3.7 McDiarmid’s Inequality

Now we generalize Hoeffding's inequality to apply to any function on $X_1, \cdots, X_n$ satisfying an appropriate bounded differences condition.

[Theorem 8] (McDiarmid's inequality)

Let f be a function satisfying: bounded difference condition

$$|f(x_1, \cdots, x_i, \cdots, x_n) - f(x_1, \cdots, x_i', \cdots, x_n)| \le c_i$$

for all $i = 1, \cdots, n$ and $x_1, \cdots, x_n, x_i'$.

Let $X_1, \cdots, X_n$ be independent random variables. Then we have

$$P[f(X_1, \cdots, X_n) - E[f(X_1, \cdots, X_n)] \ge \varepsilon] \le \exp\left(\frac{-2\varepsilon^2}{\sum_{i=1}^{n} c_i^2}\right)$$

First, we introduce martingales.

[Definition 8] martingale.

① A sequence of r.v. $Z_0, Z_1, \cdots, Z_n$ is **martingale** w.r.t. another sequence of r.v. $X_1, \cdots, X_n$ iff Z_i is a funtion of $X_{1:i}$, $E[|Z_i|] < \infty$ and $E[Z_i | X_{1:i-1}] = Z_{i-1}$

② Define $D_i := Z_i - Z_{i-1}$, $D_{1:n}$ is called a **martingale difference sequence** w.r.t. $X_{1:n}$. $E[D_i | X_{1:i-1}] = 0$.

[Lemma 4] (sub-Gaussian martingales)

Let $Z_0, \cdots, Z_n$ be a martingale w.r.t. $X_1, \cdots, X_n$.

Suppose that each $D_i = Z_i - Z_{i-1}$ is conditionally sub-Gaussian with σ_i^2, that is: $E[e^{tD_i} | X_{1:i-1}] \le \exp(\sigma_i^2 t/2)$

Then $Z_n - Z_0 = \sum_{i=1}^{n} D_i$ is sub-Gaussian with $\sigma^2 := \sum_{i=1}^{n} \sigma_i^2$

Pf: $E[e^{t(Z_n - Z_0)}] = E[e^{tD_n} e^{t(Z_{n-1} - Z_0)}]$

$= E[E[e^{tD_n} e^{t(Z_{n-1} - Z_0)} | X \quad]]$

$$= E[E[e^{tD_n} | X_{1:n-1}] e^{t(Z_{n-1}-Z_0)}]$$
$$\le \exp(\sigma_i^2 t/2) E[e^{t(Z_{n-1}-Z_0)}]$$
$$\le \cdots$$
$$\le \exp(\sigma^2 t/2) \qquad \square$$

Now we can prove Theorem 8 (McDiarmid's inequality)

Pf of Th. 8:

① : we construct a Doob martingale:

$$Z_i = E[f(X_1, \cdots, X_n) | X_{1:i}]$$

Note that $Z_0 = E[f(X_1, \cdots, X_n)]$, $Z_n = f(X_1, \cdots, X_n)$.

Question: Is $\{Z_i\}$ martingale?

(i) $E[|Z_i|] = E[|E[f(X_1, \cdots, X_n) | X_{1:i}]|]$
$$\le E[E[|f(X_1, \cdots, X_n)| | X_{1:i}]]$$
$$\le \sup |f(x_1, \cdots, x_n)| < \infty \text{ (since bounded condition)}$$

(ii) $E[Z_i | X_{1:i-1}] = E[E[f(X_1, \cdots, X_n) | X_{1:i}] | X_{1:i-1}]$
$$= E[f(X_1, \cdots, X_n) | X_{1:i-1}]$$
$$= Z_{i-1}$$

Answer: $\{Z_i\}$ is martingale.

② We show that $D_i = Z_i - Z_{i-1}$ is sub-Gaussian martingale.

$Z_{i-1} = E[f(X_1, \cdots, X_{i-1}, X_i, X_{i+1}, \cdots, X_n) | X_{1:i-1}]$ red part is fixed.

$Z_i = E[f(X_1, \cdots, X_{i-1}, X_i, X_{i+1}, \cdots, X_n) | X_{1:i}]$

Only difference, while $X_{i+1:n}$ is still random, happens on X_i.

Define: $L_i = \inf_x E[f(X_{1:n}) | X_{1:i-1}, X_i = x] - E[f(X_{1:n}) | X_{1:i-1}]$

$U_i = \sup_x E[f(X_{1:n}) | X_{1:i-1}, X_i = x] - E[f(X_{1:n}) | X_{1:i-1}]$

Note $|L_i|, |U_i|$ is finite and $L_i \le D_i \le U_i$.

Let x_L^i and x_U^i correspond to x's achieving L_i & U_i. Then

$$f(X_{1:i-1}, x_L^i, x_{i+1:n}) - f(X_{1:i-1}, x_L^i, x_{i+1:n}) \le c_i$$

Since X_i's are independent

$$\begin{aligned} U_i - L_i &= E[f(X_{1:i-1}, x_U^i, X_{i+1:n}) \mid X_{1:i-1}, X_i = x_U^i] \\ &\quad - E[f(X_{1:i-1}, x_L^i, X_{i+1:n}) \mid X_{1:i-1}, X_i = x_L^i] \\ &= E[f(X_{1:i-1}, x_U^i, X_{i+1:n}) - f(X_{1:i-1}, x_L^i, X_{i+1:n}) \mid X_{1:i-1}] \\ &\le c_i \end{aligned}$$

By the Hoeffding's Lemma, we know $D_i = Z_i - Z_{i-1}$ is sub-Gaussian with $\sigma_i^2 = \frac{1}{4} c_i^2$ conditioned on $X_{1:i-1}$.

Then by Lemma 4, we prove that:

$Z_n - Z_0$ is sub-Guassian with $\sigma^2 = \sum_{i=1}^n c_i^2 / 4$

which leads to the consequence:

$$P(Z_n - Z_0 \geq \varepsilon) \le \exp\left\{-\frac{\varepsilon^2}{2\sigma^2}\right\} = \exp\left\{-\frac{\varepsilon^2}{\sum_{i=1}^n c_i^2}\right\} \qquad \square$$