

Statistical inference practice

Joris LIMONIER

January 28, 2022

Contents

1	Inclass exercise January 12, 2022	1
1.1	Exercise 1	1
1.2	Exercise 2	2
2	Inclass exercise January 21, 2022	2
2.1	Exercise 1	2
3	Inclass exercise January 28, 2022	4
3.1	Exercise 1	4
3.1.1	Question 1	4
3.1.2	Question 2	4
3.1.3	Question 3	5
3.1.4	Question 4	6

1 Inclass exercise January 12, 2022

1.1 Exercise 1

Show that

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

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

1.2 Exercise 2

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

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

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 $h : \mathcal{X} \rightarrow \{0, 1\}$ 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

$$\mathcal{R}_D(f_D) \leq \mathcal{R}_D(h)$$

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

3 Inclass exercise January 28, 2022

3.1 Exercise 1

Let Z be a random variable with a second moment such that $\mathbb{E}[Z] = \mu$ and $\text{Var}(Z) = \sigma^2$.

3.1.1 Question 1

Let $g : t \mapsto \mathbb{E}[(Z - t)^2]$. Show that g is minimum at $t = \mu$.

$$\begin{aligned} g(t) &= \mathbb{E}[(Z - t)^2] \\ &= \mathbb{E}[Z^2 + t^2 - 2tZ] \\ &= \mathbb{E}[Z^2] + \mathbb{E}[t^2] - \mathbb{E}[2tZ] \\ &= \mathbb{E}[Z^2] + t^2 - 2t\mathbb{E}[Z] \\ &= \sigma^2 - \mu^2 + t^2 - 2t\mu \\ &= \sigma^2 - \mu^2 + t^2 - 2t\mu \end{aligned}$$

We differentiate with respect to t :

$$\begin{aligned} \frac{\partial}{\partial t} g(t) &= 0 \\ \implies \frac{\partial}{\partial t} [\sigma^2 - \mu^2 + t^2 - 2t\mu] &= 0 \\ \implies 2t - 2\mu &= 0 \\ \implies t &= \mu \end{aligned}$$

3.1.2 Question 2

Assume $Z \in [a, b]$ almost surely. Use the previous question to show that

$$\text{Var}(Z) \leq \frac{(b - a)^2}{4}$$

$$\begin{aligned}
& g(\mu) \leq g(t) \\
\implies \text{Var}(Z) & \leq \mathbb{E}[(Z - t)^2] \\
\implies \text{Var}(Z) & \leq \frac{1}{4} \mathbb{E}[(2Z - a - b)^2] \\
\implies \text{Var}(Z) & \leq \frac{1}{4} \mathbb{E}[(Z - a) + (Z - b)]^2 \\
\implies \text{Var}(Z) & \leq \frac{1}{4} \mathbb{E}[(Z - a) - (b - Z)]^2 \\
\implies \text{Var}(Z) & \leq \frac{1}{4} \mathbb{E}[|Z - a| - |b - Z|]^2 \\
\implies \text{Var}(Z) & \leq \frac{1}{4} \mathbb{E}[|(Z - a) - (Z - b)|^2] \\
\implies \text{Var}(Z) & \leq \frac{1}{4} \mathbb{E}[|b - a|^2] \\
\implies \text{Var}(Z) & \leq \frac{(b - a)^2}{4}
\end{aligned}$$

3.1.3 Question 3

Let $Z_1, \dots, Z_n \sim Z$ be i.i.d. Use Chebyshev inequality to obtain a concentration inequality for

$$Z := \frac{1}{n} \sum_{i=1}^n Z_i$$

Chebyshev inequality:

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq a) \leq \frac{\text{Var } Z}{a^2} \quad (2)$$

$$\begin{aligned}
\text{Var}(Z) &= \text{Var}\left(\frac{1}{n^2} \sum_{i=1}^n Z_i\right) \\
&= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(Z_i) \quad (Z_i \text{'s independent}) \\
&\leq \frac{1}{n^2} \sum_{i=1}^n \frac{(b - a)^2}{4} \\
&\leq \frac{(b - a)^2}{4n}
\end{aligned}$$

Then we apply (2):

$$\begin{aligned}\mathbb{P}(|Z - \mathbb{E}[Z]| \geq \varepsilon) &\leq \frac{\text{Var } Z}{\varepsilon^2} \\ \Rightarrow \mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n Z_i - \mu\right| \geq \varepsilon\right) &\leq \frac{(b-a)^2}{4n\varepsilon^2}\end{aligned}$$

3.1.4 Question 4