Empirical Processes

MAA110 - EPFL

Nikitas Georgakis*, Myrto Limnios*

01/04/2025

Notations. Without additional notice, we consider the same notations as in the lecture notes.

Exercise 1 (Massart's Lemma)**.** *Consider T to be finite of cardinality $|T| = N$ and composed of elements valued in a subset of $\mathbb{R}^n$. Let an i.i.d. sequence of Rademacher r.v.s $\varepsilon_1, \dots, \varepsilon_n$, and define the Rademacher complexity of T by $\mathcal{R}(T) = \mathbb{E}[\sup_{t \in T} R_t] = \mathbb{E}[\sup_{t \in T} \sum_{i=1}^n \varepsilon_i t_i]$. If*

$$C = \max_{t \in T} \left(\sum_{i=1}^n t_i^2 \right)^{1/2} < \infty \ ,$$

it holds true that

$$\mathcal{R}(T) \leq C\sqrt{2 \log N} \ .$$

Exercise 2. *Let Φ be $(L > 0)$-Lipschitz, and $\mathcal{H}$ a class of measurable functions. Consider the empirical Rademacher measure P_n^0. Prove that*

$$\mathbb{E}[\|P_n^0\|_{\Phi \circ \mathcal{H}}] \leq 2L\mathbb{E}[\|P_n^0\|_{\mathcal{H}}] + \frac{|\Phi(0)|}{\sqrt{n}} \ ,$$

where $\Phi \circ \mathcal{H} = \{\Phi \circ h, \ h \in \mathcal{H}\}$.

Exercise 3. *Prove Lemma 1.6. based on the notations used in the lecture. Let $u > 0$. Suppose that for any $h \in \mathcal{H}$*

$$\mathbb{P}(|(P_n - P)(h)| > u/2) \leq \frac{1}{2} \ .$$

Then,

$$\mathbb{P}(\|P_n - P\|_{\mathcal{H}} > u) \leq 4\mathbb{P}\left(\|P_n^0\|_{\mathcal{H}} > \frac{u}{4}\right) \ .$$

Hint: Prove

$$\mathbb{P}(\|P_n - P\|_{\mathcal{H}} > u) \leq 2\mathbb{P}\left(\|P_n - P_n'\|_{\mathcal{H}} > \frac{u}{2}\right) \ .$$

*{**first.last**}**@epfl.ch**, Office MA1493, CM1618