Exercise Sheet A

Machine Learning SS 2017

Gregor Bankhamer Wolfgang Kremser

Exercise 1

We bound

$$\mathbb{P}[|X - \mu| \geq 10]$$

by using Chebychev's inequality

$$\mathbb{P}[|Z - \mathbb{E}[Z]| \geq a] \leq \frac{\mathbb{V}[Z]}{a^2}, \quad \forall a > 0$$

$X = X_1 + X_2 + ... + X_{10}$ with X_i independent and uniformly distributed over 1 to 6.
$\mu_{X_i} = \frac{1+6}{2} = 3.5$ and $\sigma^2_{X_i} = \frac{(6-1)^2}{12} = \frac{25}{12} = 2.083$

Because X_i are i.d.d, $\mathbb{E}[X] = \sum \mu_{X_i} = 35$ and $\mathbb{V}[X_i] = \sum \sigma^2_{X_i} = 20.83$

Plugging the values into Chebychev's inequality with $a = 10$ we get

$$\mathbb{P}[|Z - 35| \geq 10] \leq \frac{20.83}{10^2} = 0.208$$

Exercise 2

A boolean input x_i can have 2 possible values. We have 4 boolean inputs, giving 2^4 input variations. Since the function map to boolean values, we again have 2 possible outputs for each function.
This results in a total of $2^{2^4} = 65536$ functions that map 4 boolean inputs to 1 boolean output. The number of functions we have to consider increases exponentially with the number of input variables.

Exercise 3

We show:

$$\mathbb{E}_{S|_x \sim D^m}[L_S(h)] = L_{(D,f)}(h)$$

Proof.

$$\begin{aligned}
\mathbb{E}_{S|_x \sim D^m}[L_S(h)] &= \mathbb{E}_{S|_x \sim D^m}\left[\frac{1}{m}\sum_{i=1}^{m} 1_{h(x_i)\neq f(x_i)}\right] && \text{(definition } L_S(h)\text{)} \\
&= \frac{1}{m}\sum_{i=1}^{m} \mathbb{E}_{S|_x \sim D^m}\left[1_{h(x_i)\neq f(x_i)}\right] && \text{(linearity)} \\
&= \frac{1}{m}\sum_{i=1}^{m} \mathbb{E}_{S|_x \sim D^m}\left[1_{h(x)\neq f(x)}\right] && \text{(i.d.d.)} \\
&= m \cdot \frac{1}{m} \cdot \mathbb{E}_{S|_x \sim D^m}\left[1_{h(x_i)\neq f(x_i)}\right] && \text{(sum of constants)} \\
&= \mathbb{E}_{S|_x \sim D^m}\left[1_{h(x_i)\neq f(x_i)}\right] \\
&= L_{(D,f)}(h)
\end{aligned}$$

□

Exercise 4

Exercise 5