McDiarmid's inequality

Lecturer: Dr. Hong-Gwa Yeh
Department of Mathematics
National Central University
hgyeh@math.ncu.edu.tw

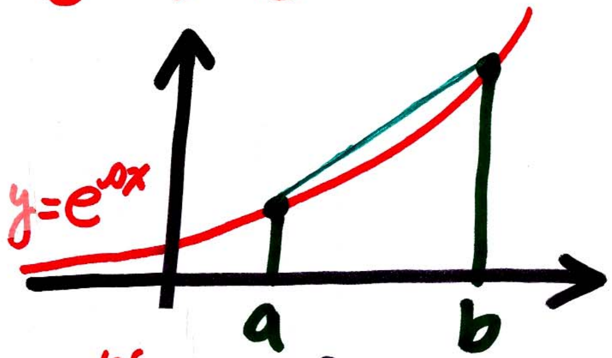
Hoeffding's Lemma

Lemma: $\mathcal{E}X=0, \quad a \le X \le b.$

Then for any $s>0$,

$$\mathcal{E}(e^{sX}) \le e^{\frac{s^2(b-a)^2}{8}}$$

pf: Note that $e^{sx} \le \left(\frac{x-a}{b-a}\right)(e^{sb}-e^{sa})+e^{sa}$ for $a \le x \le b$. It follows that



$$\mathcal{E}e^{sX} \le \frac{-a}{b-a}(e^{sb}-e^{sa})+e^{sa}$$

$$= e^{\phi(u)},$$ where $u \overset{def}{=} s(b-a)$, $\phi(u) \overset{def}{=} -pu+\ln(1-p+pe^{u})$

and $p \overset{def}{=} -\frac{a}{b-a}$

Note that $\phi'(u) = -p + \frac{p}{p+(1-p)e^{-u}}$, $\phi''(u) = \frac{p(1-p)e^{-u}}{(p+(1-p)e^{-u})^2} \le \frac{p(1-p)\frac{p}{1-p}}{\left(p+(1-p)\frac{p}{1-p}\right)^2} = \frac{1}{4}$

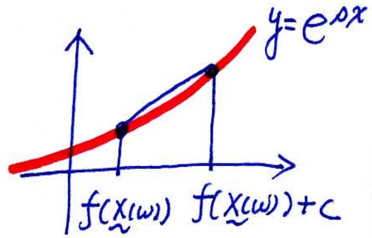
Taylor's Thm $\Rightarrow$ for some $\theta \in [0,u]$,

$$\phi(u) = \underbrace{\phi(0)}_{=0} + \underbrace{\phi'(0)}_{=0}u + \frac{\phi''(\theta)u^2}{2} \le \frac{u^2}{8} = \frac{s^2(b-a)^2}{8}$$

QED

An extension of Hoeffding's Lemma

Lemma: Suppose for rv V and r.vector $\underset{\sim}{X}$ we have
① $\mathcal{E}(V \mid \underset{\sim}{X})=0$ a.s. ② $f(\underset{\sim}{X}) \le V \le f(\underset{\sim}{X})+c$, for some fun. f and constant c.
Then for any $s>0$, $\mathcal{E}(e^{sV} \mid \underset{\sim}{X}) \le e^{\frac{s^2c^2}{8}}$.



pf: Note that

$$e^{sV} \le \frac{V-f(\underset{\sim}{X})}{c}\left(e^{s(f(\underset{\sim}{X})+c)}-e^{sf(\underset{\sim}{X})}\right)+e^{sf(\underset{\sim}{X})}$$

Thus $\mathcal{E}(e^{sV} \mid \underset{\sim}{X}) = \frac{-f(\underset{\sim}{X})}{c}\left(e^{s(f(\underset{\sim}{X})+c)}-e^{sf(\underset{\sim}{X})}\right)+e^{sf(\underset{\sim}{X})}$.

The remains is similar to Hoeffding's Lemma.

QED

McDiarmid's Inequality

Thm: Let $X_1, X_2, \cdots, X_n$ be independent rvs. Let f be a function $f: \mathbb{R}^n \to \mathbb{R}$ with a vector $(c_1, c_2, \cdots, c_n)$ s.t. $|f(\underset{\sim}{x}) - f(\underset{\sim}{y})| \le c_i$ for all $\underset{\sim}{x}, \underset{\sim}{y}$ in $\mathbb{R}^n$ that differ only at the ith coordinate, $1 \le i \le n$.

Then for any $t > 0$,

$$P\left(\left| f(X_1, \cdots, X_n) - \mathcal{E} f(X_1, \cdots, X_n) \right| \ge t \right) \le 2 e^{-\frac{2t^2}{\sum_{i=1}^n c_i^2}}$$

pf: Let $\underset{\sim}{X} = (X_1, \cdots, X_n)$, $Z_0 = \mathcal{E} f(\underset{\sim}{X})$, $Z_i = \mathcal{E}(f(\underset{\sim}{X}) | X_1, \cdots, X_i)$, $Z_n = f(\underset{\sim}{X})$.

Claim $\mathcal{E}\left(e^{s(Z_k - Z_{k-1})} \middle| X_1, \cdots, X_{k-1}\right) \le e^{\frac{s^2 c_k^2}{8}}$, $1 \le k \le n$.

pf: Let $U_k = \sup_u \left\{ \mathcal{E}\left(f(\underset{\sim}{X}) | X_1, \cdots, X_{k-1}, u\right) - \mathcal{E}\left(f(\underset{\sim}{X}) | X_1, \cdots, X_{k-1}\right) \right\}$

$L_k = \inf_l \left\{ \mathcal{E}\left(f(\underset{\sim}{X}) | X_1, \cdots, X_{k-1}, l\right) - \mathcal{E}\left(f(\underset{\sim}{X}) | X_1, \cdots, X_{k-1}\right) \right\}$

Note that $U_k - L_k \le \sup_{l,u} \left\{ \mathcal{E}\left(f(\underset{\sim}{X}) | X_1, \cdots, X_{k-1}, u\right) - \mathcal{E}\left(f(\underset{\sim}{X}) | X_1, \cdots, X_{k-1}, l\right) \right\}$

$\because X_1, \cdots, X_n$ are independent

$\le \sup_{l,u} \left\{ \sum_{y_{k+1}, \cdots, y_n} \left[f(X_1, \cdots, X_{k-1}, u, y_{k+1}, \cdots, y_n) - f(X_1, \cdots, X_{k-1}, l, y_{k+1}, \cdots, y_n) \right] \prod_{j=k+1}^{n} P(X_j = y_j) \right\}$

$\le c_k$ ($\because$ Lipschitz condition)

So $L_k \le Z_k - Z_{k-1} \le U_k \le L_k + c$ and hence the claim is true by the extension of Hoeffding's Lemma, since $\mathcal{E}(Z_k - Z_{k-1} | X_1, \cdots, X_{k-1}) = 0$ a.s.

QED of claim

Pf (continued)

$$\Pr\{ f(\underset{\sim}{X}) - \mathcal{E} f(\underset{\sim}{X}) \geq t \}$$

$$\leq e^{-st} \, \mathcal{E} e^{s[f(\underset{\sim}{X}) - \mathcal{E} f(\underset{\sim}{X})]}$$

$$= e^{-st} \, \mathcal{E} e^{s \sum_{k=1}^{n} [Z_k - Z_{k-1}]}$$

$$= e^{-st} \, \mathcal{E}\left(\mathcal{E}\left(e^{s \sum_{k=1}^{n} [Z_k - Z_{k-1}]} \,\middle|\, X_1, \dots, X_{n-1} \right)\right) \quad (\because \text{Tower property})$$

$$= e^{-st} \, \mathcal{E}\left(e^{s \sum_{k=1}^{n-1} [Z_k - Z_{k-1}]} \, \mathcal{E}\left(e^{s(Z_n - Z_{n-1})} \,\middle|\, X_1, \dots, X_{n-1} \right)\right)$$

$$\leq e^{-st} \, \mathcal{E}\left(e^{s \sum_{k=1}^{n-1} [Z_k - Z_{k-1}]} \, e^{\frac{s^2 c_n^2}{8}} \right) \quad (\text{by the claim})$$

$$\leq e^{-st} \prod_{k=1}^{n} e^{\frac{s^2 c_k^2}{8}} \quad (\text{by repeating the same argument } n \text{ times})$$

$$= e^{-st + s^2 \sum_{k=1}^{n} \frac{c_k^2}{8}} \leq e^{-\frac{2t^2}{\sum_{k=1}^{n} c_k^2}} \quad \text{by choosing } s = \frac{4t}{\sum_{k=1}^{n} c_k^2}$$

QED

Concentration of the Chromatic Number

Thm (Shamir-Spencer) Let $n \geq 2$ and $p \in (0,1)$.

Then we have

$$P\left(\left|\chi(G_{n,p}) - \mathcal{E}\chi(G_{n,p})\right| \geq t\right) \leq 2e^{-\frac{2t^2}{n-1}}$$

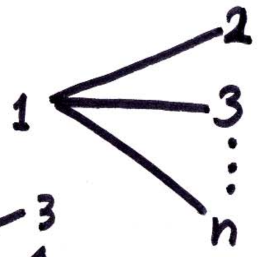
Remark: probability space $\Omega = \{\omega : \omega \text{ is a graph on the vertex set } \{1,2,\cdots,n\}\}$

with probability measure $P(\omega) = p^m (1-p)^{\binom{n}{2}-m}$ where $m = |E(\omega)|$.

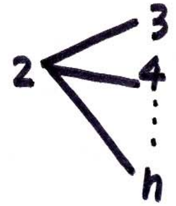
One may consider $G_{n,p}$ as a random element $G_{n,p} : \Omega \to \Omega$ s.t. $G_{n,p}(\omega) = \omega$.

Sometimes we consider $G_{n,p}$ as a prob. space.

pf: Let $\underset{\sim}{X_1} = (X_{12}, X_{13}, X_{14}, \cdots, X_{1n})$

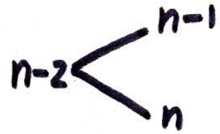


$\underset{\sim}{X_2} = (X_{23}, X_{24}, \cdots\cdots, X_{2n})$



$\vdots$

$\underset{\sim}{X_{n-2}} = (X_{n-2,n-1}, X_{n-2,n})$



$\underset{\sim}{X_{n-1}} = X_{n-1,n}$

We have $\chi(G_{n,p}) = f(\underset{\sim}{X_1}, \underset{\sim}{X_2}, \cdots, \underset{\sim}{X_{n-1}})$ for some function f.

Note that $f: S_1 \times S_2 \times \cdots \times S_{n-1} \longrightarrow \mathbb{R}$, where $S_i = \{0,1\}^{n-i}$, and

f is a 1-Lipschitz function.

By McDiarmid's Inequality, we arrive at

$$\mathbb{P}\left(|\chi(G_{n,p}) - \mathbb{E}\chi(G_{n,p})| \geq t \right) \leq 2e^{-\frac{2t^2}{n-1}}$$

QED

A Remark on Shamir-Spencer's Thm.

- Since $\mathcal{P}(|\chi(G_{n,p}) - \mathcal{E}\chi(G_{n,p})| \geq t\sqrt{n-1}) \leq 2e^{-2t^2}$, so the chromatic number is almost always concentrated on about $O(\sqrt{n})$ values.

- Note that we have no clue to what the value of $\mathcal{E}\chi(G_{n,p})$ is.