

①  $P \ll Q$  ;  $Z_i = \log\left(\frac{P(X_i)}{Q(X_i)}\right)$  ;  $W_i = \log\left(\frac{P(Y_i)}{Q(Y_i)}\right)$  ;  $X_i \stackrel{iid}{\sim} P$  ,  $Y_i \stackrel{iid}{\sim} Q$

$$P\left[\sum_{i=1}^n (W_i - Z_i) \geq nt\right] \underset{\lambda > 0}{=} P\left[e^{\lambda \sum_{i=1}^n (W_i - Z_i)} \geq e^{\lambda nt}\right] \underset{\text{Markov}}{\leq} e^{-\lambda nt} \mathbb{E}_{X, Y} \left[ e^{\lambda \sum_{i=1}^n (W_i - Z_i)} \right]$$

توابع احتمالی  $X_i$  و  $Y_i$  مستقلند و  $X_i$  و  $Y_i$  نیز مستقلند.

$$\mathbb{E}_{X, Y} \left[ e^{\lambda \sum_{i=1}^n (W_i - Z_i)} \right] = \prod_{i=1}^n \mathbb{E}_{X_i, Y_i} \left[ e^{\lambda (W_i - Z_i)} \right] = \left[ \mathbb{E}_{X, Y} e^{\lambda (W - Z)} \right]^n$$

$$\Rightarrow P\left[\sum_{i=1}^n (W_i - Z_i) \geq nt\right] \leq \left[ e^{-\lambda t} \left( \mathbb{E}_X e^{-\lambda Z} \right) \left( \mathbb{E}_Y e^{\lambda W} \right) \right]^n = \exp\left(n \left[ -\lambda t + \log(\mathbb{E}_X e^{-\lambda Z}) + \log(\mathbb{E}_Y e^{\lambda W}) \right]\right)$$

$$\Rightarrow P\left[\sum_{i=1}^n (W_i - Z_i) \geq nt\right] \leq \inf_{\lambda > 0} \exp\left(n \left[ -\lambda t + \log(\mathbb{E}_X e^{-\lambda Z}) + \log(\mathbb{E}_Y e^{\lambda W}) \right]\right)$$

در نتیجه اگر  $\inf$  از R.H.S. را بگیریم:

$$= \sup_{\lambda > 0} \exp\left(-n \left[ \lambda t - \log(\mathbb{E}_X e^{-\lambda Z}) - \log(\mathbb{E}_Y e^{\lambda W}) \right]\right) = \exp\left(-n \cdot \sup_{\lambda > 0} \left\{ \lambda t - \log(\mathbb{E}_X e^{-\lambda Z}) - \log(\mathbb{E}_Y e^{\lambda W}) \right\}\right)$$

$$P\left[\sum_{i=1}^n (W_i - Z_i) \geq nt\right] \leq \exp(-n F(t)) \quad \text{Q.E.D.}$$

نسبت به  $t$  مشتق

②  $t=0 \rightarrow F(0) = \sup_{\lambda > 0} \left\{ -\psi(-\lambda) - \psi(\lambda) \right\} = \sup_{\lambda > 0} \left\{ -\log(\mathbb{E}_X e^{-\lambda Z}) - \log(\mathbb{E}_Y e^{\lambda W}) \right\}$

$$= -\sup_{\lambda > 0} \left\{ \log\left(\int_{-\infty}^{\infty} P(x) \left(\frac{Q(x)}{P(x)}\right)^{\lambda} dx\right) + \log\left(\int_{-\infty}^{\infty} Q(y) \left(\frac{P(y)}{Q(y)}\right)^{\lambda} dy\right) \right\}$$

$$= -\sup_{\lambda > 0} \left\{ \log\left(\int_{-\infty}^{\infty} P(x)^{1-\lambda} Q(x)^{\lambda} dx\right) + \log\left(\int_{-\infty}^{\infty} P(y)^{\lambda} Q(y)^{1-\lambda} dy\right) \right\}$$

حال طبق نام Cauchy Schwartz:

$$= \sup_{\lambda > 0} -\log\left(\int_{-\infty}^{\infty} \sqrt{P(x)^{1-\lambda} Q(x)^{\lambda}} dx\right)^2 = -2 \log\left(\int_{-\infty}^{\infty} \sqrt{P(x) Q(x)} dx\right)$$

از طرف دیگر داریم:

$$F(0) = \sup_{\lambda > 0} \left\{ -2 \log\left(\int_{-\infty}^{\infty} \sqrt{P(x) Q(x)} dx\right) \right\} = -2 \log\left(\mathbb{E}_{Y \sim Q} \left[ \sqrt{\frac{P(Y)}{Q(Y)}} \right]\right)$$

حال که ثابت شد که

$$f(x) = \alpha = -2 \log B(P, Q) = -2 \log \left( \int_{\mathcal{X}} \sqrt{\frac{p(x)}{q(x)}} dx \right)$$

از طرف دیگر می بینیم که حالت متساوی برابری در اینجا وقتی است که نسبت تغییر - تغییر در این حالت ثابت باشد:

$$\forall x: \frac{\partial}{\partial \lambda} C = \frac{p(x)^{1-\lambda} q(x)^{\lambda}}{p(x)^{\lambda} q(x)^{1-\lambda}} = p(x)^{1-2\lambda} q(x)^{1-2\lambda} \rightarrow 1-2\lambda=0$$

اما بیایم بدون این تئوری بر سر صفر  $x$  را قرار دهیم چون  $p$  و  $q$  از هم متباینند و برابر هم نمی آیند.

در نتیجه برای  $p(x)$  و  $q(x)$  باید صفر باشد که نتیجه می شود  $1-2\lambda=0 \rightarrow \lambda=\frac{1}{2}$  که در نهایت نتیجه می شود:

$$\rightarrow f(x) = \alpha = -2 \log B(P, Q) = -\psi_p(-\lambda) - \psi_q(\lambda) \Big|_{\lambda=\frac{1}{2}} = -\psi_p(-1/2) - \psi_q(1/2).$$

□ Q.E.D.

$$\textcircled{II} f(t+1) = \frac{t-p}{\lambda_{t+1}} \left\{ \lambda_t - \psi_p(-\lambda) - \psi_q(\lambda) \right\} \geq f(\lambda_0, t) \Big|_{\lambda_0=\frac{1}{2}}$$

$$\forall t: f(t+1) \geq \frac{t}{2} - \underbrace{\psi_p(-1/2) - \psi_q(1/2)}_{f_{0,1}} = \frac{t}{2} + f_{0,1} \Rightarrow \forall t: f(t) \geq \frac{t}{2} + f_{0,1}$$

$$\Rightarrow \forall t: -n f_{t+1} \leq -n \left( f_{0,1} + t/2 \right) \quad (*)$$

طبق استقرا در نتیجه داریم و می توانیم نتیجه بگیریم:  $(*)$

$$P \left[ \sum_{i=1}^n (w_i - z_i) \geq nt \right] \leq \exp(-n f_{t+1}) \leq \exp \left[ -n \left( \frac{t}{2} + f_{0,1} \right) \right] \quad \square \text{ Q.E.D.}$$

حالت 2: ابتدا فرض می کنیم که  $(\epsilon_0, f_0)$  و  $(\epsilon_1, f_1)$  در سطح قابل دسترس باشند به طوریکه:  $P_{110}, P_{111}$  و  $\pi_{110}, \pi_{111}$

$$\rightarrow \left. \begin{matrix} \pi_{110} \leq 2^{-n\epsilon_0} \\ P_{110} \leq 2^{-n} \end{matrix} \right\} \text{ و } \left\{ \begin{matrix} \pi_{111} \leq 2^{-n\epsilon_1} \\ \pi_{110} \leq 2^{-nf_1} \end{matrix} \right. \quad \textcircled{I}$$

$$\begin{aligned} E_{\lambda}^* &= \lambda \epsilon_0 + (1-\lambda) \epsilon_1 \leq \max(\epsilon_0, \epsilon_1) \\ f_{\lambda}^* &= \lambda f_0 + (1-\lambda) f_1 \leq \max(f_0, f_1) \end{aligned} \Rightarrow \begin{cases} 2^{-n(\lambda \epsilon_0 + (1-\lambda) \epsilon_1)} \geq 2^{-n \max(\epsilon_0, \epsilon_1)} \geq 2^{-n\epsilon_0} \text{ و } 2^{-nf_0} \\ 2^{-n(\lambda \epsilon_1 + (1-\lambda) \epsilon_0)} \geq 2^{-n \max(\epsilon_0, \epsilon_1)} \geq 2^{-n\epsilon_1} \text{ و } 2^{-nf_1} \end{cases}$$

در نتیجه همان طاقان  $\pi$  و  $P$  نیز در مقادیر فوق حدت می باشد در نتیجه چون  $\pi$  این حالت را که زوج نزن

$(\epsilon_{\lambda}^*, f_{\lambda}^*)$  برابر هر  $\lambda \in [0, 1]$  به شرطی که زوج نزن  $(\epsilon_0, f_0)$  و  $(\epsilon_1, f_1)$  در دسترس باشند و بدین ترتیب

2 که نامیه زوج نزن را در دسترس، باید یک نامیه زوج نزن داشته باشیم. البته چون که  $\pi_{110}$  و  $\pi_{111}$  نیز زوج نزن  
و  $\pi$  طبق تبدیلی  $\frac{1}{\pi} \leq \frac{1}{\epsilon} \leq \frac{1}{\pi}$  تبدیلی می شوند، که در این است نیز بدین ترتیب (زوج نزن) نیز در دسترس است.

ادامه سوال 2: مطلوب است

$$(i) \pi_{110} = P\left(\sum_{k=1}^n \log\left(\frac{Q(X_k)}{P(X_k)}\right) \geq \tau = nt\right) \xrightarrow{\lambda > 0} P\left(e^{\sum_{k=1}^n \log\left(\frac{Q(X_k)}{P(X_k)}\right)} \geq e^{\lambda nt}\right)$$

$$\rightarrow \text{By Markov} \Rightarrow e^{-\lambda nt} E_{X_k \sim P} \left[ e^{\sum_{k=1}^n \log\left(\frac{Q(X_k)}{P(X_k)}\right)} \right] \geq P(T(X^n) \geq nt)$$

let:  $\psi_P(\lambda) = \log\left(E_{X \sim P} \left[ e^{\lambda \log\left(\frac{Q(X)}{P(X)}\right)} \right]\right)$  L.H.S. از قضیه است:

$$P(T(X^n) \geq nt) \leq \inf_{\lambda > 0} e^{-\lambda nt} E_{X \sim P} \left[ e^{\lambda \sum_{k=1}^n \log\left(\frac{Q(X_k)}{P(X_k)}\right)} \right] = \exp\left[-n \cdot \sup_{\lambda > 0} \left\{ \lambda t - \psi_P(\lambda) \right\}\right]$$

$$\Rightarrow P(T(X^n) \geq nt) \leq e^{-n \psi_P^*(t)} \quad \text{Q.E.D.} \blacksquare$$

let  $\psi_P^*(t) = \sup_{\lambda > 0} \left\{ \lambda t - \psi_P(\lambda) \right\}$  ما در بالا از فرمول اول استفاده کردیم که شرط آخر این است.

(ii)  $\pi_{011} = Q\left(\sum_{k=1}^n \log\left(\frac{P(X_k)}{Q(X_k)}\right) < \tau = nt\right) = Q\left(e^{-\sum_{k=1}^n \log\left(\frac{Q(X_k)}{P(X_k)}\right)} \geq e^{\lambda nt}\right)$

$$\rightarrow \text{By Markov} \Rightarrow e^{-\lambda nt} E_{X_k \sim Q} \left[ e^{-\sum_{k=1}^n \log\left(\frac{Q(X_k)}{P(X_k)}\right)} \right] \geq Q(T(X^n) \geq nt)$$

let  $\psi_Q(\lambda) = \log\left(E_{X \sim Q} \left[ e^{\lambda \log\left(\frac{P(X)}{Q(X)}\right)} \right]\right)$  L.H.S. از قضیه است:

$$\Rightarrow Q(T(X^n) \leq nt) \leq \sup_{\lambda > 0} e^{-\lambda nt} E_{X \sim Q} \left[ e^{+\lambda \sum_{k=1}^n \log\left(\frac{P(X_k)}{Q(X_k)}\right)} \right] = \exp\left[-n \cdot \sup_{\lambda > 0} \left\{ \lambda t - \psi_Q(\lambda) \right\}\right]$$

let  $\psi_Q^*(t) = \sup_{\lambda > 0} \left\{ \lambda t - \psi_Q(\lambda) \right\}$

$$\Rightarrow Q(T(X^n) \leq nt) \leq e^{-n \psi_Q^*(t)}$$

Q.E.D.  $\blacksquare$

$$\begin{cases} \pi_{110}^{(n)} \leq e^{-n \psi_P^*(t)} \\ \pi_{011}^{(n)} \leq e^{-n \psi_Q^*(t)} \end{cases}$$

در نهایت (نتیجه) داریم که:  $\forall t \geq E_{X \sim P} \left[ \log \frac{P(X)}{Q(X)} \right] = -D_{KL}(P \| Q)$

$\forall t \leq E_{X \sim Q} \left[ \log \frac{P(X)}{Q(X)} \right] = D_{KL}(Q \| P)$



$$\textcircled{III} \rightarrow E_0(t) = \frac{1}{n} \sup_{\pi_{110}} \log \pi_{110}^{(n)} = \frac{1}{n} \log \left[ \exp(-n \psi_P^*(t)) \right] = \psi_P^*(t) \quad \text{201-201}$$

$$\rightarrow E_1(t) = \frac{1}{n} \sup_{\pi_{011}} \log \pi_{011}^{(n)} = \frac{1}{n} \log \left[ \exp(-n \psi_Q^*(t)) \right] = \psi_Q^*(t) \quad \text{(*)}$$

:  $\psi_P^*$  و  $\psi_Q^*$  و  $\psi_P^*$  و  $\psi_Q^*$  و  $\psi_P^*$  و  $\psi_Q^*$

$$\psi_Q(\lambda) = \log \left( \mathbb{E}_{x \sim q} \left[ e^{\lambda T(x)} \right] \right) = \log \left( \sum_x e^{\lambda \log \frac{q(x)}{p(x)}} \cdot q(x) \right) = \log \left( \sum_x e^{\lambda \log \frac{q(x)}{p(x)}} \cdot q(x) \right)$$

$$T(x) = \log \frac{q(x)}{p(x)}$$

→ change of measure:

$$\rightarrow \psi_Q(\lambda) = \log \left( \sum_x e^{(\lambda+1) \log \frac{q(x)}{p(x)}} p(x) \right) = \mathbb{E}_{x \sim p} \left[ e^{(\lambda+1) T(x)} \right] = \psi_P(\lambda+1) = \psi_Q(\lambda) \quad \text{(*)}$$

$$\Rightarrow \psi_Q^*(\lambda) = \sup_{\lambda \geq 0} \{ \lambda t - \psi_Q(\lambda) \} = \sup_{\lambda \geq 0} \{ \lambda t - \psi_P(\lambda+1) \} = \sup_{\lambda \geq 0} \{ (\lambda+1)t - \psi_P(\lambda+1) \} - t$$

$$\Rightarrow \psi_Q^*(\lambda) = \psi_P^*(t) - t$$

:  $\psi_P^*$  و  $\psi_Q^*$  و  $\psi_P^*$  و  $\psi_Q^*$  و  $\psi_P^*$  و  $\psi_Q^*$

$$\begin{cases} E_1(t) = \psi_Q^*(t) = \psi_P^*(t) - t \\ E_0(t) = \psi_P^*(t) \end{cases}$$

$$: \forall D_{KL}(P||Q) \leq t \leq D_{KL}(Q||P)$$

Q.E.D. ■

IV

$$\begin{cases} \epsilon_0 = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\pi_{110}^{(n)}} \leq \inf_{Q: E_Q(X) > \tau} \{ D_{KL}(Q \| P) \} \\ \epsilon_1 = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\pi_{011}^{(n)}} \leq \inf_{Q: E_Q(X) < \tau} \{ D_{KL}(Q \| P) \} \end{cases}$$

مثل می بینیم اگر  $\pi_{110}$  و  $\pi_{011}$  کراخ بالا از بین و کراخ converse را به دست می آوریم که اگر  $\epsilon_0, \epsilon_1$  بیشتر به ترتیب شوند می بینیم که  $\frac{1}{\pi_{110}^{(n)}}$  و  $\frac{1}{\pi_{011}^{(n)}}$  را به دست می آوریم:

$$\log \left( \frac{1}{P(X \in E)} \right) = D_{KL}(P_{X|X \in E} \| P_X) \quad \text{let } \hat{P} = P_{X^n} | \sum X_i > n\tau$$

$$\rightarrow \log \left[ \frac{1}{\pi_{110}^{(n)}} \right] = D_{KL}(\hat{P}_{X^n} \| P_{X^n}) \geq \inf_{Q_{X^n}: E_Q(\sum X_i) > n\tau} \{ D(Q_{X^n} \| P_{X^n}) \} \geq \sum_{j=1}^n D(Q_{X_j} \| P) = n D(\bar{Q} \| P)$$

$X_i$ 's are iid, KL properties

Q.E.D

$$\Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\pi_{110}^{(n)}} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\pi_{110}^{(n)}} \geq \inf_{Q: E_Q(X) > \tau} \frac{1}{n} D_{KL}(Q \| P) = \psi_P^*(t) = \epsilon_0(t)$$

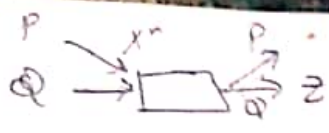
به طریق مشابه برای  $\epsilon_1(t)$  به دست می آید که:

$$\log \frac{1}{P(X \in E)} = D_{KL}(P_{X|X \in E})$$

$$\rightarrow \log \left[ \frac{1}{\pi_{011}^{(n)}} \right] = D_{KL}(\hat{P}_{X^n} \| P_{X^n}) \geq \inf_{Q_{X^n}: E_Q(\sum X_i) < n\tau} \{ D(Q_{X^n} \| P_{X^n}) \} \geq \sum_{j=1}^n D(Q_{X_j} \| P) = n D(\bar{Q} \| P)$$

$$\Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\pi_{011}^{(n)}} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\pi_{011}^{(n)}} \geq \inf_{Q: E_Q(X) < \tau} \{ D_{KL}(Q \| P) \} = \psi_P^*(H) = \psi_P^*(t) - t$$

در نتیجه می توانیم نتیجه بگیریم که  $\epsilon_0$  و  $\epsilon_1$  همان مقادیر  $\psi_P^*(t)$  و  $\psi_P^*(H)$  هستند. converse Q.E.D. است.


 چون نسیم) نتیاً منظور ال بیت مست = Achievability (یا نیز اثبات می بینیم):  
 امکان دارد منظور ال این بیت است  
 (در من بیت قبل از این نوشتیم)

$$\pi_{110} = \mathbb{P}\left[\sum_{k=1}^n X_k > n\gamma\right] = 1 - o(1)$$

$X_i$  i.i.d

$$\xrightarrow{\text{DPI}} D_{kl}(P_{X_1-X_n} \| Q_{X_1-X_n}) \geq D_{kl}(P_Z \| Q_Z)$$

$$\begin{cases} Z \sim \text{Bern}(\epsilon) \\ Q|Z=0 = \pi_1 = 1-\epsilon \\ P(Z=0) = \pi_0 \approx e^{-n\epsilon} \end{cases}$$

$$D_{kl}(P_Z \| Q_Z) = P(Z=0) \log\left(\frac{P(Z=0)}{Q(Z=0)}\right) + P(Z=1) \log\left(\frac{P(Z=1)}{Q(Z=1)}\right)$$

$$= \left[(-\epsilon) \log(1-\epsilon) + \epsilon \log \epsilon\right] + \epsilon \log \frac{1}{1-e^{-n\epsilon}} + (1-\epsilon)n\epsilon$$

$$\approx -H_b(\epsilon) + (1-\epsilon)n\epsilon$$

$$\epsilon_1 = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\pi_{110}} \leq \lim_{n \rightarrow \infty} \left( \frac{D(Q_Z \| P_Z) + o(1/n)}{1-\epsilon} \right) \xrightarrow{\text{DPI}} \frac{D(Q_X \| P_X)}{1-\epsilon} \approx D_{kl}(Q_X \| P_X)$$

حال اگر  $Q$  و  $P$  را  $Q_{X^n}$  و  $P_{X^n}$  بگیریم:

$$\epsilon_1 = \lim_{n \rightarrow \infty} \sup_{Q: E_Q(X) > \gamma} \frac{1}{n \pi_{110}} \leq \inf_{Q: E_Q(X) > \gamma} \left\{ D_{kl}(Q_{X^n} \| P_{X^n}) \right\} \quad \text{و.ع.و.}$$

Achievability

④

الامثلة: اكتب  $\pi_0, \pi_1$  في  $\mathbb{R}^n$  بحيث يكونان متعامدين.

$$; e = \frac{1}{n} \log(P_c)$$

$$P_{\mathcal{C}} = \pi_0 \pi_{1|0} + \pi_1 \pi_{0|1} \geq \min \{ \pi_{q|1}, \pi_{1|0} \}$$

حال نزدیک به غیر هاله (میر) فایم است: (معدلاً)  $\pi_{140} \sim \pi_{110}$  و  $\pi_{110} \sim \pi_{110}$  و  $\pi_{110} \sim \pi_{110}$  که حالت ابدال

$$\varepsilon = \sup_t \inf_{\pi_0, \pi_1, \dots} \{P_e\} = \sup_t \left\{ \min \{E_1, \dots, E_0\} \right\}$$

نست می نمود که تقریباً از ۷۰ سالخ تقریب اولی است که در سوال در ۱۰۰ می بود و در ۱۰۰ می

از ضرب بگیر در سمت چپ ثابت کردم که:

$$-D_{\alpha_1}(P \| Q) \leq t \leq D_{\alpha_2}(Q \| P)$$

$$\begin{cases} \varepsilon_1(t) = \psi_p^*(t) \\ \varepsilon_0(t) = \psi_a^*(t) = \psi_p^*(t) - t \end{cases} \Rightarrow \begin{cases} \varepsilon_0 > \varepsilon_1 & \text{if } t > 0 \\ \varepsilon_0 < \varepsilon_1 & \text{if } t < 0 \end{cases}$$

کہ نبی کے لئے

$$E = \sup_t \begin{cases} \psi^+(t) & t > 0 \\ \psi^-(t) - t & t < 0 \end{cases}$$

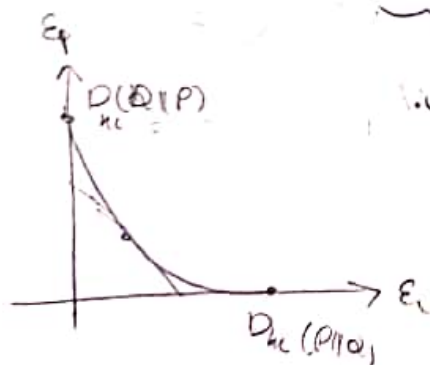
اینجا بار  $t=0$  فله اقل کرده و نشان می‌دهیم که  $t=0$  جواب میده پسندار است و به همین ترتیب

$\varphi_p^{(0)} = \varphi_p^{(1)}|_{t=0} = \varphi_p^{(1)}(t) - t|_{t=0}$  که  $t=0$  از  $t$  جدا (که)  
 و در نتیجه در هر دو حالت  $\varphi_p^{(0)}(0) = 0$

جواب کیسے لکھتے۔ انہی

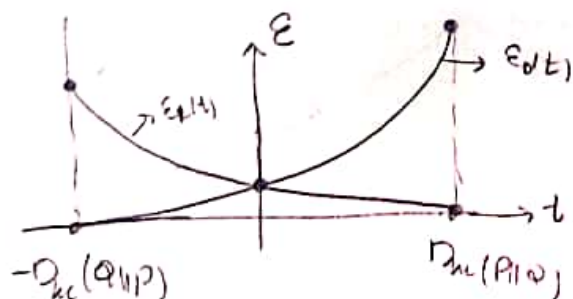
نشان شود که در اینم  $\begin{cases} t < 0 \\ t > 0 \end{cases}$  بر حسب است و برابر  $\begin{cases} t > 0 \\ t < 0 \end{cases}$  نزدیک است

در نتیجه می‌توانیم  $E(t)$  را به صورت  $E(t) = E_1(t) + E_2(t)$  بنویسیم.



۱. شرط پیر بودن در  $t=0$  :  $\Psi_Q(t)|_{t=0} = \Psi_P(t)|_{t=0}$

کے لیے ہم نے  $\min$  کا  $t=0$  پر  $u$  کا



$$\Rightarrow \epsilon = \max_t \min_{(E_0, E_1) \in \mathcal{E}} \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{p_{E_1}^{(n)}} \right\} = \psi_p^{\text{dc}}(0)$$

سوال 3: ابتدا نقاط را بصورت یکدست نمی‌کنیم. در کل  $2^n$  نقطه داریم و احتمال در یک  $\frac{1}{2^n}$  است. همچنین احتمال اینکه به نام بردار مقدار  $X$  باشد  $\frac{1}{2}$  است. حال می‌توانیم  $f(r_0, n)$  را بر حسب احتمال بنویسیم. داریم:

$$f(r_0, n) = 2^n \times \mathbb{P}[\|X\|_2 \geq r_0 \sqrt{n}] = 2^n \times \mathbb{P}[\|X\|_2^2 \geq n r_0^2] \leq \frac{2^n}{n r_0^2} \mathbb{E}[\|X\|_2^2]$$

$$\mathbb{E}[\|X\|_2^2] = \mathbb{E}\left[\sum_{i=1}^n X_i^2\right] = n \mathbb{E}[X_i^2] = \frac{5n}{2} \rightarrow f(r_0, n) \leq \frac{2^n}{n r_0^2} \times \frac{5n}{2} = \frac{2^n \cdot 5}{2 r_0^2}$$

$$\exp(f(r_0, n)) = \lim_{n \rightarrow \infty} \frac{1}{n} \log f(r_0, n) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \left( n + \log\left(\frac{5}{2 r_0^2}\right) \right) = 1 + \lim_{n \rightarrow \infty} \frac{1}{n} \log\left(\frac{5}{2 r_0^2}\right)$$

حال اگر  $r_0^* = \alpha(r_0, n)$  باشد و در این صورت  $f(r_0^*, \sqrt{n})$  را می‌توانیم بنویسیم.

$$f(\alpha(\sqrt{n}), n) \leq 1 + \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{5}{2n} = 1 - \lim_{n \rightarrow \infty} \frac{\log n}{n} = 1$$

از طرف دیگر چون  $1 = f(r_0, n)$  داریم به تبع  
 اگر  $r_0 < r_0^*$  :  $\exp[f(r_0, n)] = 1$  همچنین برای  $r_0 > r_0^*$  چون که  $r_0$  مقداری بزرگتر از  $r_0^*$  است پس  $0(r_0) > 0(\sqrt{n})$  در این صورت  $\frac{1}{n} \log\left(\frac{5}{2 r_0^2}\right)$  منفی می‌شود و در این صورت exponent  $\leq 1$  است.

$$\forall r_0 > r_0^* : \exp[f(r_0, n)] \leq 1$$



40: از تقسیم می‌توانیم به دست آوریم که  $P_2(x)$  را کمینه کردیم و باید یقین است و باید به دست بیاوریم  $P_2(x)$  را جدا کرد: ثابت.

(I)

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

$$\min_{A^{(n)} \subseteq X^n} P_E^{(n)} = \pi_1 + \min_{A^{(n)} \subseteq X^n} \{ \pi_2 P_2^{(n)}(A^{(n)}) - \pi_1 P_1^{(n)}(A^{(n)}) \} = \pi_1 + \min_{A^{(n)} \subseteq X^n} \left\{ \sum_{x \in X^n} \pi_2 P_2^{(n)}(x) - \pi_1 P_1^{(n)}(x) \right\} \times \mathbb{1}(x \in A^{(n)})$$

که مثلاً بدیم مثلاً  $A_*^{(n)} = \{x \mid \pi_1 P_1^{(n)}(x) \geq \pi_2 P_2^{(n)}(x)\}$  به سببی است:

$$\min_{A^{(n)} \subseteq X^n} P_E^{(n)} = (\pi_1 - \pi_1 P_1(A_*^{(n)})) + \pi_2 P_2(A_*^{(n)}) = \pi_1 P_1^{(n)} \left[ \frac{P_1(x)}{P_2(x)} < \frac{\pi_2}{\pi_1} \right] + \pi_2 P_2^{(n)} \left[ \frac{P_1(x)}{P_2(x)} > \frac{\pi_2}{\pi_1} \right]$$

$$\begin{cases} \psi_P^*(\lambda) = \inf_{\lambda \in \mathbb{R}} \{ -\psi_P(\lambda) \} = -\inf_{\lambda \in \mathbb{R}} \psi_P(\lambda) \\ \psi_{P_2}^*(\lambda) = -\inf_{\lambda \in \mathbb{R}} \psi_{P_2}(\lambda) \end{cases} \quad \psi_Q^*(\lambda) = \psi_P^*(\lambda) = \min_{P_2(x)} P_E \quad \text{که مثال 2 بدیم}$$

6: می‌توانیم از این تفاوت خود را جدا کنیم؛ (بدیم)

$$\begin{cases} \sup_{\lambda} \psi_P(\lambda) = \sup_{\lambda} -\log \mathbb{E}_P[e^{\lambda T}] = -\log \left[ \sup_{\lambda} \mathbb{E}_P[e^{\lambda T}] \right] \\ \frac{d}{d\lambda} \psi_P(\lambda) = \frac{d}{d\lambda} \log \frac{Q(x)}{P(x)} \end{cases} \quad \begin{aligned} &\text{اما نه از هر دو به دست می‌آید که تفاوت} \\ &\text{نسبت است و تمام نیز جمع فرق و تفاوت} \\ &\text{نسبت است پس از مشتق گرفتن بدیم:} \end{aligned}$$

$$\frac{d}{d\lambda} \psi_P(\lambda) = \frac{d}{d\lambda} \mathbb{E}_P \left[ e^{\lambda \log \frac{Q(x)}{P(x)}} \right] = \left[ \sum_x P(x) e^{\lambda \log \frac{Q(x)}{P(x)}} \left( \log \frac{Q(x)}{P(x)} \right) \right]$$

$$= \sum_x P(x) \cdot \log \left( \frac{Q(x)}{P(x)} \right) \cdot \left( \frac{Q(x)}{P(x)} \right)^\lambda = 0 \quad \text{فرض کنیم بر } \sum_{y \in X} P_1^{1-\lambda}(y) P_2^{1-\lambda}(y) :$$

$$\text{let } P_{\lambda}(x) = \frac{P_1(x)^{\lambda} P_2(x)^{1-\lambda}}{\sum_{y \in X} P_1(y)^{\lambda} P_2(y)^{1-\lambda}} \quad \sum_x \frac{P(x)^{1-\lambda} Q(x)^{\lambda} \left( \log \frac{Q(x)}{P(x)} \right)}{\sum_{y \in X} P_1(y)^{\lambda} P_2(y)^{1-\lambda}} = 0$$

$$\Downarrow \quad \mathbb{E}_{x \sim P_{\lambda}} [T] = 0$$

$$D_{KL}(P_{\lambda} \| Q) = D_{KL}(P_{\lambda} \| P) \quad \text{نویسه داشته باشیم:} \quad \mathbb{E}_{x \sim P_{\lambda}} \left[ \lambda \log \frac{Q(x)}{P(x)} \cdot \frac{dQ}{dP} \right] = 0$$

حال اگر  $P$  و  $Q$  را  $P_1$  و  $P_2$  بگیریم، داریم هر دو طبق به هم نزدیک شدن به هم می‌روند که اگر بدیم که:

$$\exists \lambda^* = \arg \max_{\lambda} \{ \psi_P^*(\lambda) \} = \arg \max_{\lambda} \{ \psi_{P_2}^*(\lambda) \}$$

$$0^* = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{P_c^{(n)}} = \inf_{\lambda} \{ \psi_P^*(\lambda) \} = \inf_{\lambda} \{ \psi_{P_2}^*(\lambda) \} = D_{KL}(P_1 \| P_1) = D_{KL}(P_1 \| P_2)$$

$\Rightarrow$  2. E. D.

II

$$C(P_1, P_2) = \min_{0 \leq \lambda \leq 1} -\log \left[ \mathbb{E}_{X \sim P_2} \left[ \left( \frac{P_1(X)}{P_2(X)} \right)^\lambda \right] \right]$$

می دانیم که برای  $0.5 \leq \lambda$ ،  $P_1^\lambda P_2^{1-\lambda}$  ممیانس است. در نتیجه چون  $\lambda = 1$  نیز ممیانس است و ممیانس را می توانیم به این نتیجه می رسیدیم که  $C(P_1, P_2)$  نسبت به  $\lambda$  ممیانس است و این را می توانیم ثابت کنیم:

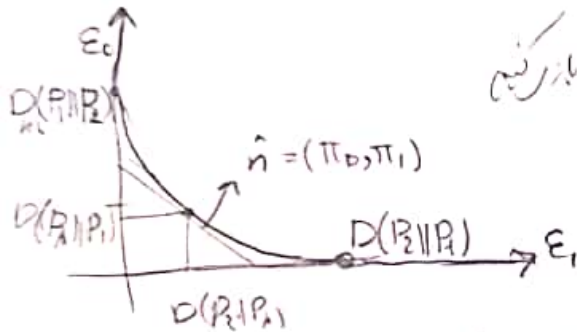
$$\frac{\partial C(\lambda)}{\partial \lambda} = \frac{\mathbb{E}_{X \sim P_2} \left[ \log \left( \frac{P_1(X)}{P_2(X)} \right) \left( \frac{P_1(X)}{P_2(X)} \right)^\lambda \right]}{\mathbb{E}_{X \sim P_2} \left[ \left( \frac{P_1(X)}{P_2(X)} \right)^\lambda \right]} = 0 \Rightarrow \sum_x P_1(x)^\lambda P_2(x)^{1-\lambda} \log \left( \frac{P_1(x)}{P_2(x)} \right) = 0$$

$$0 = \mathbb{E}_{X \sim P_2} \left[ \log \left( \frac{P_1(X)}{P_2(X)} \right) \left( \frac{P_1(X)}{P_2(X)} \right)^\lambda \right]$$

ا در کتب قبل می بینیم که این فاصله را ثابت کردیم:

$$\left( \frac{P_1(x)}{P_2(x)} \right)^\lambda = e^{\lambda \log \frac{P_1(x)}{P_2(x)}} \quad (*)$$

$$\mathbb{E}_{X \sim P_2} \left[ \log \left( \frac{P_1(X)}{P_2(X)} \right) \left( \frac{P_1(X)}{P_2(X)} \right)^\lambda \right] = 0$$



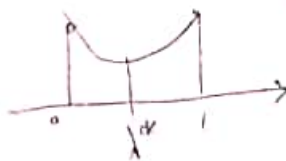
همچون هر یک از  $P_1$  و  $P_2$  ممیانس است و ما می توانیم در  $x \in \mathcal{X}$  به این نتیجه می رسیدیم که  $P_1 = P_2$  و این را می توانیم ثابت کنیم که  $P_1 = P_2$  و این را می توانیم ثابت کنیم:

$$\begin{cases} D^* = D_{KL}(P_1^* || P_1) = D_{KL}(P_1^* || P_2) = C(P_1, P_2) \\ C(P_1, P_2) = \min_{\lambda \in [0,1]} -\log \left[ \mathbb{E}_{X \sim P_2} \left( \frac{dP_1}{dP_2} \right)^\lambda \right] \end{cases}$$

Q.E.D.

وقت می بینیم که  $\lambda = 0$  یا  $\lambda = 1$  نیز نقطه ممیانس است و این را می توانیم ثابت کنیم که  $\lambda = 0$  یا  $\lambda = 1$  نیز نقطه ممیانس است و این را می توانیم ثابت کنیم:

$$C(P_1, P_2) \Big|_{\lambda=0}^{\lambda=1} = 0$$



i)  $X \sim \text{Bin}(n, p)$   
 $X_1, \dots, X_n \stackrel{iid}{\sim} p$   $\left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{P\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \gamma\right)} = \inf_{Q: E(X) \geq \gamma} \left\{ D_{KL}(P \parallel Q) \right\} \right.$

طبق قضیه  
 Sanov برای اینجمله:  
 در نتیجه داریم:

$$P\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \gamma\right) \leq \exp\left(-n \cdot \inf_{Q: E(X) \geq \gamma} \left\{ D_{KL}(P \parallel Q) \right\}\right)$$

در نتیجه اگر  $X_i \sim \text{Ber}(p)$  باشد در این صورت  $Y = \sum_{i=1}^n X_i \sim \text{Bin}(n, p)$  است.  
 که قضیه در مورد کر به  $P\left[\sum_{i=1}^n X_i \geq k\right]$  می باشد به عبارتی  $Y = k$  می باشد و داریم:

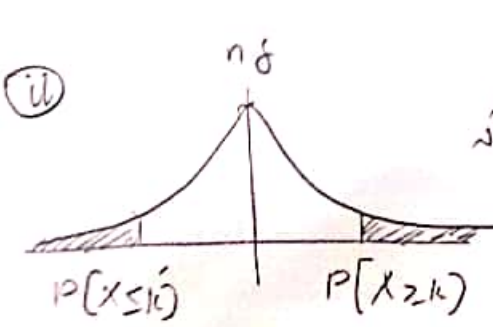
$$P\left[\sum_{i=1}^n X_i \geq k\right] = P\left[\frac{1}{n} \sum_{i=1}^n X_i \geq \frac{k}{n}\right] \leq \exp\left(-n \cdot \inf_{Q: E(X) \geq \frac{k}{n}} \left\{ D_{KL}(Q, P) \right\}\right)$$

در در این بین که  $E_Q(X) = \frac{k}{n}$  می باشد

$\left\{ \begin{array}{l} Q \sim \text{Ber}(q) \\ P \sim \text{Ber}(p) \end{array} \right.$   
 $\hookrightarrow D_{KL}(P \parallel Q) = d(P \parallel q) \rightarrow \inf_{q \geq \frac{k}{n}} q \log \frac{q}{p} + (1-q) \log \frac{1-q}{1-p}$

$\frac{\partial}{\partial q} d(P, q) = 1 + \log q - \log p + \log(1-p) - 1 - \log(1-q) = 0 \rightarrow \text{either } \frac{q}{1-q} = \frac{p}{1-p}$   
 $\rightarrow q=1 \text{ یا } q=p$   
 در نتیجه  $q \neq p$  و  $q=1$  پس به ازای  $q=0$  می باشد. در نتیجه داریم:

$$P\left[\sum_{i=1}^n X_i \geq k\right] \leq \exp\left(-n \cdot D_{KL}\left(\text{Ber}\left(\frac{k}{n}\right) \parallel \text{Ber}(p)\right)\right) \quad \forall k \geq EY = np$$



حال از نمودارهای بالا بزرگ و دقیق تر استفاده می کنیم:  
 اگر به توزیع  $Y$  وقتی که  $n \rightarrow \infty$  نگاه کنیم، چون نسبت به وسط آن که  $np$  است متقارن است  $P(X \geq k) = \lim_{n \rightarrow \infty} P(X \leq k')$   
 $k' = np + k$

$n \rightarrow \infty$   
 $\Rightarrow P[X \geq k] = P[X' \leq k'] \leq \exp\left(-n d\left(\frac{k}{n} \parallel P\right)\right) = \exp\left(-n D_{KL}\left(\text{Ber}\left(\frac{k'}{n}\right) \parallel \text{Ber}(p)\right)\right)$   
 $\forall k' \leq EY = np$



② II

الامثلة 5:  $d(\frac{k}{n} || P)$  اگر  $n$  است قبل از اینکه  $n$  را تغییر دهیم.

$$P(X \geq k) \leq \exp \left[ -n \cdot \left( \frac{k}{n} \log \frac{k}{np} + \frac{n-k}{n} \log \left( \frac{n-k}{n-p} \right) \right) \right]$$

now since

$$\forall u, v \in [0, 1]^2 : x \log \frac{x}{y} \geq x - y \Rightarrow \left. \begin{array}{l} \text{let } k = u \mathbb{E}[X] = unp \\ \text{if } k < np \Rightarrow (u < 1) \end{array} \right\} \frac{k}{n} = up$$

$$d\left(\frac{k}{n} || P\right) = d(up || P) = up \log u + (1-up) \log \left( \frac{1-up}{1-p} \right)$$

$$\text{now let } \left. \begin{array}{l} x = 1-up \\ y = 1-p \end{array} \right\} d\left(\frac{k}{n} || P\right) \geq up \log(u) + [1-up - (1-p)] = up \log u + p(1-u)$$

$$d\left(\frac{k}{n} || P\right) = P \left[ \underbrace{u \log u + (1-u)}_{f(u)} \right] = P f(u) \quad \left. \begin{array}{l} \mathbb{E}X = np \end{array} \right\} n d\left(\frac{k}{n} || P\right) \geq np f(u) = f(u) \mathbb{E}X$$

$$\Rightarrow P(X \geq k) \leq \exp(-n d\left(\frac{k}{n} || P\right)) \leq \exp(-n \cdot f(u) \mathbb{E}X) \quad (0 < u < 1)$$

$$\lim_{t \rightarrow \infty} P(X \leq k) \leq \exp(-n d\left(\frac{k}{n} || P\right)) \leq \exp(-n \cdot f(u) \mathbb{E}X)$$

$$u > 1 \quad \left\{ \begin{array}{l} k = u \mathbb{E}X = unp \\ \mathbb{E}X \geq np \end{array} \right.$$

Q.E.D.  $\left\{ \begin{array}{l} P(X \geq k) \leq \exp(-n f(u) \mathbb{E}X) : 1 > u > p \\ P(X \leq k) \leq \exp(-n f(u) \mathbb{E}X) : u > 1 \end{array} \right.$  پس به دست آمد که

$$f(u) = \int_1^u \frac{u-n}{n} du \geq \int_1^u \frac{u-x}{u} dx = \frac{1}{u} \int_1^u (u-n) dx = \frac{(u-1)^2}{2u} \quad \text{Q.E.D.}$$

حال بقیه  $f(u) \leq \frac{(u-1)^2}{2u}$  و قسمت قبل را بنویسیم

$$P(X \geq u \mathbb{E}X) \leq \exp(-f(u) \mathbb{E}X) \leq \exp\left(-\frac{(u-1)^2}{2u} \mathbb{E}X\right) \rightarrow u > 1 \Rightarrow \hat{u} = \frac{u}{\mathbb{E}X}$$

$$\text{now let } u = \mathbb{E}X + t = np + t$$

$$P(X \geq t + np) \leq \exp\left(-\frac{(t/\mathbb{E}X)^2}{2 \cdot \frac{(\mathbb{E}X + t)}{\mathbb{E}X}}\right) = \exp\left(-\frac{t^2}{2(\mathbb{E}X + t)}\right)$$

$$\mathbb{E}X = np \rightarrow P(X \geq np + t) \leq \exp\left(-\frac{t^2}{2(np + t)}\right) \quad \text{Q.E.D.}$$



(IV)  $D_n(P||Q) = \sum_n p_{(n)} \log \frac{p_{(n)}}{q_{(n)}} = 2 \sum_n p_{(n)} \log \sqrt{\frac{p_{(n)}}{q_{(n)}}} = -2 \sum_n p_{(n)} \log \sqrt{\frac{q_{(n)}}{p_{(n)}}}$  250/2 ~ 125

since  $(\log u \geq u-1) \forall u \Rightarrow D_n(P||Q) \geq 2 \sum_n p_{(n)} \left[ 1 - \sqrt{\frac{q_{(n)}}{p_{(n)}}} \right]$   
 $u = \frac{q_{(n)}}{p_{(n)}} \longrightarrow q \ll p$  از 125

$\Rightarrow D_n(P||Q) \geq 2 \sum_n \sqrt{p_{(n)}} (\sqrt{p_{(n)}} - \sqrt{q_{(n)}}) \geq 2 \sum_n (\sqrt{p_{(n)}} - \sqrt{q_{(n)}}) (\sqrt{p_{(n)}} - \sqrt{q_{(n)}})^2$

①  $\sqrt{p_{(n)}} \geq \sqrt{p_{(n)}} - \sqrt{q_{(n)}}$   
 ②  $u < 1 \Rightarrow 2u \geq u$

$\Rightarrow D_n(P||Q) \geq \sum_n (\sqrt{p_{(n)}} - \sqrt{q_{(n)}})^2 \Rightarrow \left. \begin{matrix} P \sim \text{Ber}(p) \\ Q \sim \text{Ber}(q) \end{matrix} \right\} d(P||Q) \geq (\sqrt{p} - \sqrt{q})^2$

let  $q = k/n \Rightarrow d(\frac{k}{n} || p) \geq (\sqrt{\frac{k}{n}} - \sqrt{p})^2$

$\begin{cases} IP[X \geq k] \leq \exp(-n d(\frac{k}{n} || p)) \\ IP[X \leq k] \leq \exp(-n d(\frac{k}{n} || p)) \end{cases}$  اولی نتیجه I است  
 $\sqrt{k} = t + \mathbb{E}\sqrt{X} = \sqrt{np} + t$  دوم نتیجه است  
 $t > 0 \Rightarrow \sqrt{k} > \sqrt{np}$

(i)  $IP[\sqrt{X} \geq \sqrt{k}] = IP[\sqrt{X} - \sqrt{np} \geq t] \leq \exp(-n d(\frac{k}{n} || p)) \leq \exp(-n (\sqrt{\frac{k}{n}} - \sqrt{p})^2)$

$\Rightarrow IP[X \geq k] \leq \exp(-n (\frac{\sqrt{np} + t}{\sqrt{n}} - \sqrt{p})^2) = \exp(-\frac{nt^2}{n}) = \exp(-t^2)$  Q.E.D.

(ii)  $IP[X \leq k] = IP[\sqrt{X} \leq \sqrt{k}]$  ثالث نتیجه است  
 $\sqrt{k} = -t + \mathbb{E}\sqrt{X} = -t + \sqrt{np}$   
 $\sqrt{k} < \sqrt{np}$   
 $t > 0$   
 $\Rightarrow IP[\sqrt{X} - \sqrt{np} \leq -t] \leq \exp(-n d(\frac{k}{n} || p))$   
 $\leq \exp(-n (\sqrt{\frac{k}{n}} - \sqrt{p})^2) = \exp(-\frac{nt^2}{n}) = \exp(-t^2)$  Q.E.D.

$\begin{cases} IP[\sqrt{X} - \sqrt{np} \geq t] \leq e^{-t^2} \\ IP[\sqrt{X} - \sqrt{np} \leq -t] \leq e^{-t^2} \end{cases} \quad t > 0$  نتیجه نهایی

سوال 6: فرض کنیم که  $f^*$ ، بیشینه کننده  $V_{\alpha, \beta, r, s, t}$  باشد. داریم:  $\chi^{\text{IR}} \Rightarrow (P_x \ll Q_x) : (P_x \ll Q_x) \text{ از } \chi^{\text{IR}} \text{ به } \mathbb{R}$

$$\sup_{f \in \chi^{\text{IR}}} V_{\alpha, \beta, r, s, t}(P_x \ll Q_x) = V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^*} \geq V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_f : \forall f \in \chi^{\text{IR}}$$

حال می‌دانیم که در (۱)  $V_{\alpha, \beta, r, s, t}$  تابعی است مابین  
(- عددیت بگیر کیفیت عددی تاثیر نداشته باشد.)  $V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^*} = V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^* + c}$

$$\begin{aligned} V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^*} &= \mathbb{E}_{P_x}[f^*(x) + c] - r \mathbb{E}_{Q_x}[f^*(x) + c] - s \log(\mathbb{E}_{P_x}[\exp(\alpha f^*(x) + \alpha c)]) - t \log(\mathbb{E}_{Q_x}[\exp(\beta f^*(x) + \beta c)]) \\ &= c + \mathbb{E}_{P_x}[f^*(x)] - rc - r \mathbb{E}_{Q_x}[f^*(x)] - s \alpha - s \log(\mathbb{E}_{P_x}[\exp(\alpha f^*(x))]) - t \beta - t \log(\mathbb{E}_{Q_x}[\exp(\beta f^*(x))]) \\ &= V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^*} + c(1 - r - \alpha s - \beta t) \end{aligned}$$

حال اگر  $c = 0$ ،  $V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^*} = \mathbb{E}_{P_x}[f^*(x)] - r \mathbb{E}_{Q_x}[f^*(x)] - s \alpha - t \beta$ ، در این صورت،  $c(1 - r - \alpha s - \beta t)$  را نیز می‌توان به عنوان کارزار که  $V_{\alpha, \beta, r, s, t} = 0$ ، در این صورت از آن فواید قوی نیست.

$$V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^*} = \begin{cases} \alpha & \text{if } 1 \neq r + \alpha s + \beta t \\ V_{\alpha, \beta, r, s, t}(P_x \ll Q_x)|_{f^* + c} & \text{if } r + \alpha s + \beta t = 1 \end{cases}$$

\*  $\text{A}$  می‌دانیم که تابع  $\exp$  آنقدر آهسته است در مرتبه  $\mathbb{E} \exp(x) \geq \exp(\mathbb{E} x)$ ؛ بنابراین تابع  $\exp$  را می‌توانیم در هر توابع دیگر قرار دهیم که مثبت باشد.

$$V_{\alpha, \beta, r, s, t}(P_x \ll Q_x) = \sup_{f \in \chi^{\text{IR}}} V_{\alpha, \beta, r, s, t}(P_x \ll Q_x) \geq 0$$

$$\begin{aligned} &\geq \sup_{f \in \chi^{\text{IR}}} \left\{ \mathbb{E}_{P_x}[f(x)] - r \mathbb{E}_{Q_x}[f(x)] - s \log(\exp(\alpha \mathbb{E}_{P_x} f(x))) - t \log(\exp(\beta \mathbb{E}_{Q_x} f(x))) \right\} \\ &= \sup_{f \in \chi^{\text{IR}}} \left\{ \mathbb{E}_{P_x}[f(x)] - \mathbb{E}_{Q_x}[f(x)](r + \alpha s + \beta t) \right\} \geq c(1 - (\alpha s + \beta t + r)) = 0 \end{aligned}$$

$$\sup_{f \in \chi^{\text{IR}}} \mathcal{H}(P_x, Q_x; f) \geq \mathcal{H}(P_x, Q_x; f(x) = c)$$

$$\Rightarrow V_{\alpha, \beta, r, s, t}(P_x \ll Q_x) \geq 0$$

Q.E.D

III:  $V_{\alpha, \beta, r, s, t}(P_X \| Q_X) \geq 0$  (چون  $Q_X = P_X$ )

$$V_{\alpha, \beta, r, s, t} = \sup_{f \in \mathcal{X}^{\mathbb{R}}} \left\{ \mathbb{E}_P[f(X)](r+1) - s \log(\mathbb{E}_{P_X}[e^{\alpha f(X)}]) - t \log(\mathbb{E}_{P_X}[e^{\beta f(X)}]) \right\}$$

فرض کنیم

$$\leq \mathbb{E}_{P_X}[f(X)](1-r-\alpha s-\beta t) = 0 \quad (\text{چون } \alpha + \beta + r = 1)$$

Q.E.D

$$\rightarrow (V_{\alpha, \beta, r, s, t}(P_X \| P_X) \geq 0) \wedge (V_{\alpha, \beta, r, s, t}(P_X \| P_X) \leq 0) \Rightarrow V_{\alpha, \beta, r, s, t}(P_X \| P_X) = 0$$

IV:  $f \in \mathcal{X} \times \mathcal{Y}$

ابتدا معین می‌کنیم  $\alpha, \beta, r, s, t$  را به گونه‌ای که  $\alpha + \beta + r = 1$  و  $s, t \geq 0$  باشند.

$$V_{\alpha, \beta, r, s, t}(P_{X,Y} \| Q_{X,Y}) = \sup_{f \in \mathcal{X} \times \mathcal{Y}} \left\{ \mathbb{E}_{P_{X,Y}}[f(X,Y)] - r \mathbb{E}_{Q_{X,Y}}[f(X,Y)] - s \log(\mathbb{E}_{Q_{X,Y}}[e^{\alpha f(X,Y)}]) - t \log(\mathbb{E}_{Q_{X,Y}}[e^{\beta f(X,Y)}]) \right\}$$

$$= \sup_{f \in \mathcal{X} \times \mathcal{Y}} \left\{ \mathbb{E}_P[\mathbb{E}_{P_{Y|X}}[f(X,Y)]] - r \mathbb{E}_Q[\mathbb{E}_{Q_{Y|X}}[f(X,Y)]] - s \log(\mathbb{E}_P[\mathbb{E}_{Q_{Y|X}}[e^{\alpha f(X,Y)}]]) - t \log(\mathbb{E}_Q[\mathbb{E}_{Q_{Y|X}}[e^{\beta f(X,Y)}]]) \right\}$$

$$\geq \sup_{f(X,Y) = g(X) + c} \left\{ \mathbb{E}_P[\mathbb{E}_{P_{Y|X}}[g(X) + c]] - r \mathbb{E}_Q[\mathbb{E}_{Q_{Y|X}}[g(X) + c]] - s \log(\mathbb{E}_P[\mathbb{E}_{Q_{Y|X}}[e^{\alpha(g(X) + c)}]]) - t \log(\mathbb{E}_Q[\mathbb{E}_{Q_{Y|X}}[e^{\beta(g(X) + c)}]]) \right\}$$

$$\geq \sup_{g \in \mathcal{X}^{\mathbb{R}}} \left\{ \mathbb{E}_P[g(X) + c] - r \mathbb{E}_Q[g(X) + c] - s \log(\mathbb{E}_P[e^{\alpha g(X)}]) - s \log(e^{\alpha c}) - t \log(\mathbb{E}_Q[e^{\beta g(X)}]) - t \log(e^{\beta c}) \right\}$$

$$= V_{\alpha, \beta, r, s, t}(P_X \| Q_X) + c(1-r-\alpha s-\beta t) = V_{\alpha, \beta, r, s, t}(P_X \| Q_X)$$

در نهایت ثابت شد که:

$$V_{\alpha, \beta, r, s, t}(P_{X,Y} \| Q_{X,Y}) \geq V_{\alpha, \beta, r, s, t}(P_X \| Q_X)$$

که نتیجه دیگر از این است که  $\alpha, \beta, r, s, t$  را به گونه‌ای می‌توانیم انتخاب کنیم که  $\alpha + \beta + r = 1$  و  $s, t \geq 0$  باشند.

Q.E.D



ادرس 60: بهارشی که در سمت III استفاده کردیم متولد شدیم.

از سمت قبل می بینیم که  $(*) \quad V(P_X \| Q_X) \leq V(\hat{P}_{X,Y} \| \hat{Q}_{X,Y})$

کامینیت ثابت کنیم که به ازای  $\hat{Q}_{X,Y} = Q_X W_{Y|X}$  و  $\hat{P}_{X,Y} = P_X W_{Y|X}$

تا نتیجه بگیریم  $V(P_X \| Q_X) = V(P_X W_{Y|X} \| Q_X W_{Y|X})$

نشان بدهیم  $V(P_X \| Q_X) \geq V(\hat{P}_{X,Y} \| \hat{Q}_{X,Y})$  : (-)

$$V_{\alpha, \beta, \gamma, \delta, t}(\hat{P}_{X,Y} \| \hat{Q}_{X,Y}) = V_{\alpha, \beta, \gamma, \delta, t}(P_X W_{Y|X} \| Q_X W_{Y|X}) =$$

$$= \sup_{f(x,y) \in \mathcal{R}} \left\{ \mathbb{E}_{P_X} \left[ \mathbb{E}_{W_{Y|X}} f(x,y) \right] - \gamma \mathbb{E}_{Q_X} \left[ \mathbb{E}_{W_{Y|X}} f(x,y) \right] - \delta \log \left( \mathbb{E}_{Q_X} \left[ \mathbb{E}_{W_{Y|X}} e^{\alpha f(x,y)} \right] \right) - t \log \left( \mathbb{E}_{Q_X} \left[ \mathbb{E}_{W_{Y|X}} e^{\beta f(x,y)} \right] \right) \right\}$$

$$\text{let } \mathbb{E}_{W_{Y|X}} [f(x,y)] = g(x) \Rightarrow \leq \sup_{g \in \mathcal{R}} \left\{ \mathbb{E}_{P_X} [g(x)] - \gamma \mathbb{E}_{Q_X} [g(x)] - \delta \log \left( \mathbb{E}_{Q_X} [e^{\alpha g(x)}] \right) - t \log \left( \mathbb{E}_{Q_X} [e^{\beta g(x)}] \right) \right\}$$

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

Q.E.D.

در نتیجه ثابت شد که  $V_{\alpha, \beta, \gamma, \delta, t}(P_X \| Q_X) \geq V_{\alpha, \beta, \gamma, \delta, t}(P_X W_{Y|X} \| Q_X W_{Y|X})$

همچنین  $V(P_X \| Q_X) = V(P_X W_{Y|X} \| Q_X W_{Y|X})$  (\*)

حال بایه ثابت کنیم که اگر  $P_X$  و  $Q_X$  را با یک ترکیب جدید از آنها بسازیم  $P_\lambda = \lambda P + \bar{\lambda} Q$  که  $\lambda \in (0,1)$  را نیز داریم.

let  $Z \sim \text{Ber}(\lambda)$

$$P_{X|Z} = \begin{cases} P_\lambda & Z=0 \\ P_{\bar{\lambda}} & Z=1 \end{cases} \quad Q_{X|Z} = \begin{cases} Q_\lambda & Z=0 \\ Q_{\bar{\lambda}} & Z=1 \end{cases}$$

$$\Rightarrow V(P_{X2} \| Q_{X2}) \geq \mathbb{E}_{P_Z} [V(P_{X|Z} \| Q_{X|Z})] = \lambda V(P_\lambda \| Q_\lambda) + \bar{\lambda} V(P_{\bar{\lambda}} \| Q_{\bar{\lambda}})$$

$$\stackrel{\text{DPI}}{\Rightarrow} V(P_{X2} \| Q_{X2}) \geq V(P_X \| Q_X) = V(\underbrace{\lambda P_\lambda + \bar{\lambda} P_{\bar{\lambda}}}_{\text{marginal}} \| \underbrace{\lambda Q_\lambda + \bar{\lambda} Q_{\bar{\lambda}}}_{\text{marginal}})$$

حال از بالا در می یابیم که  $V_{\alpha, \beta, \gamma, \delta, t}(P_X \| Q_X)$  با شرط  $r + s\alpha + \beta t = 1$  هر زوج توانی را می سازد  $(P, Q)$  مناسب است.

Q.E.D.



$$f(x, y) = h(x) + g(y)$$

در این تیر سه مرتبه از قبل عمل میکنیم. داریم:

$$V_{\alpha, \beta, \gamma, \delta, t}(P_{XY} \| Q_X Q_Y) = \sup_{f \in \mathcal{X} \times \mathcal{Y}^{\mathbb{R}}} \left\{ \mathbb{E}_{P_{XY}}[f(x, y)] - \gamma \mathbb{E}_{Q_X Q_Y}[f(x, y)] - \delta \log \left( \mathbb{E}_{Q_X Q_Y} \left[ e^{\alpha f(x, y)} \right] \right) - t \log \left( \mathbb{E}_{Q_X Q_Y} \left[ e^{\beta f(x, y)} \right] \right) \right\}$$

$$\begin{aligned} & \geq \sup_{f \in \mathcal{X} \times \mathcal{Y}^{\mathbb{R}}} \left\{ \mathbb{E}_{P_{XY}}[h(x) + g(y)] - \gamma \mathbb{E}_{Q_X Q_Y}[h(x) + g(y)] - \delta \log \left( \mathbb{E}_{Q_X Q_Y} \left[ e^{\alpha(h(x) + g(y))} \right] \right) - t \log \left( \mathbb{E}_{Q_X Q_Y} \left[ e^{\beta(h(x) + g(y))} \right] \right) \right\} \\ & = \sup_{f \in \mathcal{X} \times \mathcal{Y}^{\mathbb{R}}} \left\{ \underbrace{\mathbb{E}_{P_X}[h(x)] - \gamma \mathbb{E}_{Q_X}[h(x)] - \delta \log \left( \mathbb{E}_{Q_X} \left[ e^{\alpha h(x)} \right] \right)}_{\text{فصل اول و دوم از } x} + \underbrace{\mathbb{E}_{P_Y}[g(y)] - \gamma \mathbb{E}_{Q_Y}[g(y)] - \delta \log \left( \mathbb{E}_{Q_Y} \left[ e^{\alpha g(y)} \right] \right) - t \log \left( \mathbb{E}_{Q_Y} \left[ e^{\beta g(y)} \right] \right)}_{\text{فصل اول و دوم از } y} \right\} \end{aligned}$$

$$= \sup_{f \in \mathcal{X}^{\mathbb{R}}} ( \quad ) + \sup_{f \in \mathcal{Y}^{\mathbb{R}}} ( \quad )$$

هر دو فصل از دستش از هم هستند و میتوان به آرا جدا کرد

$$= V_{\alpha, \beta, \gamma, \delta, t}(P_X \| Q_X) + V_{\alpha, \beta, \gamma, \delta, t}(P_Y \| Q_Y)$$

که در نهایت نتیجه داریم:

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

که این خاصیت برای  $D_{\alpha, \beta}$  نیز برقرار بود.

$$W_{\alpha}(P_X \| Q_X) = V_{\alpha, 0, 1 - \frac{1}{\alpha}, \frac{1}{\alpha^2}, 0}(P_X \| Q_X)$$

$$= \sup_{f \in \mathcal{X}^{\mathbb{R}}} \left\{ \mathbb{E}_{P_X}[f(x)] - \left(1 - \frac{1}{\alpha}\right) \mathbb{E}_{Q_X}[f(x)] - \frac{1}{\alpha^2} \log \left( \mathbb{E}_{Q_X} \left[ e^{\alpha f(x)} \right] \right) \right\}$$

$$\mathbb{E}_{Q_X} \left[ e^{\alpha f(x)} \right] = \mathbb{E}_{Q_X} \left[ \sum_{k=0}^{\infty} \frac{\alpha^k f(x)^k}{k!} \right] = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \mathbb{E}_{Q_X} [f(x)^k] \quad \text{ابتدا ترمیم میکنیم}$$

در آنگاه  $\alpha \rightarrow 0$  قرار است. برود اینجا نیست مقصود از آنجا قبل تیر را بیاوریم:

$$\mathbb{E}_{Q_X} \left[ e^{\alpha f(x)} \right] = 1 + \alpha \mathbb{E}_{Q_X} [f(x)] + \frac{\alpha^2}{2} \mathbb{E}_{Q_X} [f(x)^2] + \alpha \alpha^3 = 1$$

$$\rightarrow \log(1+t) = \sum_{k=1}^{\infty} \frac{t^k (-1)^{k+1}}{k} \rightarrow \log(1+t) = t - \frac{t^2}{2} + o(t^3)$$

$$\Rightarrow \log \left( \mathbb{E}_{Q_X} \left[ e^{\alpha f(x)} \right] \right) = \log \left( 1 + \alpha \mathbb{E}_{Q_X} [f(x)] + \frac{\alpha^2}{2} \mathbb{E}_{Q_X} [f(x)^2] + \alpha \alpha^3 \right)$$

$$\Rightarrow \log(\mathbb{E}_{Q_x} e^{\alpha f(x)}) = \left( \alpha \mathbb{E}_{Q_x}[f(x)] + \frac{\alpha^2}{2} \mathbb{E}_{Q_x}[f(x)^2] + o(\alpha^3) \right) - \frac{1}{2} \left( \alpha \mathbb{E}_{Q_x}[f(x)] + \frac{\alpha^2}{2} \mathbb{E}_{Q_x}[f(x)^2] + o(\alpha^3) \right)^2 + o \left( \left( \alpha \mathbb{E}_{Q_x}[f(x)] + \frac{\alpha^2}{2} \mathbb{E}_{Q_x}[f(x)^2] + o(\alpha^3) \right)^3 \right)$$

$$\Rightarrow \log(\mathbb{E}_{Q_x} e^{\alpha f(x)}) = \alpha \mathbb{E}_{Q_x}[f(x)] + \frac{\alpha^2}{2} \mathbb{E}_{Q_x}[f(x)^2] + o(\alpha^3) + \frac{o(\alpha^6)}{2} - \frac{\alpha^4}{4} \mathbb{E}_{Q_x}[f(x)^2]^2 + \frac{\alpha^2}{2} \mathbb{E}_{Q_x}[f(x)]^2 + o(\alpha^9)$$

$$\Rightarrow \log(\mathbb{E}_{Q_x} e^{\alpha f(x)}) \approx \alpha \mathbb{E}_{Q_x}[f(x)] + \frac{\alpha^2}{2} \left( \mathbb{E}_{Q_x}[f(x)^2] - (\mathbb{E}_{Q_x}(f(x)))^2 \right) + o(\alpha^3) \\ = \alpha \mathbb{E}_{Q_x}[f(x)] + \frac{\alpha^2}{2} \text{Var}_{Q_x}(f(x)) + o(\alpha^3)$$

$$\lim_{\alpha \rightarrow 0} W_\alpha(P_x \| Q_x) = \sup_{f \in \mathcal{H}^{\text{Lip}}_{\alpha \rightarrow 0}} \left\{ \mathbb{E}_{P_x}[f(x)] - (1 - \frac{1}{\alpha}) \mathbb{E}_{Q_x}[f(x)] - \frac{1}{\alpha^2} \left( \alpha \mathbb{E}_{Q_x}(f(x)) + \frac{\alpha^2}{2} \text{Var}_{Q_x}(f(x)) \right) \right\}$$

$$= \lim_{\alpha \rightarrow 0} \sup_{f \in \mathcal{H}^{\text{Lip}}_{\alpha \rightarrow 0}} \left\{ \mathbb{E}_{P_x}[f(x)] - \mathbb{E}_{Q_x}[f(x)] - \frac{1}{2} \text{Var}_{Q_x}(f(x)) + o(\alpha^3) \right\}$$

حال چون می‌خواهیم که رابطه تولید نرم  $f(x)$  نیز  $\sup$  را تغییر ندهد،  $\sup$  نسبت به  $\alpha$  ثابت است و ابتدا  $f$  و  $c$  گذاشته و دوباره  $c$  را در  $\sup$  گذاشته و در نهایت  $c$  را پهن می‌کنیم:

$$\mathbb{E}_{P_x}(f(x) \cdot c) - \mathbb{E}_{Q_x}(c \cdot f(x)) - \frac{1}{2} \text{Var}_{Q_x}(f(x)) = \overbrace{c \mathbb{E}_{P_x}(f(x)) - c \mathbb{E}_{Q_x}(f(x))}^{g(c)} - \frac{c^2}{2} \text{Var}_{Q_x}(f(x))$$

$$\frac{\partial}{\partial c} g(c) = \mathbb{E}_{P_x}(f(x)) - \mathbb{E}_{Q_x}(f(x)) - c \text{Var}_{Q_x}(f(x)) = 0$$

$$\hookrightarrow c^* = \frac{\mathbb{E}_{P_x}(f(x)) - \mathbb{E}_{Q_x}(f(x))}{\text{Var}_{Q_x}(f(x))} \rightarrow \text{هنا می‌بینیم که سرتاسر } c \text{ بکار می‌آید!}$$

$$\Rightarrow \lim_{\alpha \rightarrow 0} W_\alpha(P_x \| Q_x) = \sup_{f \in \mathcal{H}^{\text{Lip}}_{\alpha \rightarrow 0}} \left\{ \mathbb{E}_{P_x}[c^* f(x)] - \mathbb{E}_{Q_x}[c^* f(x)] - \frac{1}{2} \text{Var}_{Q_x}[c^* f(x)] \right\}$$

در ادامه سوال 6 پیش از (VII)

$$\lim_{\alpha \rightarrow 0} W_\alpha(P_X \| Q_\alpha) = \sup_{f \in \mathcal{X}^{\mathcal{R}}} \left\{ C^2 \left( \mathbb{E}_{P_X}(f(X)) - \mathbb{E}_{Q_\alpha}(f(X)) \right)^2 + \frac{C^2}{2} \text{Var}_{Q_\alpha}(f(X)) \right\}$$

$$= \sup_{f \in \mathcal{X}^{\mathcal{R}}} \left\{ \frac{(\mathbb{E}_{P_X}(f(X)) - \mathbb{E}_{Q_\alpha}(f(X)))^2}{\text{Var}_{Q_\alpha}(f(X))} + \frac{1}{2} \frac{(\mathbb{E}_{P_X}(f(X)) - \mathbb{E}_{Q_\alpha}(f(X)))^2}{\text{Var}_{Q_\alpha}(f(X))} \right\}$$

$$= \frac{1}{2} \sup_{f \in \mathcal{X}^{\mathcal{R}}} \left\{ \frac{(\mathbb{E}_{P_X}(f(X)) - \mathbb{E}_{Q_\alpha}(f(X)))^2}{\text{Var}_{Q_\alpha}(f(X))} \right\} = \frac{1}{2} \chi^2(P_X \| Q_\alpha) \quad \text{Q.E.D.} \quad \blacksquare$$

سے فراموش نہ کیے کہ  $\chi^2$  بہ نسبت  $\alpha$  اور  $\beta$

صفت فراموش نہ کیے کہ 5 صاب کریم: اگر  $r + \alpha s + \beta t = 1$  اور  $\alpha, \beta, r, s, t \geq 0$

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

حال اگر  $\alpha = \alpha$ ,  $\beta = 0$ ,  $r = 1 - \frac{1}{\alpha}$ ,  $s = \frac{1}{\alpha^2}$  و  $t = 0$  رکھیں تو  $\alpha \rightarrow 0$  صفت فراموش نہ کیے کہ بالا بہ نسبت  $\alpha$  اور  $\beta$ ۔

$$\left\{ \begin{aligned} \lim_{\alpha \rightarrow 0} W_\alpha(P_{XY} \| Q_X Q_Y) &= \frac{\chi^2(P_{XY} \| Q_X Q_Y)}{2} = V_{\alpha, 0, 1 - \frac{1}{\alpha}, \frac{1}{\alpha^2}, 0}(P_{XY} \| Q_X Q_Y) \\ \lim_{\alpha \rightarrow 0} W_\alpha(P_X \| Q_X) &= \frac{\chi^2(P_X \| Q_X)}{2} = V_{\alpha, 0, 1 - \frac{1}{\alpha}, \frac{1}{\alpha^2}, 0}(P_X \| Q_X) \\ \lim_{\alpha \rightarrow 0} W_\alpha(P_Y \| Q_Y) &= \frac{\chi^2(P_Y \| Q_Y)}{2} = V_{\alpha, 0, 1 - \frac{1}{\alpha}, \frac{1}{\alpha^2}, 0}(P_Y \| Q_Y) \end{aligned} \right\} \quad (**) \quad \blacksquare$$

حال (\*) و (\*\*) با ہم یکجا لیں کہ:

$$\chi^2(P_{XY} \| Q_X Q_Y) \geq \chi^2(P_X \| Q_X) + \chi^2(P_Y \| Q_Y) \quad \text{Q.E.D.} \quad \blacksquare$$