with running time at most exponential in its space , $2^{O(\log^2 n)} \approx n^{\log n}$ which is not in polynomial time).
A candidate randomized algorithm is a random walk on G:
Start at s, at each step go to a uniformly chosen neighbor of your current vertex, repeat for at most $2n^3$ steps. If t is visited then stop and accept; else, after $2n^3$ steps stop and reject.
We'll show in the next lecture that if t is reachable from s, then with probability $\geq \frac{1}{2}$, t will be visited during this random walk.

References

[1] Papadimitriou, **Computational Complexity**, Addison Wesley 1993.

[2] Rajeev Motwani, Prabhakar Raghavan, **Randomized Algorithms**. Cambridge University Press, August 1995.

[3] W. Savitch **Relationship between nondeterministic and deterministic tape complexities**, J. Comput. Syst. Sci., 1970.