Homework Assignment 3

Due March 25th by midnight.

SDS 384-11 Theoretical Statistics

1. Let $\{X_i\}_{i=1}^n$ be an i.i.d. sequence of Bernoulli variables with parameter $\alpha \in (0, 1/2]$, and consider the binomial random variable $Z_n = \sum_i X_i$. We want to prove for any $\delta \in (0, \alpha)$,

$$P(Z_n \leq \delta n) \leq \exp(-nKL(\delta||\alpha)) \qquad KL(\delta||\alpha) := \delta \log \frac{\delta}{\alpha} + (1-\delta)\log\frac{1-\delta}{1-\alpha}$$

where $KL(p, q)$ is the Kullback-Leibler divergence between two bernoullis with parameters p, q respectively. Show that the above is strictly better than Hoeffding's inequality.

2. Now we will prove a lower bound on the binomial tail to show that indeed what you derived in the last question is sharp upto polynomial factors. Define $m = \lfloor n\delta \rfloor$ and $\delta' = \frac{m}{n}$.

   (a) Prove $\frac{1}{n}\log P(Z_n \leq \delta n) \geq \frac{1}{n}\log\binom{n}{m} + \delta' \log\alpha + (1-\delta')\log(1-\alpha)$.

   (b) Show that

$$\frac{1}{n}\log\binom{n}{m} \geq -\delta'\log\delta' - (1-\delta')\log(1-\delta') - \frac{\log(n+1)}{n}$$

   Hint: Use the fact that for $Y \sim Bin(n, m/n)$ $P(Y = k)$ is maximized at $k = m$.

   (c) Now show that

$$P(Z_n \leq \delta n) \geq \frac{1}{n+1}\exp(-nKL(\delta'||\alpha))$$

3. We will use the Efron Stein inequality to obtain bounds of variances for separately convex functions whose partial derivatives exist. A separately convex function $f(x_1, \dots, x_n)$ is a convex function of its i^{th} variable, when all else are held fixed.

   (a) Let $X_1, \dots, X_n$ be independent random variables taking values in the interval $[0, 1]$ and let $f : [0, 1]^n \to R$ be a separately convex function whose partial derivatives exist. Then $f(X) := f(X_1, \dots, X_n)$ satisfies

$$\mathrm{var}(f(X)) \leq E[\|\nabla f(X)\|^2]$$

   Hint: Recall that var$(Z) \leq \sum_i E(Z - E_i Z)^2 \leq \sum_i E(Z - Z_i)^2$, where $E_i[Z] = E[Z|X_{1:i-1}, X_{i+1:n}]$. Define $Z_i = \inf_x f(X_{1:i-1}, x, X_{i+1:n})$ and then use convexity of f.

(b) Let A be a $m \times n$ random matrix with independent entries $A_{ij} \in [0,1]$. Let

$$Z = \sqrt{\lambda_1(A^T A)} = \sqrt{\sup_{u \in R^n : \|u\|=1} u^T A^T A u} = \sup_{u \in R^n : \|u\|=1} \|Au\|$$

Show that $\mathrm{var}(Z) \leq 1$.

4. In this question we will look at the Gaussian Lipschitz theorem. Consider $X_1, \ldots, X_n \overset{iid}{\sim} N(0,1)$

(a) Prove that the order statistics are 1-Lipschitz.

(b) Now show that

$$c\sqrt{\log n} \leq E[\max_i X_i] \leq \sqrt{2 \log n}$$

where c is some universal constant.

5. In class we proved McDiarmid's inequality for bounded random variables. But now we will look at extensions for unbounded R.V's. Take a look at "Concentration in unbounded metric spaces and algorithmic stability" by Aryeh Kontorovich, `https://arxiv.org/pdf/1309.1007.pdf`. Reproduce the proof of theorem 1. The steps of this proof is very similar to the martingale based inequalities we looked at in class.