18.S997 Spring 2015: Problem Set 1

Problem 1.1

A random variable X has χ^2_n (chi-squared with n degrees of freedom) if it has the same distribution as $Z_1^2 + \ldots + Z_n^2$, where $Z_1, \ldots, Z_n$ are iid $\mathcal{N}(0,1)$.

(a) Let $Z \sim \mathcal{N}(0,1)$. Show that the moment generating function of $Y = Z^2 - 1$ satisfies

$$\phi(s) := E\big[e^{sY}\big] = \begin{cases} \dfrac{e^{-s}}{\sqrt{1-2s}} & \text{if } s < 1/2 \\ \infty & \text{otherwise} \end{cases}$$

(b) Show that for all $0 < s < 1/2$,

$$\phi(s) \le \exp\left(\frac{s^2}{1-2s}\right).$$

(c) Conclude that

$$\mathbb{P}(Y > 2t + 2\sqrt{t}) \le e^{-t}$$

`[Hint: you can use the convexity inequality` $\sqrt{1+u} \le 1 + u/2$`]`.

(d) Show that if $X \sim \chi^2_n$, then, with probability at least $1-\delta$, it holds

$$X \le n + 2\sqrt{n \log(1/\delta)} + 2\log(1/\delta).$$

Problem 1.2

Let $A = \{A_{i,j}\}_{\substack{1 \le i \le n \\ 1 \le j \le m}}$ be a random matrix such that its entries are iid sub-Gaussian random variables with variance proxy σ^2.

(a) Show that the matrix A is sub-Gaussian. What is its variance proxy?

(b) Let $\|A\|$ denote the operator norm of A defined by

$$\max_{x \in \mathbb{R}^m} \frac{|Ax|_2}{|x|_2}.$$

Show that there exits a constant $C > 0$ such that

$$\mathbb{E}\|A\| \le C(\sqrt{m} + \sqrt{n}).$$

Problem 1.3

Let K be a compact subset of the unit sphere of $\mathbb{R}^p$ that admits an ε-net $\mathcal{N}_\varepsilon$ with respect to the Euclidean distance of $\mathbb{R}^p$ that satisfies $|\mathcal{N}_\varepsilon| \le (C/\varepsilon)^d$ for all $\varepsilon \in (0,1)$. Here $C \ge 1$ and $d \le p$ are positive constants. Let $X \sim \mathsf{subG}_p(\sigma^2)$ be a centered random vector.

Show that there exists positive constants c_1 and c_2 to be made explicit such that for any $\delta \in (0,1)$, it holds

$$\max_{\theta \in K} \theta^\top X \le c_1 \sigma \sqrt{d \log(2p/d)} + c_2 \sigma \sqrt{\log(1/\delta)}$$

with probability at least $1-\delta$. Comment on the result in light of Theorem 1.19.

Problem 1.4

Let $X_1, \dots, X_n$ be n independent and random variables such that $\mathbb{E}[X_i] = \mu$ and $\mathrm{var}(X_i) \le \sigma^2$. Fix $\delta \in (0,1)$ and assume without loss of generality that n can be factored into $n = K \cdot G$ where $G = 8\log(1/\delta)$ is a positive integers.

For $g = 1, \dots, G$, let $\bar{X}_g$ denote the average over the gth group of k variables. Formally

$$\bar{X}_g = \frac{1}{k} \sum_{i=(g-1)k+1}^{gk} X_i \,.$$

1. Show that for any $g = 1, \dots, G$,
$$\mathbb{P}\big[\bar{X}_g - \mu > \frac{2\sigma}{\sqrt{k}}\big] \le \frac{1}{4}\,.$$
2. Let $\hat{\mu}$ be defined as the median of $\{\bar{X}_1, \dots, \bar{X}_G\}$. Show that
$$\mathbb{P}\big[\hat{\mu} - \mu > \frac{2\sigma}{\sqrt{k}}\big] \le \mathbb{P}\big[\mathcal{B} \ge \frac{G}{2}\big]\,,$$
where $\mathcal{B} \sim \mathsf{Bin}(G, 1/4)$.
3. Conclude that
$$\mathbb{P}\big[\hat{\mu} - \mu > 4\sigma\sqrt{\frac{2\log(1/\delta)}{n}}\big] \le \delta$$
4. Compare this result with Corollary 1.7 and Lemma 1.4. Can you conclude that $\hat{\mu} - \mu \sim \mathsf{subG}(\bar{\sigma}^2/n)$ for some $\bar{\sigma}^2$? Conclude.