Randomised Algorithms

UNIVERSITY OF PASSAU
Faculty of Computer Science and Mathematics

Winter term 2022/2023, Exercise Sheet No. 9

Hand-out:	Mon,	12. Dec.
Hand-in:	Sun,	18. Dec.

Dr. Duc-Cuong Dang
Prof. Dr. Dirk Sudholt

Exercise 1 [$4 + 2 = 6$ points]

Consider n independent tosses of a fair coin, that is, n i.i.d. indicator random variables $X_1, \dots, X_n$ with $X_i \in \{0, 1\}$ with $\text{Prob}(X_i = 1) = 1/2$. Let $X := \sum_{i=1}^{n} X_i$, eg. the number of heads (or tails).

(a) Calculate upper bounds for $\text{Prob}(X \geq (3/4)n)$ and $\text{Prob}(X \leq (1/4)n)$ using the simplified Chernoff bounds.

(b) Calculate upper bounds for $\text{Prob}(X \geq n/2 + 2\sqrt{n})$ first using the simplified Chernoff bound, then the additive version of the bound. Explain if there is any requirement on n for each bound.

(c) **(Bonus +2 points)** In this question, we focus on the opposite of concentration bounds for X, called anti-concentration. Remark that X is binomially distributed with parameters n and $1/2$, denoted $X \sim \text{Bin}(n, p)$, use the property of the binomial distribution show that the following bound holds for some constant $c \in (0, 1)$:

$$\text{Prob}\big(X \in [n/2 - \sqrt{n}/4, n/2 + \sqrt{n}/4]\big) \leq c.$$

Hint: The following inequalities $\binom{n}{k} \leq \binom{n}{\lfloor n/2 \rfloor} \leq 2^n \cdot \sqrt{2/n}$ which hold for all $0 \leq k \leq n$ can be useful. You can assume that n is sufficiently large, ie. $n \geq n_0$ for some constant n_0, to find a specific value for c.

Exercise 2 [6 points]

Consider a runoff mayor election (the last round) with two remaining candidates Alice and Bob of a large city with around few millions voters that all have made up their minds on who they will be voting for. Suppose a p-fraction of the population is supporting Alice and we want to estimate this $p \in (0, 1)$ by polling before the election day. Thus we pick k people uniformly at random from the population (with replacement) and ask who they will vote for (and we assume everyone is willing to answer and honest). If we obtain A answers voting for Alice, then the estimate of p is $\hat{p} = A/k$.
Use Chernoff bounds to derive the error of this estimate by an ε-deviation, that is an upper bound the probability $\text{Prob}(|\hat{p} - p| \geq \varepsilon)$, as a function of k and ε. Give the numerical value of this bound for $k = 5000$ and $\varepsilon = 5\%$.
Hint: Define X_i as the indicator random variable telling whether the i-th queried person is voting for Alice or not, then $A := \sum_{i=1}^{k} X_i$. You may want to multiply the both sides of the expression inside the probability statement with k to see the Chernoff bounds, and at one point the dependence on p can be removed using $p < 1$.

Exercise 3 [4 points]

Let X be the sum of n die rolls as seen on the previous Exercise 6.3. Use Hoeffding bounds to find an upper bound for $\text{Prob}(X \geq 4n)$. Compare the result with the bounds from Markov's inequality and Chebyshev's inequality from the previous sheet.

Exercise 4 [6 points]

Consider the n-dimensional hypercube graph $G = (V, E)$ and define the distance $d(u, v)$ between two vertices $u, v \in V$ as the number of edges on the shortest path between u and v. Note that this distance is also the Hamming distance between the labels of the nodes u and v.
Show the following statement: for any constant $\varepsilon > 0$, there exists a set $S \subset V$ of size $|S| = n$ such that for pairs of nodes from S, their distance lies in $[(1 - \varepsilon)n/2, (1 + \varepsilon)n/2]$:

$$\forall u, v \in S \colon (1 - \varepsilon) \cdot \frac{n}{2} \leq d(u, v) \leq (1 + \varepsilon) \cdot \frac{n}{2}.$$

Hint: The probabilistic method (see Chapter 6, slide 6) can be used to show that a uniform random choice of n nodes has the desired properties with positive probability. The Chernoff and union bounds can be useful. You can assume that n is large enough, i.e. $n \geq n_0$, where n_0 is a constant that can depend on ε.