Statistical inference practice

Joris LIMONIER

January 12, 2022

Contents

1 Inclass exercise January 12, 2022

1.1 Exercise 1

Show that

$$\mathbb{E}\left[\hat{\mathcal{R}}_S(h)\right] = \mathcal{R}_{D,f}(h) \tag{1}$$

$$\begin{aligned}
\mathbb{E}\left[\hat{\mathcal{R}}_S(h)\right] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{h(x_i)\neq y_i}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mathbf{1}_{h(x_i)\neq y_i}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}\left(h(x_i) \neq y_i\right) \\
&= \frac{1}{n} n \mathbb{P}\left(h(x_i) \neq y_i\right) \\
&= \mathbb{P}\left(h(x_i) \neq y_i\right) \\
&= \mathbb{P}\left(h(x_i) \neq f(x)\right) \\
&= \mathcal{R}_{D,f}(h)
\end{aligned}$$

1.2 Exercise 2

We must prove that the variance of $\hat{\mathcal{R}}_S(h) \to 0$

$$
\begin{aligned}
Var\left[\hat{\mathcal{R}}_S(h)\right] &= Var\left[\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{h(x_i)\neq y_i}\right] \\
&= Var\frac{1}{n^2}\left[\sum_{i=1}^{n} \mathbf{1}_{h(x_i)\neq y_i}\right]
\end{aligned}
$$

Let the Z_i be defined as follows:

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{h(x_i)\neq f(x_i)} =: \frac{1}{n}\sum_{i=1}^{n} Z_i$$

(not finished, see lecture 1 slides)