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

$$|Z(x_1, \ldots, x_n) - Z(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)| \leq c_i. \tag{25.1}$$

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

$$\begin{aligned}
Z(x_1, \ldots, x_n) - \mathbb{E}_{x'} Z(x_1', \ldots, x_n') &= (Z(x_1, \ldots, x_n) - \mathbb{E}_{x'} Z(x_1', x_2, \ldots, x_n)) \\
&+ (\mathbb{E}_{x'} Z(x_1', x_2, \ldots, x_n) - \mathbb{E}_{x'} Z(x_1', x_2', x_3, \ldots, x_n)) \\
&\ldots \\
&+ \left(\mathbb{E}_{x'} Z(x_1', \ldots, x_{n-1}', x_n) - \mathbb{E}_{x'} Z(x_1', \ldots, x_n')\right) \\
&= Z_1 + Z_2 + \ldots + Z_n
\end{aligned}$$

where

$$Z_i = \mathbb{E}_{x'} Z(x_1', \ldots, x_{i-1}', x_i, \ldots, x_n) - \mathbb{E}_{x'} Z(x_1', \ldots, x_i', x_{i+1}, \ldots, x_n).$$

Assume

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

Lemma 25.1. *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 λ, function $e^{\lambda s}$ is convex and

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

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

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

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

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

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 25.1. *If condition (25.1) is satisfied,*

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

Proof. For any $\lambda > 0$

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

Furthermore,

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

Hence,

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

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

Example 25.1. Let $\mathcal{F}$ be a class of functions: $\mathcal{X} \mapsto [a,b]$. Define the empirical process

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

Then, for any i,

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

because

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

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,

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

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

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

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

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

Then, for any i,

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

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

Hence,

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

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,

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