Homework 2: 1 March 2022
CS-526 Learning Theory

This homework goes step by step through the proof of *Hoeffding's inequality*:

Hoeffding's inequality. *Let $Z_1, \dots, Z_m$ be independent random variables. Assume that $Z_i \in [a, b]$ for every i. Then, for any $\epsilon > 0$, we have:*

$$\mathbb{P}\left(\frac{1}{m}\sum_{i=1}^{m} Z_i - \mathbb{E}[Z_i] \geq \epsilon\right) \leq \exp\left(-\frac{2m\epsilon^2}{(b-a)^2}\right).$$

1. Let $\lambda > 0$. Let X be a random variable such that $a \leq X \leq b$ and $\mathbb{E}[X] = 0$. By considering the convex function $x \mapsto e^{\lambda x}$, show that

$$\mathbb{E}[e^{\lambda X}] \leq \frac{b}{b-a}e^{\lambda a} - \frac{a}{b-a}e^{\lambda b}. \tag{1}$$

2. Let $p = -a/(b-a)$ and $h = \lambda(b-a)$. Verify that the right-hand side of (**??**) equals $e^{L(h)}$ where

$$L(h) = -hp + \log\left(1 - p + pe^h\right).$$

3. By Taylor's theorem, there exists $\xi \in (0, h)$ such that

$$L(h) = L(0) + hL'(0) + \frac{h^2}{2}L''(\xi).$$

   Show that $L(h) \leq h^2/8$ and hence $\mathbb{E}[e^{\lambda X}] \leq e^{\lambda^2(b-a)^2/8}$.

4. Let $Z_1, \dots, Z_m$ be independent random variables such that $a \leq Z_i \leq b$ for every i. Using Markov's inequality and the above, show that for every $\lambda > 0$ and $\epsilon > 0$:

$$\mathbb{P}\left(\frac{1}{m}\sum_{i=1}^{m} Z_i - \mathbb{E}[Z_i] \geq \epsilon\right) \leq \exp\left(-\lambda\epsilon + \frac{\lambda^2(b-a)^2}{8m}\right).$$

   We recall *Markov's inequality*: if X is a non-negative random variable and $c > 0$ then $\mathbb{P}(X \geq c) \leq \mathbb{E}[X]/c$.

5. Finally, show that

$$\mathbb{P}\left(\frac{1}{m}\sum_{i=1}^{m} Z_i - \mathbb{E}[Z_i] \geq \epsilon\right) \leq \exp\left(-\frac{2m\epsilon^2}{(b-a)^2}\right).$$