ECE 299: STATISTICAL LEARNING THEORY

MAXIM RAGINSKY

Homework 1

Assigned February 2, 2011; due February 16, 2011

---

(1) Let $X_1, \dots, X_n$ be n independent Bernoulli(θ) random variables. Prove the following *multiplicative Chernoff bound*:

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \ge (1+\gamma)\mu\right) \le e^{-\gamma^2 n\mu/3}$$

for any $0 \le \gamma \le 1$ and any $\theta \le \mu \le 1$. You may need the fact that

$$\sum_{k \ge (1+\gamma)\mu n} \binom{n}{k} \mu^k (1-\mu)^{n-k} \le e^{-\gamma^2 n\mu/3}.$$

(2) Let $X_1, \dots, X_n$ be $n \ge 2$ real-valued random variables (not necessarily independent). Suppose that there exists some constant $\sigma > 0$, such that for all $s > 0$ and all $i = 1, \dots, n$

$$\mathbb{E}\left[e^{sX_i}\right] \le e^{s^2\sigma^2/2}$$

Prove the following:

(a)

$$\mathbb{E}\left[\max_{1 \le i \le n} X_i\right] \le \sigma\sqrt{2\log n}.$$

Hint. Try to bound $e^{s\mathbb{E}[\max_{1 \le i \le n} X_i]}$, exploit convexity of the function $\phi(x) = e^{sx}$.

(b) Assuming that $X_i \ge 0$ for all $i = 1, \dots, n$,

$$\mathbb{P}\left(\max_{1 \le i \le n} X_i \ge \sqrt{2\sigma^2 n}\right) \le \sqrt{\frac{\log n}{n}}.$$

(3) Let $X^n = (X_1, \dots, X_n)$ be an n-tuple of independent random variables taking values in some space X. The *Hamming distance* between any n-tuples $x^n, y^n \in \mathsf{X}^n$ is defined as the number of coordinates in which x^n and y^n differ:

$$d(x^n, y^n) \triangleq \sum_{i=1}^n \mathbf{1}_{\{x_i \ne y_i\}}.$$

If B is an arbitrary (measurable) subset of X^n, then the Hamming distance between an n-tuple $x^n \in \mathsf{X}^n$ and B is defined as

$$d(x^n, B) \triangleq \min_{y^n \in B} d(x^n, y^n).$$

Use McDiarmid's inequality to prove the following fact: if

$$\varepsilon \ge \sqrt{\frac{1}{2n}\log\frac{1}{\mathbb{P}(B)}},$$

then

$$\mathbb{P}\left(d(X^n, B) \geq n\varepsilon\right) \leq \exp\left(-2n\left(\varepsilon - \sqrt{\frac{1}{2n}\log\frac{1}{\mathbb{P}(B)}}\right)^2\right),$$

where $\mathbb{P}(B) = \mathbb{P}(X^n \in B)$ is the probability of the set $B \in \mathsf{X}^n$ under the joint distribution of X^n.

Hint. You may find the following fact handy: $d(x^n, B) \leq 0$ if and only if $x^n \in B$.