School of Computer Science and Technology
University of Science and Technology of China

Exercise Sheet 1 for
Design and Analysis of Algorithms
Autumn 2022
Solution

Exercise 1 **(45 points, graded by Yinhao Dong)**

Suppose you have an unfair dice such that if you roll it, the probability that each odd number (i.e., $1, 3, 5$) appears on the top is $\frac{1}{12}$, and the probability that each even number (i.e., $2, 4, 6$) appears on the top is $\frac{1}{4}$.
(1) Suppose that you roll this dice exactly once.

- What is the expected value of the number X on top of the dice? **(5 points)**
  Solution. $\mathrm{E}[X] = \sum_i i \cdot \Pr[X = i] = 1 \times \frac{1}{12} + 2 \times \frac{1}{4} + 3 \times \frac{1}{12} + 4 \times \frac{1}{4} + 5 \times \frac{1}{12} + 6 \times \frac{1}{4} = \frac{15}{4}$.

- What is the variance of X? **(5 points)**
  Solution 1. $\mathrm{var}[X] = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \left(-\frac{11}{4}\right)^2 \times \frac{1}{12} + \left(-\frac{7}{4}\right)^2 \times \frac{1}{4} + \left(-\frac{3}{4}\right)^2 \times \frac{1}{12} + \left(\frac{1}{4}\right)^2 \times \frac{1}{4} + \left(\frac{5}{4}\right)^2 \times \frac{1}{12} + \left(\frac{9}{4}\right)^2 \times \frac{1}{4} = \frac{137}{48}$.
  Solution 2. $\mathrm{E}\left[X^2\right] = \sum_i i^2 \cdot \Pr[X = i] = 1^2 \times \frac{1}{12} + 2^2 \times \frac{1}{4} + 3^2 \times \frac{1}{12} + 4^2 \times \frac{1}{4} + 5^2 \times \frac{1}{12} + 6^2 \times \frac{1}{4} = \frac{203}{12}$, so $\mathrm{var}[X] = \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2 = \frac{203}{12} - \left(\frac{15}{4}\right)^2 = \frac{137}{48}$.

(2) Suppose that you roll this dice n times. Let Y denote the sum of the numbers that are on the top of the dice throughout all the n rolls.

- What is the expected value of Y? **(5 points)**
  Solution. Let X_j denote the number on top of the dice in the j-th roll, so $Y = \sum_{j=1}^n X_j$. According to (1), $\mathrm{E}[X_j] = \frac{15}{4}$, $\mathrm{var}[X_j] = \frac{137}{48}$. Therefore, $\mathrm{E}[Y] = \sum_{j=1}^n \mathrm{E}[X_j] = \frac{15n}{4}$.

- What is the variance of Y? **(5 points)**
  Solution. Because $X_1, X_2, \cdots, X_n$ are independent, $\mathrm{var}[Y] = \sum_{j=1}^n \mathrm{var}[X_j] = \frac{137n}{48}$.

- Give an upper bound of $\Pr[Y > 4n]$ by using Markov's inequality. **(5 points)**
  Solution. $\Pr[Y > 4n] \le \frac{\mathrm{E}[Y]}{4n} = \frac{15}{16}$.

- Give an upper bound of $\Pr[Y > 4n]$ by using Chebyshev's inequality. **(5 points)**
  Solution. $\Pr[Y > 4n] \le \Pr\left[\left|Y - \frac{15n}{4}\right| \ge \frac{n}{4}\right] = \Pr\left[|Y - \mathrm{E}[Y]| \ge \frac{n}{4}\right] \le \frac{\mathrm{var}[Y]}{(\frac{n}{4})^2} = \frac{137}{3n}$.

- Give an upper bound of $\Pr[Y > 4n]$ by using Chernoff Bound. **(15 points)**
  Solution. Let $Z_j = \frac{X_j}{6}$ so that each $Z_j \in [0, 1]$. Define $Z = \sum_{j=1}^n Z_j$. By linearity of expectation, $\mathrm{E}[Z] = \frac{1}{6} \sum_{j=1}^n \mathrm{E}[X_j] = \frac{5n}{8}$. Therefore, by using Chernoff Bound, $\Pr[Y > 4n] = \Pr\left[\frac{Y}{6} > \frac{4n}{6}\right] = \Pr\left[Z > \frac{2n}{3}\right] \le \Pr\left[\left|Z - \frac{5n}{8}\right| \ge \frac{1}{15} \cdot \frac{5n}{8}\right] = \Pr\left[|Z - \mathrm{E}[Z]| \ge \frac{1}{15}\mathrm{E}[Z]\right] \le 2\exp\left(-\frac{n}{1080}\right)$.

Exercise 2 *(An application of Markov's inequality: turning a Las Vegas algorithm to a Monte Carlo algorithm)* **(35 points, graded by Yudong Zhang)**

Let $\mathcal{A}$ be a randomized algorithm for some decision problem P (e.g., to decide if a graph is connected or not). Suppose that for any given instance of P of size n, $\mathcal{A}$ runs in *expected* $T(n)$ time and always outputs the correct answer.
Use $\mathcal{A}$ to give a new randomized algorithm NewAlg for the problem P such that NewAlg *always* runs in $100T(n)$ time and

- if the input is a 'Yes' instance[1], then it will be accepted with probability at least $\frac{5}{6}$;
- if the input is a 'No' instance, then it will always be rejected.

Describe your algorithm and justify its correctness and running time.

Hint: You may use Markov's inequality.

Solution. The algorithm NewAlg is given as follows **(15 points)**:

Algorithm 1: NewAlg

Apply $\mathcal{A}$ until $100T(n)$ time.
if $\mathcal{A}$ *terminates before* $100T(n)$ *time* **then**
    Output the answer output by $\mathcal{A}$.
else
    Terminate and reject.
end

Correctness **(15 points)**: Let $T_{\mathcal{A}}$ denote the running time of $\mathcal{A}$. If the input is a 'Yes' instance, by using Markov's inequality, Pr[A 'Yes' instance is accepted] = Pr[$\mathcal{A}$ terminates before $100T(n)$ time] $= \Pr[T_{\mathcal{A}} \leq 100T(n)] = \Pr[T_{\mathcal{A}} \leq 100\mathrm{E}[T_{\mathcal{A}}]] = 1 - \Pr[T_{\mathcal{A}} > 100\mathrm{E}[T_{\mathcal{A}}]] \geq 1 - \frac{1}{100} = \frac{99}{100} > \frac{5}{6}$. If the input is a 'No' instance, it will always be rejected whether $\mathcal{A}$ terminates before $100T(n)$ time.

Running time **(5 points)**: Obviously, the running time of NewAlg is at most $100T(n)$.

Note: It is correct as long as we apply $\mathcal{A}$ until $\alpha T(n)$ time ($6 \leq \alpha \leq 100$) to get the new randomized algorithm. The justification of its correctness and running time is similar to the above.

Exercise 3 **(20 points, graded by Di Wu)**

Take R to be the IQ of a random person you pull off the street. What is the probability that some's IQ is at least 200 given that the average IQ is 100 and the variance of R is 100?

Solution. By using Markov's inequality, $\Pr[R \geq 200] \leq \frac{\mathrm{E}[R]}{200} = \frac{100}{200} = \frac{1}{2}$.
By using Chebyshev's inequality, $\Pr[R \geq 200] \leq \Pr[|R - 100| \geq 100] = \Pr[|R - \mathrm{E}[R]| \geq 100] \leq \frac{\mathrm{var}[R]}{100^2} = \frac{1}{100}$.

---

[1] An instance is a 'Yes'-instance if the correct answer to the Problem is 'Yes'. Otherwise (the correct answer is 'No'), it is a 'No'-instance.