

Let  $Z(x_1, \dots, x_n) : \mathcal{X}^n \mapsto \mathbb{R}$ . We would like to bound  $Z - \mathbb{E}Z$ . We will be able to answer this question if for any  $x_1, \dots, x_n, x'_1, \dots, x'_n$ ,

$$(1) \quad |Z(x_1, \dots, x_n) - Z(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)| \leq c_i.$$

Decompose  $Z - \mathbb{E}Z$  as follows

$$\begin{aligned} Z(x_1, \dots, x_n) - \mathbb{E}_{x'} Z(x'_1, \dots, x'_n) &= (Z(x_1, \dots, x_n) - \mathbb{E}_{x'} Z(x'_1, x_2, \dots, x_n)) \\ &\quad + (\mathbb{E}_{x'} Z(x'_1, x_2, \dots, x_n) - \mathbb{E}_{x'} Z(x'_1, x'_2, x_3, \dots, x_n)) \\ &\quad \dots \\ &\quad + (\mathbb{E}_{x'} Z(x'_1, \dots, x'_{n-1}, x_n) - \mathbb{E}_{x'} Z(x'_1, \dots, x'_n)) \\ &= Z_1 + Z_2 + \dots + Z_n \end{aligned}$$

where

$$Z_i = \mathbb{E}_{x'} Z(x'_1, \dots, x'_{i-1}, x_i, \dots, x_n) - \mathbb{E}_{x'} Z(x'_1, \dots, x'_i, x_{i+1}, \dots, x_n).$$

Assume

- (1)  $|Z_i| \leq c_i$
- (2)  $\mathbb{E}_{X_i} Z_i = 0$
- (3)  $Z_i = Z_i(x_i, \dots, x_n)$

**Lemma 22.1.** For any  $\lambda \in \mathbb{R}$ ,

$$\mathbb{E}_{x_i} e^{\lambda Z_i} \leq e^{\lambda^2 c_i^2 / 2}.$$

*Proof.* Take any  $-1 \leq s \leq 1$ . With respect to  $\lambda$ , function  $e^{\lambda s}$  is convex and

$$e^{\lambda s} = e^{\lambda(\frac{1+s}{2}) + (-\lambda)(\frac{1-s}{2})}.$$

Then  $0 \leq \frac{1+s}{2}, \frac{1-s}{2} \leq 1$  and  $\frac{1+s}{2} + \frac{1-s}{2} = 1$  and therefore

$$e^{\lambda s} \leq \frac{1+s}{2} e^{\lambda} + \frac{1-s}{2} e^{-\lambda} = \frac{e^{\lambda} + e^{-\lambda}}{2} + s \frac{e^{\lambda} - e^{-\lambda}}{2} \leq e^{\lambda^2 / 2} + s \cdot \text{sh}(\lambda)$$

using Taylor expansion. Now use  $\frac{Z_i}{c_i} = s$ , where, by assumption,  $-1 \leq \frac{Z_i}{c_i} \leq 1$ . Then

$$e^{\lambda Z_i} = e^{\lambda c_i \cdot \frac{Z_i}{c_i}} \leq e^{\lambda^2 c_i^2 / 2} + \frac{Z_i}{c_i} \text{sh}(\lambda c_i).$$

Since  $\mathbb{E}_{x_i} Z_i = 0$ ,

$$\mathbb{E}_{x_i} e^{\lambda Z_i} \leq e^{\lambda^2 c_i^2 / 2}.$$

□

We now prove McDiarmid's inequality

**Theorem 22.1.** *If condition (1) is satisfied,*

$$\mathbb{P}(Z - \mathbb{E}Z > t) \leq e^{-\frac{t^2}{2\sum_{i=1}^n c_i^2}}.$$

*Proof.* For any  $\lambda > 0$

$$\mathbb{P}(Z - \mathbb{E}Z > t) = \mathbb{P}(e^{\lambda(Z - \mathbb{E}Z)} > e^{\lambda t}) \leq \frac{\mathbb{E}e^{\lambda(Z - \mathbb{E}Z)}}{e^{\lambda t}}.$$

Furthermore,

$$\begin{aligned} \mathbb{E}e^{\lambda(Z - \mathbb{E}Z)} &= \mathbb{E}e^{\lambda(Z_1 + \dots + Z_n)} \\ &= \mathbb{E}\mathbb{E}_{x_1} e^{\lambda(Z_1 + \dots + Z_n)} \\ &= \mathbb{E}\left[e^{\lambda(Z_2 + \dots + Z_n)} \mathbb{E}_{x_1} e^{\lambda Z_1}\right] \\ &\leq \mathbb{E}\left[e^{\lambda(Z_2 + \dots + Z_n)} e^{\lambda^2 c_1^2 / 2}\right] \\ &= e^{\lambda^2 c_1^2 / 2} \mathbb{E}\mathbb{E}_{x_2} \left[e^{\lambda(Z_2 + \dots + Z_n)}\right] \\ &= e^{\lambda^2 c_1^2 / 2} \mathbb{E}\left[e^{\lambda(Z_3 + \dots + Z_n)} \mathbb{E}_{x_2} e^{\lambda Z_2}\right] \\ &\leq e^{\lambda^2 (c_1^2 + c_2^2) / 2} \mathbb{E}e^{\lambda(Z_3 + \dots + Z_n)} \\ &\leq e^{\lambda^2 \sum_{i=1}^n c_i^2 / 2} \end{aligned}$$

Hence,

$$\mathbb{P}(Z - \mathbb{E}Z > t) \leq e^{-\lambda t + \lambda^2 \sum_{i=1}^n c_i^2 / 2}$$

and we minimize over  $\lambda > 0$  to get the result of the theorem. □

**Example 1.** *Let  $\mathcal{F}$  be a class of functions:  $\mathcal{X} \mapsto [a, b]$ . Define the empirical process*

$$Z(x_1, \dots, x_n) = \sup_{f \in \mathcal{F}} \left| \mathbb{E}f - \frac{1}{n} \sum_{i=1}^n f(x_i) \right|.$$

Then, for any  $i$ ,

$$\begin{aligned} & |Z(x_1, \dots, x'_i, \dots, x_n) - Z(x_1, \dots, x_i, \dots, x_n)| \\ &= \left| \sup_f \left| \mathbb{E}f - \frac{1}{n} (f(x_1) + \dots + f(x'_i) + \dots + f(x_n)) \right| \right. \\ &\quad \left. - \sup_f \left| \mathbb{E}f - \frac{1}{n} (f(x_1) + \dots + f(x_i) + \dots + f(x_n)) \right| \right| \\ &\leq \sup_{f \in \mathcal{F}} \frac{1}{n} |f(x_i) - f(x'_i)| \leq \frac{b-a}{n} = c_i \end{aligned}$$

because

$$\sup_t f(t) - \sup_t g(t) \leq \sup_t (f(t) - g(t))$$

and

$$|c| - |d| \leq |c - d|.$$

Thus, if  $a \leq f(x) \leq b$  for all  $f$  and  $x$ , then, setting  $c_i = \frac{b-a}{n}$  for all  $i$ ,

$$\mathbb{P}(Z - \mathbb{E}Z > t) \leq \exp\left(-\frac{t^2}{2 \sum_{i=1}^n \frac{(b-a)^2}{n^2}}\right) = e^{-\frac{nt^2}{2(b-a)^2}}.$$

By setting  $t = \sqrt{\frac{2u}{n}}(b-a)$ , we get

$$\mathbb{P}\left(Z - \mathbb{E}Z > \sqrt{\frac{2u}{n}}(b-a)\right) \leq e^{-u}.$$

**Example 2.** Let  $\varepsilon_1, \dots, \varepsilon_n$  be i.i.d. such that  $\mathbb{P}(\varepsilon = \pm 1) = \frac{1}{2}$ . Define

$$Z((\varepsilon_1, x_1), \dots, (\varepsilon_n, x_n)) = \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n \varepsilon_i f(x_i) \right|.$$

Then, for any  $i$ ,

$$\begin{aligned} & |Z((\varepsilon_1, x_1), \dots, (\varepsilon'_i, x'_i), \dots, (\varepsilon_n, x_n)) - Z((\varepsilon_1, x_1), \dots, (\varepsilon_i, x_i), \dots, (\varepsilon_n, x_n))| \\ &\leq \sup_{f \in \mathcal{F}} \left| \frac{1}{n} (\varepsilon'_i f(x'_i) - \varepsilon_i f(x_i)) \right| \leq \frac{2M}{n} = c_i \end{aligned}$$

where  $-M \leq f(x) \leq M$  for all  $f$  and  $x$ .

Hence,

$$\mathbb{P}(Z - \mathbb{E}Z > t) \leq \exp\left(-\frac{t^2}{2 \sum_{i=1}^n \frac{(2M)^2}{n^2}}\right) = e^{-\frac{nt^2}{8M^2}}.$$

By setting  $t = \sqrt{\frac{8u}{n}}M$ , we get

$$\mathbb{P}\left(Z - \mathbb{E}Z > \sqrt{\frac{8u}{n}}M\right) \leq e^{-u}.$$

Similarly,

$$\mathbb{P}\left(\mathbb{E}Z - Z > \sqrt{\frac{8u}{n}}M\right) \leq e^{-u}.$$