Exericse: 26 February 2019

CS-526 Learning Theory

In this exercise we review the proof of Hoeffding's inequality.

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 $f(X) = e^{\lambda X}$, 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 (1) equals $e^{L(h)}$, where
$$L(h) = -hp + \log(1 - p + pe^h).$$

3. Taylor's theorem tells that
$$L(h) = L(0) + hL'(0) + \frac{h^2}{2} L''(\xi)$$
for some $\xi \in (0, h)$. Show that $L(h) \leq h^2/8$ and hence $\mathbb{E}[e^{\lambda X}] \leq e^{-\lambda^2 (a-b)^2/8}$.

4. Let $Z_1, \ldots, Z_m$ be i.i.d. random variables such that $a \leq Z_i \leq b$ and $\mathbb{E}[Z_i] = \mu$. Using Markov's inequality and the above, show that for every $\lambda > 0$ and $\epsilon > 0$
$$\mathbb{P}\big(\frac{1}{m} \sum_{i=1}^{m} Z_i - \mu \geq \epsilon\big) \leq \exp\big(-\lambda\epsilon + \frac{\lambda^2 (b-a)^2}{8m}\big).$$
(Recall the statement of 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. Show that
$$\mathbb{P}\big(\frac{1}{m} \sum_{i=1}^{m} Z_i - \mu \geq \epsilon\big) \leq \exp\big(-\frac{2m\epsilon^2}{(b-a)^2}\big).$$