36-788, Fall 2015
Homework 1

Due Sep 17.

1. (Mill's ratio). Let $\Phi\colon \mathbb{R} \to [0,1]$ the c.d.f. of the standard Gaussian distribution on $\mathbb{R}$ and ϕ its p.d.f..

   (a) Prove that, for all $x > 0$,
   $$\frac{x}{1+x^2}\phi(x) \le \Phi(x) \le \frac{1}{x}\phi(x)$$

   (b) Prove that, for all $x > 0$,
   $$\Phi(x) \le \frac{1}{2}\exp\left(-x^2/2\right).$$

2. Let $X = (X_1, \ldots, X_d) \in \mathbb{R}^d$ be a random vector with covariance matrix Σ such that $\frac{X_i}{\sqrt{\Sigma_{i,i}}}$ is sub-Gaussian with parameter ν^2, for all $i = 1, \ldots, d$. Assume we observe n i.i.d. copies of X and compute the empirical covariance matrix $\widehat{\Sigma}$. Show that, for all $i, j \in \{1, \ldots, d\}$,
   $$\mathbb{P}\left(\left|\widehat{S}_{i,j} - \Sigma_{i,j}\right| > \epsilon\right) \le C_1 e^{-\epsilon^2 n C_2},$$
   for some constants C_1 and C_2. Conclude that
   $$\max_{i,j}\left|\widehat{S}_{i,j} - \Sigma_{i,j}\right| = O_P\left(\sqrt{\frac{\log d}{n}}.\right)$$
   You may want to consult these references:

   - Lemma 12 in Yuan. M. (2010). High Dimensional Inverse Covariance Matrix Estimation via Linear Programming, JMLR, 11, 2261-2286.
   - Lemma 1 in Ravikumar, P., Wainwright, M.J., Raskutti, G. and Yu, B. (2011). EJS, 5, 935-980.
   - Lemma A.3 in Bickel, P.J. and Levina, E. (2008). Regularized estimation of large covariance matrices, teh Annals of Statistics, 36(1), 199-227.

3. (Sampling with replacement). Let $\mathcal{X}$ a finite set with N elements. Let $X_1, \ldots, X_n$ be a random sample without replacement from $\mathcal{X}$ and $Y_1, \ldots, Y_n$ be a random sample with replacement from $\mathcal{X}$. Show that, for any convex function $f\colon \mathbb{R}^n \to \mathbb{R}$,
   $$\mathbb{E}\left[f\left(\sum_{i=1}^n X_i\right)\right] \le \mathbb{E}\left[f\left(\sum_{i=1}^n Y_i\right)\right].$$
   Use this result to show that all the inequalities derived for the sums of independent random variables $\{Y_1, \ldots, Y_n\}$ using Chernoff's bounding techniques remain true also for the sums of the X_i's. (see Hoeffding, W. (1963). Probility Inequalities for sums of Bounded Random Variables, by W. Hoeffding, *JASA*, 58, 13–30., 1963).

4. (Moments versus Cernoff bounds). Show that moment bounds for tail prob- abilities are always better than CrameérChernoff bounds. More precisely, let Y be a nonnegative random variable and

let $t > 0$. The best moment bound for the tail probability $\mathbb{P}Y \geq t$ is $\min_q \mathbb{E}[Y^q]t^q$ where the minimum is taken over all positive integers. The best CramérChernoff bound is $\inf_{\lambda > 0} \mathbb{E}e^{\lambda(Yt)}$. Prove that

$$\min_q \mathbb{E}[Y^q]t^q \leq \inf_{\lambda > 0} \mathbb{E}e^{\lambda(Yt)}.$$

(See Philips, T.K. and Nelson, R. (1995). The moment bound is tighter than Chernoffs bound for positive tail probabilities. *The American Statistician*, 49, 175-178.)