10/36-702 Homework 1
Due: Friday 2/8/2013

Instructions: Hand in your homework to Michelle Martin (GHC 8001) before 3:00pm on Friday 2/8/2013.

1. Suppose that $E(V|Z) = 0$ and that
$$h(Z) \leq V \leq h(Z) + c$$
for some h and some $c > 0$. Show that
$$E(e^{tV}|Z) \leq e^{t^2c^2/8}.$$

2. Suppose that $g(x_1, \ldots, x_n)$ satisfies the bounded difference property. Show that
$$\mathrm{Var}(g(X_1, \ldots, X_n)) \leq \frac{1}{4}\sum_{i=1}^{n} c_i^2.$$

3. Let $h(x, y) = \sum_j I(x_j \neq y_j)$ denote the Hamming distance between two vectors x and y. For a set $B \subset \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, let $h(B, x) = \min_{y \in B} h(y, x)$. Let P be some univariate distribution on a finite set A and suppose that for all $i = 1, \ldots, n$, $X_i \sim P$, independently. Let $B \subset A^n$. Show that for $\epsilon > \sqrt{\frac{1}{2n}\log\left(\frac{1}{P(B)}\right)}$:
$$\mathbb{P}(h(B, X) \geq n\epsilon) \leq \exp\left(-2n\left[\epsilon - \sqrt{\frac{1}{2n}\log\left(\frac{1}{P(B)}\right)}\right]^2\right).$$

4. VC Dimension:
   (a) What is the VC dimension of origin-centered radial classifiers i.e. $\mathrm{sign}(||\mathrm{x}||^2 - \mathrm{b})$ in 2-d? (*If VC dim = d, you must argue both the lower bound VC dim $\geq$ d and upper bound VC dim $<$ d + 1.*)
   (b) Does the VC dimension change if the reverse sign classifiers are also contained in the class, i.e. $\mathrm{sign}(\mathrm{b} - ||\mathrm{x}||^2)$?
   (c) What is the VC dimension of arbitrarily centered radial classifiers i.e. $\mathrm{sign}(||\mathrm{x} - \mathrm{x}_0||^2 - \mathrm{b})$ in 2-d?
   (d) What is the VC dimension of classifiers whose decision boundaries are defined by union of two arbitrarily centered circles in 2-d? Note that the indicator functions for the level sets of a mixture of two isotropic gaussians is exactly the union of two arbitarily centered circles.

5. Here we will use Rademacher Complexity to obtain generalization error bounds for finite function classes and function classes with known VC-dimension.

   (a) Use the result from question 1 (or the bound on the moment-generating function of bounded random variables from Hoeffding's inequality) to show that for any finite set $A \subset \mathbb{R}^n$:

   $$\widehat{R}_n(A) \leq \frac{cR}{n}\sqrt{\log|A| + \log n}$$

   where $R = \sup_{a\in A} ||a||_2$ and where c is some constant (independent of R, A and n. Here we define:

   $$\widehat{R}_n(A) = \mathbb{E}_\sigma \left[\sup_{a\in A} \left| \frac{1}{n} \sum_{i=1}^{n} \sigma_i a_i \right| \right]$$

   (b) When $\mathcal{F}$ is a finite set of binary classifiers, use part (a) to obtain a uniform bound on the 0-1 risk.

   (c) When $\mathcal{F}$ is a set of binary classifiers with VC-dimension d, use part(a) to obtain a uniform bound on the 0-1 risk.