Statistical inference practice

Joris LIMONIER

January 21, 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)

2 Inclass exercise January 21, 2022

2.1 Exercise 1

Set $g(x) = \mathbb{P}(Y = 1 \mid X = x)$. We define the Bayes optimal predictor as:

$$
f_{\mathcal{D}}(x) = \begin{cases} 1 & g(x) \geq 1/2 \\ 0 & \text{otherwise} \end{cases}
$$

Question 1. Let $h : \mathcal{X} \to \{0, 1\}$ be a classifier. Show that

$$
\begin{aligned}
&\mathbb{P}(h(X) \neq Y \mid X = x) \\
&= g(x) \cdot \mathbb{P}(h(X) = 0 \mid X = x)) + (1 - g(x)) \cdot \mathbb{P}(h(X) = 1 \mid X = x)
\end{aligned}
$$

$$
\begin{aligned}
&g(x) \cdot \mathbb{P}(h(X) = 0 \mid X = x)) + (1 - g(x)) \cdot \mathbb{P}(h(X) = 1 \mid X = x) \\
&= \mathbb{P}(Y = 1 \mid X = x) \cdot \mathbb{P}(h(X) = 0 \mid X = x)) \\
&+ (1 - \mathbb{P}(Y = 1 \mid X = x)) \cdot \mathbb{P}(h(X) = 1 \mid X = x) \\
&= \mathbb{P}(Y = 1 \cap h(X) = 0 \mid X = x)) \\
&+ \mathbb{P}(h(X) = 1 \mid X = x) - \mathbb{P}(Y = 1 \cap h(X) = 1 \mid X = x) \\
&= \mathbb{P}(Y = 1 \cap h(X) = 0 \mid X = x)) + \mathbb{P}(Y = 0 \cap h(X) = 1 \mid X = x) \\
&= \mathbb{P}(h(X) \neq Y \mid X = x)
\end{aligned}
$$

Question 2. Deduce that

$$
\mathbb{P}(f_D(X) \neq Y \mid X = x) = \min(g(x), 1 - g(x))
$$

$$
\begin{aligned}
&\mathbb{P}(f_D(X) \neq Y \mid X = x) \\
&= \begin{cases} \mathbb{P}(1 \neq Y \mid X = x), & g(x) \geq 1/2 \\ \mathbb{P}(0 \neq Y \mid X = x), & g(x) < 1/2 \end{cases} \\
&= \begin{cases} 1 - g(x), & g(x) \geq 1 - g(x) \\ g(x), & g(x) < 1 - g(x) \end{cases} \\
&= \min(g(x), 1 - g(x))
\end{aligned}
$$

Question 3. Show that

$$\mathbb{P}(h(X) \neq Y \mid X = x) \geq \mathbb{P}(f_D(x) \neq Y \mid X = x)$$

$$
\begin{aligned}
\mathbb{P}(f_D(x) \neq Y \mid X = x) &= \min(g(x), 1 - g(x)) \\
&= \min(g(x), 1 - g(x)) \\
&\cdot (\mathbb{P}(h(X) = 0 \mid X = x) + \mathbb{P}(h(X) = 1 \mid X = x)) \\
&\leq g(x) \cdot (\mathbb{P}(h(X) = 0 \mid X = x) \\
&+ (1 - g(x)) \cdot \mathbb{P}(h(X) = 1 \mid X = x)) \\
&= \mathbb{P}(h(X) \neq Y \mid X = x)
\end{aligned}
$$

Question 4. Prove that

$$\mathcal{R}_{\mathcal{D}}(f_{\mathcal{D}}) \leq \mathcal{R}_{\mathcal{D}}(h)$$

$$
\begin{aligned}
&\mathbb{P}(f_D(x) \neq Y \mid X = x) \leq \mathbb{P}(h(X) \neq Y \mid X = x) \\
\implies &\mathbb{E}\left[\mathbb{P}(f_D(x) \neq Y \mid X = x)\right] \leq \mathbb{E}\left[\mathbb{P}(h(X) \neq Y \mid X = x)\right] \\
\implies &\mathcal{R}_{\mathcal{D}}(f_{\mathcal{D}}) \leq \mathcal{R}_{\mathcal{D}}(h)
\end{aligned}
$$