EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät

Fachbereich Mathematik

Prof. Dr. Andreas Prohl
Dr. Abhishek Chaudhary

Statistical Learning 2

Summer-Semester 2023 **Tübingen, 11.05.2023**

Homework 2

Problem 1. In the lecture, we currently show a lemma that states the *(uniform) exponential inequality by Pollard from '84*. Recall the notations that we used in **steps 1** of the proof, *i.e.*, of $\mathbf{Z}$, and of $\mathscr{Z}_n$ for an *i.i.d.* n-sample, as well as of $\overline{\mathscr{Z}}_n := \{\overline{\mathbf{Z}}_j\}_{j=1}^n$ for an *i.i.d.* ghost n-sample. With the help of the function $g^* \equiv g^*(\cdot;\mathscr{Z}_n)$ constructed there, and by the assumptions made in the lemma we were then lead to bound

$$\mathbb{P}\left[\left|\mathbb{E}[g^*(\mathbf{Z})|\mathscr{Z}_n] - \frac{1}{n}\sum_{j=1}^n g^*(\overline{\mathbf{Z}}_j)\right| \leq \frac{\varepsilon}{2}\middle|\mathscr{Z}_n\right] \geq \frac{1}{2} \qquad \text{a.s.}$$

Detail the proof outlined in the lecture that leads to this estimation.

<u>Hint</u>: Use Tschebycheff's inequality when you derive an *upper* bound for the probability of the 'complementary set'.

Problem 2. While **Problem 1** relates to the *'symmetrization by a ghost n-sample'* as a first tool to verify the lemma mentioned, another tool is in **step 2** of its proof, which introduces *'additional randomness by random signs'*. This comes with the help of another independent *i.i.d.* n-sample $\mathscr{U}_n := \{U_j\}_{j=1}^n$, where

$$\mathbb{P}[U_j = 1] = \mathbb{P}[U_j = -1] = \frac{1}{2} \qquad (1 \leq j \leq n)\,.$$

Now, the idea was to consider randomly exchanged *same* entries of the $\mathscr{Z}_n$ and $\overline{\mathscr{Z}}_n$, for which we consider the (related) map

$$\mathscr{Y}_n = \boldsymbol{h}\big(\mathscr{Z}_n, \overline{\mathscr{Z}}_n, \mathscr{U}_n\big)\,,$$

where $\mathscr{Y}_n = \{\mathscr{Y}_n^{(j)}\}_{j=1}^n$, with components $\mathscr{Y}_n^{(j)} \equiv \mathbf{h}\big(\mathbf{Z}_j, \overline{\mathbf{Z}}_j, \mathbf{U}_j\big)$, and the function $\mathbf{h}: \mathbb{R}^{2d+1} \to \mathbb{R}^{2d}$ via

$$\mathbf{h}\big(\mathbf{z}, \overline{\mathbf{z}}, u\big) = \Big(\frac{1}{2}[1+u]\mathbf{z} + \frac{1}{2}[1-u]\overline{\mathbf{z}}, \frac{1}{2}[1-u]\mathbf{z} + \frac{1}{2}[1+u]\overline{\mathbf{z}}\Big)^\top\,.$$

Study if the induced laws coincide, *i.e.*,

$$\mathscr{L}\big(\mathscr{Y}_n\big) = \mathscr{L}\big((\mathscr{Z}_n, \overline{\mathscr{Z}}_n)\big)\,.$$

<u>Hint</u>: Consider $n = 1$ first, and then proceed for each entry independently.

Date of Submission: 12.00 on 17.05.2023.