Fall 2016

Learning Theory (048995): Exercise 1

Instructions: Please submit by 18/12/16 at the beginning of class. You may submit in pairs. We allow you to consult your peers, but ask that you mention the peers you have consulted with. Note that Question 5 is a bonus question. You may score more than 100%.

Remark: Throughout, we use c, c_1,..., etc., to denote absolute constants. These are positive numbers that are independent of the dimension or of any other parameter of the problem.

Question 1 [24%]

Let $X_1, ..., X_N$ be iid random variables distributed as X and denote by $(X_i^*)_{i=1}^N$ the non-increasing rearrangement of $(|X_i|)_{i=1}^N$.

(a) Show that for $1 \leq k \leq N$, $p \geq 1$ and $t > 0$,

$$Pr(X_k^* > t) \leq \left(\frac{eN}{k}\right)^k \cdot \left(\frac{\|X\|_{L_p}}{t}\right)^{kp}.$$

(b) Let $g_1, ..., g_n$ be independent standard Gaussian random variables. Use (a) to show that for $1 \leq k \leq N/2$, $\mathbb{E}g_k^* \leq c_1\sqrt{\log(en/k)}$.

(c) Show that the estimate from (b) is accurate, in the sense that $\mathbb{E}g_k^* \geq c_2\sqrt{\log(en/k)}$.

Question 2 [28%]

Let $G = (g_1, ..., g_n)$ be the standard Gaussian vector in $\mathbb{R}^n$ and set $X = (\varepsilon_1, ..., \varepsilon_n)$ to be a random vector whose coordinates are independent, symmetric $\{-1, 1\}$-valued random variables.

(a) If $\| \ \|_p$ denotes the ℓ_p norm, find upper and lower estimates on $\mathbb{E}\|G\|_p$.

(b) Find upper bounds on $Pr(\|G\|_2 \geq t\mathbb{E}\|G\|_2)$ for $t > 2$ and on $Pr(\|G\|_2 \leq t\mathbb{E}\|G\|_2)$ for $0 < t < 1/2$. We are particularly interested in what happens for 'large' t and for t close to 0.

(c) Let $T = \{x : \|x\|_p \leq 1\}$ for $1 \leq p \leq \infty$. Find upper and lower bounds on $\mathbb{E}\sup_{t\in T}\langle G, t\rangle$ and $\mathbb{E}\sup_{t\in T}\langle X, t\rangle$.

(d) Let T be the set of vectors in the Euclidean unit sphere S^{n-1} that are supported on at most d coordinates. Show that

$$\mathbb{E}\sup_{t\in T}\langle G, t\rangle \leq c\sqrt{d\log(en/d)}.$$

Question 3 [24%]

Let $x_1, ..., x_n$ be independent standard exponential random variables (that is, with density that is proportional to $\exp(-|t|)$) and set $X = (x_1, ..., x_n)$.

(a) Find a direction $\theta \in S^{n-1}$ for which $\|\langle X, \theta \rangle\|_{\psi_2} = \infty$.

(b) Show that there is an absolute constant c for which *any* direction $\theta \in S^{n-1}$ satisfies $\|\langle X, \theta \rangle\|_{\psi_1} \leq c$.

Question 4 [24%]

Let X be a random variable and assume that $\|X - \mathbb{E}X\|_{L_4} \leq L\|X - \mathbb{E}X\|_{L_2}$.

(a) Use the Berry-Esseen theorem to show that there is a constant m_0 that depends only on L and an absolute constant c for which, for $m \geq m_0$,

$$Pr\left(\left| \frac{1}{m} \sum_{i=1}^{m} X_i - \mathbb{E}X \right| \geq c \frac{\sigma_X}{\sqrt{m}} \right) \geq \frac{3}{4},$$

where, as always, $\sigma_X = \|X - \mathbb{E}X\|_{L_2}$.

(b) Let $N, m > m_0$, set $n = N/m$ and assume that n is an integer. Let $I_1, ..., I_n$ be a partition of $\{1, ..., N\}$ to n 'block', each one of cardinality m and set

$$Z_j = \frac{1}{m} \sum_{i \in I_j} X_i.$$

Define $\hat{Z}$ to be a median of $Z_1, ..., Z_n$; that is, for every realization of $Z_1, ..., Z_n$, $\hat{Z}$ is a number that is larger than half of the Z_i's and smaller than half of the Z_i's.

Show that with probability at least $1 - 2\exp(-c_1 n)$,

$$|\hat{Z} - \mathbb{E}X| \leq c_2 \frac{\sigma_X}{\sqrt{m}},$$

for suitable absolute constants c_1 and c_2.

(c) Prove that there is a procedure that receives as data an iid sample $X_1, ..., X_N$ and $0 < \delta < 1$ and returns a value $\hat{\mu}$, which, with probability $1 - \delta$ satisfies

$$|\mathbb{E}X - \hat{\mu}| \leq c\sqrt{\log(2/\delta)} \frac{\sigma_X}{\sqrt{N}}.$$

Question 5 [(Bonus question: 20%]

Let X be a random vector in $\mathbb{R}^n$ whose coordinates are bounded by 1 almost surely. For any $(\varepsilon_1, ..., \varepsilon_n) \in \{-1, 1\}^n$ let $\theta_\varepsilon = (\varepsilon_1/\sqrt{n}, ..., \varepsilon_n/\sqrt{n})$. Show that more than 2^{n-1} of the points in $\{-1, 1\}^n$ satisfy $\|\langle X, \theta_\varepsilon \rangle\|_{\psi_2} \leq C$, where C is an absolute constant.