۱. TV و ارتباط آن با برخی انحراف ها

۱. $$\|P - Q\|_{TV} \leq \sum_{i=1}^{n} \mathbb{E}_{X_{1:i-1}\sim P}\left[\|P_i(\cdot \mid X_{1:i-1}) - Q_i(\cdot \mid X_{1:i-1})\|_{TV}\right]$$

طرفین: $$d_{TV}(P,Q) = \frac{1}{2}\sum_{x}|P(x) - q(x)| \longrightarrow d_{TV}(P^{\otimes n}, Q^{\otimes n}) = \frac{1}{2}\sum_{x_1 \dots x_n}|P(x_{1:n}) - Q(x_{1:n})|$$

$$\Rightarrow d_{TV}(P^{\otimes n}, Q^{\otimes n}) = \frac{1}{2}\sum_{x_{1:n}}|P(x_{2:n}\mid x_1)P(x_1) - q(x_{2:n}\mid x_1)q(x_1)|$$

$$= \frac{1}{2}\sum |P(x_{2:n}\mid x_1)P(x_1) - q(x_{2:n}\mid x_1)p(x_1) + q(x_{2:n}\mid x_1)p(x_1) - q(x_{2:n}\mid x_1)q(x_1)|$$

(نامساوی مثلث) $$\leq \frac{1}{2}\sum |P(x_1)\left(P(x_{2:n}\mid x_1) - q(x_{2:n}\mid x_1)\right)| + \frac{1}{2}\sum \underbrace{\underbrace{q(x_{2:n}\mid x_1)}_{q(x_{2:n})}}_{\leq 1}|P(x_1) - q(x_1)|$$

$$\leq \frac{1}{2}\sum P(x_1)|P(x_{2:n}\mid x_1) - q(x_{2:n}\mid x_1)| + d_{TV}(P(x_1), q(x_1))$$

$$= \frac{1}{2}\sum P(x_1)|P(x_{3:n}\mid x_1 x_2)P(x_2) - q(x_{3:n}\mid x_1, x_2)q(x_2)| + d_{TV}(P(x_1), q(x_1))$$

از نامساوی مثلث $$\leq d_{TV}(P(x_1), q(x_1)) + \frac{1}{2}\sum p(x_{1:2})|P_{x_{3:n}}(\cdot \mid x_{1:2}) - q_{x_{3:n}}(\cdot \mid x_{1:2})| +$$

$$+ \frac{1}{2}\sum P(x_1)\underbrace{q(x_{3:n}\mid x_{1:2})}_{\leq 1}|P(x_2\mid x_1) - q(x_2\mid x_1)|$$

$$\Rightarrow \leq d_{TV}(P_{x_1}, q_{x_1}) + \frac{1}{2}\sum P(x_{1:2})|P_{x_{3:n}}(\cdot \mid x_{1:2}) - q_{x_{3:n}}(\cdot \mid x_{1:2})| + \underbrace{\frac{1}{2}\sum P(x_1)|P(x_2\mid x_1) - q(x_2\mid x_1)|}_{\mathbb{E}_{x_1}[d_{TV}(P_{x_2|x_1}, q_{x_2|x_1})]}$$

$$\leq d_{TV}(P_{x_1}, q_{x_1}) + \mathbb{E}_{x_1}[d_{TV}(P_{x_2|x_1}, q_{x_2|x_1})] + \frac{1}{2}\sum P(x_{1:2})|P_{x_{3:n}}(\cdot \mid x_{1:2}) - q_{x_{3:n}}(\cdot \mid x_{1:2})|$$

Subject : Date :

به طور مشابه و استقرایی هر بار در هر مرحله یک ترم $\{\mathbb{E}_{X_{1:i-1}}[d_{TV}(P_i(\cdot|X_{1:i-1}), Q_i(\cdot|X_{1:i-1}))]\}$ بیرون می‌کشیم. پس خواهیم داشت:

$$d_{TV}(P^{\otimes n}, Q^{\otimes n}) \leq \sum_{i=1}^{n} \mathbb{E}_{X_{1:i-1}\sim P}\left[d_{TV}(P_i(\cdot|X_{1:i-1}), Q_i(\cdot|X_{1:i-1}))\right] \qquad \blacksquare$$

۲. این همان نامساوی ای است که در سوال آخر تمرین قبل بدون اثبات استفاده کردیم! این نامساوی را با تعریف متغیرهای کمکی U, V, W اثبات می‌کنیم:

$$U \triangleq \frac{q(x)}{p(x)}, \quad V \triangleq \frac{q(x)-p(x)}{p(x)}\Big|_{p(x)<q(x)}, \quad W \triangleq \frac{p(x)-q(x)}{p(x)}\Big|_{q<p}$$

دقت کنید که $V := (U-1)_+$.

داریم:

$$d_{TV}(P,Q) = \frac{1}{2}\mathbb{E}_p[|U-1|] = \mathbb{E}_p[V] = \mathbb{E}_p[W]$$

داریم $\Rightarrow$ $U = (1+V)(1-W) \longrightarrow \log U = \log(1+V) + \log(1-W)$

پس $\rightarrow$ $D_{KL}(P,q) = \mathbb{E}_p\left[\log\left(\frac{p}{q}\right)\right] = \mathbb{E}_p[\log U] = \mathbb{E}_p[\log(1+V)] + \mathbb{E}_p[\log(1-W)]$

Jensen $\longrightarrow$ $D_{KL}(P,q) \leq \log(1+\underbrace{\mathbb{E}_p[V]}_{d_{TV}}) + \log(1-\underbrace{\mathbb{E}_p[W]}_{d_{TV}})$

$$\leq \log(1+d_{TV}(P,q)) + \log(1-d_{TV}(P,q))$$

پس $\Rightarrow$ $D_{KL}(P,q) \leq \log(1-d_{TV}(P,q)^2)$

$$\Rightarrow e^{-D_{KL}(P,q)} \leq 1-d_{TV}(P,q)^2$$

$$\Rightarrow d_{TV}(P,q)^2 \leq 1-e^{-D_{KL}(P,q)}$$

MICRO

Subject : Date :

$$\Rightarrow d_{TV}(P,q) \leq \sqrt{1-\exp(D_{KL}(P,q))}$$

پس کافیست نامساوی اثبات شود. برای قسمت دیگر داریم که به وضوح برای x داریم:

$$\sqrt{1-e^{-x}} \leq 1-\frac{1}{2}e^{-x}$$

$$\Leftrightarrow 1-e^{-x} \leq 1-e^{-x}+\frac{1}{4}e^{-2x}$$

$$\Leftrightarrow 0 \leq e^{-2x} \checkmark \qquad \blacksquare$$

$$H^2(P,q) = \sum(\sqrt{P}-\sqrt{q})^2 = \sum P - 2\sqrt{Pq} + q = 2\left(1-\sum\sqrt{Pq}\right) = 2(1-BC(P,Q))$$ ۳.

$$d_{TV}(P,Q)+1 = \sum_x \max(p(x),q(x)) \quad , \quad 1-d_{TV}(P,Q) = \sum \min(p(x),q(x)) \leq \sum\sqrt{p(x)q(x)} = BC(P,Q)$$

$$\Rightarrow (1+d_{TV})(1-d_{TV}) \underset{\text{کوشی شوارتز}}{\geq} \left(\sum\sqrt{p(x)q(x)}\right)^2 \Rightarrow 1-TV(P,Q)^2 \geq BC(P,Q)^2$$

حال با توجه به حالت‌های بالا داریم:

$$\left.\begin{array}{l} 1-TV \leq BC \rightarrow \frac{H^2}{2} \leq TV \\ 1-TV^2 \geq BC^2 \\ BC = 1-\frac{H^2}{2} \\ 1-TV^2 \geq 1-H^2-\frac{H^4}{4} \rightarrow H^2\left(1-\frac{H^2}{4}\right) \geq TV^2 \end{array}\right\} \rightarrow H^2(P,Q) \leq TV \leq \sqrt{\frac{H^2}{4}(4-H^2(P,Q))}$$

$$\frac{H}{2}\sqrt{4-H^2}$$

$$1-\sqrt{1-TV^2} \leq \frac{H^2}{2} \leq TV$$

MICRO

Subject : Date :

$BC(P^{\otimes n}, Q^{\otimes n}) = BC(P,Q)^n$ حال برای $P^{\otimes n}$ ، $Q^{\otimes n}$ نیز می‌توان این نامساوی‌ها را تعمیم داد. با توجه به:

خواهیم داشت:

$$\frac{1}{2} H^2(P^{\otimes n}, Q^{\otimes n}) \leq TV(P^{\otimes n}, Q^{\otimes n}) \leq \sqrt{\frac{H^2(P^{\otimes n}, Q^{\otimes n})}{4}\left(4 - H^2(P^{\otimes n}, Q^{\otimes n})\right)}$$

$$, \quad 1 - \sqrt{1 - TV^2(P^{\otimes n}, Q^{\otimes n})} \leq \frac{1}{2} H^2(P^{\otimes n}, Q^{\otimes n}) \leq TV(P^{\otimes n}, Q^{\otimes n})$$

از طرفی برای $H^2(P^{\otimes n}, Q^{\otimes n})$ داریم:

$$H^2(P^{\otimes n}, Q^{\otimes n}) = 2 - 2\sqrt{P^{\otimes n}, Q^{\otimes n}} = 2 - 2\sqrt{\prod_{i=1}^{n} P(x_i) Q(x_i)}$$

$$= 2 - 2\left(\sqrt{P(x_i) Q(x_i)}\right)^n = 2 - 2\left(1 - \frac{1}{2} H^2(P,Q)\right)^2$$

اثبات طرف اول: اگر $H^2(P^{\otimes}, Q^{\otimes}) \to o/\omega(\frac{1}{n})$ آنگاه خواهیم داشت: $TV(P^{\otimes}, Q^{\otimes}) = o/1$

$$\Rightarrow \text{if } H^2(P^{\otimes n}, Q^{\otimes n}) \to o(\tfrac{1}{n}) \xrightarrow{\text{then}} \lim_{n\to\infty} \frac{H^2(P,Q)}{\frac{1}{n}} = 0 \quad \text{and} \quad |H^2(P^{\otimes}, Q^{\otimes})| < k\frac{1}{n}, \ k>0$$

we have : $H^2(P^{\otimes n}, Q^{\otimes n}) = 2 - 2\left(1 - \frac{1}{2} H^2(P_n, Q_n)\right)^n$

$$\to \lim_{n\to\infty} \left(1 - H^2(P_n, Q_n)\right)^n = \lim_{n\to\infty} \log_2\left(1 - \frac{1}{2} H^2(P_n, Q_n)\right)^n$$

$$= \lim_{n\to\infty} n \log_2\left(1 - \frac{1}{2} H^2(P_n, Q_n)\right) = \lim_{n\to\infty} \frac{\log_2\left(1 - \frac{1}{2} H^2(P_n, Q_n)\right)}{\frac{1}{n}}$$

بسط تیلور، $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3}$ $\longrightarrow$

$$= \lim_{n\to\infty} \frac{-\frac{1}{2}H^2 - \frac{(-\frac{1}{2}H^2)^2}{2} + \frac{(-\frac{1}{2}H^2)^3}{3} + \cdots}{\frac{1}{n}} = 0$$

داریم:

$$\lim_{n\to\infty} \log_2\left(1 - \frac{1}{2} H^2(P,Q)\right)^n = 0 \to \lim_{n\to\infty} \left(1 - \frac{1}{2} H^2(P_n, Q_n)\right) = 2^0 = 1$$

MICRO

پس: $\Rightarrow \lim_{n\to\infty} H^2(P^{\otimes n}, Q^{\otimes n}) = 2 - 2\lim_{n\to\infty} \underbrace{\left(1-\frac{1}{2}H^2(P_n,Q_n)\right)^2}_{1} = 0$

حال با توجه به نامساوی: $\frac{1}{2}H^2(P^{\otimes n}, Q^{\otimes n}) \le TV(P^{\otimes n}, Q^{\otimes n}) \le \sqrt{\frac{H^2(P^{\otimes n}, Q^{\otimes n})}{4}\left(4 - H^2(P^{\otimes n}, Q^{\otimes n})\right)}$

داریم: $\lim \frac{1}{2}H^2(N) \le \lim_{n\to\infty} TV(N) \le \lim_{n\to\infty} \sqrt{N}$

$$\Rightarrow 0 \le \lim_{n\to\infty} TV(N) \le 0$$

پس طبق قضیه ساندویچ، داریم: $\lim_{n\to\infty} TV(P^{\otimes n}, Q^{\otimes n}) = 0$ ■

حال برای حالت $\text{if } H^2(P,Q) \to \omega(\frac{1}{n})$ را بررسی می‌کنیم.

$$\Rightarrow \text{if } H^2(P,Q) \to \omega(\tfrac{1}{n}) \xrightarrow{\text{then}} H^2(P_n,Q_n) \ge k\frac{1}{n} \ (k>0) \quad \text{and} \quad \lim_{n\to\infty} \frac{H^2(P_n,Q_n)}{\frac{1}{n}} = \infty$$

داریم $\Rightarrow \lim_{n\to\infty} \left(1-\frac{1}{2}H^2(P_n,Q_n)\right)^n \Rightarrow \lim_{n\to\infty} \log\left(1-\frac{1}{2}H^2(P_n,Q_n)\right)^n$

$$= \lim_{n\to\infty} \frac{\log\left(1-\frac{1}{2}H^2(P_n,Q_n)\right)}{\frac{1}{n}}$$

Taylor Series $= \lim_{n\to\infty} \frac{-\frac{1}{2}H^2(P_n,Q_n) - \frac{\left(-\frac{1}{2}H^2(P_n,Q_n)\right)^2}{2} + O(\frac{1}{n^3})}{\frac{1}{n}}$

$$= \lim_{n\to\infty} \frac{-\frac{1}{2}H^2(P_n,Q_n)}{\frac{1}{n}} = -\infty$$

پس: $\lim_{n\to\infty} \left(1-\frac{1}{2}H^2(P_n,Q_n)\right)^n = 2^{-\infty} = 0$

Subject : Date :

$$H^2(P^{\otimes n}, Q^{\otimes n}) = 2 - 2\left(1 - \frac{1}{2}H^2(P_n, Q_n)\right)^n$$

$$\rightarrow \lim_{n\to\infty} H^2(P^{\otimes n}, Q^{\otimes n}) = 2 - 2\underbrace{\lim_{n\to\infty}\left(1 - \frac{1}{2}H^2(P_n, Q_n)\right)^n}_{0} = 2$$

مانند حالت قبل داریم: $\frac{1}{2}H^2(P^{\otimes n}, Q^{\otimes n}) \leq TV(P^{\otimes n}, Q^{\otimes n}) \leq \sqrt{\frac{H^2(P^{\otimes n}, Q^{\otimes n})}{4}\left(4 - H^2(P^{\otimes n}, Q^{\otimes n})\right)}$

$$\longrightarrow \lim_{n\to\infty} \frac{1}{2}H^2(N) \leq \lim_{n\to\infty} TV(N) \leq \lim_{n\to\infty}\sqrt{N}$$

$$\longrightarrow \frac{2}{2} \leq \lim_{n\to\infty} TV(N) \leq \sqrt{\frac{2}{4}(4-2)}$$

$$\longrightarrow 1 \leq \lim_{n\to\infty} TV(P^{\otimes n}, Q^{\otimes n}) \leq 1$$

پس مانند بخش قبل، با توجه به قضیه ساندویچ: $\boxed{} \quad \lim_{n\to\infty} TV(P^{\otimes n}, Q^{\otimes n}) = 1$

حال فرض کنید که: $TV(P_n^{\otimes n}, Q_n^{\otimes n}) \to 0$ داریم:

$$1 - \sqrt{1 - TV(P^{\otimes n}, Q^{\otimes n})} \leq \frac{1}{2}H^2(P^{\otimes n}, Q^{\otimes n}) \leq TV(P^{\otimes n}, Q^{\otimes n})$$

$$\Rightarrow 0 \leq H^2(P^{\otimes n}, Q^{\otimes n}) \leq 0$$

از طرفی $\longrightarrow H^2(P^{\otimes n}, Q^{\otimes n}) = 2 - 2\left(1 - \frac{1}{2}H^2(P_n, Q_n)\right)^n = 0$

$$\Rightarrow \lim_{n\to\infty}\left(1 - \frac{1}{2}H^2(P_n, Q_n)\right)^n = 1 \longrightarrow H^2(P_n, Q_n) \to O\left(\frac{1}{n}\right) \quad \blacksquare$$

Subject :   Date :

حال فرض کنید که: $TV(P^{\otimes n}, Q^{\otimes n}) \to 1$

داریم $\Rightarrow \quad 1 - \sqrt{1 - TV(P^{\otimes n}, Q^{\otimes n})} \leq \frac{1}{2} H^2(P^{\otimes n}, Q^{\otimes n}) \leq TV(P^{\otimes n}, Q^{\otimes n})$

$$\Rightarrow \quad 1 \leq \frac{1}{2} H^2(P^{\otimes n}, Q^{\otimes n}) \leq 1$$

ساندویچ $\longrightarrow \lim_{n \to \infty} H^2(P^{\otimes n}, Q^{\otimes n}) = 2$

$$\longrightarrow \lim_{n \to \infty} H^2(P^{\otimes n}, Q^{\otimes n}) = 2 - 2\left(1 - \frac{1}{2} H^2(P_n, Q_n)\right)^n$$

$$\Rightarrow \lim_{n \to \infty} \left(1 - \frac{1}{2} H^2(P_n, Q_n)\right)^n = 0$$

$$\Longrightarrow H^2(P_n, Q_n) \longrightarrow \omega\left(\frac{1}{n}\right) \qquad \blacksquare$$

۴. $f(x) = x \log x + \underbrace{x - 1}_{\text{تغییری در مقدار f ایجاد نمی‌کند.}} = (x-1)^2 \int_0^1 \frac{1-t}{tx + (1-t)} \, dt$

$$\Rightarrow D_{KL}(P_0 \| P_1) = D_f(P_0, P_1) = \mathbb{E}_{x \sim Q}\left[\int_0^1 \frac{\left(\frac{P_0(x)}{P_1(x)} - 1\right)^2 (1-t)}{t\left(\frac{P_0(x)}{P_1(x)}\right) + (1-t)} \, dt\right]$$

$$= \sum_x \int_0^1 \frac{(P_0(x) - P_1(x))^2 (1-t)}{tP_0(x) + (1-t)P_1(x)} \, dt$$

داریم:

Subject :

Date :

$$\left(\sum_x \int_0^1 \frac{(P_0(x)-P_1(x))^2(1-t)}{tP_0(x)+(1-t)P_1(x)}\,dt\right)\times\left(\sum_x \int_0^1 \underbrace{(tP_0(x)+(1-t)P_1(x))(1-t)\,dt}_{t-\frac{t^2}{2}\big|_0^1=\frac{1}{2}}\right)$$

$$= \frac{1}{2}\left(\sum_x \int_0^1 \frac{(P_0(x)-P_1(x))^2(1-t)}{tP_0(x)+(1-t)P_1(x)}\,dt\right) \geq \left(\sum_x \int_0^1 (1-t)\,|P_0(x)-P_1(x)|\,dt\right)^2$$

کران tight است $\Rightarrow$ $$\frac{1}{2}\left(\sum_x \int_0^1 \frac{(P_0(x)-P_1(x))^2}{tP_0(x)+(1-t)P_1(x)}\,dt\right) = \left(2\,TV\times\frac{1}{2}\right)^2 = TV^2(P_0,P_1)$$

پس $\Rightarrow$ $$TV(P_0,P_1) = \sqrt{\frac{1}{2}\sum_x \int_0^1 \frac{(1-t)(P_0(x)-P_1(x))^2}{tP_0(x)+(1-t)P_1(x)}\,dt}$$

از طرفی $\frac{(1-t)(P_0(x)-P_1(x))^2}{tP_0(x)+(1-t)P_1(x)}$ که ... است پس $\frac{(1-t)(P_0(x)-P_1(x))^2}{tP_0(x)+(1-t)P_1(x)} \leq \frac{(1-t)(P_0(x)-P_1(x))^2}{g(x)}$

که $g(x)$ توزیعی بین P_0 و P_1 است.

پس:

$$\sum \int_0^1 \frac{(1-t)(P_0(x)-P_1(x))^2}{tP_0(x)+(1-t)P_1(x)}\,dt \leq \inf \sum_x \int_0^1 \frac{(1-t)(P_0(x)-P_1(x))^2}{g(x)}$$

$$\leq \inf_{g(x)} \frac{1}{2}\sum_x \frac{(P_0(x)-P_1(x))^2}{g(x)}$$

$$\Rightarrow TV(P_0,P_1) \leq \frac{1}{2}\inf_g \sqrt{\sum_x \frac{(P_0(x)-P_1(x))^2}{g(x)}}$$

tight بودن $\Rightarrow$ $$TV(P_0,P_1) = \frac{1}{2}\inf_{g(x)} \sqrt{\sum \frac{(P_0(x)-P_1(x))^2}{g(x)}}$$

Subject : Date :

$$I(X;Y) = D(P_{XY} \| P_X P_Y) = \mathbb{E}_{P_X}[D(P_{Y|X} \| P_Y)] + \underbrace{D(P_X \| P_X)}_{0}$$ ۵.

$$\Rightarrow P_X \begin{cases} 1 & \alpha \\ 0 & \bar{\alpha} \end{cases} \longrightarrow I(X;Y) = \alpha D(P_1 \| \alpha P_1 + \bar{\alpha} P_0) + \bar{\alpha} D(P_0 \| \alpha P_1 + \bar{\alpha} P_0)$$

Pinsker $\longrightarrow$ $I(X;Y) = \alpha D(P_1 \| \alpha P_1 + \bar{\alpha} P_0) + \bar{\alpha} D(P_0 \| \alpha P_1 + \bar{\alpha} P_0) \geq \alpha 2 TV^2(P_1, \alpha P_1 + \bar{\alpha} P_0)$
$+ \bar{\alpha} 2 TV^2(P_0, \alpha P_1 + \bar{\alpha} P_0)$

داریم $\Rightarrow TV(P_1, \alpha P_1 + \bar{\alpha} P_0) = \frac{1}{2} \sum_x |P_1 - \alpha P_1 - \bar{\alpha} P_0| = \frac{1}{2} \bar{\alpha} \sum |P_1 - P_0| = \bar{\alpha} TV(P_0, P_1)$

پس $\Rightarrow I(X;Y) \geq 2\alpha (\bar{\alpha} TV(P_0, P_1))^2 + 2\bar{\alpha} (\alpha TV(P_0, P_1))^2 = 2\alpha\bar{\alpha} TV(P_0, P_1)^2$

$\alpha = \frac{1}{2} \xrightarrow{\text{قرار میدهیم}} I(X;Y) \geq \frac{1}{2} TV(P_0, P_1)^2$ ∎

پس قسمت اول اثبات شد. برای قسمت دوم داریم:

$$I(X;Y) = \mathbb{E}_{P_X}\{D_{KL}(P_{Y|X} \| P_Y)\}$$

از طرفی داریم:

$$D_{KL}(P \| Q) = \sum_x P(x) \log\left(\frac{P(x)}{q(x)}\right) \overset{\log x \leq x-1}{\leq} \sum_x p(x)\left(\frac{p(x)}{q(x)} - 1\right) = \sum \frac{P(x)^2}{q(x)} - 1 = \chi^2(P,Q)$$

با استفاده از $D_{KL}(P \| Q) \leq \chi^2(P,Q)$ داریم:

$$I(X;Y) = \mathbb{E}_{P_X}[D_{KL}(P_{Y|X} \| P_Y)] \leq \mathbb{E}_{P_X}[\chi^2(P_{Y|X}, P_Y)]$$

داریم $P_Y = \alpha P_1 + \bar{\alpha} P_0$ پس:

Subject :   Date :

$$\mathbb{E}_{P_X}\left[\chi^2(P_{Y|X}, P_Y)\right] = \alpha\, \chi^2(P_1 \| P_Y) + \bar{\alpha}\, \chi^2(P_0 \| P_Y) = \alpha \sum \frac{P_1(x)^2}{\alpha P_1(x) + \bar{\alpha} P_0(x)} - 1 + \bar{\alpha} \sum \frac{P_0(x)^2}{\alpha P_1(x) + \bar{\alpha} P_0(x)}$$

$$\leq \sum_x \max(P_0, P_1) \underbrace{\frac{\alpha P_1(x) + \bar{\alpha} P_0(x)}{\alpha P_1(x) + \bar{\alpha} P_0(x)}}_{1} - 1 = TV(P_0, P_1)$$

$$\Rightarrow I(X;Y) \leq TV(P_0, P_1) \qquad \blacksquare$$

پس هر دو طرف اثبات شد.

۲ نشت اطلاعات

داریم:

هر رأس: $X_v \sim Ber(\frac{1}{2}) \longrightarrow X_v = \begin{cases} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{cases}$

هر یال: $Z_e \sim Ber(\frac{1}{2}) \longrightarrow Z_e = \begin{cases} 1 & \delta \\ 0 & 1-\delta \end{cases}$

$\xrightarrow{\text{مدل}} Y_e = X_u \oplus X_v \oplus Z_e$

احتمال حضور هر یال $\longrightarrow e = (u,v) \longrightarrow \begin{cases} 1: \text{exist} & P \\ 0: \text{not exist} & 1-P \end{cases}$

ا

$$Y_{\mathcal{E}} = Y_{e_1 e_2 \dots e_{|\mathcal{E}|}} \Rightarrow I(X_v; Y_{\mathcal{E}}) = H(Y_{\mathcal{E}}) - H(Y_{\mathcal{E}} | X_v) = H(Y_{1:|\mathcal{E}|} | X_v)$$

دقت کنید که $X_v \sim Ber(\frac{1}{2})$ است. حال مجموعه تمام یال ها را ($|\mathcal{E}|$) را به دو دسته $\mathcal{E}_v$ که متصل به رأس v هستند و $\mathcal{E}_{\bar{v}}$ متصل نیستند افراز می‌کنیم.

$$I(X_v; Y_{\mathcal{E}}) = H(Y_{\mathcal{E}_v}, Y_{\mathcal{E}_{\bar{v}}}) - H(Y_{\mathcal{E}_v} Y_{\mathcal{E}_{\bar{v}}} | X_v)$$

از طرفی داریم:

$$\forall e \in Y_{\mathcal{E}_{\bar{v}}}: \ Y_e = X_v \oplus X_u \oplus Z_e, \ X_u \perp\!\!\!\perp X_v, \ Z_e \perp\!\!\!\perp X_v \Rightarrow Y_e \perp\!\!\!\perp X_v$$

$$\forall e \in Y_{\bar{\mathcal{E}}_v} : Y_e = X_{u_1} \oplus X_{u_2} \oplus Z_e , \quad X_{v_1} \perp\!\!\!\perp X_u , \quad X_{u_2} \perp\!\!\!\perp X_v , \quad Z_e \perp\!\!\!\perp X_v \implies X_v \perp\!\!\!\perp Y_e$$

پس در هر دو حالت $X_v \perp\!\!\!\perp Y_e$ پس: $H(Y_{\mathcal{E}_v}, Y_{\bar{\mathcal{E}}_v}) = H(Y_{\mathcal{E}_v}, Y_{\bar{\mathcal{E}}_v} | X_v)$

$$\Rightarrow I(Y_{\mathcal{E}}, X_v) = H(Y_{\mathcal{E}_v}, Y_{\bar{\mathcal{E}}_v}) - H(Y_{\mathcal{E}_v}, Y_{\bar{\mathcal{E}}_v} | X_v) = 0$$

۲. اگر $v \in S$ داریم: $\mathbb{P}_{(g,\eta)}(v \leadsto S) = 1$

$$\Rightarrow I(X_v ; X_S, Y_{\mathcal{E}}) = H(X_v) - H(X_v | X_S, Y_{\mathcal{E}}) = H(X_v) - H(X_v | X_v, X_{S-\{v\}}, Y_{\mathcal{E}}) = H(X_v) = \log 2$$

$$\Rightarrow I(X_v ; X_S, Y_{\mathcal{E}}) = \log 2$$

حال فرض کنید که $v \notin S$. برای این حالت روی $|\mathcal{E}|$ استقرا می‌زنیم:

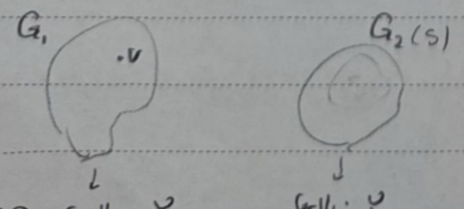
پایه استقرا: $|\mathcal{E}| = 0 \rightarrow X_v \perp\!\!\!\perp X_S \rightarrow I(X_S, X_v) = 0 \rightarrow 0 \leq \mathbb{P}_{(g,\eta)}[v \leadsto S] \log 2$

گام استقرا: حال فرض کنید برای $|\mathcal{E}| = k$ حکم درست باشد و صدق کند. حال باید درستی حکم را برای $|\mathcal{E}| = k+1$ ثابت کنیم:

در اینجا دو حالت داریم:

حالت اول: از مجموعه S به بیرون آن هیچ یالی نباشد. در این حالت می‌توان گراف را متعلق به دو زیرگراف افراز کرد که با توجه به فرض v در خارج S قرار دارد.

این دو زیرگراف را G_1 و G_2 می‌نامیم:



یال‌ها $\rightarrow Y_{G_1}$ — یال‌ها : Y_{G_2}

داریم: $X_v \perp X_S$ ، $X_v \perp Y_{G_2}$

$$\Rightarrow I(X_v ; X_S, Y_{\mathcal{E}}) = H(X_v) - H(X_v \mid X_S, X_{G_1}, X_{G_2}) = \underbrace{H(X_v) - H(X_v \mid Y_{G_1})}_{I(X_v, Y_{G_1})}$$

$$\Rightarrow I(X_v ; X_S, Y_{\mathcal{E}}) = I(X_v, Y_{G_1})$$

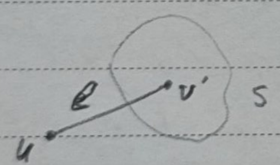
حال برای $I(X_v, Y_{G_1})$ بازهم مانند بخش قبل با افزایش یال‌ها به مجموعه که به v وصل هستند یا نه، می‌توان نتیجه گرفت $I(X_v, Y_G) = 0$

پس: $I(X_v ; X_S, Y_{\mathcal{E}}) = I(X_v, Y_{G_1}) = 0 \leq \underbrace{\mathbb{P}[v \rightsquigarrow S]}_{0} \log 2$ ✓

چون مسیری نیست.

در این حالت تساوی برقرار است.

حالت دوم: حال فرض کنید که از مجموعه S ما (G_2) یالی به بیرون آن داشته باشیم. مثلاً از v' درون آن به u خارج آن:



تعریف می‌کنیم: $\tilde{\mathcal{E}} = \mathcal{E} - \{e\} \Rightarrow \tilde{g} = \{\mathcal{V}, \tilde{\mathcal{E}}\}$

داریم $\Rightarrow$ $\mathbb{P}_{[g,\eta]}\{v \rightsquigarrow S\} = (1-\eta)\, \mathbb{P}_{[\tilde{g},\eta]}\{v \rightsquigarrow S\} + \eta \left(\mathbb{P}_{[\tilde{g},\eta]}\{v \rightsquigarrow S\} + \mathbb{P}_{[\tilde{g},\eta]}\{v \rightsquigarrow u\}\right)$

$$= (1-\eta)\, \mathbb{P}_{[\tilde{g},\eta]}(v \rightsquigarrow S) + \eta\, \mathbb{P}_{[\tilde{g},\eta]}[v \rightsquigarrow (S \cup \{u\})]$$

دقت کنید که $\tilde{g}$ یک یال کمتر دارد پس فرض استقرا برای آن برقرار است:

Subject : Date :

$$\mathbb{P}_{(g,n)}[V \rightsquigarrow S] \log 2 \geq I(X_V ; X_S, Y_{\mathcal{E}})$$ پس داریم :

$$\mathbb{P}_{[g,n]}[V \rightsquigarrow S \cup \{u\}] \log 2 \geq I(X_V ; X_S, X_u, Y_{\mathcal{E}})$$

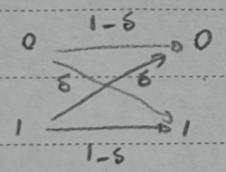
$$\Rightarrow \mathbb{P}_{g,n}[V \rightsquigarrow S] \log 2 = (1-n)\,\mathbb{P}_{g,n}[V \rightsquigarrow S] \log 2 + n\,\mathbb{P}_{g,n}[V \rightsquigarrow S \cup \{u\}] \log 2$$

$$\geq (1-n)\, I(X_V ; X_S, Y_{\mathcal{E}}) + n\, I(X_V ; X_S, X_u, Y_{\mathcal{E}})$$

$$\geq (1-n)\, I(X_V ; X_S, Y_{\mathcal{E}}) + n\left(I(X_V ; X_S, Y_{\mathcal{E}}) + I(X_V, X_u \mid X_S, Y_{\mathcal{E}}) \right)$$

$$= I(X_V ; X_S, Y_{\mathcal{E}}) + n\, I(X_V, X_u \mid X_S, Y_{\mathcal{E}})$$

now we have conditioned Markov chain : $X_V \to X_u \to Y_e \mid X_S, Y_{\mathcal{E}}$

یعنی داریم $\longrightarrow$ X_u  $Y_e \mid X_S, Y_{\mathcal{E}}$

$$I(X_V ; Y_e \mid X_S, Y_{\mathcal{E}}) \leq n\, I(X_V ; X_u \mid X_S, Y_{\mathcal{E}})$$ حال با توجه به نامساوی تعمیم پردازش داده ها داریم :

پس داریم :

$$\mathbb{P}_{g,n}[V \rightsquigarrow S] \log 2 \geq I(X_V ; X_S, Y_{\mathcal{E}}) + n\, I(X_V, X_u \mid X_S, Y_{\mathcal{E}})$$

$$\geq I(X_V ; X_S, Y_{\mathcal{E}}) + I(X_V ; Y_e \mid X_S, Y_{\mathcal{E}})$$

$$= I(X_V ; X_S, Y_{\mathcal{E}}, Y_e) = I(X_V ; X_S, Y_{\mathcal{E}}) \quad \blacksquare$$

پس گام استراتژی اثبات شد هلینگر نتیجه می‌شود

۳ تمرین در اثبات نامساوی‌ها!

۱. حالا باید اثبات کنیم:

$$P\{\frac{p}{q} > M\} \leq \frac{D_f(P,Q)}{f'(M)}$$

داریم:

$$\circledast\quad X \geq 0 \rightarrow E[X] = \int_0^\infty x f(x) = \int_0^\infty \underbrace{\int_t^\infty f(x)\,dx}_{P\{X \geq t\}}\,dt = \int_0^\infty \mathbb{P}\{X \geq t\}\,dt$$

همچنین چون f محدب است ← $f'' \geq 0$ ← f' صعودی است.

از طرفی $f(1)=0$ ← $\forall x \geq 1:\ f'(x) \geq f'(1)=0$ ← $\forall x \geq 1: f'(x) \geq 0$ ← f صعودی است.

حال داریم:

$$P(\frac{P}{Q} > M) = E_P[\mathbb{1}(\frac{P}{Q} > M)] = E_Q[\frac{P}{Q}\mathbb{1}(\frac{P}{Q} > M)]$$

$$\circledast \xrightarrow{\text{طبق}} \int_0^\infty \mathbb{P}\{\frac{P}{Q}\mathbb{1}(\frac{P}{Q} > M) \geq t\} = \int_M^\infty \mathbb{P}\{\frac{P}{Q} \geq t\}\,dt \xrightarrow[t > M \geq 1]{f \text{ صعودی}} \int_M^\infty P(f(\frac{P}{Q}) \geq f(t))\,dt$$

$$\underset{f(t)=\alpha}{\Longrightarrow} \int_{f(M)}^\infty P(f(\frac{P}{Q}) \geq \alpha)\,\frac{d\alpha}{f'(t)} \xrightarrow{f' \text{ صعودی}} \leq \int_{f(M)}^\infty P(f(\frac{P}{Q}) \geq \alpha)\,\frac{d\alpha}{f'(M)}$$

f' صعودی

Subject :   Date :

$f(m) \geq f(1) = 0$ چون $\longrightarrow \displaystyle\int_{f(m)}^{\infty} \mathbb{P}[f(\tfrac{P}{Q}) \geq \alpha]\, \frac{d\alpha}{f'(m)} \leq \int_{0}^{\infty} \mathbb{P}[f(\tfrac{P}{Q}) \geq \alpha]\, \frac{d\alpha}{f'(m)}$

⊛ با توجه به $\longrightarrow \; = \dfrac{1}{f'(m)} \mathbb{E}_Q[f(\tfrac{P}{Q})] = \dfrac{D_f(P,Q)}{f'(m)} \quad \blacksquare$

۲.

$$P_Y = (1-\delta) P_{Y|\mathcal{E}} + \delta P_{Y|\bar{\mathcal{E}}}$$

convexity of KL $\rightarrow D_{KL}(P_Y \| Q_Y) \leq (1-\delta) D_{KL}(P_{Y|\mathcal{E}} \| Q_Y) + \delta D_{KL}(P_{Y|\bar{\mathcal{E}}} \| Q_Y)$

$$\leq D_{KL}(P_{Y|\mathcal{E}} \| Q_Y) + \delta D_{KL}(P_{Y|\bar{\mathcal{E}}} \| Q_Y)$$

$$\leq \log(1 + D_{\chi^2}(P_{Y|\mathcal{E}} \| Q_Y)) + \delta D_{KL}(P_{Y|\bar{\mathcal{E}}} \| Q_Y)$$

حالا یک ترم نامساوی را ثابت کردیم . برای ترم دوم داریم :

$X \perp\!\!\!\perp \mathcal{E} \;,\; X \perp\!\!\!\perp \bar{\mathcal{E}}$

$$\rightarrow D_{KL}(P_{XY|\bar{\mathcal{E}}} \| P_X Q_Y) = \underbrace{D_{KL}(P_X \| P_X)}_{0} + \mathbb{E}_{P_X}[D_{KL}(P_{Y|X,\bar{\mathcal{E}}} \| Q_Y)]$$

$$= D_{KL}(P_{Y|\bar{\mathcal{E}}} \| Q_Y) + \underbrace{\mathbb{E}_{P_{Y|\bar{\mathcal{E}}}}[D_{KL}(P_{X|Y,\bar{\mathcal{E}}} \| P_X)]}_{\geq 0}$$

$$\Rightarrow D_{KL}(P_{Y|\bar{\mathcal{E}}} \| Q_Y) \leq \mathbb{E}_{P_X}[D_{KL}(P_{Y|X,\bar{\mathcal{E}}} \| Q_Y)]$$

پس داریم :

Subject : Date :

$$D_{KL}(P_Y\|Q_Y) \leq \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \delta\, \mathbb{E}_{P_X}[D_{KL}(P_{Y|X,\bar{\mathcal{E}}}\|Q_Y)]$$

$$= \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \delta\, \mathbb{E}_{P_X}[\mathbb{E}_{P_{Y|X,\bar{\mathcal{E}}}}\{\log \frac{P_{Y|X,\bar{\mathcal{E}}}}{Q_Y}\}]$$

$$= \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \delta\, \mathbb{E}_{P_X}[\mathbb{E}_{P_{Y|X}}[\frac{1}{\delta}\mathbb{1}\{Y\in\bar{\mathcal{E}}\}\log(\frac{1}{\delta}\mathbb{1}(Y\in\bar{\mathcal{E}})\frac{P_{Y|X}}{Q_Y}]]$$

$$= \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \mathbb{E}_{P_X}[\mathbb{E}_{P_{Y|X}}[\mathbb{1}\{Y\in\bar{\mathcal{E}}\}\log(\frac{1}{\delta}) + \log(\frac{P_{Y|X}}{Q_Y})]]$$

$$= \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \log\frac{1}{\delta}\mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})] + \mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})\log(\frac{P_{Y|X}}{Q_Y})]$$

$$= \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})(\log(\frac{P_{Y|X}}{Q_Y}) - \mathbb{E}_{P_{XY}}[\log(\frac{P_{Y|X}}{Q_Y})]] + \mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})\mathbb{E}_{P_{XY}}[\log(\frac{P_{Y|X}}{Q_Y})]]$$

$$= \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \delta\log(\frac{1}{\delta}) + \underbrace{\mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})]}_{\delta}\mathbb{E}_{P_{XY}}[\log\frac{P_{Y|X}}{Q_Y}] + \mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})(\log\frac{P_{Y|X}}{Q_Y} - \mathbb{E}_{P_{XY}}[\log(\frac{P_{Y|X}}{Q_Y})]]$$

پس! $\Rightarrow D_{KL}(P_Y\|Q_Y) \leq \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \delta\log\frac{1}{\delta} + \overbrace{\mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})]}^{\delta}\mathbb{E}_{P_{XY}}[\log(\frac{P_{Y|X}}{Q_Y})]$

$$+ \mathbb{E}_{P_{XY}}\{\mathbb{1}(Y\in\bar{\mathcal{E}})(\log(\frac{P_{Y|X}}{Q_Y}) - \mathbb{E}_{P_{XY}}[\log(\frac{P_{Y|X}}{Q_Y})]]$$

حال برای تعابیر ترم $\sqrt{\delta\, var(\log(\frac{dP_{Y|X}}{dQ_Y})}$ از کوشی استفاده می‌کنیم:

$$D_{KL}(P_Y\|Q_Y) \leq \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \delta(\log(\frac{1}{\delta}) + \mathbb{E}_{P_X}[\log(\frac{P_{Y|X}}{Q_Y})]) + \sqrt{\mathbb{E}_{P_{XY}}[\mathbb{1}(Y\in\bar{\mathcal{E}})]}\sqrt{\mathbb{E}_{P_{XY}}[\log\frac{P_{Y|X}}{Q_Y} - \mathbb{E}_{P_{XY}}(\log\frac{P_{Y|X}}{Q_Y})]}$$

$$\leq \log(1+D_{\chi^2}(P_{Y|\mathcal{E}}\|Q_Y)) + \delta(\log(\frac{1}{\delta}) + \mathbb{E}_{P_X}[\log(\frac{P_{Y|X}}{Q_Y})]) + \sqrt{\delta}\sqrt{Var(\log(\frac{P_{Y|X}}{Q_Y}))}$$ ✓

MICRO

۵. اطلاعات متقابل و خطای تخمین

۱.

$$E_{XY}[(X-f(Y))^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-f(y))^2 P_{XY}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\underbrace{x-E[X|Y=y]}_{A} + \underbrace{E[X|Y=y]-f(y)}_{B})^2 P_{XY}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A^2 P_{XY}\,dx\,dy + \iint B^2 P_{XY}\,dx\,dy + 2\iint AB\,P_{XY}\,dx\,dy$$

داریم $\Rightarrow$ $\iint AB\,P_{XY}\,dx\,dy = \iint AB\,P_{X|Y}P_Y\,dx\,dy = \int B P_Y\,dy \underbrace{\int (x-E[X|Y=y])P_{X|Y}(x|y)\,dx}_{0}$

پس $\Rightarrow$ $E_{XY}[(X-f(Y))^2] = \iint (x-E[X|Y=y])^2 P_{XY}\,dx\,dy + \iint (E[X|Y=y]-f(y))^2 P_{XY}\,dx\,dy$

ترم اول که تابعی از $f(y)$ نیست. پس بهترین حالت این است که ترم دوم را که همیشه نامنفی است صفر کنیم پس $f(Y)=E[X|Y]$

$$\rightarrow f(Y) = E[X|Y]$$

۲.

Subject : Date :

$X \to Y_1 \to Y_2$.۳

$\to I(X;Y_1,Y_2) = I(X;Y_2) + I(X;Y_1|Y_2) = H(X) - H(X|Y_1,Y_2) = H(X) - H(X|Y_1) = I(X;Y_1)$

پس $\Rightarrow \quad I(X;Y_1) - I(X;Y_2) = I(X;Y_1|Y_2) \quad \blacksquare$

$\to \frac{1}{2} M_E(A) = \frac{d}{dA} I(A) = \lim_{\delta \to 0} \frac{I(A+\delta) - I(A)}{\delta} = \lim_{\delta \to 0} \frac{I(X;\sqrt{A+\delta}\,X+Z) - I(X;\sqrt{A}\,X+Z)}{\delta}$

$\Leftrightarrow I(X;\sqrt{A+\delta}\,X+Z) - I(X;\sqrt{A}\,X+Z) = \frac{\delta}{2} M_E(A) + O(\delta)$

we know that : $I(X;cY) = H(cY) - H(cY|X) = H(Y) - H(Y|X) = I(X;Y)$

پس $\Rightarrow \quad I(X; X + \frac{1}{\sqrt{A+\delta}} Z) - I(X; X + \frac{1}{\sqrt{A}} Z) = \frac{\delta}{2} M_E(A) + O(\delta)$

با توجه به کانال ها :

$Y_1 = X + \sigma_1 Z_1 = X + \frac{1}{\sqrt{A+\delta}} Z_1 \quad , \quad Z_1 \perp\!\!\!\perp X$

$Y_2 = X + \sigma_1 Z_1 + \sigma_2 Z_2 \xrightarrow[\text{با توزیع}]{} = X + \sigma_t Z_t = X + \frac{1}{\sqrt{A}} Z_t \quad , \quad Z_t \perp\!\!\!\perp X$

$\Rightarrow I(X; X + \frac{1}{\sqrt{A+\delta}} Z_1) - I(X; X + \frac{1}{\sqrt{A}} Z_t) = I(X;Y_1) - I(X;Y_2) = \frac{\delta}{2} M_E(A) + O(\delta)$

Subject : Date :

$$Y_1 = X + \frac{1}{\sqrt{A+\delta}} Z_1 = X + \sigma_1 Z_1$$ 4

$$Y_2 = X + \sigma_1 Z_1 + \sigma_2 Z_2 \longrightarrow SNR = \frac{1}{\sigma_1^2 + \sigma_2^2} = A$$

$$\longrightarrow \sigma_1^2 + \sigma_2^2 = \frac{1}{A}$$

$$\longrightarrow \frac{1}{A+\delta} + \sigma_2^2 = \frac{1}{A} \longrightarrow \sigma_2^2 = \frac{\delta}{A(A+\delta)}$$

$$\longrightarrow Y_2 = X + \frac{1}{\sqrt{A+\delta}} Z_1 + \frac{\sqrt{\delta}}{\sqrt{A(A+\delta)}} Z_2$$

داریم $\Rightarrow$ $$(A+\delta) Y_1 - A Y_2 - \delta X = (A+\delta)\left(X + \frac{1}{\sqrt{A+\delta}} Z_1\right) - A\left(X + \frac{1}{\sqrt{\delta+A}} Z_1 + \frac{\sqrt{\delta}}{\sqrt{A(A+\delta)}} Z_2\right) - \delta X$$

$$= \underset{\sim}{AX} + \underset{\sim}{\delta X} + \sqrt{A+\delta}\, Z_1 - \underset{\sim}{AX} - \frac{A}{\sqrt{\delta+A}} Z_1 - \frac{A\sqrt{\delta}}{\sqrt{A(A+\delta)}} Z_2 - \underset{\sim}{\delta X}$$

$$= \frac{\delta}{\sqrt{A+\delta}} Z_1 - \frac{\sqrt{A\delta}}{\sqrt{A+\delta}} Z_2 = \sqrt{\delta}\left(\frac{\sqrt{\delta}}{\sqrt{A+\delta}} Z_1 - \frac{\sqrt{A}}{\sqrt{A+\delta}} Z_2\right)$$

$$= \sqrt{\delta}\; Z$$

$$Z_1, Z_2 \sim N(0,1) \longrightarrow E[Z] = 0 \;,\; Var(Z) = \frac{\delta}{A+\delta} + \frac{A}{A+\delta} = 1 \Rightarrow Z \sim N(0,1)$$

$$Z_1, Z_2 \perp\!\!\!\perp X \longrightarrow Z \perp\!\!\!\perp X \longrightarrow (\delta+A) Y_1 - A Y_2 - \delta X = \sqrt{\delta}\, Z \;,\; Z \sim N(0,1) \;,\; Z \perp\!\!\!\perp X$$

Subject :

Date :

$I(X;Y_1) - I(X;Y_2) = I(X;Y_1|Y_2) = \mathbb{E}_{Y_1 X}[D(P_{Y_1|Y_2,X} \| P_X)] - \mathbb{E}_{Y_2}[D(P_{Y_1|Y_2} \| P_X)]$ .a

ميلا $\rightarrow (A+\delta)Y_1 = AY_2 + \delta A + \sqrt{\delta}Z$

$\rightarrow I(X;Y_1|Y_2) = H(X|Y_2) - H(X|Y_1,Y_2)$

$$W = \frac{\sqrt{\delta}Z + \delta\mathbb{E}[X|Y_2] + AY_2}{A+\delta} \quad , \quad \sigma_2^2 = \frac{\delta}{A(A+\delta)}$$

$$I(X;Y_1|Y_2) = \mathbb{E}_{Y_2}[I(X;Y_1|Y_2)] = \underbrace{\mathbb{E}_{Y_1,X}[D(P_{\frac{AY_2+\sqrt{\delta}Z+\delta A}{A+\delta}|Y_2,X} \| P_{\frac{AY_2+\sqrt{\delta}Z+\delta\mathbb{E}[X|Y_2]}{A+\delta}})]}$$

$$- \mathbb{E}_{Y_2}[D(P_{\frac{AY_2+\sqrt{\delta}Z+\delta\mathbb{E}[X|Y_2]}{A+\delta}|Y_2} \| P_{\frac{AY_2+\sqrt{\delta}Z+\delta\mathbb{E}[X|Y_2]}{A+\delta}})]$$

$$\rightarrow \mathbb{E}_{XY_2}[N(\frac{AY_2+\delta X}{A+\delta}, \frac{\delta}{(A+\delta)^2}) \| N(\frac{AY_2+\delta\mathbb{E}[X|Y_2]}{A+\delta}, \frac{\delta}{(A+\delta)^2})]$$

$$= \mathbb{E}_{XY_2}[\frac{1}{2}(\frac{(A+\delta)^2}{\delta})(\frac{\delta}{A+\delta}(X-\mathbb{E}[X|Y_2]))^2 = \frac{1}{2}\delta\,\mathbb{E}_{XY_2}[(X-\mathbb{E}[X|Y_2])^2]$$

$$= \frac{1}{2}\delta\, M_E(A)$$

$$\rightsquigarrow \underbrace{\frac{1}{A+\delta}(AY_2+\sqrt{\delta}Z+\delta\mathbb{E}[X|Y_2])}_{a_1} = \underbrace{\frac{1}{A+\delta}(AY_2+\sqrt{\delta}Z+\delta X)}_{a_2} - \frac{\delta}{A+\delta}(X-\mathbb{E}[X|Y_2])$$

$\rightarrow a_1|Y_1 = a_2|Y_2 + \delta^2 W$

$\rightarrow D_{KL}(P_{a_1|Y_1} \| P_{a_2|Y_2}) = O(\delta) \rightarrow \mathbb{E}_{Y_2}[D_{KL}(P_{a_1|Y_1} \| P_{a_2|Y_2})] = O(\delta)$

MICRO

پس $\rightarrow I(X;Y_1|Y_2) = \frac{1}{2}\delta M_E(A) + o(\delta) \xrightarrow{\text{با مشتق}} \frac{d}{dA} I(A) = \frac{1}{2} M_E(A)$ ☒

۶ مسئله تشخیص در SBM

۱. $P_\sigma = G(\sigma, P, Q)$

$$A \in \mathbb{R}^{n\times n} \rightarrow P_\sigma(A) = \prod_{\substack{i<j\\ \sigma_i\sigma_j=1}} p(A_{ij}) \prod_{i<j} q(A_{ij}) = \prod_{i<j}\left(P(A_{ij})\frac{1+\sigma_i\sigma_j}{2} + q(A_{ij})\frac{1-\sigma_{ij}}{2}\right)$$

$$= \prod_{i<j} \frac{1}{2}\left(P(A_{ij}) + q(A_{ij}) + \sigma_i\sigma_j\left(P(A_{ij}) - q(A_{ij})\right)\right)$$

$$P_{\hat\sigma} \xrightarrow{\text{به طور مشابه}} P_{\hat\sigma} = \prod_{i<j}\left(P(A_{ij}) + q(A_{ij}) + \hat\sigma_i\hat\sigma_j\left(P(A_{ij}) - q(A_{ij})\right)\right)$$

$$\rightarrow \gamma_0 = \prod_{i<j}\frac{1}{2}\left(P(A_{ij}) + q(A_{ij})\right)$$

$$\rightarrow \mathbb{E}_{R_0}\left[\frac{P_\sigma P_{\hat\sigma}}{\gamma_0^2}\right] = \int \frac{P_\sigma P_{\hat\sigma}}{\gamma_0}\, dA$$

داریم $\rightarrow$ $$\frac{P_\sigma P_{\hat\sigma}}{\gamma_0} = \prod_{i<j}\frac{1}{2}\,\frac{\left(P(A_{ij}) + q(A_{ij}) + \sigma_i\sigma_j\left(P(A_{ij}) - q(A_{ij})\right)\right)\left(P(A_{ij}) + q(A_{ij}) + \hat\sigma_i\hat\sigma_j\left(P(A_{ij}) - q(A_{ij})\right)\right)}{P(A_{ij}) + q(A_{ij})}$$

$$\Rightarrow \frac{P_\sigma P_{\hat\sigma}}{\gamma_0} = \prod_{i<j}\frac{1}{2}\underbrace{\left(P(A_{ij}) + q(A_{ij}) + (\sigma_i\sigma_j + \hat\sigma_i\hat\sigma_j)\left(P(A_{ij}) - q(A_{ij})\right) + \sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\frac{\left(P(A_{ij}) - q(A_{ij})\right)^2}{P(A_{ij}) + q(A_{ij})}\right)}$$

$$\mathbb{E}_{R_0}\left[\frac{P_\sigma P_{\hat\sigma}}{\gamma_0^2}\right] = \int \frac{P_\sigma P_{\hat\sigma}}{\gamma_0}\, dA = \int \prod \quad dA = \prod_{i<j}\Big[\overbrace{\int \frac{P(A_{ij}) + q(A_{ij})}{2}\, dA_{ij}}^{1} + \overbrace{\int \frac{\sigma_i\sigma_j + \hat\sigma_i\hat\sigma_j}{2}\left(P(A_{ij}) - q(A_{ij})\right) dA_{ij}}^{0}$$

$$+ \sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\Big(\underbrace{\frac{1}{2}\int \frac{\left(P(A_{ij}) - q(A_{ij})\right)^2}{P(A_{ij}) + q(A_{ij})}\, dA_{ij}}_{P}\Big)$$

Subject : Date :

$$\Rightarrow \mathbb{E}_{R_0}\left[\frac{P_\sigma P_{\hat\sigma}}{P_0^2}\right] = \prod_{i<j}\left[1+\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j \rho\right] \leq \prod_{i<j} \exp(\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\rho)$$

$$e^x = 1 + x + \frac{x^2}{2} + \cdots$$

$$= \exp\left(\rho \sum_{i<j} \sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\right) = \exp\left(\frac{\rho}{2}\langle\sigma,\hat\sigma\rangle^2\right)$$

$$\Rightarrow W(\sigma,\hat\sigma) \leq \exp\left(\frac{\rho}{2}\langle\sigma,\hat\sigma\rangle\right)$$

$P \sim Ber\left(\frac{a}{n}\right)$, $Q \sim Ber\left(\frac{b}{n}\right)$ ۲.

$$\rightarrow \rho = \sum_{x=0}^{1} \frac{(P(x)-q(x))^2}{2(P(x)+q(x))} = \frac{\left(\frac{a-b}{n}\right)^2}{2\left(\frac{a+b}{n}\right)} + \frac{\left(2-\left(\frac{a+b}{n}\right)\right)^2}{2\left(2-\frac{a+b}{n}\right)} = \frac{(a-b)^2}{2n(2n-a-b)} + \frac{1}{n}\frac{(a-b)^2}{2(a+b)}$$

$$\Rightarrow \frac{O(1)}{n} + \frac{\Gamma}{n}$$

$$\Rightarrow \rho = \frac{\Gamma + O(1)}{n} \qquad \blacksquare$$

$P(\sigma_i\hat\sigma_i = 1) = P(\sigma_i\hat\sigma_i = -1) = \frac{1}{2} \Rightarrow E[\sigma_i\hat\sigma_i] = 0$, $Var(\sigma_i\hat\sigma_i) = 1$ ۳.

$$CLT \longrightarrow Z = \lim_{n\to\infty} \frac{1}{\sqrt n}\langle\sigma,\hat\sigma\rangle \sim N(0,1)$$

$$D_{\chi^2}(P_\sigma, P_{\hat\sigma}) \leq \mathbb{E}_{\sigma,\hat\sigma}\left[\exp\left(\frac{\rho}{2}\langle\sigma,\hat\sigma\rangle^2\right)\right] - 1 = \mathbb{E}_{\sigma,\hat\sigma}\left[\exp\left(\frac{n\rho}{2}\left(\frac{\langle\sigma,\hat\sigma\rangle}{\sqrt n}\right)^2\right)\right] - 1$$ داریم:

$$= \mathbb{E}_Z\left[\exp\left(\frac{n\rho}{2}Z^2\right)\right] - 1 = \mathbb{E}_Z\left[\exp\left(\frac{\Gamma + O(1)}{2}Z^2\right)\right] - 1$$

MICRO

Subject : Date :

$$\to n \to \infty \to \frac{1}{\sqrt{2\pi}} \int \exp\left(-\frac{z^2}{2}\right) \exp\left(\frac{\tau + o(1)}{2} z^2\right) dz - 1$$

$$= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{2\pi}{1-\tau-o(1)}} - 1 = \frac{1}{\sqrt{1-\tau-o(1)}} - 1 < \infty$$

پس $\to$ if $n \to \infty \to D_{\chi^2}(P_\sigma, P_{\hat{\sigma}}) < \dfrac{1-\sqrt{1-\tau-o(1)}}{\sqrt{1-\tau-o(1)}} < \infty$

پس $P_\sigma \Delta P_{\hat{\sigma}}$ ← پس به هم چسبیده هستند، و تشخیص بین آنها نمی‌تواند به صفر برسد!

۷. تزویج بهینه

۱. $P(E) = \sum_{x \in E} \sum_{\hat{x}} P_{x\hat{x}}$, $Q(E) = \sum_{\hat{x} \in E} \sum_{x} P_{x\hat{x}}$

$$\Rightarrow \{P(E) - Q(E)\} = \left\{ \sum_{x \in E} \sum_{\hat{x} \notin E} P_{x\hat{x}} - \sum_{\hat{x} \in E} \sum_{x \notin E} P_{x\hat{x}} \right\}$$

$$\leq \sum_{x \in E} \sum_{\hat{x} \notin E} P_{x\hat{x}} + \sum_{\hat{x} \in E} \sum_{x \notin E} P_{x\hat{x}} \leq \sum_{x \neq \hat{x}} P_{x\hat{x}}$$

$$\Rightarrow \sum_{x \neq \hat{x}} P_{x\hat{x}} = \sum_{x \in E} \sum_{\hat{x} \notin E} P_{x\hat{x}} + \sum_{\hat{x} \in E} \sum_{x \notin E} P_{x\hat{x}} + \sum_{x \in E} \sum_{\substack{\hat{x} \in E \\ \hat{x} \neq x}} P_{x\hat{x}} + \sum_{\hat{x} \notin E} \sum_{\substack{x \notin E \\ x \neq \hat{x}}} P_{x\hat{x}}$$

پس داریم $\to \sup\{P(E) - Q(E)\} \leq \inf(P_{x\hat{x}}(x \neq \hat{x}))$

برای اثبات اینکه حتما دوطرف برابر شدند کافی است حالت تساوی ای نشان دهیم و تعریف کنیم: $r(x) = \min(P(x), Q(x))$

$$P(X=\hat{X}) = \sum_x r(x) \quad , \quad P(X \neq \hat{X}) = 1 - \sum_x r(x)$$

داریم در کلاس اثبات کردیم ⇒ $\sum_x r(x) = \sum_x \min(P(x), q(x)) = 1 - d_{TV}(P,Q)$

$$\Rightarrow P(\hat{X} \neq X) = 1 - \sum r(x) = d_{TV}(P,Q)$$

دقت شود که تعریف $r(x)$ به گونه‌ای است که با احتمال $r(x)$ داریم $X = \hat{X} = x$.

۲. $\frac{1}{n} \sum_{i=1}^{n} P_{X^n \hat{X}^n}(X_i \neq \hat{X}_i) \leq \frac{1}{n} \sum d_{TV}(P_{X_i}, P_{\hat{X}_i})$

where ⇒ $\sum d_{TV}(P_{X_i}, P_{\hat{X}_i}) \underset{\text{Pinsker}}{\leq} \sum \sqrt{\frac{1}{2} D_{KL}(P_{X_i} \| P_{\hat{X}_i})} = \frac{1}{\sqrt{2}} \sum \sqrt{D_{KL}(P_{\hat{X}_i} \| P_{X_i})}$

we have → $\left(\sum_i \sqrt{D_{KL}(P_{X_i} \| P_{\hat{X}_i})}\right)^2 \underset{\text{کوشی شوارز}}{\leq} n \sum_{i=1}^{n} D_{KL}(P_{X_i} \| P_{\hat{X}_i})$

$$\Rightarrow \sum \sqrt{D_{KL}(P_{\hat{X}_i} \| P_{X_i})} \leq \sqrt{n \sum D_{KL}(P_{\hat{X}_i} \| P_{X_i})}$$

از طرفی با توجه به خاصیت زنجیره‌ای KL داریم: $\sum D_{KL}(P_{X_i} \| P_{\hat{X}_i}) \leq D_{KL}(P_{\hat{X}^n} \| P_{X^n})$

پس داریم ⇒ $\frac{1}{n} \sum P_{X^n \hat{X}^n}(X_i \neq \hat{X}_i) \leq \frac{1}{n} \sum d_{TV}(P_{X_i}, P_{\hat{X}_i}) \leq \frac{1}{n} \times \frac{1}{\sqrt{2}} \times \sqrt{n D_{KL}(P_{\hat{X}^n} \| P_{X^n})}$

$$\Rightarrow \frac{1}{n} \sum P_{X^n \hat{X}^n}(X_i \neq \hat{X}_i) \leq \sqrt{\frac{1}{2n} D_{KL}(P_{\hat{X}^n} \| P_{X^n})} \quad \blacksquare$$

۸. انحراف Marton

۱. ابتدا حالتی که در آن تساوی رخ می‌دهد را بررسی می‌کنیم، تعریف می‌کنیم:

$$f(y) = \overset{\text{پیوسته}}{\frac{dP}{dQ}(y)} = \overset{\text{گسسته}}{\frac{P(X=y)}{Q(X=y)}}$$

حال P_{XY} را با توجه به f به این صورت تعریف می‌کنیم: $P_{X|Y}(X=y \mid Y=y) = \min(1, f(y))$

پس داریم $\Rightarrow$ $P_{XY}(X \neq Y \mid Y=y) = 1 - \min(1, f(y)) = (1 - f(y))_+$

$$\Rightarrow \mathbb{E}[(P_{XY}(X \neq Y \mid Y)^2)] = \int (1 - f(y))_+^2 \, dQ(y) = \int \left(1 - \frac{dP}{dQ}\right)_+^2 dQ = D_m(P \| Q)$$

پس حالتی که در آن تساوی رخ می‌دهد است را آوردیم. حال مانند سوال ۷ ـ بخش اول، می‌خواهیم اثبات کنیم که $\mathbb{E}[P_{XY}(X \neq Y \mid Y)^2]$ همیشه بزرگتر مساوی $D_m(P\|Q)$ است. داریم:

$$Q(y) = P_{XY}(X=y, Y=y) + \sum_{y' \neq y} P_{XY}(X=y, Y=y') \geq P_{XY}(X=y, Y=y) = \underbrace{P_{XY}(X=y \mid Y=y)}_{a(y)} Q(y)$$

از طرفی تعریف می‌کنیم $\Rightarrow g(y) = P_{XY}(X \neq Y \mid Y=y) = 1 - P_{XY}(X=y \mid Y=y) = 1 - a(y)$

از طرفی اگر y در فضای حالت ما باشد، داریم: $a(y) \leq f(y)$ و چون همیشه $a(y) \leq 1$ پس: $a(y) \leq \min(f(y), 1)$

حال داریم: $g(y) = 1 - a(y) \geq 1 - \min(f(y), 1) = (1 - f(y))_+$

$$\Rightarrow (g(y))^2 \geq (1 - f(y))_+^2$$

دقت کنید که تعریف کرده بودیم، $g(y) = P_{XY}(X \neq Y \mid Y=y)$ پس با انتگرال گیری (یا سیگما برای حالت گسسته) داریم:

Subject : Date :

پس داریم:

$$E[(P_{XY}(X\neq Y|Y))^2] = \int g(y)^2\, dQ(y) \geq \int (1-f(y))_+^2\, dQ(y) = D_m(P\|Q)$$

پس $D_m(P\|Q)$ همیشه کوچک‌تر مساوی $E[(P_{XY}(X\neq Y|Y))^2]$ است، و اثبات ما کامل می‌شود. ☒

۲. با توجه به بخش قبل اثبات کردیم:

$$E[P_{XY}(X\neq Y|Y)^2] \geq D_m(P\|Q)$$

به صورت مشابه می‌توان نشان داد:

$$E[P_{XY}(X\neq Y|X)^2] \geq D_m(Q\|P)$$

پس به طور کلی داریم:

$$E[P_{XY}(X\neq Y|Y)^2] + E[P_{XY}(X\neq Y|X)^2] \geq D_m(P\|Q) + D_m(Q\|P) = D_{SM}(P\|Q)$$

حال برای اثبات حکم کافیست نشان دهیم برای P_{XY} ای تساوی رخ می‌دهد. کافی است P_{XY} را به صورت زیر تعریف کنیم:

$$P_{XY}(X=y, Y=y) = \min(P(y), Q(y)) \quad , \quad P(X=y|Y=y) = \min\left(1, \frac{dP}{dQ}(y)\right)$$

$$P(Y=x|X=x) = \min\left(1, \frac{dQ}{dP}(x)\right)$$

در نتیجه: $P_{XY}(X\neq Y|Y=y) = \left(1-\frac{dP}{dQ}(y)\right)_+ \quad , \quad P_{XY}(X\neq Y|X) = \left(1-\frac{dQ}{dP}(x)\right)_+$

پس خواهیم داشت:

$$\Rightarrow \mathbb{E}\left[P_{XY}(X \neq Y \mid Y)^2 \right] = \int \left(1 - \frac{dP}{dQ}(y)\right)_+^2 dQ(y) = D_m(P \| Q)$$

$$\Rightarrow \mathbb{E}\left[P_{XY}(X \neq Y \mid X)^2 \right] = \int \left(1 - \frac{dQ}{dP}(x)\right)_+^2 dP(x) = D_m(Q \| P)$$

$$\xrightarrow{\text{پس}} \mathbb{E}\left[P_{XY}(X \neq Y \mid Y)^2 \right] + \mathbb{E}\left[P_{XY}(X \neq Y \mid X)^2 \right] = D_m(P \| Q) + D_m(Q \| P) = D_{SM}(P \| Q)$$

پس برای P_{XY} ای تساوی رخ می‌دهد را هم نشان دادیم و حکم مسئله ثابت می‌شود. ■