ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Prof. Dr. Helmut Bölcskei
Chair for Mathematical Information Science

Handout
Examination on Mathematics of Information
February 8, 2021

Theorem 1 (Bernstein's inequality). *Let $X_1,\dots,X_m$ be independent complex-valued random variables with zero mean such that $|X_\ell| \le B$, for $\ell \in [m]$ and some constant $B > 0$. Furthermore assume that $\mathbb{E}\left[|X_\ell|^2\right] \le \sigma_\ell^2$, for constants $\sigma_\ell > 0$, $\ell \in [m]$. Then, for all $t > 0$,*

$$\mathbb{P}\left(\left|\sum_{\ell=1}^{m} X_\ell\right| \ge t\right) \le 2\exp\left(-\frac{t^2/2}{\sigma^2 + Bt/3}\right),$$

where $\sigma^2 = \sum_{\ell=1}^{m} \sigma_\ell^2$.

Lemma (One sided bounded difference inequality). *Suppose that $f\colon \left(\mathbb{R}^d\right)^n \to \mathbb{R}$ is such that*

$$|f(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n) - f(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_n)| \le L,$$

for every $i = 1,\dots,n$, every $x_1^n \coloneqq (x_1,\dots,x_n)$, and every $y \in \mathbb{R}^d$. Also suppose that the random vector $X = (X_1,\dots,X_n)$ has i.i.d. components. Then, we have

$$\mathbb{P}\big[\mathbb{E}[f(X)] - f(X) > \epsilon\big] \le e^{-\frac{2\epsilon^2}{nL^2}}, \quad \forall \epsilon \ge 0.$$

Lemma (Ledoux-Talagrand contraction). *Let $\phi\colon \mathbb{R} \to \mathbb{R}$ be an L-Lipschitz function with $\phi(0) = 0$ and $\mathcal{F}$ a class of functions. Let $\phi \circ \mathcal{F} \coloneqq \{\phi \circ f \mid f \in \mathcal{F}\}$. Then, we have*

$$\mathcal{R}_n\left(\left(\phi \circ \mathcal{F}\right)\left(x_1^n\right)/n\right) \le 2L\mathcal{R}_n\left(\mathcal{F}\left(x_1^n\right)/n\right).$$

Lemma (Vapnik-Chervonenkis, Sauer-Shelah). *Consider a set class $\mathcal{S}$ with finite VC-dimension $\nu < \infty$. Then, for any collection of points $P = (x_1,\dots,x_n)$ with $n \ge \nu$, we have*

$$|\mathcal{S}(P)| \le \sum_{i=0}^{\nu} \binom{n}{i} \le (n+1)^\nu,$$

where $\mathcal{S}(P) \coloneqq \{(\mathbb{1}_S(x_1),\dots,\mathbb{1}_S(x_n)) \mid S \in \mathcal{S}\}$.