Problem Set 3

- This problem set is due on **April 20, 2021** in the class.
- Each problem carries 10 points.
- No collaboration allowed. Each student must write his/her own solution.

1. **(Moments Vs. Chernoff Bounds)** Show that moment bounds for tail probabilities are always better than Chernoff bounds. More precisely, let X be a nonnegative random variable and let $t > 0$. The best moment bound for the tail probability $\mathbb{P}(X \geq t)$ is $\min_q \mathbb{E}(X^q)t^{-q}$ where the minimum is taken over all positive integers. The best Chernoff bound is $\inf_{s>0} \mathbb{E}\big(\exp(s(X-t))\big)$. Prove that

$$\min_q \mathbb{E}(X^q)t^{-q} \leq \inf_{s>0} \mathbb{E}\big(\exp(s(X-t))\big).$$

2. **(Spectrum of random matrices)** Let $A \in \mathbb{R}^{n\times n}$ be a random symmetric matrix with independent upper-triangular part. Assume that the absolute value of each entry is bounded by 1. Then show that

$$\text{Var}[\sigma_{\max}(A)] \leq 8,$$

where $\sigma_{\max}(A) = \max\{x^TAx : x^Tx = 1\}$.

3. **(Configuration functions)** Frequently, one studies functions of $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ defined as

$$f(x) = \sup\{|S| : x_S \in \mathcal{P}, S \subseteq [n]\}, \tag{1}$$

where $x_S = (x_{i_1}, x_{i_2}, \ldots)$ with $i_1 < i_2 < \ldots \in S$ is a *subsequence* and $\mathcal{P} \subseteq \mathcal{X}^*$ is a set of (variable-length) $\mathcal{X}$-valued sequence. We say that $\mathcal{P}$ is *hereditary* if it is closed under operation of taking a subsequence. Functions f defined as (1) for a hereditary $\mathcal{P}$ were christened *configuration functions* by Talagrand. They possess a notable self-bounding property discussed below.

   (a) Prove that each of the following are configuration functions by identifying the appropriate $\mathcal{P}$: (a) longest increasing subsequence of x, (b) longest common subsequence of x and a fixed $y = (y_1, y_2, \ldots, y_{n'})$, clique number of a graph, (d) maximal degree of a graph, (e) number of distinct values occurring in x.

   (b) Show that for any configuration function f and X with independent components we have:

$$\text{Var}(f(X)) \leq \mathbb{E}[f(X)].$$

(c) Show that for any configuration function f, for all x and $a \in \mathbb{R}$ we have

$$f(x) \leq a + d_c(x, \{f \leq a\}),$$

where $d_c(x, A) = \inf_{y \in A} \sup_{\alpha:||\alpha||_2=1} \sum_{i=1}^{n} \alpha_i \mathbb{1}\{y_i \neq x_i\}$ is the Talagrand's convex distance. This implies, via Talagrand's inequality, good tail bounds on $f(X)$.

4. **(Size of the maximum matching in a random bipartite graph)** Given $1 \leq d \leq n$, let $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, v_2, \ldots, v_n\}$ be disjoint sets of cardinality n, and let G be a bipartite random graph with vertex set $U \cup V$, such that if V_i denotes the set of neighbors of u_i, then $V_1, V_2, \ldots, V_n$ are independent, and each is uniformly distributed over the set of all $\binom{n}{d}$ subsets of V of cardinality d. A matching for G is a subset of edges M such that no two edges in M have a common vertex. Let Z denote the maximum of cardinalities of the matchings in G.
(a) Show that $\mathbb{E}(Z) = \Theta(n)$.

   (b) Give an upper bound on $\mathbb{P}(|Z - \mathbb{E}[Z]| \geq \gamma\sqrt{n})$, for $\gamma > 0$, showing that for fixed d, the distribution of Z is concentrated about its mean as $n \to \infty$.

5. **(Disjoint Triangles in Random Graphs)** Consider a random graph G on n vertices in which edges appear independently each with probability $p(n) = \frac{\lambda}{n}$. Let $\Delta(G)$ denote the maximum number of edge-disjoint triangles in G. Using Talagrand's concentration inequality show that:

$$\mathbb{P}(\Delta(G) \leq m - t\sqrt{3m}) \leq 2e^{-t^2/4},$$

   where m is the median of $\Delta(G)$.