Homework Assignment 2
Due Friday March 3rd midnight

SDS 384-11 Theoretical Statistics

1. Remember Hoeffding's Lemma? We proved it with a weaker constant in class using a symmetrization type argument. Now we will prove the original version. Let X be a bounded r.v. in $[a, b]$ such that $E[X] = \mu$. Let $f(\lambda) = \log E[e^{\lambda(X-\mu)}]$. Show that $f''(\lambda) \leq (b - a)^2/4$. Now use the fundamental theorem of calculus to write $f(\lambda)$ in terms of $f''(\lambda)$ and finish the argument.

2. Bernstein's inequality for bounded i.i.d sequences of random variables $\{X_i\}$ with $|X_i| \leq M$ gives: $P(|\sum_i (X_i - E[X_i])| \geq t) \leq 2 \exp \left(\frac{-t^2/2}{\sum_i \text{var}(X_i) + Mt/3} \right)$. There is another better inequality called Bennett's inequality, which we will prove here.

   (a) Consider zero mean r.v.s X_i such that $|X_i| \leq b$ and $\text{var}(X_i) = \sigma_i^2$. Prove that

   $$\log E[\exp(\lambda X_i)] \leq \sigma_i^2 \lambda^2 \left(\frac{e^{\lambda b} - 1 - \lambda b}{(\lambda b)^2} \right) \quad \forall \lambda \in \mathbb{R}.$$

   (b) Given independent r.v.s $X_i, i = 1, \ldots, n$ satisfying the above condition prove

   $$\text{(Bennett's inequality)} \qquad P \left(\sum_i X_i \geq n\delta \right) \leq \exp \left(-\frac{n\sigma^2}{b^2} h(b\delta/\sigma^2) \right),$$

   where $\sigma^2 = \frac{\sum_i \sigma_i^2}{n}$ and $h(t) := (1 + t) \log(1 + t) - t$ for $t \geq 0$.

   (c) Show that Bennett's inequality is at least as good as Bernstein's inequality.

3. Given a scalar random variable X, suppose that there are positive constants c_1, c_2 such that,
   $$P(|X - E[X]| \geq t) \leq c_1 \exp(-c_2 t^2) \qquad \forall t \geq 0.$$

   (a) Prove that $\text{var}(X) \leq \frac{c_1}{c_2}$

   (b) A median m_X is any number such that $P(X \geq m_X) \geq 1/2$ and $P(X \leq m_X) \geq 1/2$. Show by example that the median does not need to be unique.

   (c) Show that if the mean concentration bound holds, then for any median m_X, $\exists$ positive constants c_3, c_4 such that

   $$P(|X - m_X| \geq t) \leq c_3 \exp(-c_4 t^2),$$

Extra credit Conversely, show that whenever the above median concentration holds, then mean concentration holds with positive constants $c_1 = 2c_3$ and $c_2 = c_4/8$. (Note the change from $c_4/4$ to $c_4/8$). We will prove this using the following steps.

i. Create a iid copy Y of X. Now show that $\forall \lambda > 0$,

$$P(|X - EX| \geq t) \leq \exp(-\lambda^2 t^2) E_{X,Y} \exp(\lambda^2 (X - Y)^2)$$

   Hint: use Jensen's inequality

ii. Now show that for a suitable choice of λ,

$$E[\exp(\lambda^2 (X - Y)^2)] \leq 2c_3$$

   Hint: use the trick to go from tail bound to moment bound. Use $P(|X - Y| \geq t) \leq P(|X - m_X| \geq t/2) + P(|Y - m_Y| \geq t/2)$, where $m_X = m_Y$ are medians of X and Y.

iii. Now finally put the above two parts to show the final result.

4. Given a positive semidefinite matrix $Q \in \mathbb{R}^{n \times n}$, consider $Z = \sum_{i,j} Q_{ij} X_i X_j$. When $X_i \sim N(0,1)$, prove the Hanson-Wright inequality.

$$P\left(Z \geq \text{trace}(Q) + t\right) \leq \exp\left(- \min\left\{c_1 t / \|Q\|_{op}, c_2 t^2 / \|Q\|_F^2\right\}\right),$$

where $\|Q\|_{op}$ and $\|Q\|_F$ denote the operator and frobenius norms respectively. *Hint: The rotation-invariance of the Gaussian distribution and sub-exponential nature of χ^2-variables could be useful.*