Exercise Sheet 5

Exercise 1

Let $\mathcal{X} \times \mathcal{Y}$ be the input-output space with $\mathcal{Y} = \{\pm 1\}$. Suppose that

$$P(y = -1) = P(y = +1) = 0.5.$$

Let $\mathcal{F}$ consist of all (measurable) functions $f : \mathcal{X} \to \{\pm 1\}$ Let $\mathbb{S}$ be the set of all finite training sets. Consider the hypothesis class $\mathcal{H} = \{f_{\mathcal{S}} \in \mathcal{F} : \mathcal{S} \in \mathbb{S}\}$, where

$$f_{\mathcal{S}}(x) = \begin{cases} y & : & (x, y) \in \mathcal{S} \\ 1 & : & \text{otherwise} \end{cases}.$$

Show that $\mathcal{H}$ is not uniformly convergent.

Exercise 2

Consider the following setting: The input space is $\mathcal{X} = [0, 1]$ and the output space is $\mathcal{Y} = \{\pm 1\}$. The input data $x \in \mathcal{X}$ is uniformly distributed. The Bayes classifier

$$f^*(x) = \begin{cases} +1 & : & x \geq 0.3 \\ -1 & : & x < 0.3 \end{cases}.$$

has zero Bayes risk $R^* = R(f^*) = 0$. The hypothesis class

$$\mathcal{H} = \left\{f_+, f_-\right\}$$

consists of two classifiers $f_+(x) = +1$ and $f_-(x) = -1$ for all $x \in \mathcal{X}$.

Is the Empirical Risk Minimization principle consistent with respect to $\mathcal{H}$ and the uniform distribution over $\mathcal{X}$ in the sense of lecture 1, slide 35?

Exercise 3

Construct a classification problem and a finite hypothesis class $\mathcal{H}$ to illustrate that the probability

$$p(\varepsilon) = \mathbb{P}\left(\sup_{f \in \mathcal{H}} |R_n(f) - R(f)| \geq \varepsilon\right)$$

violates Hoeffding's inequality but satisfies the union bound. Proceed as follows:

1. Generate a large population dataset of size N as ground truth
2. Construct a finite hypothesis class $\mathcal{H}$
3. Compute the true risk for all $f \in \mathcal{H}$
4. Choose a small sample size n
5. Repeat T times:

   (a) generate a training set of size n
   (b) compute the empirical risk for all $f \in \mathcal{H}$

6. Estimate the probabilities $p(\varepsilon)$ as a function of ε.
7. Plot the distribution of deviations $|R_n(f) - R(f)|$ using a violin plot.
8. Plot the Hoeffding and the union bound together with the estimates of $p(\varepsilon)$ as a function of ε.