Institut for Matematik og Datalogi
Syddansk Universitet

October 29, 2018
JFB

# DM551 – Algorithms and Probability – 2018
# Lecture 10

## Lecture, October 8

We finished section 13.4 and covered section 13.5 in Kleinberg and Tardos.

## Lecture, October 9

We will analyze the expected number of comparisons done by Randomized Quicksort, using section 7.4.2 of *Introduction to Algorithms*, 3rd edition, by Cormen, Leiserson, Rivest, and Stein (CLRS). We will cover section 13.9 (without proofs) and section 13.10 in Kleinberg and Tardos.

## Lecture, October 22

We will cover sections 5.1, 5.2, 5.3, 5.4.3, and 5.4.4 in CLRS.

## Problems to be discussed on October 24

1. Compare the Chernoff bounds to what you would get using Chebyshev's inequality on the same random variables. Just do the case where you want to find out the probability of being less than the expected value. Note that you need to find an upper bound on the variance.

2. Exercise 2 on page 782 of Kleinberg and Tardos. Then compute an upper bound on the probability of at least 1000 Democrats voting for candidate R.

3. Exercises 13 and 15 on pages 790–791 of Kleinberg and Tardos.

4. This problem considers robustness of a score (grade) given to students based on a multiple choice test. The problem set-up is based on the paper Frandsen and Schwartzbach, A singular choice for multiple choice,

   http://dl.acm.org/citation.cfm?id=1189164.

   Assume a multiple choice test consists of $n$ questions, each having 4 choices. For each question precisely one choice is correct. Students are allowed to make 0 or 1 check (cross) for each question. The score for a question is 1 if the student has checked the correct choice, $-1/3$ if the student has checked a wrong choice and 0 if no choices are checked. The score for the test is computed as the sum of the scores for all questions. The maximal score is therefore $n$. We assume that the test is used only to decide pass/fail, and the threshold for passing is a 50% score, i.e. a score $\geq n/2$. Define a challenged student to be a student that knows the answers to at most 40% of the questions. Let us assume that a challenged student leaves no questions unanswered. Then clearly he has nothing to lose by guessing the answers to the questions he does not know. So assume that he accordingly puts down checks uniformly at random choices (one per question). Define a multiple choice test to be good, if the probability that a challenged student passes is at most 5%. A teacher has to make a test, and naturally he wants it to be good. He suspects that if he has enough questions in the test, then it will be good. This is indeed correct as we shall see below. Consider a challenged student and assume that he guesses uniformly at random one of the 4 possible answers to each of the $m = 3n/5$ questions to which he does not know the answer. Define

$$X_i = \begin{cases} 1, \text{if the ith guess is correct} \\ 0, \text{otherwise} \end{cases}$$

Define $X = \sum_{i=1}^{m} X_i$.

(a) Determine $E[X]$.

(b) Show that the challenged student only passes if $X \geq \frac{3}{2}E[X]$.

(c) Using the Chernoff bound technique, determine a size $n$ of the test for which the challenged student only passes with a probability at most .05.