E2 207 – Concentration Inequalities, Aug-Dec 2017
Homework #3

---

(1) *Quantiles of a monotone set*
Let $A \subseteq \{-1,1\}^n$ be a monotone set. Denote $p_\theta = \inf\{p : P_p(A) \geq \theta\}$. Show that if $p_{1/2} = 1/2$, there exists a universal constant $c > 0$ such that $p_{3/4} - p_{1/4} \geq c/\sqrt{n}$.

(2) *Concentration and influence*
Let $A \subseteq \{-1,1\}^n$ and let X be uniformly distributed on $\{-1,1\}^n$. Let $d_H(x,A)$ denote the Hamming distance of $x \in \{-1,1\}^n$ to the set A. Prove that

$$\mathbb{E} d_H(X,A) \leq \frac{I(A)}{2P(A)}.$$

(3) *Talagrand's convex distance – equivalent definitions*
Recall the definition of Talagrand's distance between a point $x \in \mathcal{X}^n$ and $A \subseteq \mathcal{X}^n$: $d_T(x,A) := \sup_{\|\alpha\|=1} \inf_{y \in A} d_\alpha(x,y) = \sup_{\|\alpha\|=1} \inf_{y \in A} \sum_{i=1}^n \alpha_i \, \mathbb{1}_{\{x_i \neq y_i\}}$. Prove the following equivalent characterization of d_T.

Let $U(A,x) := \{(s_1,\ldots,s_n) \in \{0,1\}^n : \exists y \in A \; \forall i \; x_i \neq y_i \Rightarrow s_i = 1\}$. (What do you think $U(A,x)$ represents?). Then, $d_T(x,A)$ is the least distance between the origin and a point in the convex hull of $U(A,x)$. (This is the reason for the name "convex distance".)

(4) *Concentration of Lipchitz functions in set-distance*
Let $d(x,A)$ denote the distance of x from a set A, namely d is a nonnegative function of $x \in \mathcal{X}$ and $A \subseteq \mathcal{X}$ such that $d(x,A) = 0$ iff $x \in A$. We say that a function $f : \mathcal{X} \to \mathbb{R}$ is Lipchitz in d if $f(x) \leq d(x,A) + \sup_{y \in A} f(y)$. Show that Levy's inequalities hold for any function f Lipchitz in d. Furthermore, show that Talagrand's convex distance constitutes a set distance in the sense described above. Finally, derive the concentration bounds for Lipchitz functions in d_T around their medians.
(We did most of this in class; just some clean-up is required.)

(5) *Distance between sets*
Let P be a product probability measure on $\mathcal{X}^n$. Let $d_H(A,B) = \min_{x \in A, y \in B} d_H(x,y)$ denote the Hamming distance between two sets $A, B \subseteq \mathcal{X}^n$. Show that

$$d_H(A,B) \leq \sqrt{\frac{n}{2} \log \frac{1}{P(A)}} + \sqrt{\frac{n}{2} \log \frac{1}{P(B)}}.$$

(6) *Maximum measurement*
Let $X_1, .., X_n$ be iid with common distribution $\texttt{unif}([0,a])$. Let A be a fixed $m \times n$ matrix consisting of nonnegative real entries. Derive concentration bounds for $\max_{1 \leq i \leq m} (Ax)_i$ using McDiarmid's inequality, the Enropy method, Talagrand's convex distance inequality, and the Transportation cost method. Derive bounds for both upper and lower tails.