Date: ....

۱. انحراف بزرگ برای log likelihood

۱. $P[\sum_{i=1}^{n}(w_i - z_i) \geqslant nt] \leqslant \exp(-n \sup_{\lambda \geqslant 0} \{\lambda t - \log E[e^{-\lambda z_i}] - \log E[e^{\lambda w_i}]\})$

داریم: $P[\sum_{i=1}^{n}(w_i - z_i) \geqslant nt] = P[e^{\lambda \sum_{i=1}^{n}(w_i - z_i)} \geqslant e^{\lambda nt}] \underset{\text{markov}}{\leqslant} \frac{E[e^{\lambda \sum (w_i - z_i)}]}{e^{n\lambda t}}$

$\Rightarrow \quad \bigcirc \leqslant \inf_{\lambda \geqslant 0} \frac{\exp(\log(E[e^{\lambda \sum (w_i - z_i)}]))}{e^{n\lambda t}}$

از طرفی داریم: $E[e^{\lambda \sum (w_i - z_i)}] = E[e^{\lambda w_1} e^{\lambda w_2} \cdots e^{-\lambda z_n}] = E[e^{\lambda w_i}]^n E[e^{-\lambda z_i}]^n$

پس:

$\Rightarrow \quad \bigcirc \leqslant \inf_{\lambda \geqslant 0} \exp(-n\lambda t + \log E[e^{\lambda w_i}]^n E[e^{-\lambda z_i}]^n)$

$\leqslant \inf_{\lambda \geqslant 0} \exp(-n\lambda t + n \log E_Q[e^{\lambda w_i}] + n \log E[e^{-\lambda z_i}])$

$\leqslant \exp(-n \sup_{\lambda \geqslant 0} (\lambda t - \log E_Q[e^{\lambda w_i}] - \log E[e^{-\lambda z_i}]))$

$= \exp(-n \sup_{\lambda \geqslant 0} (\lambda t - \psi_Q(\lambda) - \psi_P(-\lambda))) = \exp(-n \cdot F(t)) \blacksquare$

Date: ....

۲. هدف: اثبات رابطه روبرو:

$$F(0) = -\log E_P[e^{-\frac{z_i}{2}}] - \log E_Q[e^{\frac{w_i}{2}}]$$

$$F(0) = \sup_{\lambda \geq 0} \{ -\log E_P[e^{-\lambda z_i}] - \log[e^{\lambda w_i}] \} = \sup_{\lambda \geq 0} \{ -\log(E_P[e^{-\lambda \log(\frac{P(x_i)}{q(x_i)})}]) - \log(E_Q[e^{+\lambda \log(\frac{P(y_i)}{q(y_i)})}]) \}$$

$$\Rightarrow F(0) = \sup_{\lambda \geq 0} \{ -\log \int_x p^{1-\lambda} q^{\lambda} dx - \log \int_y p^{+\lambda} q^{1-\lambda} dy \}$$

$$= \sup_{\lambda \geq 0} \{ -\log (\int_x p^{1-\lambda} q^{\lambda} dx \times \int_y p^{\lambda} q^{1-\lambda} dy) \}$$

با توجه به کوشی شوارتز داریم:

$$\underbrace{\int_x p^{1-\lambda} q^{\lambda} dx \times \int_y p^{\lambda} q^{1-\lambda} dy}_{A} \geq \underbrace{\left(\int \sqrt{p.q} \right)^2}_{B(P,Q)^2}$$

get log

$$\log(A) \geq 2 \log B(P,Q) \rightarrow -\log(A) \leq -2 \log(B(P,Q))$$

با توجه به رابطه ۲

$$F(0) \leq -2 \log(B(P,Q))$$

از این حالت تساوی نامساوی کوشی شوارتز در $\lambda = \frac{1}{2}$ رخ می‌دهد و در آن حالت:

$$\lambda = \frac{1}{2} \rightarrow F(0) = -\psi_P(-\tfrac{1}{2}) - \psi_Q(\tfrac{1}{2}) = -2 \log B(P,Q) = \alpha$$

۳. از قسمت قبل، با توجه به $\sup_{\lambda\geq 0}\{-\psi_P(-\lambda)-\psi_Q(\lambda)\}$، داریم (با توجه به تساوی در $\lambda=\frac{1}{2}$)

$$\sup_{\lambda\geq 0}\{\lambda t-\psi_P(-\lambda)-\psi_Q(\lambda)\}\geq \frac{1}{2}t+F(0)\rightarrow F(t)\geq \frac{t}{2}+F(0)$$

حال طبق:

$$F(t)\geq \frac{t}{2}+F(0)\rightarrow -nF(t)\leq -n\left(\frac{t}{2}+F(0)\right)$$

$$\rightarrow e^{-nF(t)}\leq e^{-n\left(\frac{t}{2}+F(0)\right)}$$

حال با توجه به بخش ۱، داریم:

$$P\left[\sum_{i=1}^{n}(w_i-z_i)\geq nt\right]\leq \exp(-n.F(t))\leq \exp\left(-n\left(\frac{t}{2}+F(0)\right)\right) \quad \blacksquare$$

۲. زوج خطاهای قابل دسترس

۱. به طور شهودی، می‌توانیم توسط یک سکه، تصمیم گیری بین هر دو نقطه، روی خط آن‌ها حرکت کنیم؛ فرض کنید داریم:

$$\text{test1}\rightarrow P^{(1)}_{Z|X}\rightarrow\left(\pi^{(1)}_{1|0},\pi^{(1)}_{0|1}\right)$$

$$\text{test 2}\rightarrow P^{(2)}_{Z|X}\rightarrow\left(\pi^{(2)}_{1|0},\pi^{(2)}_{0|1}\right)$$

حال اگر در هر تصمیم گیری با $Ber(\lambda)$ اقدام کنیم، داریم:

$$\text{test 3}\rightarrow Ber(\lambda)\begin{cases}1\rightarrow \text{test1}\\ 0\rightarrow \text{test2}\end{cases}\rightarrow \begin{aligned}\pi^{(3)}_{1|0}&=\lambda\pi^{(1)}_{1|0}+\bar{\lambda}\pi^{(2)}_{1|0}\leq e^{-nE_0}\\ \pi^{(3)}_{0|1}&=\underbrace{\lambda\pi^{(1)}_{0|1}+\bar{\lambda}\pi^{(2)}_{0|1}}_{\text{convex}}\leq e^{-nE_1}\end{aligned}$$

Date:

۲

تست اول $\rightarrow \begin{cases} H_0 \rightarrow \log\left(\frac{P^{\otimes n}_{(x)}}{Q^{\otimes n}_{(x)}}\right) \geq T \\ H_1 \rightarrow \log\left(\frac{P^{\otimes n}_{(x)}}{Q^{\otimes n}_{(x)}}\right) < T \end{cases}$ $\longrightarrow$ $H_0 : Q(Z=0) = Q\left(\log\left(\frac{P^{\otimes n}_{(x)}}{Q^{\otimes n}_{(x)}}\right) \geq T\right) = \pi_{0|1}$

$H_1 : P(Z=1) = P\left(\log\left(\frac{P^{\otimes n}_{(x)}}{Q^{\otimes n}_{(x)}}\right) < T\right) = \pi_{1|0}$

$$\Rightarrow \log\left(\frac{P^{\otimes n}_{(x)}}{Q^{\otimes n}_{(x)}}\right) = \sum_{i=1}^{n} \log\left(\frac{P(x_i)}{Q(x_i)}\right) \quad , \quad T = nt$$

$$\longrightarrow Q(Z=0) = Q\left(\sum \frac{P(x_i)}{Q(x_i)} \geq nt\right) \leq \exp\Big(-n\big(\lambda t - \underbrace{\log E_{x\sim Q}\left[\exp\left(\lambda \log\left(\tfrac{P(x)}{Q(x)}\right)\right)\right]}_{\psi_Q(\lambda)}\big)\Big)$$

$$= \exp\left(-n\left(\lambda t - \psi_Q(\lambda)\right)\right) \rightarrow (\forall \lambda \geq 0)$$

$$\leq \inf_{\lambda \geq 0} \exp\left(-n\left(\lambda t - \psi_Q(\lambda)\right)\right) =$$

$$= \exp\left(-n \sup_{\lambda \geq 0} \{\lambda t - \psi_Q(\lambda)\}\right)$$

$$\longrightarrow P(Z=1) = P\left(\sum \log\left(\frac{Q(x)}{P(x)}\right) \geq nt\right) \leq \exp\Big(-n\big(-\lambda t - \underbrace{\log E_{x\sim P}\left[\exp\left(\lambda \log \tfrac{Q(x)}{P(x)}\right)\right]}_{\psi_P(\lambda)}\big)\Big)$$

$$= \exp\left(-n\left(-\lambda t - \psi_P(\lambda)\right)\right)$$

$$\leq \inf_{\lambda \leq 0} \exp\left(-n\left(-\lambda t - \psi_P(\lambda)\right)\right) = \exp\left(-n \sup_{\lambda \leq 0} \{-\lambda t - \psi_P(\lambda)\}\right)$$

پس داریم: $\pi_{1|0} = P(Z=1) \le \exp(-n\psi_P^*(t))$ ، $\pi_{0|1} = Q(Z=0) \le \exp(-n\psi_Q^*(t))$

حال کران ها را تعیین می کنیم:

$$\psi_P(\lambda) = \log E_{x\sim P}\left[\exp\left(\lambda\log\left(\frac{Q(x)}{P(x)}\right)\right)\right] \ge E_{x\sim P}\left[\log\left(\exp\left(\lambda\log\left(\frac{Q(x)}{P(x)}\right)\right)\right)\right] = \lambda E_{x\sim P}\left[\log\left(\frac{Q(x)}{P(x)}\right)\right]$$

$$\Rightarrow \psi_P(\lambda) \ge -\lambda D(P\|Q)$$

$$\to -\lambda t - \psi_P(\lambda) \le -\lambda(t - D(P\|Q)) \to \psi_P^*(t) \le \sup_{\lambda\le 0}\{-\lambda(t-D(P\|Q))\} \to \begin{cases}0 & t\ge D(P\|Q)\\ \infty & o.w.\end{cases}$$

$$\psi_Q(\lambda) = \log E_{x\sim Q}\left[\exp\left(\lambda\log\left(\frac{P(x)}{Q(x)}\right)\right)\right] \ge E_{x\sim Q}\left[\log\exp\left(\lambda\log\frac{P(x)}{Q(x)}\right)\right] = \lambda E_{x\sim Q}\left[\log\left(\frac{P(x)}{Q(x)}\right)\right]$$

$$\Rightarrow \psi_Q(\lambda) \ge -\lambda D(Q\|P)$$

$$\to \lambda t - \psi_Q(\lambda) \le \lambda(t + D(P\|Q)) \to \psi_Q^*(t) \le \sup_{\lambda\ge 0}\{\lambda(t+D(P\|Q))\} \to \begin{cases}0 & t\le -D(Q\|P)\\ \infty & o.w.\end{cases}$$

پس اگر $-D(Q\|P) \le t \le D(P\|Q) \leftarrow \sup_{\lambda\ge 0}$, $\sup_{\lambda\le 0}$ بس $\sup_{\lambda\in\mathbb{R}}$

درنتیجه داریم: $\text{if } -D(Q\|P) \le t \le D(P\|Q) \longrightarrow$

$$\pi_{1|0} \le e^{-n\psi_Q^*(t)}$$

$$\pi_{0|1} \le e^{-n\psi_P^*(t)}$$

Date: ....

$$\forall \lambda: \quad \psi_Q(\lambda) = \log \mathop{E}_{x\sim Q}\left[\exp\left(\lambda \log \frac{P(x)}{Q(x)}\right)\right] = \log \mathop{E}_{x\sim Q}\left[\left(\frac{P(x)}{Q(x)}\right)^{\lambda}\right] = \log \sum_x P(x)^{\lambda}\, q(x)^{1-\lambda}$$ ۳

$$\psi_P(\lambda) = \log \mathop{E}_{x\sim P}\left[\exp\left(\lambda \log \frac{Q(x)}{P(x)}\right)\right] = \log \mathop{E}_{x\sim P}\left[\left(\frac{Q(x)}{P(x)}\right)^{\lambda}\right] = \log \sum_x P(x)^{1-\lambda}\, q(x)^{\lambda}$$

پس: $\psi_Q(\lambda) = \psi_P(1-\lambda)$

$$\psi_Q^*(t) = \sup_{\lambda\in\mathbb{R}} \{\lambda t - \psi_Q(\lambda)\} = \sup_{\lambda\in\mathbb{R}} \{\lambda t - \psi_P(1-\lambda)\} = \sup_{\bar\lambda\in\mathbb{R}} \{(1-\bar\lambda)t - \psi_P(\bar\lambda)\}$$

$$= t + \sup_{\bar\lambda\in\mathbb{R}} \{-\bar\lambda t - \psi_P(\bar\lambda)\} = \psi_P^*(t) + t$$

$$\Rightarrow \underbrace{\psi_P^*(t)}_{E_0(t)} + t = \underbrace{\psi_Q^*(t)}_{E_1(t)}$$

$$\text{if } -D(Q\|P) \le t \le D(P\|Q) \longrightarrow \begin{array}{l} \pi_{1|0}^{(n)} \le e^{-nE_0} \le e^{-n\psi_P^*(t)} \\ \pi_{0|1}^{(n)} \le e^{-nE_1} \le e^{-n\psi_Q^*(t)} \end{array}$$

$$\Rightarrow \psi_P^*(t) \le E_0 \quad , \quad \psi_Q^*(t) = \psi_P^*(t) + t \le E_1$$

پس اگر شرط $-D(Q\|P) \le t \le D(P\|Q)$ برقرار باشد، زوج $(E_0, E_1) = (\psi_P^*(t), \psi_P^*(t)+t)$ قابل حصول خواهد بود.

۴. با توجه به قضیه Sanov می‌دانیم در صورتی که $n \to \infty$، کران چرنف یک کران تنگ خواهد بود. پس زوج‌های قابل حصول (E_0, E_1) تماماً زیر

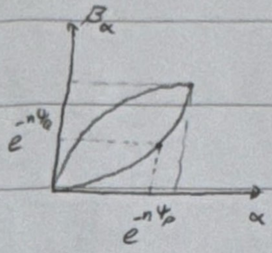
منحنی بین دو منحنی های تشکیل دهنده خواهند بود. پس اگر $n \to \infty$ داریم:

$$\forall t: \quad E_0(t) < E_0$$
$$E_1(t) < E_1$$

درنتیجه (E_0, E_1) همان منحنی مرزی زوج‌های قابل حصول هستند.

۵.

$$\min_{P_{Z|X^n}(z|x^n)} \left\{ \pi_0 \pi_{1|0} + \pi_1 \pi_{0|1} \right\}$$



با توجه به قسمت قبل داریم: $\pi_{0|1} = e^{-nE_1}$, $\pi_{1|0} = e^{-nE_0}$

$$\forall n \in [0, \infty) \to P_E = \pi_0 e^{-nE_0} + \pi_1 e^{-nE_1} \leq e^{-nE^*} \to n \to \infty \to \min\{E_0, E_1\} \leq E_p$$

تساوی زمانی رخ می‌دهد که روی مرز باشیم

$$P_E \leq e^{-nE_p} \to E_p = \max_{(E_0, E_1) \in R(P,Q)} \min\{E_0, E_1\}$$

مرز ناحیه $(E_0(t), E_1(t))$ است.

اگر طبق قسمت قبل باشیم: $-D(Q||P) \leq t \leq D(P||Q)$ $\leftarrow$ $E_p = \max\min = \max_t \min\{E_0(t), E_1(t)\}$

$\to E_0(t) = E_1(t)$ زمانی رخ می‌دهد $\max\min \to \varphi_P^*(t) + t = \varphi_Q^*(t) \to t = 0 \to E_p = \varphi_P^*(0)$

$$\to \varphi_P^*(0) = \sup_{\lambda \in \mathbb{R}} \{-\varphi_P(\lambda)\} = -\inf_{\lambda \in \mathbb{R}} \{\varphi_P(\lambda)\}$$

Date: ....

۳ اطلاعات جزئی

$$P_E^{(n)} = \pi_1 \alpha_n + \pi_2 \beta_n = \pi_1 P_1^{(n)}(X^{(n)} / A^{(n)}) + \pi_2 P_2^{(n)}(A^{(n)}) = \pi_1 + \pi_2 P_2^{(n)}(A^{(n)}) + \pi_1 P_1^{(n)}(A^{(n)})$$

$$\rightarrow \min_{A^{(n)} \in X^{(n)}} P_E^{(n)} = \min_{A^{(n)} \in X^{(n)}} \{\pi_1 + \pi_2 P_2^{(n)}(A^{(n)}) - \pi_1 P_1^{(n)}(A^{(n)})\}$$

$$= \pi_1 + \min_{A^{(n)} \in X^{(n)}} \{\pi_2 P_2^{(n)}(A^{(n)}) - \pi_1 P_1^{(n)}(A^{(n)})\}$$

$$= \pi_1 + \min_{A^{(n)} \in X^{(n)}} \{(\pi_2 P_2^{(n)}(A^{(n)}) - \pi_1 P_1^{(n)}(A^{(n)}))\, \mathbb{1}(x^{(n)} \in A^{(n)})\}$$

حالت کمینه سازی خطا $\longrightarrow A^{(n)} = \{x^{(n)} \mid \pi_1 P_1^{(n)}(A^{(n)}) \geq \pi_2 P_2^{(n)}(A^{(n)})\}$

$$\rightarrow \min P_E^{(n)} = \pi_1 - \pi_1 P_1^{(n)}\left[\frac{P_1(x^{(n)})}{P_2^{(n)}(x^{(n)})} > \frac{\pi_2}{\pi_1}\right] + \pi_2 P_2^{(n)}\left[\frac{P_1(x^{(n)})}{P_2^{(n)}(x^{(n)})} > \frac{\pi_2}{\pi_1}\right]$$

$$= \pi_1 P_1^{(n)}\left[\frac{P_1^{(n)}(x^{(n)})}{P_2^{(n)}(x^{(n)})} < \frac{\pi_2}{\pi_1}\right] + \pi_2 P_2\left[\frac{P_1^{(n)}(x^{(n)})}{P_2^{(n)}(x^{(n)})} > \frac{\pi_2}{\pi_1}\right]$$

$E_1 = \psi_{P_1}^{*}(0) = \psi_{P_2}^{*}(0) = \sup_{\lambda \in \mathbb{R}} \{-\psi_{P_2}(\lambda)\} = -\inf_{\lambda \in \mathbb{R}} \{\psi_{P_2}(\lambda)\} = \psi_{P_2}(\lambda^{*})$: از سوال ۲ داریم

$$\rightarrow \psi_{P_2}(\lambda) = \log \mathbb{E}_{P_2}\left[\exp\left(\lambda \log \frac{P_1(x)}{P_2(x)}\right)\right] \rightarrow \inf_{\lambda \in \mathbb{R}} \mathbb{E}_{P_2}\left[\exp\left(\lambda \log \left(\frac{P_1(x)}{P_2(x)}\right)\right)\right]$$

Date:

$$\to \frac{d}{d\lambda} = 0 \to E_{P_2}\left[\log\left(\frac{P_1(x)}{P_2(x)}\right)\exp\left(\lambda^* \log\left(\frac{P_1(x)}{P_2(x)}\right)\right)\right] = 0$$

$$\to \sum_{x \in \mathcal{X}} P_1^{\lambda^*}(x)\, P_2^{1-\lambda^*}(x) \log\left(\frac{P_1(x)}{P_2(x)}\right) = 0$$

$$\to \frac{\sum_{x \in \mathcal{X}} P_1^{\lambda^*}(x)\, P_2^{1-\lambda^*}(x) \log\left(\frac{P_1(x)}{P_2(x)}\right)}{\sum_x P_1^{\lambda^*}(x)\, P_2^{1-\lambda^*}(x)} = 0 \to P_{(\lambda^*)} = \frac{P_1^{\lambda^*}(x)\, P_2^{1-\lambda^*}(x)}{\sum_x P_1^{\lambda^*}(x)\, P_2^{1-\lambda^*}(x)}$$

مستقیم $\to P(\lambda) = \dfrac{P_1^{\lambda}(x)\, P_2^{1-\lambda}(x)}{\sum P_1^{\lambda}(x)\, P_2^{1-\lambda}(x)}$

$$\to f(y) = E_{P_\lambda}\left[\log\left(\frac{P_1(x)}{P_2(x)}\right)\right] \to f(\lambda^*) = 0\,?$$

$$\left.\begin{aligned} &\to f(0) = E_{P_2}\left[\log\left(\frac{P_2(x)}{P_1(x)}\right)\right] = D(P_2 \| P_1) \geq 0 \\ &\to f(1) = E_{P_1}\left[\log \frac{P_2(x)}{P_1(x)}\right] = -D(P_1 \| P_2) \leq 0 \end{aligned}\right\} \longrightarrow$$

پس $1 \geq \lambda^* \geq 0$ وجود دارد که $f(\lambda^*) = 0$

$$\to D^* = -\psi_{P_1}(\lambda^*) = \psi_{P_2}(\lambda^*) \quad \Big| \quad \exists \lambda^* \in [0,1] \to f(\lambda^*) = 0$$

$$\to D^* = -\psi_{P_2}^*(\lambda^*) = -\log_{P_2} E_{P_2}\left[\exp\left(\lambda^* \log\left(\frac{P_1(x)}{P_2(x)}\right)\right)\right] = -\log E_{x \sim P_2}\left[\left(\frac{P_1(x)}{P_2(x)}\right)^{\lambda^*}\right]$$

$$= \lambda^* E_{P_{\lambda^*}}\left[\log \frac{P_1(x)}{P_2(x)}\right] - \log\left(\sum_{y \in \mathcal{X}} P_1^{\lambda^*}(y)\, P_2^{1-\lambda^*}(y)\right)$$

SCBO

Date: ....

$$= E_{P_{\lambda^*}}\left[\log\left(\frac{P_1(x)}{P_2(x)}\right)^{\lambda^*}\right] - \log\left(\sum_{y\in X} P_1^{\lambda^*}(y)\, P_2^{1-\lambda^*}(y)\right)$$

$$= E_{P_{\lambda^*}}\left[\log \frac{P_1^{\lambda^*}(x)\, P_2^{1-\lambda^*}(x)}{\sum_{y\in X} P_1^{\lambda^*}(y)\, P_2^{1-\lambda^*}(y)}\right] = E_{P_{\lambda^*}}\left[\log \frac{P_{\lambda^*}(x)}{P_2(x)}\right] = D(P_{\lambda^*}\|P_2)$$

$$\rightarrow D(P_{\lambda^*}\|P_2) = E_{P_{\lambda^*}}\left[\log \frac{P_{\lambda^*}}{P_2}\right] = E_{P_{\lambda^*}}\left[\log\left(\frac{P_{\lambda^*}(x)}{P_1(x)}\right) - \underbrace{\log\left(\frac{P_1(x)}{P_2(x)}\right)}_{0}\right] = D(P_{\lambda^*}\|P_1)$$

$$\Rightarrow D^* = D(P_{\lambda^*}\|P_1) = D(P_{\lambda^*}\|P_2)$$

$$D(P_{\lambda^*}\|P_2) = E_{P_{\lambda^*}}\left[\log\left(\frac{P_{\lambda^*}(x)}{P_2(x)}\right)\right] = E_{P_{\lambda^*}}\left[\log\left(\frac{P_1^{\lambda}(x)\, P_2^{1-\lambda}(x)}{\sum P_1^{\lambda}(x)\, P_2^{1-\lambda}(x)}\right)\right]$$ ٢

$$= E_{P_{\lambda^*}}\left[\lambda \log\left(\frac{P_1(x)}{P_2(x)}\right) - \log\left(\sum P_2(x)\left(\frac{P_1(x)}{P_2(x)}\right)^{\lambda}\right)\right]$$

$$= \lambda E_{P_{\lambda^*}}\left[\log \frac{P_1}{P_1}\right] - \lambda^* \log E_{P_2}\left[\left(\frac{P_1(x)}{P_2(x)}\right)\right]$$

$$\rightarrow D^* = -\lambda^* \log E_{P_2}\left[P_1(x) P_2(x)^{-1}\right]$$

حالا $\rightarrow C_\lambda(P_1, P_2) = -\log E_{x\sim P_2}\left[\exp\left(\lambda \log \frac{P_1}{P_2}\right)\right] \rightarrow \frac{d}{d\lambda} = 0 \rightarrow$

Date:

$$\rightarrow \frac{dC_\lambda}{d\lambda} = \frac{E_{x\sim P_2}\{\log \frac{P_1}{P_2} (\frac{P_1(x)}{P_2(x)})^\lambda\}}{E_{x\sim P_2}[P_1^\lambda(x) P_2^{1-\lambda}(x)]} = \frac{\sum P_1^\lambda(x) P_2^{1-\lambda}(x) \log(\frac{P_1(x)}{P_2(x)})}{\sum_x P_1^\lambda(x) P_2^{1-\lambda}(x)} = -\sum P_\lambda(x) \log \frac{P_1}{P_2}$$

$$= f(\lambda) = 0$$

داشتیم $\rightarrow f(0) \geq 0 ,\ f(1) \leq 0 \rightarrow \exists\, \lambda^* \in [0,1]$

$$\rightarrow D^* = \max_{0\leq\lambda\leq 1} \left\{ -\log\left(E_{P_2}\left[\left(\frac{P_1(x)}{P_2(x)}\right)^\lambda\right]\right)\right\} = -\min_{0\leq\lambda\leq 1} \log\left(E_{P_2}\left\{\frac{P_1(x)}{P_2(x)}\right\}^\lambda\right) = C(P_1,P_2)$$

٤ f - انحراف و آزمون فرض

۱. محدب $\rightarrow$ $f'_\alpha(x) = \frac{\alpha x^{\alpha-1}}{\alpha-1} \rightarrow f''_\alpha(x) = \alpha x^{\alpha-2} > 0 \rightarrow f''_\alpha(x) \in R^+$

۲. $D_{f_\alpha}(P\|Q) = \sum_{x\in\mathcal{X}} Q(x) f_\alpha\left(\frac{P(x)}{Q(x)}\right) = \sum_{x\in\mathcal{X}} Q(x) \left(\frac{(\frac{P(x)}{Q(x)})^\alpha - 1}{\alpha-1}\right) = \sum_{x\in\mathcal{X}} \frac{P(x)^\alpha - Q(x)^\alpha}{(\alpha-1)\, Q(x)^{\alpha-1}}$

برای حالت توزیع های پیوسته به صورت انتگرال.

۳. داریم: $D_{f_\alpha}(P_{X,Y}\|Q_{X,Y}) = \int_X Q_X(x)\,dx \int_Y Q_{Y|X}(y|x) f_\alpha\left(\frac{P_X(P_{Y|X})}{Q_X(Q_{Y|X})}\right) dy \quad (*)$

حل با توجه به اینکه f_α محدب است، با توجه به Jensen داریم:

Date:

$$(*) \underset{\text{Jensen}}{\geq} \int_{\mathcal{X}} Q_X(x)\,dx\, f_\alpha\left(\int \cancel{Q_{Y|X}(y|x)} \times \frac{P_X P_{Y|X}}{Q_X \cancel{Q_{Y|X}}}\,dy\right)$$

$$= \int Q_X(x)\,dx\, f_\alpha\Big(\frac{P_X}{Q_X} \underbrace{\int_{\mathcal{Y}} P_{Y|X}\,dy}_{1}\Big) = \int_{\mathcal{X}} Q_X(x)\, f_\alpha\left(\frac{P(x)}{Q(x)}\right) dx$$

$$= D_{f_\alpha}(P_X \| Q_X)$$

۴. ابتدا کمی تابع انحراف را بررسی می‌کنیم، تابع مقابل را در نظر بگیرید: $g_\alpha(x) = \frac{1}{1-\alpha}\log\left(1+(\alpha-1)x\right)$

(دقت کنید که : $g_\alpha(D_{f_\alpha}(P\|Q)) = R_\alpha(P\|Q)$) داریم: $g_\alpha'(x) = \frac{-(\alpha-1)}{(1-\alpha)(1+(\alpha-1)x)} = \frac{1}{1+(\alpha-1)x}$

دقت کنید که اگر $g_\alpha'(x)$ در نقطه‌ای که انحراف f (رنیی) حساب شود مثبت باشد، می‌توان از کمک قبل استفاده کرد و نامساوی پردازش داده را برای آن اثبات کرد

در نقطه مورد نظر داریم:

$$g_\alpha'(D_{f_\alpha}(P\|Q)) = \frac{1}{1+(\alpha-1)D_{f_\alpha}(P\|Q)} \geq 0 \iff (\alpha-1)D_{f_\alpha}(P\|Q) \geq -1$$

$$\iff (\alpha-1)\left(\sum \frac{P^\alpha(x) - Q^\alpha(x)}{(\alpha-1)\,Q^{\alpha-1}(x)}\right) = \sum_x \frac{P^\alpha(x)}{Q^{\alpha-1}(x)} - Q(x) = -1 + \sum \frac{P^\alpha(x)}{Q^{\alpha-1}(x)} \geq -1$$

$$\iff \sum_{x\in\mathcal{X}} \frac{P^\alpha(x)}{Q^{\alpha-1}(x)} \geq 0 \quad \checkmark$$

پس در این نقطه مثبت است و با توجه به نامساوی پردازش داده $D_{f_\alpha}(P_{X,Y}\|Q_{X,Y}) \geq D_{f_\alpha}(P_X\|Q_X)$ برای انحراف رنیی نیز درست است:

Scbo

Date: ....

$$R_\alpha(P_{X,Y}\|Q_{X,Y}) \geq R_\alpha(P_X\|Q_X) \quad \blacksquare$$

۵. از Hopital استفاده می‌کنیم:

$$D_{f_\alpha}(P\|Q) = \sum_x Q(x)\left(\frac{\left(\frac{P(x)}{Q(x)}\right)^\alpha - 1}{\alpha - 1}\right) \xrightarrow{\text{Hopital}} \sum_x Q(x)\left(\frac{\frac{\partial}{\partial\alpha}\left(e^{\alpha\log\frac{P(x)}{Q(x)}} - 1\right)}{\frac{\partial}{\partial\alpha}(\alpha-1)}\right)$$

$$= \sum_x Q(x)\log\left(\frac{P(x)}{Q(x)}\right) \times e^{\alpha\log\left(\frac{P(x)}{Q(x)}\right)}$$

$$\alpha \to 1 \quad \Rightarrow \sum_x P(x)\log\left(\frac{P(x)}{Q(x)}\right) = D_{KL}(P\|Q) \quad \checkmark \; \blacksquare$$

برای انحراف رینی هم از Hopital استفاده می‌کنیم.

با توجه به سوال قبل

$$R_\alpha(P\|Q) \underset{\alpha\to 1}{\xrightarrow{\text{Hopital}}} = \frac{\frac{\partial}{\partial\alpha}\log\left(1+(\alpha-1)D_{f_\alpha}(P\|Q)\right)}{\frac{\partial}{\partial\alpha}(1-\alpha)} = \frac{\frac{\partial}{\partial\alpha}\log\left(\sum \frac{P(x)^\alpha}{Q(x)^{\alpha-1}}\right)}{\frac{\partial}{\partial\alpha}(1-\alpha)} =$$

$$= \frac{\sum_x P(x)^\alpha Q(x)^{1-\alpha}\log\left(\frac{P(x)}{Q(x)}\right)}{\sum \frac{P(x)^\alpha}{Q(x)^{\alpha-1}}} \overset{\alpha\to 1}{=} \frac{\sum_x P(x)\log\left(\frac{P(x)}{Q(x)}\right)}{\underbrace{\sum_x P(x)}_{1}}$$

$$= \sum P(x)\log\left(\frac{P(x)}{Q(x)}\right) = D_{KL}(P\|Q) \quad \checkmark \blacksquare$$

Date: ....

۶. برای $R_\alpha(P_{X^n} \| Q_{X^n})$ داریم:

$$R_\alpha(P_{X^n} \| Q_{X^n}) = \frac{1}{\alpha-1} \log\Big(\sum_{x_1 \dots x_n} \Big(\frac{P_{x_1} \dots P_{x_n}}{Q_{x_1} \dots Q_{x_n}}\Big)^\alpha (Q_{x_1} \dots Q_{x_n})\Big)$$

$$= \frac{1}{\alpha-1} \log\Big(\sum P_{x_1}^\alpha Q_{x_1}^{1-\alpha} \times \sum P_{x_2}^\alpha Q_{x_2}^{1-\alpha} \times \dots \times \sum_{x_n} P_{x_n}^\alpha Q_{x_n}^{1-\alpha}\Big)$$

$$= \sum_{i=1}^n \frac{1}{\alpha-1} \log\Big(\sum_{x_i} P_{x_i}^\alpha Q_{x_i}^{1-\alpha}\Big) = \sum_{i=1}^n R_\alpha(P \| Q) = n R_\alpha(P \| Q)$$

از طرفی با توجه به نامساوی پردازش داده داریم:

$$R_\alpha(P_{X^n} \| Q_{X^n}) = R_\alpha(P_{X^n,Z} \| Q_{X^n,Z}) \geq R_\alpha(P_Z \| Q_Z)$$

پس با توجه به در نکته بالا کافیست ثابت کنیم:

$$R_\alpha(P_Z \| Q_Z) = R_\alpha(Ber(P(z=1)) \| Ber(1-e^{-nE_n})) \geq R_\alpha(Ber(\varepsilon) \| Ber(1-e^{-nE_n}))$$

دقت کنید که $P(z=1) < \varepsilon$. داریم:

$$R_\alpha(Ber(P(z=1)) \| Ber(1-e^{-nE_n})) = \frac{1}{\alpha-1} \log\Big(P(z=1)^\alpha (1-e^{-nE_n})^{1-\alpha} + (1-P(z=1))^\alpha (e^{-nE_n})^{1-\alpha}\Big)$$

$$R_\alpha(Ber(\varepsilon) \| Ber(1-e^{-nE_n})) = \frac{1}{\alpha-1} \log\Big(\varepsilon^\alpha (1-e^{-nE_n})^{1-\alpha} + (1-\varepsilon)^\alpha (e^{-nE_n})^{1-\alpha}\Big)$$

اگر $1 < \alpha$ باید ثابت کنیم:

$$\varepsilon^\alpha (1-e^{-nE_n})^{1-\alpha} + (1-\varepsilon)^\alpha (e^{-nE_n})^{1-\alpha} \overset{?}{\geq} P(z=1)^\alpha (1-e^{-nE_n})^{1-\alpha} + (1-P(z=1))^\alpha (e^{-nE_n})^{1-\alpha}$$

Date:

داریم:

$$\left(\varepsilon^{\alpha} - P(z=1)^{\alpha}\right)\left(1-e^{-nE_n}\right)^{1-\alpha} \overset{?}{\geq} \left((1-P(z=1))^{\alpha} - (1-\varepsilon)^{\alpha}\right)\left(e^{-nE_n}\right)^{1-\alpha}$$

$$\rightarrow \left(\frac{1-e^{-nE_n}}{e^{-nE_n}}\right)^{1-\alpha} \overset{?}{\geq} \frac{(1-P(z=1))^{\alpha} - (1-\varepsilon)^{\alpha}}{\varepsilon^{\alpha} - P(z=1)^{\alpha}}$$

کافی است که $\frac{1-e^{-nE_n}}{e^{-nE_n}}$ به اندازه کافی بزرگ باشد. یعنی e^{-nE_n} به اندازه کافی کوچک باشد. برای $\alpha<1$ نیز همین‌طور اثبات می‌شود. ■

۷- در مسئله قبل قرار می‌دهیم:

$$nR_{1+\frac{1}{\sqrt{n}}}(P_x\|Q_x) \geq \frac{1}{\frac{1}{\sqrt{n}}}\log\left(\left(\tfrac{1}{3}\right)^{1+\frac{1}{\sqrt{n}}}\left(1-e^{-nE_n}\right)^{\frac{1}{\sqrt{n}}} + \left(\tfrac{2}{3}\right)^{1+\frac{1}{\sqrt{n}}}\left(e^{-nE_n}\right)^{\frac{1}{\sqrt{n}}}\right)$$

$$\geq \frac{1}{\frac{1}{\sqrt{n}}}\log\left(\left(\tfrac{1}{3}\right)^{1+\frac{1}{\sqrt{n}}}\left(1-e^{-nE_n}\right)^{\frac{1}{\sqrt{n}}}\right) = \log\left(\left(\tfrac{1}{3}\right)^{\frac{1}{\sqrt{n}}+\frac{1}{n}}\left(1-e^{-nE_n}\right)^{\frac{1}{n}}\right)$$

$$= \frac{1}{\sqrt{n}}\log\left(\left(\tfrac{1}{3}\right)^{1+\frac{1}{\sqrt{n}}}\right) + \frac{1}{n}\ln\left(1-e^{-nE_n}\right)$$

برای n های بزرگ عبارت اول صفر می‌شود و برای عبارت دوم طبق هوپیتال:

$$\lim_{n\to\infty}\frac{1}{n}\ln\left(1-e^{-nE_n}\right) = \frac{E_n}{1-e^{-nE_n}} \simeq E_n$$

پس برای n های به اندازه کافی بزرگ:

$$nR_{1+\frac{1}{\sqrt{n}}}(P_x\|Q_x) \geq E_n \quad ■$$

Date: ....

۸. داریم:

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta_{1-\varepsilon}(P,Q)} = E_n \leq \underbrace{R_{1+\frac{1}{\sqrt{n}}} (P_x \| Q_x)}_{\lim_{n\to\infty}} = R_{\alpha\to 1} (P\|Q) = D_{KL}(P\|Q)$$

پس داریم:

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta_{1-\varepsilon}(P,Q)} \leq D_{KL}(P\|Q) \qquad \blacksquare$$

Subject :   Date :

۵. انحراف من درآوردی!

۱. $V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X) = \sup_{f: \mathcal{X} \to \mathbb{R}} \left\{ \mathbb{E}_{P_X}[f(x)] - \gamma \mathbb{E}_{Q_X}[f(x)] - s \log \mathbb{E}_{Q_X}[e^{\alpha f(x)}] - t \log \mathbb{E}_{Q}[e^{\beta f(x)}] \right\}$

با توجه به اکیداً محدب بودن exp داریم: $\underset{\text{Jensen}}{\mathbb{E}[e^{x}] \geq \exp(\mathbb{E}[x])}$

پس داریم:

$$V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X) \geq \mathbb{E}_{P_X}[f(x)] - (\gamma + s\alpha + \beta t)\mathbb{E}_{Q_X}[f(x)]$$

فرض کنید تابع را ثابت بگیریم: $f(x) = c$. پس داریم: $\mathbb{E}_{Q_X}[f(x)] = \mathbb{E}_{P_X}[f(x)] = c$

داریم: $V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X) \geq c(1 - (\gamma + s\alpha + \beta t))$

حال اگر $\gamma + s\alpha + \beta t \neq 1$ باشد، می‌توانیم c را بی‌نهایت بزرگ کنیم تا عملا $V_{\alpha,\beta,\gamma,s,t}$ به بی‌نهایت می‌رود!

اما اگر $\gamma + s\alpha + \beta t = 1$ باشد تنها خواهیم داشت: $V_{\alpha,\beta,\gamma,s,t} \geq 0$.

۲. طبق سؤال قبل داریم: $V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X) \geq \mathbb{E}_{P_X}[f(x)] - (\gamma + s\alpha + \beta t)\mathbb{E}_{Q_X}[f(x)]$

حال کافی است $f(x) = c$ قرار دهیم، در این حالت خواهیم داشت: $V_{\alpha,\beta,\gamma,s,t} \geq 0$ پس با توجه به sup و اینکه برای

هر تابعی نامنفی است، پس این مقدار در هر صورت نامنفی است!

MICRO

۳. در این قسمت، از محدب بودن $-\log(x)$ استفاده می‌کنیم، داریم: $\mathbb{E}[-\log(X)] \geq -\log \mathbb{E}[X]$

پس داریم: $V_{\alpha,\beta,\gamma,s,t}(P_x \| Q_x) \leq \mathbb{E}_{P_x}[f(x)] - \gamma \mathbb{E}_{Q_x}[f(x)] + s\mathbb{E}_{Q_x}[-\log \exp(\alpha f(x))] + t\mathbb{E}[-\log \exp(\beta f(x))]$

$$\Rightarrow V_{\alpha,\beta,\gamma,s,t}(P_x \| Q_x) \leq \mathbb{E}_{P_x}[f(x)] - (\gamma + s\alpha + \beta t)\mathbb{E}_{Q_x}[f(x)]$$

$P_x = Q_x$ اگر $\longrightarrow \mathbb{E}_{P_x}[f(x)] = \mathbb{E}_{Q_x}[f(x)] \Longrightarrow V_{\alpha,\beta,\gamma,s,t}(P_x \| Q_x) \leq \underbrace{(1 - \underbrace{(\gamma + s\alpha + \beta t)}_{1})}_{0} \mathbb{E}_{P_x}[f(x)]$

پس:

$$V_{\alpha,\beta,\gamma,s,t}(P_x \| Q_x) \leq 0$$

از طرفی از قسمت قبل داشتیم که نامنفی است! پس در مجموع برای $P_x = Q_x$ داریم: $V_{\alpha,\beta,\gamma,s,t}(P_x \| Q_x) = 0$

برای اثبات عکس قضیه، تابع f را به این صورت تعریف می‌کنیم که برای تنها یک x^* در همه‌ی حالات c باشد و در بقیه حالات 0. داریم:

$$\mathbb{E}_{P_x}[f(x)] - \gamma \mathbb{E}_{Q_x}[f(x)] - s\log(\mathbb{E}_{Q_x}[\exp(\alpha f(x))]) - t\log(\mathbb{E}_{Q_x}[\exp(\beta f(x))])$$

$$c\,p(x^*) - c\gamma q(x^*) - s\log(e^{\alpha c} q(x^*)) - t\log(q(x^*)\, e^{\beta c})$$

$$= \underbrace{c(p(x^*) - (\underbrace{\gamma + s\alpha + t\beta}_{1}) q(x^*))}_{p(x^*) - q(x^*)} - s\log(q(x^*)) - t\log(q(x^*))$$

دقت کنید که اگر مقادیر $p(x^*), q(x^*)$ در این x^* خاص مساوی نشوند،

می‌توان c را به گونه‌ای تغییر داد که تابع V بی‌نهایت زیاد شود که با فرض $V(P_x \| Q_x) = 0$ در تناقض است. پس برای همه x^*ها این دو

برابرند. پس $P_x = Q_x$.

Subject : Date :

$$V_{\alpha,\beta,\gamma,s,t}(P_{XY}\|Q_{XY}) = \sup_{f:\mathcal{X}\to R}\{\mathbb{E}_{P_{XY}}[f(X,Y)] - \gamma\mathbb{E}_{Q_{XY}}[f(X,Y)] - s\log(\mathbb{E}_{Q_{XY}}[e^{\alpha f(X,Y)}]) - t\log(\mathbb{E}_{Q_{XY}}[e^{\beta f(X,Y)}]\}$$.ε

$$= \sup_{f:\mathcal{X},\mathcal{Y}\to R}\{\mathbb{E}_{P_X}[\mathbb{E}_{P_{Y|X}}[f(X,Y)]] - \gamma\mathbb{E}_{Q_X}[\mathbb{E}_{Q_{Y|X}}[f(X,Y)]] - s\log(\mathbb{E}_{Q_X}[\mathbb{E}_{Q_{Y|X}}[e^{\alpha f(X,Y)}]])$$

$$- t\log(\mathbb{E}_{Q_X}[\mathbb{E}_{Q_{Y|X}}[\exp(\beta f(X,Y))]])\}$$

$$\geq \sup_{f:\mathcal{X}\to R}\{\mathbb{E}_{P_X}[\mathbb{E}_{P_{Y|X}}[f(X)]] - \gamma\mathbb{E}_{Q_X}[\mathbb{E}_{Q_{Y|X}}[f(X)]] - s\log(\mathbb{E}_{Q_X}(\mathbb{E}_{Q_{Y|X}}[\exp(\alpha f(X))]))$$

$$- t\log(\mathbb{E}_{Q_X}(\mathbb{E}_{Q_{Y|X}}[\exp(\beta f(X))]))\}$$

$$\geq \sup_{f:\mathcal{X}\to R}\{\mathbb{E}_{P_X}[f(X)] - \gamma\mathbb{E}_{Q_X}[f(X)] - s\log\mathbb{E}_{Q_X}[\exp(\alpha f(X))] - t\log(\mathbb{E}_{Q_X}[\exp(\beta f(X))])\}$$

$$= V_{\alpha,\beta,\gamma,s,t}(P_X\|Q_X) \quad \blacksquare$$

$$V_{\alpha,\beta,\gamma,s,t}(P_{XY}\|Q_{XY}) = V_{\alpha,\beta,\gamma,s,t}(P_X W_{Y|X}\|Q_X W_{Y|X})$$.ω

$$= \sup\{\mathbb{E}_{P_{XY}}[f(X,Y)] - \gamma\mathbb{E}_{Q_{XY}}[f(X,Y)] - s\log\mathbb{E}_{Q_{XY}}[\exp(\alpha f(X,Y))] - t\log\mathbb{E}_{Q_{XY}}[\exp(\beta f(X,Y))]\}$$

$$= \sup\{\mathbb{E}_{P_X}[\mathbb{E}_{W_{Y|X}}[f(X,Y)]] - \gamma\mathbb{E}_{Q_X}[\mathbb{E}_{W_{Y|X}}[f(X,Y)]] - s\log\mathbb{E}_{Q_X}[\mathbb{E}_{W_{Y|X}}[\exp(\alpha f(X,Y))]] - t\log\mathbb{E}_{Q_X}[\mathbb{E}_{W_{Y|X}}[\exp(\beta f(X,Y))]]\}$$

$$\leq \sup_{f^{\beta}_{\gamma}:\mathcal{X}\to R}\{\mathbb{E}_{P_X}[f^{\alpha}_{\gamma}(X)] - \gamma\mathbb{E}_{Q_X}[f^{\alpha}_{\gamma}(X)] - s\log\mathbb{E}_{Q_X}[\exp(\alpha f^{\alpha}_{\gamma}(X))] - t\log\mathbb{E}_{Q_X}[\exp(\beta f^{\alpha}_{\gamma}(X))]\}$$

$$= V_{\alpha,\beta,\gamma,s,t}(P_X\|Q_X)$$

Subject : Date :

پس داریم: $$V_{\alpha,\beta,\gamma,s,t}(P_X W_{Y|X} \| Q_X W_{Y|X}) \leq V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X)$$

از طرفی از بخش قبل داشتیم: $$V_{\alpha,\beta,\gamma,s,t}(P_{XY} \| Q_{XY}) \geq V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X)$$

پس با تلفیق این دو، داریم: $$V_{\alpha,\beta,\gamma,s,t}(P_X W_{Y|X} \| Q_X W_{Y|X}) = V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X) \quad \blacksquare$$

حال توجه کنید که داریم: $$\underbrace{V_{\alpha,\beta,\gamma,s,t}(P_{XY} \| Q_{XY})} \geq V_{\alpha,\beta,\gamma,s,t}(P_Y \| Q_Y) \longrightarrow \underbrace{V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X)} \geq V_{\alpha,\beta,\gamma,s,t}(P_Y \| Q_Y)$$

حال برای اثبات محدب بودن فرض کنید که داریم:

$$P(X|Z) = \begin{cases} P_\lambda & Z=0 \\ P_{\bar\lambda} & Z=1 \end{cases} , \quad Q(X|Z) = \begin{cases} Q_\lambda & Z=0 \\ Q_{\bar\lambda} & Z=1 \end{cases}$$

که $Z = Ber(\lambda)$ است، داریم:

$$V(P_{XZ} \| Q_{XZ}) \geq V(P_X \| Q_X)$$

$$V(P_{XZ} \| Q_{XZ}) = \mathbb{E}_{P_Z}\left[V(P_{X|Z} \| Q_{X|Z}) \right] = \lambda V(P_\lambda \| Q_\lambda) + \bar\lambda V(P_{\bar\lambda} \| Q_{\bar\lambda}) \geq V(P_X \| Q_X)$$

$$\geq V(\lambda P_\lambda + \bar\lambda P_{\bar\lambda} \| \lambda Q_\lambda + \bar\lambda Q_{\bar\lambda}) \longrightarrow \text{پس محدب است} \quad \blacksquare$$

$$V_{\alpha,\beta,\gamma,s,t}(P_{XY} \| Q_X Q_Y) = \sup_{f: XY \to \mathbb{R}} \left\{ \mathbb{E}_{P_{XY}}[f(x,y)] - \gamma \mathbb{E}_{Q_X Q_Y}[f(x,y)] - s \log \mathbb{E}_{Q_X Q_Y}[\exp(\alpha f(x,y))] - t \log \mathbb{E}_{Q_X Q_Y}[\exp(\beta f(x,y))] \right\}$$

$$\geq \sup_{\substack{g: X \to \mathbb{R} \\ h: Y \to \mathbb{R}}} \left\{ \mathbb{E}_{P_{XY}}[g(x)+h(y)] - \gamma \mathbb{E}_{Q_X Q_Y}[g(x)+h(y)] - s \log \mathbb{E}_{Q_X Q_Y}[\exp(\alpha(g(x)+h(y)))] - t \log \mathbb{E}_{Q_X Q_Y}[\exp(\beta(g(x)+h(y)))] \right\}$$

Subject : Date :

$$\geq \sup \Big\{ \mathbb{E}_{P_X}[g(X)] + \mathbb{E}_{P_Y}[h(Y)] - r\,\mathbb{E}_{Q_X}[g(X)] - r\,\mathbb{E}_{Q_Y}[h(Y)] - s \log \mathbb{E}_{Q_X}[\exp(\alpha g(X))] - s \log \mathbb{E}_{Q_Y}[\exp(\alpha h(Y))]$$

$$- t \log \mathbb{E}_{Q_X}[\exp(\beta g(X))] - t \log \mathbb{E}_{Q_Y}[\exp(\beta h(Y))] \Big\}$$

$$\geq V_{\alpha,\beta,r,s,t}(P_X \| Q_X) + V_{\alpha,\beta,r,s,t}(P_Y \| Q_Y) \qquad \blacksquare$$

$$W_\alpha(P_X \| Q_X) = \sup_{f: \mathcal{X} \to \mathbb{R}} \Big\{ \mathbb{E}_{P_X}[f(X)] - \left(1 - \frac{1}{\alpha}\right) \mathbb{E}_{Q_X}[f(X)] - \frac{1}{\alpha^2} \log \mathbb{E}_{Q_X}[\exp(\alpha f(X))] \Big\}$$

داریم:

$$\mathbb{E}_{Q_X}[e^{\alpha f(X)}] = 1 + \alpha\, \mathbb{E}_{Q_X}[f(X)] + \frac{1}{2} \alpha^2\, \mathbb{E}_{Q_X}[f(X)^2] + O(\alpha^3)$$

$$\Rightarrow \log \mathbb{E}_{Q_X}[e^{\alpha f(X)}] = \alpha\, \mathbb{E}_{Q_X}[f(X)] + \frac{1}{2}\alpha^2\, \mathbb{E}_{Q_X}[f(X)^2] - \frac{1}{2}\alpha^2\, \mathbb{E}[f(X)]^2 + O(\alpha^3)$$

$$= \alpha\, \mathbb{E}_{Q_X}[f(X)] + \frac{\alpha^2}{2} \mathrm{Var}_{Q_X}(f(X)) + O(\alpha^3)$$

$$\Rightarrow W_\alpha(P_X \| Q_X) = \sup_{f: \mathcal{X} \to \mathbb{R}} \Big\{ \mathbb{E}_{P_X}[f(X)] - \Big(1 - \underbrace{\cancel{\tfrac{1}{\alpha}}}_{\times}\Big) \mathbb{E}_{Q_X}[f(X)] \underbrace{- \frac{1}{\alpha} \mathbb{E}_{Q_X}[f(X)]}_{\times} - \frac{1}{2} \mathrm{Var}(f(X)) + O(\alpha) \Big\}$$

$$\rightarrow \lim_{\alpha \to 0} W_\alpha(P_X \| Q_X) = \sup_{f: \mathcal{X} \to \mathbb{R}} \Big\{ \mathbb{E}_{P_X}[f(X)] - \mathbb{E}_{Q_X}[f(X)] - \frac{1}{2} \mathrm{Var}_{Q_X}(f(X)) + \underbrace{O(\alpha)}_{\alpha \to 0:\ 0} \Big\}$$

$$= \sup_{f: \mathcal{X} \to \mathbb{R}} \Big\{ \mathbb{E}_{P_X}[f(X)] - \mathbb{E}_{Q_X}[f(X)] - \frac{1}{2} \mathrm{Var}_{Q_X}(f(X)) \Big\}$$

حال برای به دست آوردن این مقدار، از بسط استفاده می‌کنیم تا بتوان تفاوت‌های لحظه‌ای دو توزیع را بسنجیم، داریم:

$$\lim_{\alpha\to 0} W_\alpha(P_X \| Q_X) = \sup_{c} \sup_{cf} \left\{ \mathbb{E}_{P_X}[c f(x)] - \mathbb{E}_{Q_X}[c f(x)] - \frac{c^2}{2} \mathrm{Var}(f(x)) \right\}$$

$$\frac{\partial}{\partial c} = 0 \qquad C_{\text{optimum}} = \frac{\mathbb{E}_{P_X}[f(x)] - \mathbb{E}_{Q_X}[f(x)]}{\mathrm{Var}_{Q_X}(f(x))}$$

گذاردن بهینه در رابطه $\longrightarrow$ $$\lim_{\alpha\to 0} W_\alpha(P_X \| Q_X) = \frac{1}{2} \sup_{f: X\to R} \frac{\left(\mathbb{E}_{P_X}[f(x)] - \mathbb{E}_{Q_X}[f(x)]\right)^2}{\mathrm{Var}(f(x))} = \frac{1}{2}\chi^2(P_X, Q_X)$$

۸. از قسمت های قبل داریم:

$$V_{\alpha,\beta,\gamma,s,t}(P_{XY} \| Q_X Q_Y) \geq V_{\alpha,\beta,\gamma,s,t}(P_X \| Q_X) + V_{\alpha,\beta,\gamma,s,t}(P_Y \| Q_Y)$$

$$\Big\downarrow \; \alpha=\alpha,\ \beta=0,\ \gamma = 1-\frac{1}{\alpha},\ s = \frac{1}{\alpha^2},\ t=0$$

$$V_{\alpha,0,1-\frac{1}{\alpha},\frac{1}{\alpha^2},0}(P_{XY} \| Q_X Q_Y) \geq V_{\alpha,0,1-\frac{1}{\alpha},\frac{1}{\alpha^2},0}(P_X \| Q_X) + V_{\alpha,0,1-\frac{1}{\alpha},\frac{1}{\alpha^2},0}(P_Y \| Q_Y)$$

$$\Big\downarrow \lim_{\alpha\to 0}$$

$$\chi^2(P_{XY}, Q_X Q_Y) \geq \chi^2(P_X, Q_X) + \chi^2(P_Y, Q_Y)$$

۶. نامساوی‌های اطلاعاتی

۱. داریم:

$$D_{\chi^2}(P_X \| Q_X) = \sup_{f} \left(2\mathbb{E}_P[f] - 2\mathbb{E}_Q[f] - \mathrm{Var}_Q(f) \right)$$

$$\Rightarrow D_{\chi^2}(P_{XY} \| Q_X Q_Y) = \sup_{T: X, Y \to \mathbb{R}} \left(2\mathbb{E}_{P_{XY}}[T] - 2\mathbb{E}_{Q_X Q_Y}[T(X,Y)] - \mathrm{Var}_{Q_X Q_Y}(T(X,Y)) \right)$$

حال تابع $T(X,Y)$ را می‌توان به صورت $f_1(X) + f_2(Y)$ نیز نوشت، و علاوه برای این حالت کوچک‌تر است! داریم:

$$\mathbb{E}_{P_{XY}}(f_1(X) + f_2(Y)) = \mathbb{E}_{P_X}(f_1(X)) + \mathbb{E}_{P_Y}[f_2(Y)]$$

$$\mathbb{E}_{Q_X Q_Y}[f_1(X) + f_2(Y)] = \mathbb{E}_{Q_X}[f_1(X)] + \mathbb{E}_{Q_Y}[f_2(Y)]$$

$$\mathrm{Var}_{Q_X Q_Y}(f_1 + f_2) = \mathbb{E}_{Q_X Q_Y}[(f_1 + f_2)^2] - \left(\mathbb{E}_{Q_X}[f_1] + \mathbb{E}_{Q_Y}[f_2]\right)^2 = \mathbb{E}_{Q_X Q_Y}[f_1^2 + f_2^2 + 2f_1 f_2] - \left(\mathbb{E}_{Q_X}[f_1] + \mathbb{E}_{Q_Y}[f_2]\right)^2$$

$$= \mathbb{E}_{Q_X}[f_1^2] - (\mathbb{E}_{Q_X}[f_1])^2 + \mathbb{E}_{Q_Y}[f_2^2] - (\mathbb{E}_{Q_Y}[f_2])^2 + \underbrace{2\mathbb{E}_{Q_X Q_Y}[f_1 f_2] - 2\mathbb{E}_{Q_X}[f_1]\,\mathbb{E}_{Q_Y}[f_2]}_{0}$$

$$= \mathrm{Var}_{Q_X}(f_1) + \mathrm{Var}_{Q_Y}(f_2)$$

پس داریم:

$$D_{\chi^2}(P_{XY} \| Q_X Q_Y) \geq \sup_{\substack{f_1: X \to \mathbb{R} \\ f_2: Y \to \mathbb{R}}} \left(2\mathbb{E}_{P_X}(f_1) + 2\mathbb{E}_{P_Y}[f_2(Y)] - 2\mathbb{E}_{Q_X}[f_1] - 2\mathbb{E}_{Q_Y}[f_2] - \mathrm{Var}_{Q_X}(f_1) - \mathrm{Var}_{Q_Y}(f_2) \right)$$

$$= \sup_{f_1: X \to \mathbb{R}} \left(2\mathbb{E}_{P_X}[f_1] - 2\mathbb{E}_{Q_X}[f_1] - \mathrm{Var}_{Q_X}(f_1) \right) + \sup_{f_2: Y \to \mathbb{R}} \left(2\mathbb{E}_{P_Y}[f_2] - 2\mathbb{E}_{Q_Y}[f_2] - \mathrm{Var}_{Q_Y}(f_2) \right)$$

$$= D_{\chi^2}(P_X \| Q_X) + D_{\chi^2}(P_Y \| Q_Y) \quad \blacksquare$$

۲.

$$D_{KL}(P_{XYZ} \| P_X P_{YZ}) = \sum_{x,y,z} P_{XYZ} \log\left(\frac{P_X P_{Y|X} P_{Z|X}}{P_X P_Y P_{Z|Y}}\right)$$

$$= \sum_{x,y,z} P_{XYZ} \log\left(\frac{P_{XY}}{P_Y P_X}\right) + \sum_{x,y,z} P_{XYZ} \log\left(\frac{P_{Z|X}}{P_{Z|Y}}\right)$$

$$= \sum_{x,y} P_{XY} \log\left(\frac{P_{XY}}{P_X P_Y}\right) + \sum_{x,y,z} P_{XYZ}\left(\log\left(\frac{P_{XZ}}{P_X P_Z}\right) - \log\left(\frac{P_{Z|Y} P_{Y|Z}}{P_Y}\right)\right)$$

$$= D_{KL}(P_{XY} \| P_X P_Y) + D_{KL}(P_{XZ} \| P_X P_Z) - \sum P_{XYZ} \log\left(\frac{P_{ZY}}{P_Z P_Y}\right)$$

$$= D_{KL}(P_{XY} \| P_X P_Y) + D_{KL}(P_{XZ} \| P_X P_Z) - \underbrace{D_{KL}(P_{ZY} \| P_Z P_Y)}_{\geq 0}$$

$$\Longrightarrow D_{KL}(P_{XYZ} \| P_X P_{YZ}) \leq D_{KL}(P_{XY} \| P_X P_Y) + D_{KL}(P_{XZ} \| P_X P_Z) \quad \blacksquare$$

۳. داریم:

$$D_{sKL}(P \| Q) = \sum q\left(\frac{p}{q} - 1\right)\left(\log\left(\frac{p}{q}\right)\right) = \sum p \log\left(\frac{p}{q}\right) + \sum q \log\left(\frac{q}{p}\right)$$

$$\Longrightarrow D_{sKL}(P \| Q) = D_{KL}(P \| Q) + D_{KL}(Q \| P)$$

$$D_{KL}(P_X P_{YZ} \| P_{XYZ}) = \sum_{x,y,z} P_X P_{YZ} \log\left(\frac{P_X P_Y P_{Z|Y}}{P_X P_{Z|X} P_{Y|X}}\right) = \sum P_X P_Y P_{Z|Y}\left(\log\left(\frac{P_X P_Y}{P_{XY}}\right) - \log\left(\frac{P_{Z|X}}{P_{Z|Y}}\right)\right)$$

$$= \sum_{x,y} P_X P_Y \log\left(\frac{P_X P_Y}{P_{XY}}\right) - \sum P_X P_Y P_{Z|Y}\left(-\log\left(\frac{P_X P_Z}{P_{XZ}}\right) - \log\left(\frac{P_{Z|Y} P_{Y|Z}}{P_Y}\right)\right)$$

$$= D_{KL}(P_X P_Y \| P_{XY}) + \sum_{x,z} P_X P_Z \log\left(\frac{P_X P_Z}{P_{XZ}}\right) + \sum_{x,y,z} P_X P_Y P_{Z|Y} \log\left(\frac{P_{Z|Y} P_{Y|Z}}{P_Y}\right)$$

$$= D_{KL}(P_X P_Y \| P_{XY}) + D_{KL}(P_X P_Z \| P_{XZ}) + \sum_{y,z} P_{YZ} \log\left(\frac{P_{ZY}}{P_Y P_Z}\right)$$

$$= D_{KL}(P_X P_Y \| P_{XY}) + D_{KL}(P_X P_Z \| P_{XZ}) + D_{KL}(P_{YZ} \| P_Y P_Z)$$

Subject : Date :

دقت کنید که این حکم آخر با توجه به عبارتی که در بخش قبل به دست آوردیم، نتیجه را به ما می‌دهد:

$$D_{SKL}(P_{XYZ} \| P_X P_{YZ}) = D_{SKL}(P_{XY} \| P_X P_Y) + D_{SKL}(P_{XZ} \| P_X P_Z) \quad \blacksquare$$

۴. باید ثابت کنیم:

$$1-\sum\sqrt{P_{XYZ}Q_{XYZ}} \le 1-\sum\sqrt{P_{XY}Q_{XY}} + 1-\sum\sqrt{P_{XZ}Q_{XZ}}$$

$$\Longleftrightarrow \sum\sqrt{P_{XY}Q_{XY}} - \sum\sqrt{P_{XYZ}Q_{XYZ}} \le 1-\sum\sqrt{P_{XZ}Q_{XZ}}$$

$$\Longleftrightarrow \sum_X\sqrt{P_XQ_X}\sum_Y\sqrt{P_{Y|X}Q_{Y|X}}\left(1-\sum_Z\sqrt{P_{Z|X}Q_{Z|X}}\right) \le \sum_X\sqrt{P_XQ_X}\left(1-\sum_Z\sqrt{P_{Z|X}Q_{Z|X}}\right)+1-\sum\sqrt{P_XQ_X}$$

$$\Longleftrightarrow \sum_X\sqrt{P_XQ_X}\left(1-\sum_Z\sqrt{P_{Z|X}Q_{Z|X}}\right)\left(\sum_Y\sqrt{P_{Y|X}Q_{Y|X}}\right) \le \sum_X\sqrt{P_XQ_X}\left(1-\sum_Z\sqrt{P_{Z|X}Q_{Z|X}}\right)+1-\sum\sqrt{P_XQ_X}$$

حال دقت کنید که طبق کوشی شوارتز داریم:

$$\sum_Y\sqrt{P_{Y|X}Q_{Y|X}} \le 1 \Longleftarrow \left(\sum_Y\sqrt{P_{Y|X}Q_{Y|X}}\right)^2 \le \left(\sum_Y P_{Y|X}\right)\left(\sum Q_{Y|X}\right)$$

پس طبق نامساوی اخیر:

$$\sum\sqrt{P_XQ_X}\left(1-\sum_Z\sqrt{P_{Z|X}Q_{Z|X}}\right)\left(\sum_Y\sqrt{P_{Y|X}Q_{Y|X}}\right) \le \sum_X\sqrt{P_XQ_X}\left(1-\sum_Z\sqrt{P_{Z|X}Q_{Z|X}}\right)$$

پس برای اثبات کل کافی است ثابت کنیم: $1-\sum_X\sqrt{P_XQ_X}$ که طبق کوشی شوارتز داریم:

$$\left(\sum_X\sqrt{P_XQ_X}\right)^2 \le \left(\sum P_X\right)\left(\sum Q_X\right) = 1\times 1 = 1$$

پس حکم به طور کل اثبات شد. $\blacksquare$

Subject :　　　　　　　　　　Date :

۵. کافی است در نامساوی اخیر قرار دهیم: $Q_{XYZ} = P_X P_{YZ}$ آنگاه داریم:

$$Q_{XY} = \sum_{z} P_X P_{YZ} = P_X P_Y$$

$$Q_{XZ} = \sum_{y} P_X P_{YZ} = P_X P_Z$$

پس نامساوی اخیر می‌شود:

$$H^2(P_{XYZ} \| P_X P_{YZ}) \leq H^2(P_{XY} \| P_X P_Y) + H^2(P_{XZ} \| P_X P_Z)$$

که همان زیرجمعی بودن را نتیجه می‌دهد.

۶. برای TV داریم:

$$TV(P, Q) = \frac{1}{2} \sum |P_x - Q_x|$$

داریم:

$$TV(P_{XYZ}, Q_{XYZ}) = \frac{1}{2} \sum |P_{XYZ} - Q_{XYZ}| \underset{\text{نامساوی مثلث}}{\leq} \frac{1}{2} \sum |P_{XYZ} - P_X P_{Y|X} Q_{Z|X}| + |P_X P_{Y|X} Q_{Z|X} - Q_{XYZ}|$$

پس

$$\Rightarrow TV(P_{XYZ}, Q_{XYZ}) \leq TV(P_{XYZ}, P_X P_{Y|X} Q_{Z|X}) + TV(P_X P_{Y|X} Q_{Z|X}, Q_{XYZ})$$

با توجه به نامساوی پردازش داده داریم:

$$TV(P_{XYZ}, P_X P_{Y|X} Q_{Z|X}) = TV(P_{XY}, Q_{XY})$$

چون فقط به X, Y بستگی دارد.

$$TV(P_X P_{Y|X} Q_{Z|X}, Q_{XYZ}) = TV(P_{XZ}, Q_{XZ})$$

فقط به XZ بستگی دارد.

پس با اعمال این دو، حکم اثبات می‌شود. ■

Subject :　　　　　　　　　　　　　　　　　　　　Date :

برای χ^2 نیز داریم:

$$\chi^2(P\|Q) = \sum Q_x \left(\frac{P_x}{Q_x}-1\right)^2 = \mathbb{E}_{Q_x}\left[\left(\frac{P_x}{Q_x}-1\right)^2\right]$$

$$\Rightarrow \chi^2(P_{XYZ}\|Q_{XYZ}) = \mathbb{E}_{Q_{XYZ}}\left[\left(\frac{P_{XYZ}}{Q_{XYZ}}-1\right)^2\right] = \mathbb{E}_{Q_X Q_{Y|X} Q_{Z|X}}\left[\left(\frac{P_X}{Q_X}\times\frac{P_{Y|X}}{Q_{Y|X}}\times\frac{P_{Z|X}}{Q_{Z|X}}-1\right)^2\right]$$

داریم $\Rightarrow \left(\frac{P_X}{Q_X}\times\frac{P_{Y|X}}{Q_{Y|X}}\times\frac{P_{Z|X}}{Q_{Z|X}}-1\right)^2 = \left(\frac{P_X}{Q_X}\times\frac{P_{Y|X}}{Q_{Y|X}}\times\frac{P_{Z|X}}{Q_{Z|X}}\right)^2 - 2\,\frac{P_X}{Q_X}\times\frac{P_{Y|X}}{Q_{Y|X}}\times\frac{P_{Z|X}}{Q_{Z|X}} + 1$

$$\Rightarrow \chi^2(P_{XYZ}\|Q_{XYZ}) = \mathbb{E}_{Q_X}\left[\left(\frac{P_X}{Q_X}\right)^2 \mathbb{E}_{Q_{Y|X}}\left[\left(\frac{P_{Y|X}}{Q_{Y|X}}\right)^2\right] \mathbb{E}_{Q_{Z|X}}\left[\left(\frac{P_{Z|X}}{Q_{Z|X}}\right)^2\right]\right] - 2\mathbb{E}_{P_X}[1] + 1$$

$$\Rightarrow \chi^2(P_{XYZ}\|Q_{XYZ}) = \mathbb{E}_{Q_X}\left[\left(\frac{P_X}{Q_X}\right)^2 \left(1+\chi^2(P_{Y|X}\|Q_{Y|X})\right)\left(1+\chi^2(P_{Z|X}\|Q_{Z|X})\right)\right] - 1$$

پس داریم:

$$\chi^2(P_{XYZ}\|Q_{XYZ}) \leq \mathbb{E}_{Q_X}\left[\left(\frac{P_X}{Q_X}\right)^2 \left(\chi^2(P_{Y|X}\|Q_{Y|X}) + \chi^2(P_{Z|X}\|Q_{Z|X})\right)\right]$$

همچنین داریم:

$$\chi^2(P_{XY}\|Q_{XY}) = \mathbb{E}_{Q_X}\left[\left(\frac{P_X}{Q_X}\right)^2 \chi^2(P_{Y|X}\|Q_{Y|X})\right]$$

$$\chi^2(P_{XZ}\|Q_{XZ}) = \mathbb{E}_{Q_X}\left[\left(\frac{P_X}{Q_X}\right)^2 \chi^2(P_{Z|X}\|Q_{Z|X})\right]$$

پس با توجه به دو تساوی اخیر داریم:

$$\chi^2(P_{XYZ}\|Q_{XYZ}) \leq \chi^2(P_{XY}\|Q_{XY}) + \chi^2(P_{XZ}\|Q_{XZ})$$

* دقت کنید که باز هم با قرار دادن $Q_{XYZ} = P_X P_{YZ}$ حکم حالت زیرجمعی بودن نتیجه می‌شود ■

MICRO

TV در برابر KL

با توجه به نامساوی ~~Bret~~ Bretagnolle-Huber (1978) داریم:

$$D_{KL}(P\|Q) \geq -\log\left(1-d_{TV}(P,Q)^2\right)$$

از طرفی با توجه به بسط تیلور داریم:

$$\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right) = 2\left(\varepsilon + \frac{\varepsilon^3}{3} + \frac{\varepsilon^5}{5} + \dots\right)$$

حال با قرار دادن $\delta = d_{TV}(P,Q)$ داریم:

$$-\log(1-\delta^2) = \log\left(\frac{1}{1-\delta^2}\right) = \log\left(\frac{1}{1+\delta}\right) + \log\left(\frac{1}{1-\delta}\right)$$

برای اثبات حکم کافی است نشان دهیم:

$$-\log(1-\delta^2) \geq \log\left(\frac{1+\delta}{1-\delta}\right) - \frac{2\delta}{1+\delta}$$

قرار دهیم: $f(\delta) = \log\left(\frac{1+\delta}{1-\delta}\right) - \frac{2\delta}{1+\delta}$ داریم:

$$f'(\delta) = \frac{2}{1-\delta^2} - \frac{2}{(1+\delta)^2} = \frac{2\left((1+\delta)^2 - (1-\delta^2)\right)}{(1-\delta^2)(1+\delta)^2} = \frac{2\delta^2+4\delta}{(1+\delta^2)(1+\delta)^2} \geq 0 \quad \forall \delta \in [0,1]$$

حال با توجه به اینکه در بازه‌ای که d_{TV} قرار دارد، f' مثبت است.

پس دقت کنید که برای مقایسه این نامساوی (Vadja) با Pinker، داریم که برای δهای کوچک، این دو با توجه به بسط تیلور تقریبا یکی هستند. اما وقتی d_{TV} نسبتا زیاد باشد، Vadja عملا باند بهتر و دقیق‌تری ارائه می‌دهد.

Subject :   Date :

$P = Ber(p)$, $Q = Ber(q)$ ، $p > q$ — برای اثبات نامساوی برای حالتی که فضای حالت ما 2 است داریم:

باید ثابت کنیم: $d = p - q \longrightarrow p \log \frac{p}{q} + (1-p)\log\left(\frac{1-p}{1-q}\right) \geq \log\left(\frac{1+d}{1-d}\right) - \frac{2d}{1+d}$

$$\longrightarrow q = p - d \longrightarrow p \log \frac{p}{p-d} + (1-p)\log\left(\frac{1-p}{1-p+d}\right) \geq \log\left(\frac{1+d}{1-d}\right) - \frac{2d}{1+d}$$

$$\log\left(\frac{1-p}{1-p+d}\right) - p\log\left(\frac{1-p}{p} \times \frac{p-d}{1-p+d}\right) = d\log\left(\frac{1+d}{1-d}\right)$$

باید ثابت کنیم $\longrightarrow$ $\underbrace{(d-1)\log\left(\frac{1+d}{1-d}\right) + \frac{2d}{1+d}}_{g(d)} \geq 0$

$$\longrightarrow g'(d) = \log\left(\frac{1+d}{1-d}\right) + (d-1) \times \frac{2}{1-d^2} + \frac{2}{(1+d)^2} \geq 0 \quad \text{for } d \in [0,1)$$

از طرفی چون $g(0)=0$ و $g' \geq 0$ پس این تابع صعودی و نامنفی است. پس حکم اثبات می‌شود. ∎

MICRO