Homework #2

Due: 2025-4-3 23:59 | 7 Questions, 120 Pts

Name: 方嘉聪 ID: 2200017849

Note: The total points of this homework is $15 + 10 + 10 + 10 + 15 + 10 + 50 = 120$, with 50 points being the *Challenge Problem*.

Question 1 (15') (Primality Test with Square Root Oracle). Suppose you are given a black-box algorithm (known as "oracle") $\mathcal{S}(b, n)$ for computing square roots of b modulo n. In other words, the algorithm may return *one* a such that $a^2 \equiv b \pmod{n}$ in each invocation, or output $\perp$ if there is no root. Using this algorithm as a black box, design an RP algorithm (i.e., algorithm with one-side error) for compositeness, and analyze its error bound.

[Hint: You do not know the behavior of algorithm $\mathcal{S}$. For example, the solution to $a^2 \equiv 1 \pmod{12}$ is $a \equiv 1, 5, 7, 11 \pmod{12}$, but $\mathcal{S}(1, 12)$ may return 1 all the time. You can never expect $\mathcal{S}$ to return a root randomly; its output can be adversarial. Therefore, in this question you must use some randomness in your algorithm.] ◀

Question 2 (10') (Ramsey Number).

Define the non-diagonal Ramsey number $R(k,t)$ as the minimum number n such that, in every 2-coloring of K_n, there exists either a red k-clique, or a blue t-clique.

Prove that, if there is a real $p, 0 \le p \le 1$ such that

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t} (1-p)^{\binom{t}{2}} < 1,$$

then $R(k,t) > n$. Using this, show that

$$R(4,t) = \Omega\left(\frac{t^{3/2}}{(\ln t)^{3/2}}\right).$$

[Hint: $\binom{n}{t} \le \frac{n^t}{t!}, t! \sim \sqrt{2\pi t}(\frac{t}{e})^t, (1-p)^x \le e^{-px}$. Note that the last inequality here will be frequently used in our lecture.] ◀

Question 3 (10') (Conditional Expectation).

Let $v_1, \cdots, v_n \in \mathbb{R}^n$, all with $\|v_i\|_2 = 1$.

a. (5') Prove that there exist $\epsilon_1, \ldots, \epsilon_n = \pm 1$ such that

$$\|\epsilon_1 v_1 + \cdots + \epsilon_n v_n\|_2 \leq \sqrt{n}.$$

b. (5') Please provide a polynomial-time *deterministic* algorithm for finding an assignment of $\epsilon_1, \ldots, \epsilon_n$ in ±1 such that $\|\epsilon_1 v_1 + \cdots + \epsilon_n v_n\|_2 \leq \sqrt{n}$, and analyze its time complexity.
[Hint: To derandomize the algorithm, consider how you can leverage the result from a. *in an inductive manner to fix the values of $\epsilon_1, \cdots, \epsilon_n$ one by one.]*

◀

Question 4 (10') (*d*-wise Independent Random Variables). Show how to construct 2^m m-bits random variables which are d-wise independent, using only dm random bits.

[Hint: Consider the finite field $\mathbb{F}_{2^m} \cong \mathbb{F}_2^m$.] ◀

Question 5 (15') (Coupon Collector). Farmer John is deploying sprinklers in a square field with side length 1. Each sprinkler has a cover radius of r ($r < 1$), meaning that all crops within distance r of the sprinkler can be watered. The sprinklers are dropped independently and uniformly in the square field at random. Farmer John wants to know how many sprinklers are needed to ensure that the whole area of the field is covered with probability at least $(1 - \epsilon)$. Let m be the minimum number of sprinklers required. Prove that,

$$m = \mathcal{O}\left(\frac{1}{r^2} \log \frac{1}{\epsilon r^2}\right).$$

[Hint: To show the required bound you only need a weak coupon collector bound; the result can be made stronger. In your solution, prove the coupon collector bound you use.] ◀

Question 6 (10') (Median-of-Mean Trick). Let $X_i, i = 1, 2, \cdots, 2s+1, s > 0$ be i.i.d random variables with $E[X_i] = \mu$. Assume each X_i is concentrated well, i.e., $\Pr[|X_i - \mu| \geq \epsilon\mu] \leq \frac{1}{4}$. Prove

$$\Pr[|\mathrm{MED}(X_1, \cdots, X_{2s+1}) - \mu| \geq \epsilon\mu] \leq \exp\left(-\frac{s}{4}\right),$$

where $\mathrm{MED}(X_1, \cdots, X_{2s+1})$ is the median value of $X_1, \cdots, X_{2s+1}$.

[Hint: If the medium is biased, then more than half are biased.] ◀

Question 7 (50') (A Combinatorial Proof of Chernoff Bound). In the lecture, you learned how to prove Chernoff Bound by Generating Function. In this problem, you will learn another way to prove Chernoff Bound. Our goal is to prove the following theorem:

Theorem 1. *Suppose $X_1, \dots, X_n$ are i.i.d random variables from $\{-1, 1\}$, each w.p. $\frac{1}{2}$. Then*

$$\Pr\left[\sum_{i=1}^{n} X_i \geq k\sqrt{n}\right] \leq e^{-\Theta(k^2)}$$

You will achieve this goal step by step.

a. (10') Prove the following lemma:

Lemma 2. *(Poor Man Chernoff Bound) Suppose $X_1, \dots, X_n$ are i.i.d random variables from $\{-1, 1\}$ each w.p. $\frac{1}{2}$. Then*

$$\Pr\left[\sum_{i=1}^{n} X_i \geq 2k\sqrt{n}\right] \leq 2^{-k}$$

Hint: First, you can use Chebeyshev's Inequality to prove the following fact:

Fact 3. *Suppose $X_1, \dots, X_n$ are i.i.d random variables from $\{-1, 1\}$ each w.p. $\frac{1}{2}$. Then*

$$\Pr\left[\sum_{i=1}^{n} X_i \geq 2\sqrt{n}\right] \leq \frac{1}{4}$$

Then, consider $S_i = \sum_{j=1}^{i} X_j$, let p be the first point where $S_p \geq 2\sqrt{n}$, then $\Pr[S_n \geq 2k\sqrt{n}] = \Pr[p \text{ exists}] \cdot \Pr[S_n - S_p \geq 2(k-1)\sqrt{n} \mid p \text{ exists}]$. If you can prove $\Pr[p \text{ exists}] \leq \frac{1}{2}$, you can prove the lemma by induction.

b. (10') Prove the following lemma:

Lemma 4. *(Chernoff Bound for Geometric Distribution)*
Suppose $X_1, \dots, X_n$ are i.i.d random variables, that $X_i \geq 0, \Pr[X_i \geq j] \leq p^j, \forall j = 1, 2, \dots$ for a $p < \frac{1}{4}$. Then

$$\Pr\left[\sum_{i=1}^{n} X_i \geq 2n\right] \leq (4p)^{-n}$$

.

Hint: If we can prove $\Pr[\sum_{i=1}^{n} \lfloor X_i \rfloor \geq n] \leq (4p)^{-n}$, then it's easy to see the lemma will hold. Suppose $\sum_{i=1}^{n} \lfloor X_i \rfloor \geq n$, then there exist $Y_1, \dots, Y_n$, that $\forall 1 \leq i \leq n, X_i \geq Y_i$ and $\sum_{i=1}^{n} Y_i = n$. Fix the sequence $Y_1, \dots, Y_n$, calculate the probability of sequence X_i satisfies $\forall i, X_i \geq Y_i$. Then use union bound for all the posible sequence of Y.

c. (10') Prove the following lemma:

Lemma 5. *(Lowerbound for Chernoff Bound)*
Suppose $X_1, \dots, X_n$ are i.i.d random variables from $\{-1, 1\}$ each w.p. $\frac{1}{2}$. Then

$$\Pr\left[\sum_{i=1}^{n} X_i \geq \frac{k}{2}\sqrt{n}\right] \geq (\frac{1}{4})^{k^2}$$

.

You can use the fact:

Fact 6. *Suppose $X_1, \ldots, X_n$ are i.i.d random variables from $\{-1, 1\}$ each w.p. $\frac{1}{2}$. Then*

$$\Pr\left[\sum_{i=1}^{n} X_i \geq \frac{1}{2}\sqrt{n}\right] \geq \frac{1}{4}$$

.

Hint: Divide $X_1, \ldots, X_n$ into $m = k^2$ groups, use the Fact 6 on each group.

d. (20') Prove Theorem 1. ◀