Probabilistic Methods in Combinatorics

Solutions to Assignment 12

Problem 1. Let $G = (V, E)$ be a graph with maximum degree Δ. Show that there exists a colouring of the vertices of G with at most $100\Delta^4$ colours such that, for each vertex v, no two vertices in $N(v)$ have same colour.

Solution. We independently give to each vertex $v \in V$ a colour uniformly at random in the set $\{1, \ldots, 100\Delta^4\}$. Let A_v be the probability that at least two vertices in $N(v)$ have same colour. We have

$$\mathbb{P}(A_v) \leq 1 - \prod_{i=1}^{\Delta}(1 - \frac{i}{100\Delta^4}) \leq 1 - \exp\left(-\frac{2\Delta(\Delta+1)}{100\Delta^4}\right) \leq \frac{1}{25\Delta^2}.$$

Every event A_v is mutually independent from all events A_u such that $|N[u] \cap N[v]| = 0$. Therefore, A_v is mutually independent of a collection of all but at most $\Delta + \Delta^2$ events A_u. As

$$e(\Delta^2 + \Delta + 1)\frac{1}{25\Delta^2} \leq 1,$$

a simple application of LLL concludes the proof.

Problem 2. Let $A = (A_1, A_2, \ldots, A_n)$ and $B = (B_1, B_2, \ldots, B_n)$ be two sequences over a finite alphabet Σ. A *common subsequence* of A and B is a sequence $(C_1, C_2, \ldots, C_k)$ such that $C_1, \ldots, C_k$ appear in A in order (not necessarily contiguously), and $C_1, \ldots, C_k$ appear in B in order (again, not necessarily contiguously). The *Longest Common Subsequence (LCS)* of A and B is a common subsequence of A and B of maximum possible length. Let $A = (A_1, A_2, \ldots, A_n)$ and $B = (B_1, B_2, \ldots, B_n)$ be two independent uniformly random sequences of length n over the alphabet $\{0, 1\}$. Let L be the LCS of A and B. Show that

$$\mathbb{P}(|L - \mathbb{E}[L]| \geq 100\sqrt{n}) \leq 1/100.$$

Solution. Let $\Omega = \prod_{i=1}^{2n} \Omega_i$, where $\Omega_i = \{0, 1\}$, and the Ω_i for $1 \leq i \leq n$ represents A, and the Ω_i for $n + 1 \leq i \leq 2n$ represents B. Note that L is 1-Lipschitz with respect to the product probability space Ω. Therefore an application of Azuma-Hoeffding gives

$$\mathbb{P}(|L-\mathbb{E}[L]| \geq 100\sqrt{n}) \leq \exp\left(-\frac{(100\sqrt{n})^2}{4n}\right) \leq 1/100,$$

as wanted.

Problem 3. Let $G_i = (V, E_i)$ for $i = 1, 2$ be two graphs on same vertex set, and let $e_i = |E_i|$ for $i = 1, 2$. Show that there exists a partition $V = X \cup Y$ such that the number of cross edges (i.e. having exactly one endpoint in X and Y) in G_i is at least $e_i/2 - 10\sqrt{e_i}$ for each $i = 1, 2$.

Solution. Let $X \cup Y$ be a uniform random partition of V, that is, for every $v \in V$, we put v in X with probability $1/2$, and in Y with probability $1/2$, independently of the other vertices. Let Z_i be the number of cross edges in G_i. We have $\mathbb{E}[Z_i] = e_i/2$, and if $e, f \in E_i$, then

$$\mathbb{P}(e, f \text{ are cross edges}) = \begin{cases} 1/2 & \text{if } e = f, \\ 1/4 & \text{if } e \neq f. \end{cases}$$

Thus, $\mathrm{Var}[Z_i] = \sum_{e \in E_i} 1/2 - 1/4 = e_i/4$. Therefore, by Chebychev's inequality:

$$\mathbb{P}(Z_i - e_i/2 \leq -10\sqrt{e_i}) \leq \mathbb{P}(|Z_i - \mathbb{E}[Z_i]| \geq 2\sqrt{\mathrm{Var}[Z_i]}) \leq 1/4.$$

A simple union bound over $i = 1, 2$, then shows that there exists a partition $X \cup Y$ such that $Z_i \geq e_i/2 - 10\sqrt{e_i}$ for $i = 1, 2$.

Problem 4. Let $G = (V, E)$ be the graph whose vertices are all 7^n vectors of length n over $\mathbb{Z}_7$, in which two vertices are adjacent if and only if they differ in precisely one coordinate. Let $U \subseteq V$ be a set of 7^{n-1} vertices of G, and let W be the set of all vertices of G whose distance from U exceeds $(c+2)\sqrt{n}$, where $c > 0$ is a constant. Prove that $|W| \leq 7^n \cdot e^{-c^2/2}$.

Solution. Choose a random vector $y \in \mathbb{Z}_7^n$ as follows: for each coordinate i independently, let $y_i = 0, 1, \ldots, 6$ with equal probability $1/7$. Note that this implies that every point of $\mathbb{Z}_7^n$ is equally likely to be chosen.

Define the random variable Y as the distance of the randomly chosen point y to U.

Observe that $Y = 0$ if and only if $y \in U$, thus, since the distribution is uniform, $\mathbb{P}(Y = 0) = \frac{|U|}{|V|} = 1/7$.

Moreover, we have $Y > (c+2)\sqrt{n}$ if and only if $y \in W$. Hence, our goal is to show that

$\mathbb{P}(Y > (c+2)\sqrt{n}) \le e^{-c^2/2}$.

Since changing one coordinate of y can change its distance to any point by at most 1, we have that Y is 1-Lipschitz. Thus, by the Azuma-Hoeffding inequality,

$$\mathbb{P}(Y - \mathbb{E}[Y] < -t\sqrt{n}) < e^{-t^2/2} \tag{1}$$

and

$$\mathbb{P}(Y - \mathbb{E}[Y] > t\sqrt{n}) < e^{-t^2/2} \tag{2}$$

for any $t > 0$.

Taking $t = 2$ in (1) gives

$$\mathbb{P}(Y < \mathbb{E}[Y] - 2\sqrt{n}) < e^{-2} < 1/7.$$

Since $\mathbb{P}(Y = 0) = 1/7$, we must have $\mathbb{E}[Y] \le 2\sqrt{n}$.

Taking $t = c$ in (2) now gives

$$\mathbb{P}(Y > (c+2)\sqrt{n}) \le \mathbb{P}(Y > \mathbb{E}[Y] + c\sqrt{n}) < e^{-c^2/2}.$$