Exam: Randomized Algorithms & Probabilistic Methods

12.12.2015, 10:00-13:00

Name:

Stud.-Number:

<table>
<thead>
<tr>
<th>Exercise</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>max. Points</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>Points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I declare that the exam took place under regular conditions, and that I have read and understood the general remarks below.

Signature: .................................................................

General remarks and instructions:

a) No electronic equipment may be used.
b) The use of pencils is not allowed. Solutions written in pencil will not be graded.
c) Attempts to cheat will result in immediate disqualification and may result in legal consequences.
d) Provide only one solution to each question. Clearly cross out any invalid solution.
e) All answers should be clearly written and fully justified. Write down the important steps in complete sentences. Incomprehensible or unjustified solutions will be penalized. You may write your solution in either English or German.
f) You may solve the problems in any order. Concentrate on one question at a time, but make sure you organize your time.
g) If you finish early, report to an assistant and quietly leave the room. You may not leave the room within the last 20 minutes of the exam. Make sure you have written your name on each sheet of paper.
Exam: Randomized Algorithms and Probabilistic Methods

Exercise 1

(a) The negative binomial distribution models the number of successes in a sequence of independent and identically distributed bernoulli trials until $r$ unsuccessful attempts have been made for some $r \in \mathbb{N}$. Let $X$ be distributed according to a negative binomial distribution where the success probability is $p$ and the number of failures is $r$. Compute $E[X]$. Your solution should not contain any sum symbols.

(b) In part (a) the number of unsuccessful attempts, $r$, was fixed. In this part $r$ is also a random variable and follows the distribution $\Pr[r = i] = 2^{-i}$ for $i \in \mathbb{N}$. Let $Y$ be distributed according the negative binomial distribution where $r$ follows the aforementioned distribution. Compute $E[Y]$ and simplify the expression so it does not include infinite sums.

Hint: You may use that for $r = 1$ the solution is $p/(1-p)$ and that $\sum_{i=1}^{\infty} i \cdot 2^{-i} = 2$.

Exercise 2

Consider the random walk on the two dimensional integer grid $\mathbb{Z}^2$, where we start at the origin and in each step move in a direction (up, down, left, right) chosen uniformly at random. Denote with $(x, y)$ our position after $n$ steps and write $X = \max \{|x|, |y|\}$.

Show that for $n$ large enough

$$\Pr \left[ |X - E[X]| \geq \log n \sqrt{n} \right] \leq \frac{1}{n}.$$ 

Exercise 3

(a) Prove that for a random coloring the discrepancy for a single set exceeds $\sqrt{6n \ln(2m)}$ with probability at most $1/2m$.

(b) Prove that the discrepancy of every family with at most $m$ members does not exceed $\sqrt{6n \ln(2m)}$.

(please turn over)
Exercise 4

Consider a set of positive integers \( \{x_1, \ldots, x_k\} \subseteq \{1, \ldots, n\} \). Let \( f(n) \) denote the maximal \( k \) such that there exists a set \( \{x_1, \ldots, x_k\} \) with distinct sums\(^1\). One such set is for example \( \{2^i | i \leq \log_2 n\} \). In this exercise we will show that a distinct sum set cannot be much larger than that.

Let \( \{x_1, \ldots, x_k\} \) be a distinct sum set and let \( B_i \) be Bernoulli random variables which are 1 with probability \( \frac{1}{2} \). Let \( X = \sum_{i=1}^k B_i x_i \). Observe that \( \mathbb{E}[X] = \frac{1}{2} \sum_{i=1}^k x_i \) and \( \text{Var}[X] = \frac{1}{4} \sum_{i=1}^k x_i^2 \leq \frac{k n^2}{4} \).

(a) Show using Chebyshev’s inequality that for \( \lambda > 1 \)

\[
1 - \frac{1}{\lambda^2} \leq \Pr\left[ |X - \mathbb{E}[X]| \leq \lambda \sqrt{k/2} \right]
\]

(b) Show that

\[
\Pr\left[ |X - \mathbb{E}[X]| \leq \lambda \sqrt{k/2} \right] \leq 2^{-k(\lambda \sqrt{k} + 1)}.
\]

Hint: Show that the probability of \( X \) having a particular value is either \( 2^{-k} \) or 0.

(c) Show that \( f(n) \leq C \cdot \log_2 n \) for some constant \( C \).

Exercise 5

We call a set \( S \subseteq \mathbb{Z} \) a Sidon set if it does not contain four elements \( a, b, c, d \) such that \( a < b < c < d \) and \( a - b = c - d \). Let \( n \) be a positive integer and let \( p \in (0, 1) \) and denote by \( [n]_p \) the random subset of \( [n] \), where each element in \( [n] \) is chosen independently with probability \( p \). Additionally, suppose that \( pn^{2/3} \to 0 \) as \( n \to \infty \). Let

\[
A = \{(a, b, c, d) \in [n]^4 \mid a < b < c < d \text{ and } a - b = c - d\}.
\]

Now let \( N = |A| \). Prove that \( \mathbb{P}([n]_p \text{ is a Sidon set}) \leq e^{-\frac{4N}{p}} \) for \( n \) sufficiently large.

\( ^1 \)A set has distinct sums if you cannot find two distinct subsets \( A, B \subseteq \{x_1, \ldots, x_k\} \) such that \( \sum_{a \in A} a = \sum_{b \in B} b \).
Exercise 6

(10 points)

Algorithm 1.1 EnclosingCircle(\(P\))

Input: Point set \(P\)
Output: Smallest enclosing circle containing \(P\)

\[
\text{return } \text{Help}(P, \emptyset)
\]

Algorithm 1.2 Help(\(P, X\))

Input: Point sets \(P\) and \(X\)
Output: Smallest enclosing circle containing \(P\) with \(X\) on the boundary

\[
\begin{align*}
\text{if } |X| &= 3 \\
\text{return } C &= C(X)
\end{align*}
\]

\[
\begin{align*}
\text{else} \\
\text{denote by } p_1, \ldots, p_n &\text{ an ordering of the points from } P \text{ chosen uniformly at random} \\
\text{compute } C &= C(X \cup \{p_1, \ldots, p_{3-|X|}\}) \\
\text{for } i = 4 - |X| \text{ to } n &\text{ do} \\
\text{if } p_i &\text{ lies outside of } C \text{ then} \\
\text{let } C &= \text{Help}([p_1, \ldots, p_{i-1}], X \cup \{p_i\}) \\
\text{return } C
\end{align*}
\]

In the lecture you saw an algorithm for computing the smallest enclosing circle of a point set. Algorithm 1.1 is an alternative way of computing such smallest enclosing circle. In the algorithm \(C(X)\) is the smallest circle containing the point set \(X\) on the boundary. In this exercise you do not need to prove correctness of the algorithm but you should compute its running time.

Denote by \(t(n, x)\) the maximum expected running time of Help maximized over all point sets \(P\) of size \(n\) and sets \(X\) of size \(x\). Compute the running time of EnclosingCircle, i.e. compute the order of \(t(n, 0)\).

**Hint:** You can use that the smallest enclosing circle is defined by at most three points.

Exercise 7

(5+10+5 points)

We consider a deck of \(n\) cards and perform the following shuffle. At each step, we label card \(i\) of the deck with a random variable \(\xi_i\) where \(P(\xi_i = 0) = P(\xi_i = 1) = 1/2\). Cards labeled with a zero go to the top of the deck, and cards labeled with a one go to the bottom, always preserving their relative order.

**Example:** Suppose \(n = 6\) and the order of the cards is \(abcdef\). If we make a step with random sequence 010111, then the 0-labeled cards are \(abc\) and the order after the shuffle is \(acdef\).

We are interested in studying the behaviour of \(\tau_{\text{mix}} = \tau_{TV}(1/4)\). Solve the following:

a) We now store the labelings given to each card at the steps of the shuffle. Then after \(k\) steps of the shuffle, each card has a sequence of \(k\) zeros or ones. Let \(T\) be the first time that all the cards have different sequences of marks. Prove that \(P(T \geq 2 \log_2 n + 1) \leq 1/4\).

b) Prove that \(\tau_{\text{mix}} \leq 2 \log_2 n + 1\).

c) Prove that \(\tau_{\text{mix}} \geq (1 + o(1)) \log_2 n\).

**Hint 1:** You may use the following asymptotic: \(n! = 2^{(1+o(1))n} \log_2 n\).

**Hint 2:** You may want to use a) to prove b).

**Good Luck!**