# Computational and Statistical Learning theory

## Problem set 1

Recommended Completion: January 5th; Due: January 10th
Email solutions to : karthik *at* ttic *dot* edu

1. **(Do, but do not submit)** $\{0,1\}$-valued random variables $X_1, \ldots, X_n$ are drawn independently each from Bernoullis distribution with parameter $p = 0.1$. Define $P_n := \mathbb{P}(\frac{1}{n}\sum_{i=1}^n X_i \leq 0.2)$.

   (a) For $n = 1$ to $30$ calculate and plot the below in the same plot :

      i. Exact value of $P_n$ (binomial distribution).

      ii. Normal approximation for $P_n$.

      iii. Hoeffding inequality bound on $P_n$.

      iv. Bernstein inequality bound on $P_n$.

   (b) For $n = 30$ to $300$ calculate and plot the below in the same plot :

      i. Normal approximation for $P_n$.

      ii. Hoeffding inequality bound on $P_n$.

      iii. Bernstein inequality bound on $P_n$.

2. Consider i.i.d. random variables $X_1, \ldots, X_n$ distributed according to each of the distributions specified below. For each distribution, write down the bound on $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}[X]\right| > \theta\right)$ implied by each of: (1) Markov inequality (2) Chebychev inequality (3) Hoeffding inequality and (4) Bernstein inequality (some of the inequalities might not be inapplicable for some of the distributions)

   (a) $\{1, \ldots, 4\}$-valued random variables $X_i$ such that $\mathbb{P}(X_i = a) \propto a^2$.

   (b) $X_i$ is distributed uniformly on interval $[-1, 1]$.

   (c) $X_i$ drawn from normal distribution $N(1, 4)$.

   (d) $X_i$ drawn from exponential distribution with parameter 2.

3. (a) $X_1, \ldots, X_n$ are i.i.d. random variables in $[0, 1]$ with mean $\mathbb{E}[X_i] \geq \mu$. Show that:

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i = 0\right) \leq e^{-\mu n}$$

   (Hint : for $t > 1$, $(1 - 1/t)^t \leq 1/e$).

(b) $X_1, \dots, X_n$ are $[0,1]$-valued random variables drawn i.i.d. from some arbitrary distribution. Using each of the inequalities below, write down the implied bound $\epsilon(n, \hat{\mu}, \delta)$ such that, for any $\delta > 0$:

$$\mathbb{P}\left(\mathbb{E}\left[X\right] \leq \hat{\mu} + \epsilon(n, \hat{\mu}, \delta)\right) \geq 1 - \delta$$

where $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
(1) Hoeffding inequality
(2) Markov inequality
(3*) Bernstein Inequality
(4) The inequality from part (a). Note that in this case we only get a bound when $\hat{\mu} =$ and so $\epsilon(n, \hat{\mu}, \delta)$ will be of the form:

$$\epsilon(n, \hat{\mu}, \delta) = \begin{cases} 1 & \text{if } \hat{\mu} > 0 \\ \textit{something} & \text{if } \hat{\mu} = 0 \end{cases}$$

Optional: (5) Chebychev Inequality
Observe the differences in the dependence on $\delta$ when using Markov (or Chebychev) as opposed to the exponential inequalities. Also observe how Bernstein's inequality interpolates nicely between the case $\hat{\mu} = 0$ and the simpler Hoeffding inequality.