Randomised Algorithms
Winter term 2022/2023, Exercise Sheet No. 9

Authors:
Ben Ayad, Mohamed Ayoub
Kamzon, Noureddine

December 18, 2022

Exercise 1.

(a) As $\mathbb{E}[X] = \frac{n}{2}$, we have:

$$\mathbb{P}[X \geq \frac{3}{4}n] = \mathbb{P}[X \geq (1 + \frac{1}{2})\frac{n}{2}] = \mathbb{P}[X \geq (1 + \frac{1}{2})\mathbb{E}[X])]$$

And similarily:

$$\mathbb{P}[X \leq \frac{1}{4}n] = \mathbb{P}[X \leq (1 - \frac{1}{2})\mathbb{E}[X])]$$

Using the Chernoff bounds, (Ineq 2 then 3, from Slides 4), we conclude that (for $\delta = 1/2 < 1$):

$$\mathbb{P}[X \leq \frac{1}{4}n] \leq e^{-\frac{n}{24}} \quad and \quad \mathbb{P}[X \geq \frac{3}{4}n] \leq e^{-\frac{n}{24}}$$

(b) Using the the simplified Chernoff bounds:

$$\mathbb{P}[X \geq n/2 + 2\sqrt{n}] = \mathbb{P}[X \geq \frac{n}{2}(1 + \frac{4}{\sqrt{n}})]$$

$$\leq e^{-\frac{n}{2}\left(\frac{4}{\sqrt{n}}\right)^2 \frac{1}{3}} \quad (Requires \ \frac{4}{\sqrt{n}} < 1 \implies n > 16)$$

$$= e^{-8/3}$$

Using the additive Chernoff (No extra requirements needed):

$$\mathbb{P}[X \geq n/2 + 2\sqrt{n}] \leq e^{-2\left(2\sqrt{n}\right)^2 \frac{1}{n}}$$

$$= e^{-8}$$

(c) We let $n_1 = \lfloor n/2 - \sqrt{n}/4 \rfloor$ and $n_2 = \lfloor n/2 + \sqrt{n}/4 \rfloor$, we have:

$$\mathbb{P}[X \in [n/2 - \sqrt{n}/4, n/2 + \sqrt{n}/4]] \le \mathbb{P}[X \in [n_1, n_2]]$$

$$= \sum_{k=n_1}^{n_2} \mathbb{P}[X = k]$$

$$= \sum_{k=n_1}^{n_2} \binom{n}{k} \frac{1}{2^n}$$

$$\le \frac{1}{2^n} \sum_{k=n_1}^{n_2} 2^n \frac{\sqrt{2}}{\sqrt{n}} \quad \textit{(Using the hint)}$$

$$= \frac{\sqrt{2}}{\sqrt{n}} \sum_{k=n_1}^{n_2} 1 = \frac{\sqrt{2}}{\sqrt{n}}(n_2 - n_2 + 1)$$

$$\le \frac{\sqrt{2}}{\sqrt{n}} \left(n/2 + \sqrt{n}/4 - (n/2 - \sqrt{n}/4 - 1)\right) \quad \textit{(using: } \lfloor x \rfloor - \lfloor y \rfloor \le x - (y-1))$$

$$= \frac{\sqrt{2}}{\sqrt{n}}(1 + \sqrt{n}/2)$$

$$= \frac{\sqrt{2}}{\sqrt{n}} + \frac{1}{\sqrt{2}}$$

For $n_0 = 24$, we have for $n \ge n_0$: $\frac{\sqrt{2}}{\sqrt{n}} + \frac{1}{\sqrt{2}} \le \frac{\sqrt{2}}{\sqrt{24}} + \frac{1}{\sqrt{2}} \approx 0.995$

Exercise 2.

Let X_i be the indicator random variable telling whether the i-th queried person is voting for Alice or not, we have $A = \sum_{i=1}^{k} X_i$. As X_i's are drawn randomly with replacement, they are considered i.i.d realisations of the true distribution, i.e., $X_i \sim Bern(p)$.

We also have: $\mathbb{E}[A] = kp$, hence, $\hat{p} = A/k$ is an unbiased estimate of p.

We have:

$$\mathbb{P}[|\hat{p} - p| > \epsilon] = \mathbb{P}[|k\hat{p} - kp| > k\epsilon]$$

$$= \mathbb{P}\left[|A - \mathbb{E}[A]| > \mathbb{E}[A]\frac{k\epsilon}{\mathbb{E}[A]}\right]$$

$$= \mathbb{P}\left[|A - \mathbb{E}[A]| > \mathbb{E}[A]\frac{\epsilon}{p}\right]$$

$$\le 2\exp\left(-\mathbb{E}[A]\left(\frac{\epsilon}{p}\right)^2 \frac{1}{3}\right) \quad \textit{(using Chernoff bounds)}$$

$$= 2\exp\left(-kp\left(\frac{\epsilon}{p}\right)^2 \frac{1}{3}\right)$$

$$= 2\exp\left(-\frac{\epsilon^2 k}{3p}\right) \le 2\exp\left(-\frac{\epsilon^2 k}{3}\right)$$

For $k = 5000$ and $\epsilon = 0.05$, we get and upperbound of: 0.031

Exercise 3.

Let the RV X represent the sum of the n rolls, we have established previously that: $\mathbb{E}[X] = \frac{7n}{2}$

Using Hoeffding's inequality leads to:

$$\mathbb{P}[X \geq 4n] = \mathbb{P}[X \geq E[X] + \frac{n}{2}]$$

$$\leq \exp\left(-\frac{2(n/2)^2}{\sum_{i=1}^{n}(6-1)^2}\right)$$

$$= e^{-n/50}$$

This tail that we obtained using Hoeffding's inequality decreases way faster than what we have obtained with Marokov and Chebychev's inequalities (respectively $\frac{7}{8}$ and $\frac{35}{3n}$).

Exercise 4.

Let's first establish a few results on $d(u,v)$ for two random vertices of the hypercube. $d(u,v)$ represents the number of bits v_i in v that are diffrent than u_i, hence, $d(u,v) \in \{0, \ldots, n\}$, and $\mathbb{P}[d(u,v) = k] = \frac{\binom{n}{k}}{2^n}$.(Why: How many ways can we pick k positions so we can flip their bits in v divided by the number of all configurations.)
By noticing that $\sum_{i=1}^{n} k\binom{n}{k} = n2^{n-1}$, we can conclude that:

$$\mathbb{E}[d(u,v)] = \sum_{i=1}^{n} k\frac{\binom{n}{k}}{2^n} = \frac{n}{2}$$

Let's prove the following result, for any two random vertices $u, v \in V$, $\epsilon > 0$:

$$\mathbb{P}\left[(1-\epsilon)\frac{n}{2} \leq d(u,v) \leq (1+\epsilon)\frac{n}{2}\right] \geq 1 - 2e^{-\frac{n\epsilon^2}{6}}$$

Proof:
Let's define the events E_1 and E_2 as follows:

$$E_1 = \{(1-\epsilon)\frac{n}{2} \leq d(u,v)\} = \{(1-\epsilon)\mathbb{E}[d(u,v)] \leq d(u,v)\}$$

And:

$$E_2 = \{(1+\epsilon)\frac{n}{2} \geq d(u,v)\} = \{(1+\epsilon)\mathbb{E}[d(u,v)] \geq d(u,v)\}$$

We are looking for to bound the probablity: $\mathbb{P}[E_1 \cap E_2]$

$$\mathbb{P}[E_1 \cap E_2] = 1 - \mathbb{P}[\overline{E_1} \cup \overline{E_2}]$$
$$\geq 1 - (\mathbb{P}[\overline{E_1}] + \mathbb{P}[\overline{E_2}])$$
$$\geq 1 - 2e^{-\mathbb{E}[d(u,v)]\epsilon^2/3}$$
$$\geq 1 - 2e^{-\frac{n\epsilon^2}{6}}$$

Here we used the union bound followed by Chernoff (Ineq 2. Slide 4 to bound $\mathbb{P}[\overline{E_2}]$, and Ineq 3 for $\mathbb{P}[\overline{E_1}]$), assumig $0 < \epsilon < 1$.

Let $V = \{v_1, \ldots, v_{2^n}\}$, **we uniformely random pick n vertices from V**. Now, let's define the RVs $\{X_i\}$, $\{Y_{i,j}\}$ and $V(v_1, \ldots, v_{2^n})$ as follows:

$$X_i = \begin{cases} 1 & \textit{if } v_i \textit{ was picked} \\ 0 & \textit{Otherwise.} \end{cases} \qquad Y_{i,j} = \begin{cases} 1 & \textit{if} \quad (1-\epsilon)\frac{n}{2} \leq d(v_i, v_j) \leq (1+\epsilon)\frac{n}{2} \\ 0 & \textit{Otherwise.} \end{cases}$$

And finally:

$$S(v_1, \ldots, v_{2^n}) = \sum_{i<j} X_i X_j Y_{i,j}$$

We need to proof that:

> **Proof:** *We have:*
>
> $$\mathbb{E}[S(\{X_i\}_i, \{Y_{i,j}\}_{i,j})] = \sum_{i<j} \mathbb{E}[X_i X_j Y_{i,j}]$$
>
> $$= \sum_{i<j} \mathbb{E}[X_i]\mathbb{E}[X_j]\mathbb{E}[Y_{i,j}] \quad \textit{(every rv is independent from the other)}$$
>
> $$= \mathbb{E}[X_1]^2 \sum_{i<j} \mathbb{E}[Y_{i,j}]$$
>
> $$\geq \mathbb{E}[X_1]^2 (1 - 2e^{-\frac{n\epsilon^2}{6}}) \sum_{i<j} 1$$
>
> $$= \mathbb{P}[\{Probb. \ to \ pick \ v_2\}]^2 (1 - 2e^{-\frac{n\epsilon^2}{6}}) \frac{2^{2n} - 2^n}{2}$$
>
> $$= \left(\frac{n}{2^n}\right)^2 (1 - 2e^{-\frac{n\epsilon^2}{6}}) \frac{2^{2n} - 2^n}{2} = \frac{n^2}{2}(1 - 2e^{-\frac{n\epsilon^2}{6}})(1 - \frac{1}{2^n})$$
>
> $$= \frac{n(n-1)}{2} \frac{n}{n-1}(1 - 2e^{-\frac{n\epsilon^2}{6}})(1 - \frac{1}{2^n})$$
>
> $$= \frac{n(n-1)}{2}(1 + \frac{1}{n-1})(1 - 2e^{-\frac{n\epsilon^2}{6}})(1 - \frac{1}{2^n})$$
>
> *For n big enough, $(1 + \frac{1}{n-1})(1 - 2e^{-\frac{n\epsilon^2}{6}})(1 - \frac{1}{2^n})$ is guarenteed to be ≥ 1*
>
> *Hence, $\forall \epsilon > 0$, and with n big enough, $\mathbb{E}[S(\{X_i\}_i, \{Y_{i,j}\}_{i,j})] \geq \frac{n(n-1)}{2}$, hence there exist a configuration of $(\{X_i\}_i, \{Y_{i,j}\}_{i,j})$ s.t $S(\{X_i\}_i, \{Y_{i,j}\}_{i,j}) = \frac{n(n-1)}{2}$, which can only happen if all the n vertices that were picked had $Y_{i,j} = 1$ for $v_i \neq v_j$.*