Homework 4 40pt (Math 466, Spring 2025)

Due Date: Apr. 17, Thur, 10 pm

Instructions

- Submissions must be made in PDF format. If there are coding problems, please provide source code that directly generates the reported results.
- In written problems, if no derivation is asked for, just give the answer. Otherwise, try to give the derivation in a clear and mathematically rigorous way.
- For coding problems, please include a mini-report summarizing your experiment, along with any relevant figures or tables.

  When creating plots or tables, ensure they are clear and easy to interpret by using different colors or markers, and by adding axis labels, titles, captions, etc. to enhance readability.

  You are also required to submit the code used to generate the results presented in your mini-report. This can be provided as either a .ipynb or .py file.
- Collaboration and AI tools are allowed, but you must write your solutions independently and acknowledge collaborators/AI tools.

Questions

1. **(20 pt) This problem is about the proof of concentration inequality.**

   Consider identical independent Bernoulli variables $X_1, X_2, \cdots, X_n$ with $\mathbb{P}(X_i = 1) = p$ and $\mathbb{P}(X_i = 0) = 1 - p$. Let $X = \sum_{i=1}^{n} X_i$.

   Follow (a)-(d) to prove the upper tail of the Chernoff bound.

   Credits are assigned independently for (a)-(d). As an example, incorrect/incomplete proof of (a) will not result in points loss in (b) if you prove (b) correctly. Nontrivial steps in reasoning must be explained.

   (a) (5pt) Prove that for any $x \in \mathbb{R}$, $1 + x \leq e^x$. As a consequence, prove that $\mathbb{E}(e^{sX_i}) \leq \exp\left(p(e^s - 1)\right)$ for any $s \in \mathbb{R}$.

   (b) (6pt) Based on (a), prove that for any $\lambda > 0$, $\mathbb{P}(X \geq np+\lambda) \leq \exp\left(np(e^s - 1) - s(np + \lambda)\right)$ holds for any $s > 0$.

(c) (6pt) Denote $\mu = np, \delta = \frac{\lambda}{\mu}$. Based on (b), prove that $\mathbb{P}(X \geq np + \lambda) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu$.

(d) (3pt) In class we proved that for any $x > 0$, $(1+x)\ln(1+x) \geq x\left(1 + \frac{1}{\frac{2}{x}+\frac{2}{3}}\right)$. Use this inequality and (c), prove that $\mathbb{P}(X \geq np + \lambda) \leq \exp\left(-\frac{\lambda^2}{2np+\frac{2}{3}\lambda}\right)$.

2. **(5 pt) This problem is about Hoeffding and Chernoff bounds.**

Consider $Y_1, \cdots, Y_n$ be independent Bernoulli random variables with $\Pr(Y_i = 1) = p$. Define $S_n = \sum_{i=1}^n Y_n$. In class, we show a version of Chernoff bound. Since Bernoulli random variable is bounded, one can apply Hoeffding's inequality to obtain a large deviation bound of S_n.

Please give that bound, and compare this Hoeffding's bound to the Chernoff bound. Which bound is better, and under what condition? Give your answer and explain.

Hint: it may be helpful to consider the smaller p *and larger* p *and the near (smaller* α*) and far (larger* α*) tail regimes.*

3. **(15 pt) This problem is about VC dimension.**

(a) (2pt) Let $\mathcal{A}$ be finite. Show that $\mathrm{VC}(\mathcal{A}) \leq \log_2 \operatorname{card} \mathcal{A}$.

(b) (3pt) Given an example of $\mathcal{A}$ with $VC(\mathcal{A}) = 1$ and $\operatorname{card} \mathcal{A} = +\infty$.

(c) (10pt) Let $\mathcal{A}$ be the set of all half-spaces of $\mathbb{R}^d$, i.e. any $A \in \mathcal{A}$ can be represented by $\{x \in \mathbb{R}^d : w^\top x + b \geq 0\}$ for some $w \in \mathbb{R}^d, b \in \mathbb{R}$. Show that $\mathrm{VC}(\mathcal{A}) = d + 1$.