Sublinear Algorithms January 19, 2012

Lecture 4

Lecturer: Sofya Raskhodnikova *Scribe(s): Dragos Nistor*

1 Probability Background

This lecture is mainly focused on three useful probability inequalities (Markov, Chebyshev, and Chernoff), and a few examples to go along. A little bit of proving lower bounds is included, too.

1.1 The Basics

Theorem 1 (Union bound)**.** *Given two events e_1 and e_2 over the same sample space,*

$$\Pr[e_1 \cup e_2] \leq Pr[e_1] + Pr[e_2].$$

Theorem 2 (Linearity of Expectation)**.** *For all random variables R_1 and R_2, the expectation*

$$E[R_1 + R_2] = E[R_1] + E[R_2].$$

1.2 Markov's Inequality

Theorem 3 (Markov's Inequality)**.** *If R is a non-negative random variable, then for all $x > 0$,*

$$Pr[R \geq x] \leq \frac{E[R]}{x}.$$

Proof.

$E[R] = E[R|R \geq x]\Pr[R \geq x] + E[R|R \leq x]\Pr[R \leq x]$, rewriting E[R]
$E[R|R \geq x]\Pr[R \geq x] + E[R|R \leq x]\Pr[R \leq x] \geq E[R|R \geq x]\Pr[R \geq x]$, a + b $\geq$ a when a, b ≥ 0
$E[R|R \geq x]\Pr[R \geq x] \geq x * \Pr[R \geq x]$, because of the conditional expectation
$\Pr[R \geq x] \leq \dfrac{E[R]}{x}$, rearranging both sides

□

Corollary 4. *If R is a non-negative random variable, then for all $c > 0$,*

$$\Pr[R \geq cE[R]] \leq \frac{1}{c}.$$

Proof. Substitute $x = cE[R]$ into the Markov's Inequality. Then

$$\Pr[R \geq cE[R]] \leq \frac{E[R]}{cE[R]} = \frac{1}{c}.$$

□

Exercise 4.1. *If we pull some random person aside, what is the probability their weight R is at least 200 given that the average weight (expected value of R) is 100? What is the probability they weigh at least 300?*

Solution: Using Corollary 2, $\Pr[R \geq 200] \leq \frac{1}{2}$ and $\Pr[R \geq 300] \leq \frac{1}{3}$
Markov's Inequality is useful if you have very little information about a variable.

Exercise 4.2. Chinese appetizer problem: *There are n people eating appetizers in a Chinese restaurant, and there is a spinner on the table holding n appetizers, one for each person. Someone antisocial decides to spin it. What is the expected amount of people that get their dishes back in front of them? What is the probability that all people get their dish back?*

Solution: Using the definition of expectation and the fact that either everyone gets their dish (with probability $\frac{1}{n}$) or no one gets their dish (with probability $\frac{n-1}{n}$, $E[C] = \Pr[\text{everyone gets their dish}] * n + \Pr[\text{no one gets their dish}] * 0 = \frac{1}{n} * n = 1$.
Using Markov's Inequality, we find that $\Pr[c \geq n] \leq \frac{1}{n}$ and we know that $\Pr[c = n] = \frac{1}{n}$, so the bound is tight in this case.

Exercise 4.3. Hat check problem: *n people checked their hat, and when they went to take their hats back the hatter went mad and started giving random people random hats. What is the expected amount of people who get their hat back? What is the probability that everyone gets their hat back?*

Solution: Let's look at each individual person and whether they get their hat back or not. So, if person i gets their hat, $c_i = 1$, otherwise $c_i = 0$. So, the count of people who get their hat back is $C = \sum_{i=1}^{n} c_i$, and $E[C] = E[\sum_{i=1}^{n} c_i] = \sum_{i=1}^{n} E[c_i]$ by the linearity of expectation. Since $E[c_i] = \Pr[\text{wrong hat}]*0 + Pr[\text{right hat}]*1$, and every person is just as likely to get any hat back, $\Pr[\text{right hat}] = \frac{1}{n}$ and $E[c_i] = \frac{1}{n}$, so $E[C] = 1$.
So, using Markov's Inequality, $\Pr[C \geq n] \leq \frac{1}{n}$, but we know that $\Pr[C = n] = \frac{1}{n!}$, so the bound is extremely loose in this case.

Both of these examples show that the bound from Markov's Inequality can be either extremely loose or extremely tight, and without further information about a variable we can't tell how tight the bound is.

Exercise 4.4. Why R must be positive: *Let's say R takes value 1 with probability .5, and value -1 with probability .5. What is the probability that R is at least 0?*

$E[R]$ is clearly 0, and with Markov's Inequality we would get that $\Pr[R \geq 0] \leq 0$, but we know that's not true. If we know that R has a lower bound, we can try to shift the variable instead. So, in the previous case, we could look at $R' = R + 1$. The following corollary uses this property.

Corollary 5. *Suppose we want to know* $\Pr[R \leq x]$*. If* $R \leq u$ *for some u, then for all* $x < w$*,*

$$\Pr[R \leq x] \leq \frac{u - E[R]}{u - x}.$$

Proof.

$$\Pr[R \leq x] = \Pr[u - R \geq u - x], \text{ so we take } R' = u - R, andx' = u - x$$
$$\Pr[R' \geq x'] \leq \frac{E[R']}{x'}, \text{ using Markov's Inequality}$$
$$\Pr[u - R \geq u - x] \leq \frac{E[u - R]}{u - x}, \text{ substituting back in}$$
$$\frac{E[u - R]}{u - x} = \frac{u - E[R]}{u - x}, \text{ u is a constant}$$
$$\Pr[R \leq x] \leq \frac{u - E[R]}{u - x}$$

□

Exercise 4.5 (Quiz scores)**.** *Take R to be the score of a random student, and suppose all scores* $\leq$ *100 and* $E[score] = 75$*. What is the probability that a student gets a score of at most 50?*

Solution: Using the inequality from Corollary 4, $\Pr[R \leq 50] \leq \frac{100-75}{100-50} = \frac{1}{2}$. Notice how now we can ask questions like "How many people could have obtained 50 or less on this exam?"

1.3 Chebyshev's Inequality

Recall that the variance of a variable R is $Var[R] = E[(R - E[R])^2] = E[R^2]0E[R]^2$, and the standard deviation of a variable $\sigma[R] = \sqrt{Var[R]}$.

Theorem 6 (Chebyshev's Inequality)**.** *For all $x > 0$ and all random variables R,*

$$\Pr[|R - E[R]| \geq x] \leq \frac{Var[R]}{x^2}.$$

Proof.

$$\Pr[|R - E[R]| \geq x] = \Pr[(R - E[R])^2 \geq x^2], \text{ squaring both sides}$$
$$\Pr[(R - E[R])^2 \geq x^2] \leq \frac{E[(R - E[R])^2]}{x^2}, \text{ by Markov's Inequality}$$
$$\frac{E[(R - E[R])^2]}{x^2} = \frac{Var[R]}{x^2}, \text{ by the definition of variance}$$
$$\Pr[|R - E[R]| \geq x] \leq \frac{Var[R]}{x^2}.$$

□

Corollary 7. $\Pr[|R - E[R]| \geq c * \sigma[R]] \leq \frac{1}{c^2}$.

Proof. We can take $x = c * \sigma[R]$ in Chebyshev's inequality, and we get

$$\begin{aligned} \Pr[|R - E[R]| \geq c * \sigma[R]] &\leq \frac{Var[R]}{(c * \sigma[R])^2} \\ &= \frac{Var[R]}{c^2 * Var[R]} \\ &= \frac{1}{c^2}. \end{aligned}$$

□

Exercise 4.6 (IQ)**.** *Take R to be the IQ of a random person you pull off the street. What is the probability that someone's IQ is at least 200 given that the average IQ is 100 and the standard deviation of IQ is 10?*

We know that the IQ is at least 0, so plugging in to Markov's Inequality, we get that $\Pr[R \geq 200] \leq \frac{100}{200} = \frac{1}{2}$. Plugging in to the formula in corollary 5, we get that

$$\begin{aligned} \Pr[R \geq 200] &= \Pr[R - 100 \geq 100] \\ &\leq \Pr[|R - 100| \geq 100 = 10 * \sigma[R]] \\ &\leq \frac{1}{10^2} \end{aligned}$$

Additionally, there exists a Chebyshev bound for one-sided analysis if you are interested, but we will mainly use the two-sided version in this course.

1.4 Chernoff bound

Theorem 8 (Chernoff bound)**.** *Let $T_1, T_2, ...T_n$ be any mutually independent random variables such that for all i, $0 \leq T_i \leq 1$. Taking $T = \sum_{i=1}^{n} T_i$, then for all $c > 1, \Pr[T \geq cE[T]] \leq e^{-\alpha * E[T]}$, where $\alpha = c * \ln(c) - 1$.*

Exercise 4.7. *Given $E[T] = 100$ and that the assumptions of all T_i as described above are true, what is the probability that T is at least 200?*

Solution: Using Markov's inequality, we get that $\Pr[T \geq 200] \leq \frac{1}{2}$. However, using the Chernoff bound, we get that $\Pr[T \geq 200] \leq e^{-.38*100} = e^{-38}$. This bound is much tighter than using Markov's inequality.

Exercise 4.8. *If 10 million people pick a 4-digit number at random and only one of those numbers is a win, what is the probability that at least 1100 people win?*

Solution: First, we consider an indicator variable for each person which is 0 if they lose, and 1 if they win. $\Pr[\text{win}] = \frac{1}{10000}$, since only one number will be a winner. Therefore, $E[\text{winners}] = \frac{10000000}{10000} = 1000$. So, using Markov's inequality, $\Pr[\text{winners} \geq 1100] \leq \frac{1000}{1100} = \frac{10}{11}$ which is pretty useless. However, using the Chernoff bound, we get that $\Pr[\text{winners} \geq 1100] \leq e^{-4.8} \leq 0.01$, so it is extremely useful even for small deviations.

1.5 Query complexity for lower bounds

We are interested in figuring out whether or not algorithms can be improved. While it is very hard to give lower bounds for deterministic algorithms, it is much easier to do so for sublinear algorithms, since there is a certain amount of queries we must perform on the input to be able to tell anything of value. The query complexity is denoted as q(P), the query complexity of problem P.

We will see that the following two statements are equivalent: 1) For every randomized algorithm A, there exists an input on which A goes as slow as possible. 2) There exists a probability distribution on an input x such that the probability that a deterministic algorithm fails on x is more than $\frac{1}{3}$.