Practice Problems on Color Coding and Concentration Inequalities

Palash Dey
Indian Institute of Technology, Kharagpur

1. Design a randomized algorithm for computing if a given directed graph contains a cycle of length at least k. Your algorithm should run in time $\mathcal{O}(c^k \text{poly}(n))$ where c is some constant.

2. In the Triangle Packing problem, we are given an undirected graph G and a positive integer k, and the objective is to test whether G has k-vertex disjoint triangles. Using color coding show that the problem admits an algorithm with running time $2^{\mathcal{O}(k)} n^{\mathcal{O}(1)}$.

3. Prove that the condition in the Markov's inequality that the random variable under consideration must be non-negative is necessary.

4. Let $A_i, i \in [n]$ be n objects each having two attributes A_i^x and A_i^y. The attribute y is 0 for every A_i. Suppose we have a deterministic quick sort algorithm that can sort $A_i, i \in [n]$ on the attribute x or on the attribute y. Can you use this deterministic quick sort algorithm to design a randomized algorithm to sort $A_i, i \in [n]$ on the attribute x which makes an expected $\mathcal{O}(n \log n)$ comparisons? Please prove that your algorithm indeed makes $\mathcal{O}(n \log n)$ comparisons on expectation.

5. Let $\mathcal{X}_i, i \in [n]$ be n pairwise independent random variables each taking values in $\{0, 1\}$ with expectation μ and $\mathcal{S} = \sum_{i=1}^{n} \mathcal{X}_i$. Then for any positive real number δ we have the following.

$$\Pr\left[\mathcal{S} \leqslant (1-\delta)\mu\right] \leqslant \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$$

6. Show that the expected number of balls one needs to through randomly into m bins to have every bin at least one ball is $\mathcal{O}(m \log m)$.

7. Give an example of a random variable whose expectation exists but variance does not exist.

8. Prove the weak law of large numbers using Chebyshev inequality. The weak law of large number states that, for random variables $X_i, i \in \mathbb{N}$ which are distributes independently and identically with mean μ and variance σ^2, we have the following for any constant $\varepsilon > 0$

$$\lim_{n \to \infty} \Pr\left[\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right] = 0$$

9. Let $\mathcal{X}_i, i \in [n]$ be n independent random variables each taking values in $\{0, 2\}$ with expectation μ and $\mathcal{S} = \sum_{i=1}^{n} \mathcal{X}_i$. Use standard Chernoff bound proved in class to upper bound the probability that $\mathcal{S}$ takes value more than $(1+\delta)\mu$.

10. Let $\mathcal{X}$ be a random variable with expectation μ and variance σ^2. Then for any $t \in \mathbb{R}_{\geqslant 0}$, prove the following.

$$\Pr\left[\mathcal{X} - \mu \geqslant t\sigma\right] \leqslant \frac{1}{1+t^2} \text{ and } \Pr\left[|\mathcal{X} - \mu| \geqslant t\sigma\right] \leqslant \frac{2}{1+t^2}$$

11. Let $\mathcal{X}$ be a non-negative integer valued random variable with positive expectation. Then prove the following.

$$\Pr\left[\mathcal{X} = 0\right] \leqslant \frac{\mathbb{E}[\mathcal{X}^2] - \mathbb{E}[\mathcal{X}]^2}{\mathbb{E}[\mathcal{X}]^2} \text{ and } \frac{\mathbb{E}[\mathcal{X}]^2}{\mathbb{E}[\mathcal{X}^2]} \leqslant \Pr[\mathcal{X} \neq 0] \leqslant \mathbb{E}[\mathcal{X}]$$