Empirical Processes

MAA110 - EPFL

Nikitas Georgakis*, Myrto Limnios*

27/05/2025

Notations. Without additional notice, we consider the same notations as in the lecture notes. We use the sup-norm for any function $h : \mathcal{X} \to \mathbb{R}$, defined by $\|h\|_\infty = \sup_{t\in\mathcal{X}} |h(t)|$.

Exercise 1. *Consider two independent and i.i.d. sequences $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ resp. drawn from a probability distribution P, Q on a measurable space $\mathcal{X}$ and $k, \ell : \mathcal{X}^2 \to \mathbb{R}$ a square integrable function resp. w.r.t. $P \otimes P$ and $P \otimes Q$.*

1. *Derive the Hoeffding decomposition of the U-statistics defined below*

$$\begin{aligned} U_n(k) &= \frac{1}{n(n-1)} \sum_{1\leq i\neq j\leq n} k(X_i, X_j) \\ U_{n,m}(\ell) &= \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \ell(X_i, Y_j) \end{aligned}$$

2. *Suppose that the class $\mathcal{K}$ of kernel functions $k(x, x')$ is* VC*-type with parameters $(A, \mathcal{V})$, with bounded sup-norm and second moment by a finite constant. Prove a maximal inequality for $U_n(k)$ uniformly over $\mathcal{K}$.*

Exercise 2 (2*nd*-order Gaussian chaos)**.** *Consider $X = (X_1, \ldots, X_n)$ to be a centered Gaussian vector of covariance matrix I_n, and let $A = (a_{i,j})_{i,j\leq n}$ be a symmetric real-valued matrix, s.t., $a_{ii} = 0$ for all $i \leq n$. Define the quadratic form*

$$Z = X^T AX = \sum_{i,j\leq n} a_{i,j} X_i X_j$$

1. *Prove that for all $\lambda \in (0, 1/2)$*

$$\log \mathbb{E}e^{\lambda(X_1^2-1)} \leq \frac{\lambda^2}{1-2\lambda} \tag{1}$$

2. *Prove that Z is a zero-mean r.v. that can be written as a weighted sum of standard Gaussian r.v.s (Hint: Use the spectral decomposition of A)*

3. *Prove that for all $\lambda \in (0, 1/(2\alpha))$*

$$\psi_Z(\lambda) \leq \frac{\lambda^2\|A\|}{1-2\lambda\alpha} \tag{2}$$

   where $\alpha = \max_i |\alpha_i|$ and $\|A\|^2 = \sum_{i\leq n} \alpha_i^2$.

   Hint: Use the rotational invariance of standard Gaussian r.v.s and notice that Z has the same distribution as $\sum_{i=1}^n \alpha_i^2(X_i^2 - 1)$ where the α_i-s are inherited from Q2

4. *Conclude that for any $u > 0$*

$$\mathbb{P}(Z > 2\|A\|\sqrt{u} + 2\alpha u) \leq e^{-t} \tag{3}$$

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