Statistics 593A, Problem Set 2

Wellner; 4/17/2014

Due: Thursday, May 1, 2014

1. BLM, page 79, problem 3.7: Show that the conditional Rademacher average Z satisfies the self-bounding property. Here Z is defined by

$$Z \equiv E\left\{ \max_{1\le j\le d} \sum_{i=1}^n \epsilon_i X_{i,j} \middle| X_1,\ldots,X_n \right\}$$

where $X_1,\ldots,X_n$ are independent random variables taking values in $[-1,1]^d$ and $\epsilon_1,\ldots,\epsilon_n$ are independent Rademacher random variables which are independent of the X_i's

2. BLM, page 78, problem 3.5: Consider the class $\mathcal{F}$ of functions $f:\mathbb{R}^n \to \mathbb{R}$ that are Lipschitz with respect to the ℓ^1 distance: i.e.

$$|f(x_1,\ldots,x_n) - f(y_1,\ldots,y_n)| \le \sum_{i=1}^n |x_i - y_i|.$$

Let $X = (X_1,\ldots,X_n)$ be a vector of independent random variables with finite variance. Use the Efron - Stein inequality to show that the maximal value of $Var(f(X))$ over $f \in \mathcal{F}$ is attained by the function $f(x) = \sum_{i=1}^n x_i$. (This is from Bobkov and Houdré (1996).)

3. BLM, page 114, problem 4.11: prove that for any fixed probability measure P on $\mathcal{X}$, the function $Q \mapsto D(Q\|P)$ is convex on the set of probability distributions over $\mathcal{X}$. *Hint:* Use the duality representation.

4. BLM, page 114, problem 4.13: Let Z be a real-valued random variable. Recall that $\psi_Z(\lambda) = \log Ee^{\lambda Z}$ for $\lambda \in \mathbb{R}$. Let $\psi^*(t) = \sup_{\lambda\in\mathbb{R}}\{\lambda t - \psi_{Z-E(Z)}(\lambda)\}$. Prove that for all $t > 0$

$$\psi^*(t) = \inf\{D(Q\|P) : \ E_Q(Z) - E(Z) \ge t\}.$$

5. **Bonus problem:** BLM, page 115, problem 4.17: Let C be a convex body (a compact convex set with nonempty interior) in $\mathbb{R}^n$, and let P be the uniform probability distribution over C. Prove Borell's lemma that states the following: if A is a symmetric convex subset of C with $P(A) > 1/2$, then for any $t > 1$,

$$P((tA)^c) \le P(A)\left(\frac{1-P(A)}{P(A)}\right)^{(t+1)/2}.$$

Hint: Prove first that for $t > 1$

$$\frac{2}{t+1}(tA)^c + \frac{t-1}{t+1}A \subset A^c$$

where $A^c = C \setminus A = C \cap A^c$ where the complement on the right side is the usual complement in $\mathbb{R}^n$. Then use the Brunn-Minkowski inequality. (This is an example of the concentration of measure phenomenon: note that the inequality does not depend on the ambient dimension n.)