Talagrand's Inequality

1. notation:
   ① $S_1, S_2, \cdots, S_N$ are sets
   ② $\mu_1, \cdots, \mu_N$ are probability measure on $S_1, \cdots, S_N$ respectively.
   ③ let $\mathcal{P}$ be the product measure $\mu_1 \times \cdots \times \mu_N$ on $S = S_1 \times \cdots \times S_N$

2. a distance: For $\underset{\sim}{x} \in S$ and $A \subseteq S$ define

$$U_A(\underset{\sim}{x}) = \{ (s_1, \cdots, s_N) \in \{0,1\}^N : \exists\, \underset{\sim}{y} \in A \text{ s.t. } x_i = y_i \text{ for all } i \text{ with } s_i = 0 \}$$

$$V_A(\underset{\sim}{x}) = \text{the convex hull of } U_A(\underset{\sim}{x})$$
$$= \{ \sum \lambda_i \underset{\sim}{s}_i : \lambda_i \geq 0, \sum \lambda_i = 1, \underset{\sim}{s}_i \in U_A(\underset{\sim}{x}) \}$$

Remark:
1. $U_A(\underset{\sim}{x}) = \bigcup_{\underset{\sim}{y} \in A} \left\{ (s_1, \cdots, s_N) : s_i = \begin{cases} 1 & \text{if } x_i \neq y_i \\ 0, 1 & \text{if o.w.} \end{cases} \right\}$

2. $V_A(\underset{\sim}{x}) \subseteq \overbrace{[0,1] \times \cdots \times [0,1]}^{N}$

3. $d(A, \underset{\sim}{x}) = d_E(V_A(\underset{\sim}{x}), 0) = \inf \left\{ \sqrt{u_1^2 + \cdots + u_N^2} : (u_1, \ldots, u_N) \in V_A(\underset{\sim}{x}) \right\}$

Euclidean norm in $\mathbb{R}^N$

Remark:
① If $A = \phi$, then $U_A(\underset{\sim}{x}) = V_A(\underset{\sim}{x}) = \phi$ and we set $d(A, \underset{\sim}{x}) = \infty$
② If $A \neq \phi$, then $d(A, x) < \infty$, since $(1, \cdots, 1) \in U_A(\underset{\sim}{x}) \subseteq V_A(\underset{\sim}{x})$.
Moreover, the infimum in the def. of $V_A(\underset{\sim}{x})$ is attained, since $U_A(x)$ is a finite set and thus $V_A(\underset{\sim}{x})$ is compact.

note: $d(A, \underset{\sim}{x}) = 0 \Leftrightarrow d_E(V_A(\underset{\sim}{x}), \underset{\sim}{0}) = 0 \Leftrightarrow \underset{\sim}{0} \in V_A(\underset{\sim}{x})$
$\Leftrightarrow \underset{\sim}{0} \in U_A(\underset{\sim}{x}) \quad (\because U_A(\underset{\sim}{x}) \subseteq \{0,1\}^N)$
$\Leftrightarrow \underset{\sim}{x} \in A$

3. Thm (General form of Talagrand's Inequality) For every (measurable) set $A \subseteq S = S_1 \times \cdots \times S_N$,

$$\int_S e^{\frac{1}{4} d^2(A, \underset{\sim}{x})} \, d\mathcal{P}(\underset{\sim}{x}) \leq \frac{1}{\mathcal{P}(A)}$$

pf: Induction on N. (The bold student may start with the really trivial case $N=0$ instead)

[Base] $N=1$. For nonempty set $A \subseteq S = S_1$,

note that $x \notin A \Rightarrow U_A(x) = \{(1)\}$ and $V_A(x) = \{(1)\} \Rightarrow d(A,x) = 1$

Consequently

$$\int_S e^{\frac{1}{4}d^2(A,x)}\, d\mathcal{P}(x)$$

$$= \int_A e^{\frac{1}{4}d^2(A,x)}\, d\mathcal{P}(x) + \int_{S \setminus A} e^{\frac{1}{4}d^2(A,x)}\, d\mathcal{P}(x)$$

$$= \mathcal{P}(A) + e^{\frac{1}{4}}(1-\mathcal{P}(A)) \quad (\because [x \in A \Rightarrow d(A,x)=0] \text{ and } [\mathcal{P}(S)=1])$$

$$\le \mathcal{P}(A) + 2(1-\mathcal{P}(A)) = 2 - \mathcal{P}(A) \le \frac{1}{\mathcal{P}(A)} \quad \left(\because \begin{array}{l} t(2-t) \le 1 \\ \text{for all } 0 \le t \le 1 \end{array}\right)$$

[Induction hypothesis] Assume that the result holds for $N \ge 1$.

- Write $S^{(k)} = S_1 \times \cdots \times S_k$, $\mathcal{P}_k = \mu_1 \times \cdots \times \mu_k$.
- Denote element in $S^{(N+1)}$ by $(\underset{\sim}{x}, \lambda)$ with $\underset{\sim}{x} \in S^{(N)}$, $\lambda \in S_{N+1}$.

[Induction part] let $A \subseteq S^{(N+1)}$ be a measurable subset.

- For $\lambda \in S_{N+1}$, define $A(\lambda) = \{\underset{\sim}{x} \in S^{(N)} : (\underset{\sim}{x}, \lambda) \in A\}$ (a section of A)

  define $B = \bigcup_{\lambda \in S_{N+1}} A(\lambda)$ (the projection of A on $S^{(N)}$)

- Remark: Each $A(\lambda)$ is measurable, but, in general, B is not; thus we select a measurable subset $B_0 \subseteq B$ of maximal $\mathcal{P}_N$ measure.

  Thus $\mathcal{P}_N(B_0) \ge \mathcal{P}_N(A(\lambda))$ for every λ ($\because A(\lambda) \subseteq B$ is measurable)

  We may assume that $\mathcal{P}_N(B_0) > 0$, since otherwise $\mathcal{P}_{N+1}(A) = 0$ and the result is trivial.

- Observation:

① For any $\underset{\sim}{x} \in S^{(N)}$ and $\lambda \in S_{N+1}$,

$$\underset{\sim}{s} \in U_{A(\lambda)}(\underset{\sim}{x}) \Rightarrow \exists\, \underset{\sim}{y} \in A(\lambda) \text{ s.t. } s_i = \begin{cases} 1 & x_i \ne y_i \\ 0,1 & \text{o.w.} \end{cases} \quad i = 1, 2, \ldots N$$

$$\Rightarrow \exists\, (\underset{\sim}{y}, \lambda) \in A \text{ s.t. } s_i = \begin{cases} 1 & x_i \ne y_i \\ 0,1 & \text{o.w.} \end{cases} \quad i = 1, 2, \ldots N+1, \text{ where } \begin{array}{l} x_{N+1} = \lambda \\ y_{N+1} = \lambda \end{array}$$

$$\Rightarrow (\underset{\sim}{s}, 0) \in U_A((\underset{\sim}{x}, \lambda))$$

②
$$\underset{\sim}{t} \in U_B(\underset{\sim}{x}) \Rightarrow \exists\, \underset{\sim}{y} \in B \text{ s.t. } t_i = \begin{cases} 1 & \text{if } x_i \ne y_i \\ 0,1 & \text{o.w.} \end{cases} \quad i = 1, 2, \ldots N$$

$$\Rightarrow \exists\, \underset{\sim}{y} \in A(\lambda') \text{ for some } \lambda' \in S_{N+1} \text{ s.t. } t_i = \begin{cases} 1 & \text{if } x_i \ne y_i \\ 0,1 & \text{o.w.} \end{cases} \quad i = 1, 2, \ldots N$$

$$\Rightarrow \exists\, (\underset{\sim}{y}, \lambda') \in A \text{ s.t. } t_i = \begin{cases} 1 & \text{if } x_i \ne y_i \\ 0,1 & \text{o.w.} \end{cases} \quad i = 1, 2, \ldots N.$$

$$\Rightarrow (\underset{\sim}{t}, 1) \in U_A((\underset{\sim}{x}, \lambda)) \text{ for any } \lambda \in S_{N+1}$$

③
$$\underset{\sim}{s} \in V_{A(\lambda)}(\underset{\sim}{x}) \Rightarrow \underset{\sim}{s} = \sum \lambda_i \underset{\sim}{s_i},\ \underset{\sim}{s_i} \in U_{A(\lambda)}(\underset{\sim}{x}) \Rightarrow (\underset{\sim}{s}, 0) = \sum \lambda_i (\underset{\sim}{s_i}, 0)$$

$$\Rightarrow (\underset{\sim}{s}, 0) \in V_A((\underset{\sim}{x}, \lambda)) \qquad \because ① \ \forall (\underset{\sim}{s_i}, 0) \in U_A((\underset{\sim}{x}, \lambda))$$

④ $\underset{\sim}{t} \in V_{B_0}(\underset{\sim}{x}) \subseteq V_B(\underset{\sim}{x}) \Rightarrow \underset{\sim}{t} = \sum \lambda_i \underset{\sim}{t_i}$, $\underset{\sim}{t_i} \in U_B(\underset{\sim}{x})$

$\Rightarrow (\underset{\sim}{t}, 1) = \sum \lambda_i (\underset{\sim}{t_i}, 1)$, $(\underset{\sim}{t_i}, 1) \in U_A((\underset{\sim}{x}, \lambda))$ $\because$ ② for any $\lambda \in S_{N+1}$

$\Rightarrow (\underset{\sim}{t}, 1) \in V_A((\underset{\sim}{x}, \lambda))$ for any $\lambda \in S_{N+1}$

⑤ If $\underset{\sim}{s} \in V_{A(\lambda)}(\underset{\sim}{x})$, $\underset{\sim}{t} \in V_{B_0}(\underset{\sim}{x})$ and $0 \le \tau \le 1$

then $(\underset{\sim}{s}, 0) \in$, $(\underset{\sim}{t}, 1) \in V_A((\underset{\sim}{x}, \lambda))$ and $0 \le \tau \le 1$

and hence

$$(1-\tau)(\underset{\sim}{s}, 0) + \tau(\underset{\sim}{t}, 1) \in V_A((\underset{\sim}{x}, \lambda))$$

i.e. $$((1-\tau)\underset{\sim}{s} + \tau\underset{\sim}{t}, \tau) \in V_A((\underset{\sim}{x}, \lambda))$$

Claim A For any $\lambda \in S_{N+1}$ and $\tau \in [0,1]$,

$$d^2(A, (\underset{\sim}{x}, \lambda)) \le \tau^2 + (1-\tau) d^2(A(\lambda), \underset{\sim}{x}) + \tau d^2(B_0, \underset{\sim}{x})$$

pf: $$d^2(A, (\underset{\sim}{x}, \lambda)) \le \inf_{\substack{\underset{\sim}{s} \in V_{A(\lambda)}(\underset{\sim}{x}) \\ \underset{\sim}{t} \in V_{B_0}(\underset{\sim}{x})}} \|((1-\tau)\underset{\sim}{s} + \tau\underset{\sim}{t}, \tau)\|^2$$ ($\because$ ⑤ and the def. of $d(\cdot)$)

$$= \inf_{\substack{\underset{\sim}{s} \in V_{A(\lambda)}(\underset{\sim}{x}) \\ \underset{\sim}{t} \in V_{B_0}(\underset{\sim}{x})}} \left[\tau^2 + \sum_{i=1}^{N} [(1-\tau)s_i + \tau t_i]^2\right]$$

$$\le \inf_{\underset{\sim}{s} \in V_{A(\lambda)}(\underset{\sim}{x}),\ \underset{\sim}{t} \in V_{B_0}(\underset{\sim}{x})} \left[\tau^2 + (1-\tau)\|\underset{\sim}{s}\|^2 + \tau\|\underset{\sim}{t}\|^2\right]$$ ($\because$ the convexity of the function $f(u) = \|u\|^2$)

$$= \tau^2 + (1-\tau) d^2(A(\lambda), \underset{\sim}{x}) + \tau d^2(B_0, \underset{\sim}{x})$$ QED

Hölder's Ineq.: For $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, if $f \in L^p = \{h : \int_\Omega |h|^p d\mu < \infty\}$ and $g \in L^q$ then $fg \in L^1$ and $\|fg\|_1 \le \|f\|_p \|g\|_q$ i.e.

$$\int_\Omega |fg| d\mu \le \left(\int_\Omega |f|^p d\mu\right)^{\frac{1}{p}} \left(\int_\Omega |g|^q d\mu\right)^{\frac{1}{q}}$$

claim β

Suppose $0 \le l \le 1$. Then $\inf_{0 \le \tau \le 1} e^{\frac{\tau^2}{4}} (l)^{\tau - 1} \le 2 - l$.

pf:

Finally, $\int_{S^{(N+1)}} e^{\frac{1}{4}d^2(A,\underset{\sim}{y})} d\mathcal{P}_{N+1}(\underset{\sim}{y})$

$$= \int_{S_{N+1}} \int_{S^{(N)}} e^{\frac{1}{4}d^2(A,(\underset{\sim}{x},\lambda))} d\mathcal{P}_N(\underset{\sim}{x})\, d\mu_{N+1}(\lambda)$$

$$\leq \int_{S_{N+1}} \int_{S^{(N)}} e^{\frac{1}{4}\left[\tau^2+(1-\tau)d^2(A(\lambda),\underset{\sim}{x})+\tau d^2(B_0,\underset{\sim}{x})\right]} d\mathcal{P}_N(\underset{\sim}{x})\, d\mu_{N+1}(\lambda) \quad (\because \text{Claim } A)$$

for any $\tau \in [0,1]$, $\lambda \in S_{N+1}$.

$$= \int_{S_{N+1}} e^{\frac{1}{4}\tau^2} \int_{S^{(N)}} \left[e^{\frac{d^2(A(\lambda),\underset{\sim}{x})}{4}}\right]^{1-\tau} \left[e^{\frac{d^2(B_0,\underset{\sim}{x})}{4}}\right]^{\tau} d\mathcal{P}_N(\underset{\sim}{x})\, d\mu_{N+1}(\lambda)$$

$$\leq \int_{S_{N+1}} e^{\frac{1}{4}\tau^2} \left[\int_{S^{(N)}} e^{\frac{d^2(A(\lambda),\underset{\sim}{x})}{4}} d\mathcal{P}_N(\underset{\sim}{x})\right]^{1-\tau} \left[\int_{S^{(N)}} e^{\frac{d^2(B_0,\underset{\sim}{x})}{4}} d\mathcal{P}_N(\underset{\sim}{x})\right]^{\tau} d\mu_{N+1}(\lambda)$$

Hölder's inequality

$$\leq \int_{S_{N+1}} e^{\frac{1}{4}\tau^2} \left[\frac{1}{\mathcal{P}_N(A(\lambda))}\right]^{1-\tau} \left[\frac{1}{\mathcal{P}_N(B_0)}\right]^{\tau} d\mu_{N+1}(\lambda)$$

induction hypothesis

$$= \int_{S_{N+1}} \frac{1}{\mathcal{P}_N(B_0)} e^{\frac{\tau^2}{4}} \left[\frac{\mathcal{P}_N(A(\lambda))}{\mathcal{P}_N(B_0)}\right]^{\tau-1} d\mu_{N+1}(\lambda)$$

for any $\tau \in [0,1]$, $\lambda \in S_{N+1}$

◉ Thus $\int_{S^{(N+1)}} e^{\frac{1}{4}d^2(A,\underset{\sim}{y})} d\mathcal{P}_{N+1}(\underset{\sim}{y})$

$$\leq \int_{S_{N+1}} \frac{1}{\mathcal{P}_N(B_0)} \inf_{0\leq\tau\leq 1} e^{\frac{\tau^2}{4}} \left(\frac{\mathcal{P}_N(A(\lambda))}{\mathcal{P}_N(B_0)}\right)^{\tau-1} d\mu_{N+1}(\lambda)$$

$$\leq \int_{S_{N+1}} \frac{1}{\mathcal{P}_N(B_0)} \left(2 - \frac{\mathcal{P}_N(A(\lambda))}{\mathcal{P}_N(B_0)}\right) d\mu_{N+1}(\lambda) \quad (\because \text{Claim } B)$$

$$= \frac{1}{\mathcal{P}_N(B_0)} \left(2 - \frac{\mathcal{P}_{N+1}(A)}{\mathcal{P}_N(B_0)}\right)$$

$\left(\because A(\lambda) = \{\underset{\sim}{x} \in S^{(N)} : (\underset{\sim}{x},\lambda) \in A\}\right.$ so $\int_{S_{N+1}} \int_{A(\lambda)} 1\, d\mathcal{P}_N(\underset{\sim}{x})\, d\mu_{N+1}(\lambda) = \int_A 1\, d\mathcal{P}_{N+1}(\underset{\sim}{x},\lambda) = \mathcal{P}_{N+1}(A)$ by Fubini's theorem $\left.\right)$

$$\leq \frac{1}{\mathcal{P}_{N+1}(A)}$$

$(\because 2-t \leq \frac{1}{t}$ for all $t \geq 0)$

i.e. $t(2-t) \leq 1$ for all $t \geq 0$

4. <u>Thm</u> (<u>Talagrand 1995</u>) Suppose A and $B \subseteq S_1 \times \cdots \times S_N$ are two measurable subsets s.t. for some $t \geq 0$ the following separation condition holds

(S) For every $\underset{\sim}{z} = (z_1, \dots, z_N) \in B$ there $\exists$ a $\underset{\sim}{\alpha} = (\alpha_1, \dots, \alpha_N) \in \mathbb{R}^N \setminus \{0\}$ ($\underset{\sim}{\alpha}$ may depend on $\underset{\sim}{z}$) s.t. for every $(y_1, \dots, y_N) \in A$,

$$\sum_{i \in [N] \text{ with } y_i \neq z_i} \alpha_i \geq t \left(\sum_{i=1}^{N} \alpha_i^2 \right)^{\frac{1}{2}} = t \|\underset{\sim}{\alpha}\|$$

Then

$$\mathcal{P}(A)\mathcal{P}(B) \leq e^{-\frac{t^2}{4}}$$

<u>pf</u>: Given $\underset{\sim}{z} = (z_1, \dots, z_N) \in B$. Let $\underset{\sim}{\alpha} \in \mathbb{R}^N \setminus \{\underset{\sim}{0}\}$ be as in condition (S). Assume $\alpha_i \geq 0$, otherwise we replace α_i by $|\alpha_i|$.

Let $\langle \cdot, \cdot \rangle$ be the inner product in $\mathbb{R}^N$.

<u>Claim A</u>: $\underset{\sim}{s} \in U_A(\underset{\sim}{z}) \Rightarrow \langle \underset{\sim}{\alpha}, \underset{\sim}{s} \rangle \geq t\|\alpha\|$

<u>pf</u>: $\underset{\sim}{s} = (s_1, s_2, \dots, s_N) \in U_A(\underset{\sim}{z})$

$\Rightarrow \exists\, \underset{\sim}{y} = (y_1, \dots, y_N) \in A$ such that $s_i = \begin{cases} 1 & \text{if } z_i \neq y_i \\ 0, 1 & \text{o.w.} \end{cases}$

$\Rightarrow \langle \underset{\sim}{\alpha}, \underset{\sim}{s} \rangle \geq t\|\alpha\|$

<u>Claim B</u>: $\underset{\sim}{s} \in V_A(\underset{\sim}{z}) \Rightarrow t \leq \|\underset{\sim}{s}\|$

<u>pf</u>: $\underset{\sim}{s} \in V_A(\underset{\sim}{z})$

$\Rightarrow \underset{\sim}{s} = \sum \lambda_i \underset{\sim}{s_i}$ where $\sum \lambda_i = 1$, $\underset{\sim}{s_i} \in U_A(\underset{\sim}{z})$

$\Rightarrow \langle \underset{\sim}{\alpha}, \underset{\sim}{s} \rangle = \sum \lambda_i \langle \underset{\sim}{\alpha}, \underset{\sim}{s_i} \rangle \geq \sum \lambda_i t\|\alpha\| = t\|\alpha\|$ ($\because$ claim A)

$\Rightarrow t\|\alpha\| \leq \langle \underset{\sim}{\alpha}, \underset{\sim}{s} \rangle \leq \|\underset{\sim}{\alpha}\| \|\underset{\sim}{s}\|$ by cauchy-schwarz ineq.

$\Rightarrow t \leq \|\underset{\sim}{s}\|$ ($\because \|\underset{\sim}{\alpha}\| > 0$)

<u>Claim C</u>: $\underset{\sim}{z} \in B \Rightarrow d(A, \underset{\sim}{z}) \geq t$.

<u>pf</u>: $\underset{\sim}{z} \in B \Rightarrow d(A, \underset{\sim}{z}) = d_E(V_A(\underset{\sim}{z}), \underset{\sim}{0})$ the Euclidean distance from 0 to $V_A(\underset{\sim}{z})$.

$\Rightarrow d(A, \underset{\sim}{z}) \geq t$ Done!

Talagrand's ineq. (general form)

$$\Rightarrow \frac{1}{\mathcal{P}(A)} \geq \int_{S_1 \times \cdots \times S_N} e^{\frac{1}{4} d^2(A, \underset{\sim}{x})} d\mathcal{P}(x) \geq \int_B e^{\frac{1}{4} d^2(A, \underset{\sim}{x})} d\mathcal{P}(\underset{\sim}{x})$$

$$\geq \int_B e^{\frac{1}{4} t^2} d\mathcal{P}(x) \geq e^{\frac{t^2}{4}} \mathcal{P}(B)$$

QED.

5. Thm. • $z_1, \dots, z_N$ are indep. rvs taking values in $S_1, \dots, S_N$ resp.
• $f: S_1 \times \dots \times S_N \to \mathbb{R}$ is Lipschitz with $\underset{\sim}{c} = (c_1, \dots, c_N)$
such that for some function $\varphi: \mathbb{R} \to \mathbb{R}$ the following condition holds

i.e. if $\underset{\sim}{z}, \underset{\sim}{z}' \in S_1 \times \dots \times S_N$ differ only in the kth coordinate then $|f(\underset{\sim}{z}) - f(\underset{\sim}{z}')| \le c_k$.

(C) If $\underset{\sim}{z} \in S_1 \times \dots \times S_N$ and $r \in \mathbb{R}$ with $f(\underset{\sim}{z}) \ge r$
then $\exists J \subseteq [N]$ having $\sum_{j \in J} c_j^2 \le \varphi(r)$
such that for all $\underset{\sim}{y} \in S_1 \times \dots \times S_N$ with $y_i = \begin{cases} z & \text{if } i \in J \\ * & \text{o.w.} \leftarrow 任意 \end{cases}$
we have $f(\underset{\sim}{y}) \ge r$.

Then, for every $r \in \mathbb{R}$ and $t \ge 0$,

$$\mathcal{P}\{ f(Z_1, \dots, Z_N) \le r - t \} \mathcal{P}\{ f(Z_1, \dots, Z_N) \ge r \} \le e^{-\frac{t^2}{4\varphi(r)}}$$

pf: Let $A = \{ \underset{\sim}{y} \in S_1 \times \dots \times S_N : f(\underset{\sim}{y}) \le r - t \}$
$B = \{ \underset{\sim}{z} \in S_1 \times \dots \times S_N : f(\underset{\sim}{z}) \ge r \}$. Assume $t > 0$ ($\because t = 0$ is trivial)

Claim: For every $\underset{\sim}{z} = (z_1, \dots, z_N) \in B$, there $\exists$ a $\underset{\sim}{\alpha} = (\alpha_1, \dots, \alpha_N) \in \mathbb{R}^N \setminus \{\underset{\sim}{0}\}$
($\underset{\sim}{\alpha}$ may depend on $\underset{\sim}{z}$) s.t. for every $\underset{\sim}{y} = (y_1, \dots, y_N) \in A$

$$\sum_{i \in [N] \text{ with } y_i \ne z_i} \alpha_i \ge \frac{t}{\sqrt{\varphi(r)}} \|\underset{\sim}{\alpha}\|$$

pf: $\underset{\sim}{z} = (z_1, \dots, z_N) \in B$
$\Rightarrow f(\underset{\sim}{z}) \ge r \Rightarrow \exists J \subseteq [N]$ having $\sum_{j \in J} c_j^2 \le \varphi(r)$ ($\because$ (C))
Define $\underset{\sim}{\alpha} = (\alpha_1, \dots, \alpha_N)$ with $\alpha_i = \begin{cases} c_i & i \in J \\ 0 & \text{o.w.} \end{cases}$
Then $\|\underset{\sim}{\alpha}\|^2 = \sum_{j \in J} c_j^2 \le \varphi(r)$. ☽
For all $\underset{\sim}{y} = (y_1, \dots, y_N) \in A$, define $\underset{\sim}{y}' = (y_1', \dots, y_N') \in S_1 \times \dots \times S_N$ by
$y_i' = \begin{cases} z_i & \text{if } i \in J \\ y_i & \text{if } i \notin J \end{cases}$
Then (C) 的後半段 $\Rightarrow f(\underset{\sim}{y}') \ge r \Rightarrow f(\underset{\sim}{y}') - t \ge r - t \ge f(\underset{\sim}{y})$ ($\because \underset{\sim}{y} \in A$)
$\Rightarrow t \le |f(\underset{\sim}{y}') - f(\underset{\sim}{y})| \le \sum_{i \in J : y_i \ne z_i} c_i$ ($\because$ Lipschitz condition + 三角不等式)
$= \sum_{i \in [N] : y_i \ne z_i} \alpha_i$

Since $t \ge \frac{t}{\sqrt{\varphi(r)}} \|\alpha\|$ ($\because$ ☽), we have $\sum_{i \in [N] : y_i \ne z_i} \alpha_i \ge \frac{t}{\sqrt{\varphi(r)}} \|\underset{\sim}{\alpha}\|$ 且 $\underset{\sim}{\alpha} \ne 0$ ($\because t > 0$)
end of claim

Note that $\mathcal{P}_{Z_1}, \dots, \mathcal{P}_{Z_N}$ are prob. measures on $S_1, \dots, S_N$ resp.

i.e. $\mathcal{P}_{Z_i}(C) = \mathcal{P}(Z_i \in C) \xrightarrow{} C$ $i = 1, 2 \dots N$. Assume $Z_1 \dots Z_N$ are r.v.s on prob. space $(\Omega, \mathcal{F}, \mathcal{P})$.

Then $\mathcal{P}_{(Z_1, \dots, Z_N)}$ is the product measure on $S_1 \times \dots \times S_N$

i.e.

$$\mathcal{P}_{(Z_1 \dots Z_N)}(C) = \mathcal{P}((Z_1, \dots, Z_N) \in C)$$

$$\text{Item 4} \Rightarrow \mathcal{P}_{(Z_1 \dots Z_N)}(A)\,\mathcal{P}_{(Z_1 \dots Z_N)}(B) \leq e^{-\frac{\left(\frac{t}{\sqrt{\varphi(r)}}\right)^2}{4}} = e^{-\frac{t^2}{4\varphi(r)}}$$

$$\Rightarrow \mathcal{P}((Z_1 \dots Z_N) \in A)\,\mathcal{P}((Z_1 \dots Z_N) \in B) \leq e^{-\frac{t^2}{4\varphi(r)}}$$

$$\Rightarrow \mathcal{P}(f(Z_1 \dots Z_N) \leq r - t)\,\mathcal{P}(f(Z_1, \dots Z_N) \geq r) \leq e^{-\frac{t^2}{4\varphi(r)}}$$

QED.

Talagrand's first Inequality

1. $(\Omega, \mathcal{F}, P)$: a probability space. i.e. $P(A) = P_r(X \in A)$

$X_1, \dots, X_n$ iid random variables with common distribution P taking its values in Ω.

$\Omega^n = \underbrace{\Omega \times \dots \times \Omega}$

$\mathcal{F}^n = \underbrace{\mathcal{F} \times \dots \times \mathcal{F}}_{n}$ = the smallest σ-algebra containing $\{A_1 \times \dots \times A_n : A_i \in \mathcal{F},\ i = 1, 2, \dots n\}$

$\mathbb{P} = \mu = \underbrace{P \times \dots \times P}$ = the product of probability measures $\underbrace{P \times \dots \times P}_{n}$

remark: we have $\mathbb{P}(A \times B) = P(A)P(B)$ for all $A, B \in \mathcal{F}$.

2. Hamming distance $d_H(x, y)$: For $x = (x_1 \dots x_n)$, $y = (y_1 \dots y_n) \in \Omega^n$,

$$d_H(x, y) = \#\{ i \mid x_i \neq y_i \}$$

$$d_H(A, x) = \min_{y \in A} d_H(x, y) \quad \text{where } A \in \mathcal{F}^n \text{ a measurable set.}$$

3 Lemma 5: For any r.v. $Y \in [0, 1]$ and $a > 1$,

$$E\{\min(a, \tfrac{1}{Y})\} E\{Y\} \leq \frac{1}{2} + \frac{a + \frac{1}{a}}{4}$$

pf:

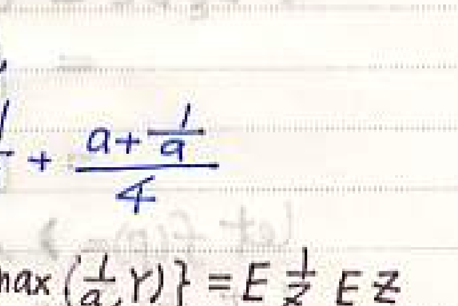
$$E\{\min(a, \tfrac{1}{Y})\} EY \leq E\left\{\frac{1}{\max(\frac{1}{a}, Y)}\right\} E\{\max(\tfrac{1}{a}, Y)\} = E\frac{1}{Z} EZ$$

where $Z = \max(\frac{1}{a}, Y) \in [\frac{1}{a}, 1]$.

Claim: If Z is a r.v. with values in $[\frac{1}{a}, 1]$ then $E\{\frac{1}{Z}\} EZ \leq \frac{1}{2} + \frac{a + \frac{1}{a}}{4}$.

pf: We want to show that if $EZ' = EZ$ with $Z \in \{\frac{1}{a}, 1\}$ for some r.v. Z' then we have $E\frac{1}{Z} \leq E\frac{1}{Z'}$.

Indeed, the right figure shows that



$$\frac{y-1}{x-1} = \frac{1-a}{1-\frac{1}{a}} = -a$$

$$\frac{1}{Z} \leq -Za + a + 1$$

$$\begin{aligned}
\Rightarrow E\frac{1}{Z} &\leq -aEZ + a + 1 \\
&= -aEZ' + a + 1 \\
&= -a\{P(Z'=1) + \tfrac{1}{a} P(Z' = \tfrac{1}{a})\} \\
&\quad + (a+1)\{P(Z'=1) + P(Z' = \tfrac{1}{a})\} \\
&= P(Z'=1) + aP(Z' = \tfrac{1}{a}) = E\frac{1}{Z'} \quad *
\end{aligned}$$

Since $EZ \in [\frac{1}{a}, 1]$ ($\because Z \in [\frac{1}{a}, 1]$), we can find $p \in [0,1]$ s.t.

$EZ = p + \frac{1}{a}(1-p)$ ($\because$ 

and we can find p s.t. $EZ = \frac{1}{a} + p(1-\frac{1}{a})$
$= p + \frac{1}{a}(1-p)$)

i.e. the desired r.v. Z' exists.

i.e. Let $P(Z'=1) = p$

$P(Z' = \frac{1}{a}) = 1-p$. Then we have

$$EZ' = EZ \text{ and } Z' \in \{1, \tfrac{1}{a}\}.$$

Next $*$

$$\Rightarrow E\frac{1}{Z} \le E\frac{1}{Z'}$$

$$\Rightarrow E\frac{1}{Z} EZ \le E\frac{1}{Z'} EZ' \quad (\because EZ = EZ')$$

$$\Rightarrow E\frac{1}{Z}EZ \le (p + a(1-p))(p + \tfrac{1}{a}(1-p))$$

$$= p^2 + \tfrac{1}{a}p - \tfrac{1}{a}p^2 + ap - ap^2 + 1 - 2p + p^2$$

$$= (2 - a - \tfrac{1}{a})p^2 + (\tfrac{1}{a} + a - 2)p + 1 =)$$

let $f(p) =)$, then $f'(p) = (4 - 2a - \frac{2}{a})p + (\frac{1}{a} + a - 2)$

let $f'(\hat{p}) = 0$, we get $\hat{p} = \frac{1}{2}$.

Also

$$f''(p) = 4 - 2(a + \tfrac{1}{a}) < 0 \quad (\because \tfrac{a + \frac{1}{a}}{2} > \sqrt{a \cdot \tfrac{1}{a}})$$

Therefore $E\frac{1}{Z}EZ \le f(\hat{p}) = f(\frac{1}{2}) = (\frac{1}{2} + \frac{a}{2})(\frac{1}{2} + \frac{1}{2a})$

$$= \frac{1}{4} + \frac{1}{4a} + \frac{a}{4} + \frac{1}{4}$$

$$= \frac{1}{2} + \frac{a + \frac{1}{a}}{4}.$$

QED.

4 Lemma 5' For any $s>0$,

$$E\{e^{-s d_H(A,X)}\} \le \frac{1}{P(X\in A)}\left(\frac{1}{2}+\frac{e^{s}+e^{-s}}{4}\right)^n$$

where $X=(X_1,\dots X_n)\in\Omega^n$, $X_1,\dots X_n$ are iid r.vs

$A\subset\Omega^n$ an arbitrary measurable set.

note: 证明过程中并没有用到 $X_1 \dots X_n$ 为独立.iid

pf: prove lemma 5' by induction on n.

$n=1$: i.e. r.v. X takes its value in Ω, $A\subset\Omega$.

$$E\{e^{s d_H(A,X)}\}$$

$$=E\{e^{s I_{\{X\notin A\}}}\}=P_r\{X\in A\}+e^{s}P_r\{X\notin A\}$$

$$=\min\{e^{s},\tfrac{1}{1}\}P\{I_{\{X\in A\}}=1\}+\min\{e^{s},\tfrac{1}{0}\}P_r\{I_{\{X\in A\}}=0\}$$

$$=E\left\{\min\left(e^{s},\frac{1}{I_{\{X\in A\}}}\right)\right\}$$

$$=\frac{1}{E\,I_{\{X\in A\}}}\,E I_{\{X\in A\}}\,E\left\{\min\left(e^{s},\frac{1}{I_{\{X\in A\}}}\right)\right\}$$

$$\le \frac{1}{P(X\in A)}\left(\frac{1}{2}+\frac{e^{s}+e^{-s}}{4}\right) \text{ by lemma 5.}$$

$(\because I_{\{X\in A\}}\in[0,1])$

Next assume lemma 5' to true for all $A\subset\Omega^n$. Then we prove that it is also true for every $A\subset\Omega^{n+1}$.

For $x_{n+1}\in\Omega$, denote $A(x_{n+1})=\{x\in\Omega^n : (x,x_{n+1})\in A\}$.

Claim: $d_H(A,(x,x_{n+1}))\le d_H(A(x_{n+1}),x)$

pf: Say $y\in A(x_{n+1})$ with $d_H(y,x)=d_H(A(x_{n+1}),x)$.

Thus $(y,x_{n+1})\in A$. Since $d_H((y,x_{n+1}),(x,x_{n+1}))=d_H(y,x)$,

we have $d_H(A,(x,x_{n+1}))\le d_H(A(x_{n+1}),x)$. end of claim

Induction hypothesis $\Rightarrow$

$$E\{e^{s d_H(A,(X,X_{n+1}))}\mid X_{n+1}=x_{n+1}\}$$

$$=E\{e^{s d_H(A,(X,x_{n+1}))}\}$$

$$\le E\{e^{s d_H(A(x_{n+1}),X)}\} \text{ by above claim}$$

$$\le \frac{1}{P\{X\in A(x_{n+1})\}}\left(\frac{1}{2}+\frac{e^{s}+e^{-s}}{4}\right)^n$$

let $B = \{x \in \Omega^n : \exists x_{n+1} \in \Omega : (x, x_{n+1}) \in A\}$ the projection of A on Ω^n.

Claim: $d_H(A, (x, x_{n+1})) \le d_H(B, x) + 1$.

pf: Say $y \in B$ with $d(y, x) = d(B, x)$. Thus we have $(y, x'_{n+1}) \in A$ for some $x'_{n+1} \in \Omega$.

Note that $d_H((y, x'_{n+1}), (x, x_{n+1})) \le d_H(y, x) + 1$

Thus done!

(B)
$$
\begin{aligned}
&E\{e^{\rho d_H(A, (X, X_{n+1}))} \mid X_{n+1} = x_{n+1}\} \\
&= E\{e^{\rho d_H(A, (X, x_{n+1}))}\} \\
&\le E\{e^{\rho d_H(B, X) + \rho}\} \\
&\le \frac{e^{\rho}}{P(X \in B)} \left(\frac{1}{2} + \frac{e^{\rho} + e^{-\rho}}{4}\right)^n
\end{aligned}
$$

(C)
$$
\begin{aligned}
(A) + (B) \Longrightarrow\ &E\{e^{\rho d_H(A, (X, X_{n+1}))} \mid X_{n+1} = x_{n+1}\} \\
&\le \min\left\{\frac{1}{P(X \in A(x_{n+1}))}, \frac{e^{\rho}}{P(X \in B)}\right\} \left(\frac{1}{2} + \frac{e^{\rho} + e^{-\rho}}{4}\right)^n
\end{aligned}
$$

Therefore
$$
\begin{aligned}
&E\{e^{\rho d_H(A, (X, X_{n+1}))}\} \\
&= E\{E\{e^{\rho d_H(A, (X, X_{n+1}))} \mid X_{n+1}\}\} \\
&= E \min\left\{\frac{1}{P(X \in A(X_{n+1}) \mid X_{n+1})}, \frac{e^{\rho}}{P(X \in B)}\right\} \left(\frac{1}{2} + \frac{e^{\rho} + e^{-\rho}}{4}\right)^n \\
&= \frac{1}{P(X \in B)} \frac{1}{\underbrace{E\left\{\frac{P\{X \in A(X_{n+1}) \mid X_{n+1}\}}{P(X \in B)}\right\}}_{\star}} \star E\left\{\min\left(e^{\rho}, \frac{P(X \in B)}{P(X \in A(X_{n+1}) \mid X_{n+1})}\right)\right\} \\
&\le \frac{1}{E\{P\{X \in A(X_{n+1}) \mid X_{n+1}\}\}} \left(\frac{1}{2} + \frac{e^{\rho} + e^{-\rho}}{4}\right) \\
&\qquad (\because \text{lemma 5 and } P(X \in B) = P(X \in \textstyle\bigcup_{x_{n+1} \in \Omega} A(x_{n+1})) \ge P\{X \in A(x_{n+1})\} \text{ for any } x_{n+1}.) \\
&= \frac{1}{P\{X \in A(X_{n+1})\}} \left(\frac{1}{2} + \frac{e^{\rho} + e^{-\rho}}{4}\right) \\
&= \frac{1}{P((X, X_{n+1}) \in A)} \left(\frac{1}{2} + \frac{e^{\rho} + e^{-\rho}}{4}\right) \quad \text{done!}
\end{aligned}
$$

寫成這樣的原因為若只寫成 $P(X \in A(X_{n+1}))$ 會變成一個值，故用 $P(\cdot \mid X_{n+1})$ 來強調此為一 X_{n+1} 的 r.v.

in fact 这时 $t \in \mathbb{R}^+$ 也对!

5 Thm 10 (Talagrand's first inequality): For any positive integer t,

$$\Pr\{d_H(A,X) \ge t\}\, P\{X \in A\} \le e^{-\frac{t^2}{n}}$$

where $X = (X_1, \dots, X_n) \in \Omega^n$, $X_1, \dots, X_n$ iid rvs. note: 证明中没有用到 iid.

$A \subset \Omega^n$ an arbitrary measurable set.

pf: For $s > 0$

$$P\{d_H(A,X) \ge t\} = \Pr\{s d_H(A,X) \ge s t\}$$

$$\le e^{-st} E\{e^{s d_H(A,X)}\}$$

Item 4

$$\le \frac{e^{-st}}{P(X \in A)} \left(\frac{1}{2} + \frac{e^s + e^{-s}}{4}\right)^n = \star$$

Note that

$$1 + \frac{e^s + e^{-s}}{2} = 2 + \sum_{n=1}^{\infty} \frac{s^{2n}}{(2n)!}$$

$$\le 2 + \sum_{n=1}^{\infty} \frac{s^{2n}}{(4^n n!)/2} = \sum_{n=0}^{\infty} 2 \frac{\left(\frac{s^2}{4}\right)^n}{n!} = 2 e^{\frac{s^2}{4}}$$

Thus

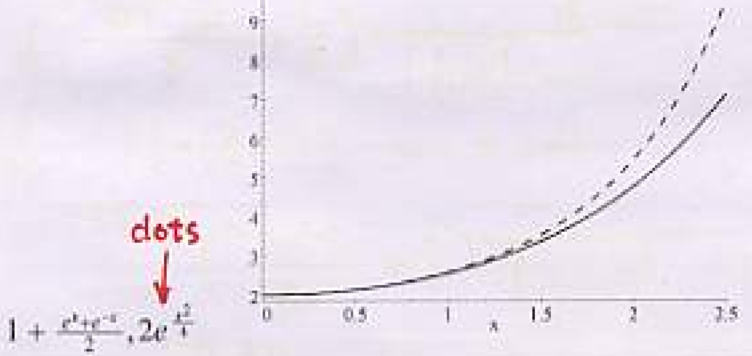
$$\star \le \frac{e^{-st} e^{\frac{n s^2}{4}}}{P(X \in A)}$$

$$= \frac{e^{-\frac{2t}{n} t} e^{\frac{n}{4} \frac{4t^2}{n^2}}}{P(X \in A)} \quad \text{by taking } s = \frac{2t}{n}$$

$$= \frac{1}{P(X \in A)} e^{-\frac{2t^2}{n} + \frac{t^2}{n}} = \frac{1}{P(X \in A)} e^{-\frac{t^2}{n}}$$

QED.

Remark:



$$2e^{\frac{x^2}{4}} = 2 + \frac{1}{2}x^2 + \frac{1}{16}x^4 + \frac{1}{192}x^6 + \frac{1}{3072}x^8 + O(x^{10})$$

$$1 + \frac{e^x + e^{-x}}{2} = 2 + \frac{1}{2}x^2 + \frac{1}{24}x^4 + \frac{1}{720}x^6 + \frac{1}{40320}x^8 + O(x^{10})$$

6. remark: Observe that on the left hand side we have the measure of the set of points whose Hamming distance from A is at least t.
If we consider a set with $P_r\{A\} \approx \frac{1}{2}$, we see something very surprising: The measure of the set of points whose Hamming distance to A is more than $10\sqrt{n}$ is smaller than e^{-100}!
In other words, product measure are concentrated on extremely small sets — hence the name "concentration of measure".

McDiarmid's ineq = Talagrand's ~~first~~ ineq

1. McDiarmid's Ineq.:
   - $X_1, X_2, \dots, X_n$ are independent rvs $X_i \in S_i$, $i = 1, 2, \dots, n$.
   - $f: S_1 \times S_2 \times \cdots \times S_n \to \mathbb{R}$ satisfies $|f(\underset{\sim}{x}) - f(\underset{\sim}{y})| \le c_i$ for all $\underset{\sim}{x}, \underset{\sim}{y}$ that differ only in the ith coordinate.

   Then for any $t > 0$,

$$P\{f(X_1, \dots, X_n) - Ef(X_1, \dots, X_n) \ge t\} \le e^{-\frac{2t^2}{\sum_{i=1}^n c_i^2}}$$

$$P\{f(X_1, \dots, X_n) - Ef(X_1, \dots, X_n) \le -t\} \le e^{-\frac{2t^2}{\sum_{i=1}^n c_i^2}}$$

Talagrand's first Inequality

identically dist.
後來沒用到，可去掉，亦成立

- Let $X_1, \dots, X_n$ be iid rvs, $X = (X_1, \dots, X_n)$
- $A \subset S^n$ an arbitrary measurable set

Then for any $t > 0$,

$$P\{d_H(A, (X_1, \dots, X_n)) \ge t\} \le \frac{1}{P\{(X_1, \dots, X_n) \in A\}} \cdot e^{-\frac{t^2}{n}}$$

2. Fact: The above two inequalities are equivalent.

pf: (Talagrand $\Rightarrow$ McDiarmid)

median: The median of a r.v. Y is the set of all y s.t. $P(Y \le y) \ge \frac{1}{2}$ and $P(Y \ge y) \ge \frac{1}{2}$ i.e. the place where $F(y) = P(Y \le y)$ crosses $\frac{1}{2}$. ∎

- Let $f: \underbrace{S \times \cdots \times S}_{S^n} \to \mathbb{R}$ s.t. $d_H(\underset{\sim}{x}, \underset{\sim}{y}) \le 1 \Rightarrow |f(\underset{\sim}{x}) - f(\underset{\sim}{y})| \le 1$, $\underset{\sim}{x}, \underset{\sim}{y} \in S^n$
- Let $A = \{\underset{\sim}{x} \in S^n : f(\underset{\sim}{x}) \le M[f]\}$, where $M[f]$ is the median of r.v. $f(X_1, \dots, X_n)$ i.e. $P\{f(X_1, \dots, X_n) \ge M[f]\} \ge \frac{1}{2}$ and, $P\{f(X_1, \dots, X_n) \le M[f]\} \ge \frac{1}{2}$

Claim: $d_H(A, (x_1, \dots, x_n)) < t \Rightarrow f(x_1, \dots, x_n) - M[f] < t$

pf: Let $\underset{\sim}{x}_0 = (x_1, \dots, x_n)$, then $\exists\, \underset{\sim}{x}_1, \dots, \underset{\sim}{x}_{t-1} \in S^n$ s.t.

$\underset{\sim}{x}_0 - \underset{\sim}{x}_1 - \underset{\sim}{x}_2 - \cdots - \underset{\sim}{x}_{t-1} \in A$ with $d_H(\underset{\sim}{x}_i, \underset{\sim}{x}_{i+1}) \le 1$, $i = 0, \dots, t-2$

$|f(\underset{\sim}{x}_0) - f(\underset{\sim}{x}_{t-1})| \le \sum_{i=0}^{t-2} |f(x_i) - f(x_{i+1})| \le t-1$ ($\because f$ is Lipschitz)

$\Rightarrow f(x_1, x_2, \dots, x_n) = f(\underset{\sim}{x}_0) \le f(\underset{\sim}{x}_{t-1}) + t - 1 \le M[f] + t - 1$

$\Rightarrow f(x_1, \dots, x_n) - M[f] \le t - 1 < t$. ∎

Date
No.

Therefore Talagrand's ineq

$$\Rightarrow P\{d_H(A,(x_1\cdots x_n))\ge t\}\le \frac{e^{-\frac{t^2}{n}}}{P\{(x_1,\cdots x_n)\in A\}}$$

$$\Rightarrow P\{d_H(A,(x_1\cdots x_n))\ge t\}\le \frac{e^{-\frac{t^2}{n}}}{P\{f(x_1\cdots x_n)\le M[f]\}}$$

$$\Rightarrow P\{d_H(A,(x_1\cdots x_n))\ge t\}\le 2e^{-\frac{t^2}{n}}$$

$$\Rightarrow P\{f(X_1\cdots X_n)-M[f]\ge t\}\le 2e^{-\frac{t^2}{n}} \quad (\because \text{前頁 claim})$$

This has the same form as McDiarmid's ineq, except that the expected value of $f(X_1,\ldots X_n)$ is replaced by its median. (The constants are also a bit worse, but that can be fixed.)

This difference is usually negligible, since

$$|Ef-Mf|\le E|f-Mf|=\int_0^\infty P\{|f-Mf|>t\}\,dt$$

Remark: For any nonnegative r.v. X

$$EX=\int_0^\infty P(X>t)\,dt=\int_0^\infty (1-F_X(t))\,dt$$

so whenever the deviation of f from its mean is small, its expected value must be close to its median.

(Talagrand's ineq ⟸ McDiarmid's)

For an arbitrary measurable subset $A\subseteq S^n$. Let $\underset{\sim}{X}=(X_1,\cdots X_n)$

Let $f(\underset{\sim}{x}): S^n\to\mathbb{R}$ s.t. $f(\underset{\sim}{x})=d_H(A,\underset{\sim}{x})$

Note that $d_H(\underset{\sim}{x},\underset{\sim}{y})\le 1 \Rightarrow \begin{cases} d_H(A,\underset{\sim}{x})\le 1+d_H(A,\underset{\sim}{y})\\ d_H(A,\underset{\sim}{y})\le 1+d_H(A,\underset{\sim}{x})\end{cases}$

$$\Rightarrow |d_H(A,\underset{\sim}{x})-d_H(A,\underset{\sim}{y})|\le 1$$

$$\Rightarrow |f(\underset{\sim}{x})-f(\underset{\sim}{y})|\le 1$$

McDiarmid's ineq $\Rightarrow \Pr\{f(\underset{\sim}{X})-Ef(\underset{\sim}{X})\le -t\}\le e^{-\frac{2t^2}{n}}$

$$\Rightarrow \Pr\{Ed_H(A,\underset{\sim}{X})-d_H(A,\underset{\sim}{X})\ge t\}\le e^{-\frac{2t^2}{n}}$$

$$\Rightarrow \Pr\{d_H(A,\underset{\sim}{X})\le 0\}\le \exp\left\{-\frac{2E^2 d_H(A,\underset{\sim}{X})}{n}\right\}$$

$$\Rightarrow \Pr\{\underset{\sim}{X}\in A\}\le \exp\left\{-\frac{2E^2 d_H(A,\underset{\sim}{X})}{n}\right\}$$

$$\Rightarrow Ed_H(A,\underset{\sim}{X})\le \sqrt{\frac{n}{2}\ln\frac{1}{P(\underset{\sim}{X}\in A)}} \quad \bigstar$$

Thus

$$P\left\{ d_H(A, \underset{\sim}{X}) \geq t + \sqrt{\frac{n}{2} \ln \frac{1}{P(\underset{\sim}{X} \in A)}} \right\}$$

$$= P\left\{ f(\underset{\sim}{X}) \geq t + \sqrt{\frac{n}{2} \ln \frac{1}{P(\underset{\sim}{X} \in A)}} \right\}$$

$$\leq P\left\{ f(\underset{\sim}{X}) \geq t + E d_H(A, \underset{\sim}{X}) \right\} \quad (\because \bigstar)$$

$$= P\left\{ f(\underset{\sim}{X}) \geq t + E f(\underset{\sim}{X}) \right\} \leq e^{-\frac{2t^2}{n}}$$ (useing McDiarmid's again)

or in other words,

$$P\{ d_H(A, \underset{\sim}{X}) \geq n\varepsilon \} \leq e^{-2n\left(\varepsilon - \sqrt{\frac{1}{2n} \ln \frac{1}{P(\underset{\sim}{X} \in A)}}\right)^2}$$

whenever $\varepsilon \geq \sqrt{\frac{1}{2n} \ln \frac{1}{P_r(\underset{\sim}{X} \in A)}}$,

which actually gives the optimal constant for Talagrand's first inequality.