1 20/12/2019 - High-dimensional probability. An Introduction with applications in Data Science (Vershynin R.)

На семинаре были обсуждены различные результаты связанные с неравенствами концентрации. В частности обсуждалась следующая теорема:

Theorem 1. (Hoeffding's inequality, two-sided) Let $X_1, \ldots, X_N$ be independent symmetric Bernoulli random variables, and $a = (a_1, \ldots, a_N) \in \mathbb{R}$. Then, for any $t > 0$, we have

$$\mathrm{P}\left\{\left|\sum_{i=1}^{N} a_i X_i\right| \geq t\right\} \leq 2\exp\left(-\frac{t^2}{2\|a\|_2^2}\right)$$

Автор также предлагает доказать аналогичный результат для ограниченных случайных величин, а также прооптимизировать константы внутри экспоненты.

Также были обсуждены следующий результат:

Theorem 2. (Chernoff's inequality) Let X_i be independent Bernoulli random variables with parameters p_i. Consider their sum $S_N = \sum_{i=1}^{N} X_i$ and denote its mean by $\mu = \mathrm{E}S_N$. Then, for any $t > \mu$, we have

$$\mathbb{P}\left\{S_N \geq t\right\} \leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$

Для того, чтобы получить оценку для $t < \mu$ было приведено следующее

Theorem 3. (Chernoff's inequality: small deviations) In the setting of Theorem 2.3.1, for any $t \in (0, \mu]$ we have

$$\mathbb{P}\left\{|S_N - \mu| \geq t\right\} \leq \exp\left(-\frac{ct^2}{\mu}\right)$$

На основе Теоремы 2 было показано, что существует $c > 0$, такое что для любых $\delta \in [0, 1]$ выполнено

$$\mathbb{P}\left\{|S_N - \mu| \geq \delta\mu\right\} \leq 2e^{-c\mu\delta^2},$$

что использовалось при анализе графа Эрдоша-Рени $G(n, p_n)$.

Упражнение: Доказать, что если средняя степень $d = (n-1)p_n$ вершины графа равна $O(1)$, то с большой вероятностью найдётся изолированная вершина при $n \to \infty$.