# Probabilistic techniques - tutorials

Problem set 3 – Problem set #3 - Markov and Chebyshev inequalities

Release: **November 9, 2023**. Hints: **November 23, 2023**. Deadline: **November 30, 2023**. Send solutions to honst+pt23@iuuk.mff.cuni.cz.

1. Let $n \geq 2$ be a positive integer and $v_1 = (x_1, y_1), \ldots, v_n = (x_n, y_n)$ vectors with $x_i$ and $y_i$ integers such that $x_i^2, y_i^2 \leq \frac{1}{10000}\frac{2^n}{n}$. Prove that there exist two non-empty disjoint subsets $I, J \subset [n]$ such that

$$\sum_{i \in I} v_i = \sum_{j \in J} v_j.$$

**[3]**

*Hint:* An easier instance of the problem was explained in the tutorial. In the case of dimension 2 consider a random ser $I$ indexes from $[n]$, and $X$ and $Y$ are the sums of the coordinates $x$ and $y$ of the vectors chosen with $I$. Prove that there are many options for $I$ in a square centered around $(\mathbb{E}[X], \mathbb{E}[Y])$ the same way as we did in the tutorial. However, notice that if we can pick $I \not\subseteq J$ we are done, otherwise we need to pick three sets $I, J, K$ to workaround the case where $I \subset J$.

2. Let $X$ denote the number of isolated vertices in $G(n, p(n))$ with $p(n) = c\frac{\ln(n)}{n}$. Show that

   (a) $\lim_{n\to\infty} \Pr[X = 0] = 1$ for $c > 1$. **[1]**

   *Hint:* Use Markov inequality.

   (b) $\lim_{n\to\infty} \Pr[X \geq 1] = 1$ for $0 \leq c < 1$. **[3]**

   *Hint:* Use that if $(X_n)_{n\in\mathbb{N}}$ is a sequence of random variables such that $\lim_{n\to\infty} \frac{\mathrm{Var}[X_n]}{\mathbb{E}[X_n]^2} = 0$ then $\lim_{n\to\infty} \Pr[X_n > 0] = 1$.

3. Let $X$ be a non-negative integer random variable such that $\mathbb{E}[X^2]$ is finite and nonzero. Prove that

$$\Pr[X = 0] \leq \frac{\mathrm{Var}[X]}{\mathbb{E}[X^2]}.$$

**[2]**

*Hint:* First, notice that as $\mathbb{E}[X^2]$ is bounded, then $\mathbb{E}[X]$ as well. Then write the definition of variance and use Cauchy–Schwarz inequality for series. This works as the $\mathbb{E}[X^2]$ is convergent.

4. Prove that

$$\lim_{n\to\infty} \Pr[G(n, 1/2) \text{ has an induced cycle of lenght} > 3\log_2 n] = 0.$$

Remember that $(v_1, \ldots, v_k)$ is an induced cycle of length $k$ if it satisfies that $v_i v_j$ is an edge if and only if $j = i + 1 \mod (k)$. **[3]**

*Hint:* Let $X$ denote the number of induced cycles of length greater than $3\log_2(n)$ in $G(n, 1/2)$. Prove that $\lim_{n\to\infty} \mathbb{E}[X] = 0$. Then apply Markov inequality to $X$.

5. Let $X$ be a real random variable with $\mathrm{Var}[X] = \sigma^2$ and $\mathbb{E}[X] = 0$. For every real number $\lambda > 0$ prove the inequality

$$\Pr[X \geq \lambda] \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

**[2]**

*Hint:* Assume that $\sigma \neq 0$. Consider the change of variable $Y = \frac{X}{\sigma}$ to reduce the problem to $\sigma = 1$. Now you just have to prove $\Pr[Y \geq t] \leq \frac{1}{t^2+1}$ where t is such that $\sigma t = \lambda$. Pick $c > 0$ and use Markov inequality on $\Pr[(Y + c)^2 \geq (t + c)^2]$ and choose the right $c$.

6. Show that there is a positive constant $c$ such that the following holds: For any $n$ vectors $a_1, \ldots, a_n \in \mathbb{R}^2$ satisfying $\sum_{i=1}^{n} \|a_i\|^2 = 1$ and $\|a_i\| \leq 1/10$, where $\|\cdot\|$ denotes the usual Euclidean norm. If $(\epsilon_1, ..., \epsilon_n)$ is a $\{-1, +1\}$-random vector obtained by choosing each $\epsilon_i$ randomly and independently with uniform distribution, then

$$\Pr\left[\left\|\sum_{i=1}^{n} \epsilon_i a_i\right\| \leq 1/3\right] \geq c.$$

**[3]**

*Hint:* Partition $[n]$ into $k$ parts such that each part has roughly the same square sum. Use Chebyshev for vectors, that is if $(X, Y)$, $(E[X], E[Y])$ and $\sigma = (\sigma(X), \sigma(Y))$ are the respective random variable, expectation and standard deviation, then $\Pr[\|(X, Y) - (E[X], E[Y])\|_2 \geq \lambda] \leq \frac{\|\sigma\|_2^2}{\lambda^2}$.

7. Prove that for every set $X$ of at least $4k^2$ distinct residue classes modulo a prime $p$, there is an integer $a$ such that the set $\{ax \mod p : x \in X\}$ intersects every interval in $\{0, 1, \ldots, p-1\}$ of length at least $p/k$. **[3]**

*Hint:* Partition $\mathbb{Z}_p$ into $2k$ intervals $I_j$ of nearly equal size. Choose $a$ and $b$ randomly and independently in $\mathbb{Z}_p$ , and let $z_x = ax + b (\mod p)$. Show that the random variables $z_x$ for $x \in X$ are pairwise independent, and apply Chebyshev to show that with positive probability at least one of them falls into each of the intervals $I_j$.

---

Information about tutorials `https://kam.mff.cuni.cz/~dbulavka/teaching/ws2324/pt.html`