تمرین سری یک JTS2 — [illegible] 400101204 — [illegible]

(I) ابتدا به عنوان یک لم اثبات می‌کنیم که برای $a_i, b_i \in \mathbb{R}^{[0,1]}$ که $1 \le i \le n$ داریم: $\left|\prod_{i=1}^{n} a_i - \prod_{i=1}^{n} b_i\right| \le \sum_{i=1}^{n} |a_i - b_i|$

اثبات:

$$\left| a_1 a_2 \cdots a_n - b_1 a_2 \cdots a_n + b_1 a_2 \cdots a_n - b_1 b_2 \cdots b_n \right| = \left|\prod_{i=1}^{n} a_i - \prod_{i=1}^{n} b_i\right|$$

طبق نامساوی مثلث:

$$\le \left| a_2 \cdots a_n (a_1 - b_1) + b_1 (a_2 \cdots a_n - b_2 \cdots b_n)\right| \le a_2 \cdots a_n |a_1 - b_1| + b_1 |a_2 \cdots a_n - b_2 \cdots b_n|$$

$$\xrightarrow[1 \le a_i \le n]{0 \le b_1 \le 1} \left|\prod a_i - \prod b_i\right| \le |a_1 - b_1| + |a_2 \cdots a_n - b_2 \cdots b_n|$$

در نتیجه مشخص است که این قضیه را برای حالت $n=1$ که پایه استقرا است ثابت کرد. بخش فوق نشان می‌دهد که اگر قضیه برای $P(1)$ و $P(2)$ و ... $P(n-1)$ برقرار باشد، در این صورت برای $P(n)$ نیز درست است چرا که: (گام استقرا)

$$\left|\prod_{i=1}^{n} a_i - \prod_{i=1}^{n} b_i\right| \le |a_1 - b_1| + \left|\prod_{i=2}^{n} a_i - \prod_{i=2}^{n} b_i\right| \le |a_1 - b_1| + \sum_{i=2}^{n} |a_i - b_i| = \sum_{i=1}^{n} |a_i - b_i| \quad \text{Q.E.D.}$$

برای $n=1$ حکم بدیهی است چرا که $|a_1 - b_1| \le |a_1 - b_1|$. حال که نامساوی اثبات شد حکم نیز نتیجه می‌شود چرا که

$$\text{let } \begin{cases} a_i = P_i \\ b_i = q_i \end{cases}$$

$$d_{TV}\left(\prod_{i=1}^{n} P_i, \prod_{i=1}^{n} q_i\right) = \frac{1}{2}\left|\prod_{i=1}^{n} P_i - \prod_{i=1}^{n} q_i\right| \le \frac{1}{2}\sum_{i=1}^{n} |P_i - q_i| = \sum_{i=1}^{n} d_{TV}(P_i, q_i)$$

Q.E.D. ■

(II) می‌دانیم برای متغیرهای تصادفی Y و X که $Y = g(X)$ که $g(\cdot)$ تابعی یک به یک بود. حال اگر تابع f، فاصله را حساب کنیم در حالت کلی ثابت می‌کنیم. برای حالت خاص total variation، $f(x) = \frac{1}{2}|x-1|$. تابع فاصله به صورت مقابل تعریف می‌شود (همانا f باید محدب باشد)

$$d_f(P, Q) = \int_{\mathcal{X}} f\left(\frac{dP}{dQ}\right) dQ = \mathbb{E}_Q\left[f\left(\frac{dP}{dQ}\right)\right]$$

$$Y = g(X) \rightarrow \begin{cases} P_Y(y) = \dfrac{P_X(g^{-1}(y))}{g'(g^{-1}(y))} = dP_Y \\[2ex] Q_Y(y) = \dfrac{Q_X(g^{-1}(y))}{g'(g^{-1}(y))} = dQ_Y \end{cases}$$

$$d_f(P_Y, Q_Y) = \int_{y} f\left(\frac{P_X(g^{-1}(y))}{Q_X(g^{-1}(y))}\right) \frac{Q_X(g^{-1}(y))}{g'(g^{-1}(y))}\, dy$$

$$\hookrightarrow \frac{dP_Y}{dQ_Y} = \frac{P_X(g^{-1}(y))}{Q_X(g^{-1}(y))} = \frac{P_X(z)}{Q_X(z)} = \frac{dP_X}{dQ_X}$$

حال تغییر متغیر می‌دهیم: $z = g^{-1}(y) \xleftarrow{\text{و بنابراین}} y = g(z) \rightarrow dy = g'(z)\,dz$

$$\Rightarrow d_f(P_Y, Q_Y) = \int_{Z} f\left(\frac{P_X(z)}{Q_X(z)}\right) \frac{Q_X(z)}{\cancel{g'(z)}} \cancel{g'(z)}\, dz \xrightarrow{\text{بنابر } (*)} = \int_{\mathcal{X}} f\left(\frac{dP}{dQ}\right) dQ = d_f(P_X, Q_X)$$

حال اگر قرار دهیم $f(x) = \frac{|x-1|}{2}$، طبق قضیه فوق ثابت می‌شود که:

$$d_{TV}(P_X, Q_X) = d_{TV}(P_Y, Q_Y) = d_{TV}(P_{g(X)}, Q_{g(X)})$$

Q.E.D. ■

(III) تساوی را برای $f(x)=\frac{|x-1|}{2}$ حال داریم:

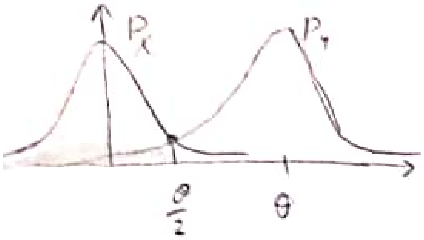
$$d_f(P_0\otimes Q, P_1\otimes Q) = \mathbb{E}_{X\sim P_0\otimes Q}\left[f\left(\frac{P_1(x)Q(x)}{P_0(x)Q(x)}\right)\right] = \mathbb{E}_{P_1\otimes Q}\left[f\left(\frac{dP_1}{dP_0}\right)\right] = \mathbb{E}_Q\left[\mathbb{E}_{P_1} f\left(\frac{dP_1}{dP_0}\right)\right] = \mathbb{E}_Q\left[d_f(P_0,P_1)\right]$$

حال چون که $d_f(P_0,P_1)$ به Q بستگی ندارد از $\mathbb{E}$ خارج می‌شود و داریم:

$$d_f(P_0\otimes Q, P_1\otimes Q) = d_f(P_0,P_1) \longrightarrow d_{TV}(P_0\otimes Q, P_1\otimes Q) = d_{TV}(P_0,P_1) \quad \blacksquare \text{ Q.E.D.}$$

$X\sim N(0,C)$, $N(\theta,C)\sim Y$

(IV) ابتدا مسئله را در حالت تک‌بعدی اثبات می‌کنیم:



$\longrightarrow X\sim N(0,\sigma)$ و $Y\sim N(\theta,\sigma)$

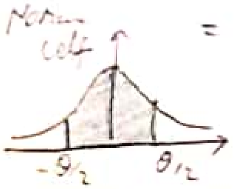
$$d_{TV}(P_X,P_Y) = \frac{1}{2}\int_{\mathbb{R}} |P_X(x)-P_Y(x)|\,dx = \int_{P_X>P_Y:=S} (P_X(x)-P_Y(x))\,dx$$

$S = \{a \mid P_X(a)\geq P_Y(a)\}.$

$$\Rightarrow d_{TV}(P_X,P_Y) = \int_{-\infty}^{\theta/2} P_X(x)-P_Y(x)\,dx$$

$$\Rightarrow d_{TV}(P_X,P_Y) = \int_{-\infty}^{\theta/2}\frac{\exp\left(\frac{-x^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}}dx - \int_{-\infty}^{\theta/2}\frac{\exp\left(\frac{-(y-\theta)^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}}dy = 1-\Phi\left(\frac{\theta/2}{\sigma}\right)-\left(1-\Phi\left(\frac{-\theta/2}{\sigma}\right)\right)$$

Normal cdf



$$= -\Phi\left(\frac{\theta}{2}\sigma^{-1}\right)+\Phi\left(-\frac{\theta}{2}\sigma^{-1}\right) = \int_{-\frac{\theta}{2}}^{\frac{\theta}{2}}\frac{\exp(-x^2/2)}{\sqrt{2\pi}}dx = 1-2\Phi\left(\frac{\theta}{2}\sigma^{-1}\right)$$

می‌دانیم که $X\sim N(0,C)$ ، حال در متغیر را سفید کنیم؛ $Y=C^{-1/2}X$ داریم:

$(C\geq 0)$

$$\mathbb{E}Y=0 \longrightarrow \mathbb{E}(YY^T) = \mathbb{E}(C^{-1/2}\underbrace{XX^T}_{C}C^{-1/2}) = \mathbb{E}I = I$$

اگر $X\sim N(\mu,C)$ در این صورت

$$\mathbb{E}Y = C^{-1/2}\mu \longrightarrow \mathbb{E}(YY^T) = \mathbb{E}(C^{-1/2}\underbrace{XX^T}_{C}C^{-1/2}) = I$$

حال یک ماتریس A تعریف می‌کنیم به طوری که اگر $Z=AY$ یعنی $X\sim N(\mu,C)$ ، در این صورت

حال باید A را به دست بیاوریم (پس سطر اول A، مشخص است)

$$\mathbb{E}Z = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}\times a$$

$$\mathbb{E}Z = AC^{-1/2}\mu \longrightarrow a = \|C^{-1/2}\mu\|$$

$$\Rightarrow A = \begin{bmatrix} \text{—} \frac{C^{-1/2}\mu}{\|C^{-1/2}\mu\|} \text{—} \\ \text{—}\ v_2\ \text{—} \\ \vdots \\ \text{—}\ v_n\ \text{—}\end{bmatrix}$$

$$\mathbb{E}(ZZ^T) = A\,\mathbb{E}(YY^T)A^T = AA^T = I$$

درنتیجه $\{v_i\}$ از الگوریتم گرام اشمیت به دست می‌آیند.

اما در قسمت قبل اثبات کردیم که اگر $g(x)=C^{-1/2}x$:

$$d_{TV}(P(x),Q(x)) = d_{TV}(P_{g(x)},Q_{g(x)})$$

$$\left.\begin{array}{ll} Z = AC^{-1/2}X & X\sim N(0,C) \\ Z' = AC^{-1/2}X & X\sim N(\mu,C)\end{array}\right\}$$ زمانی که

حال اگر به صورت تک‌بعدی نگاه کنیم دو متغیر برابر با $N(0,1)$ است و $N(\|C^{-1/2}\mu\|,1)$

$$d_{TV}(P_Z,P_{Z'}) = d_{TV}(N(0,C),N(\mu,C)) = 1-2\Phi\left(\frac{1}{2}\|C^{-1/2}\mu\|\right)$$

در نتیجه حکم ثابت شد...

$\blacksquare$ Q.E.D.

Scanned by CamScanner

سوال 2C:

(I) حکم اول قضیه‌ی factorization Neyman Fisher را بیان می‌کند.

(i) ابتدا فرض می‌کنیم که T آماره‌ی بسنده از X با پارامتر θ باشد. داریم:

(*) چون که اگر X را داشته باشیم به‌طور قطع مقدار $T=t(x)$ را نیز می‌دانیم

$$\mathbb{P}(X=x|\theta) \xrightarrow{(*)} \mathbb{P}((X=x)\wedge(T=t)|\theta) = \mathbb{P}(X=x|T=t,\theta)\,\mathbb{P}(T=t|\theta)$$

حال چون که $P_{X|T}$ مستقل از θ است داریم $\mathbb{P}(X|T,\theta)=\mathbb{P}(X|T)$ و در نتیجه خواهیم داشت.

$$\mathbb{P}(X=x|\theta) = \underbrace{\mathbb{P}(X=x|T=t)}_{h}\,\overbrace{\mathbb{P}(T=t|\theta)}^{g} \xrightarrow{(*)} = h(x)\,g(t|\theta) \xRightarrow{\exists g,h} \blacksquare \ \text{Q.E.D.}$$

$\mathbb{P}(X=x|T=t)$ صرفاً تابعی از X است چرا که خود T نیز تابعی از X است. ← صحت یک سمت قضیه اثبات شد.

(ii) حال می‌دانیم که $\mathbb{P}(X=x|\theta)=h(x)g(t|\theta)$ $^{(**)}$ در تلاشیم ثابت کنیم که T آماره‌ی بسنده از X است با پارامتر θ:

$$\mathbb{P}(X=x|T=t,\theta) = \frac{\mathbb{P}((X=x)\wedge(T=t)|\theta)}{\mathbb{P}(T=t|\theta)} \xrightarrow{(*)} = \frac{\mathbb{P}(X=x|\theta)}{\mathbb{P}(T=t|\theta)} \xrightarrow{(**)} = \frac{h(x)\,g(t|\theta)}{\mathbb{P}(T=t|\theta)} \qquad (a)$$

$$\mathbb{P}(T=t|\theta) = \sum_{X:T(X)=t} \mathbb{P}((X=x)\wedge(T=t)|\theta) = \sum_{X:T(X)=t} g(T=t|\theta)\,h(X) = g(T=t|\theta)\Big(\sum_{X:T(X)=t} h(X)\Big) \qquad (b)$$

$$\xRightarrow{(a),(b)} \mathbb{P}(X=x|T=t,\theta) = \frac{h(x)\,g(t|\theta)}{g(t|\theta)\times\Big(\sum_{X:T(X)=t} h(x)\Big)} = \frac{h(x)}{\sum_{X:T(X)=t} h(x)} \Rightarrow \blacksquare \ \text{Q.E.D.}$$

حال همانطور که دیده می‌شود، تابع توزیع احتمال $\mathbb{P}_{X|T}$ تابعی از θ نیست و به عبارت دیگر $P_{X|T}$ مستقل از θ است و ← در نتیجه دو طرف قضیه اثبات شد.

$$\exists g,h \Rightarrow P(X|\theta) = g(T(x)|\theta)\,h(x) \iff T \text{ is sufficient statistic}$$

ادامه سوال ۲:

$$\begin{cases} H_0 & X \sim P_X \\ H_1 & X \sim q_X \end{cases}$$

(II)

(i)

$$D_{kl}(P_{XY} \| q_{XY}) = \underbrace{\mathbb{E}_{P_X}\left[\log \frac{P_X}{q_X}\right]}_{D_{kl}(P_X \| q_X)} + \mathbb{E}_{P_X}\left(\cancel{D_{kl}(P_{Y|X} \| q_{Y|X})}\right)^{0} = \underbrace{\mathbb{E}_{P_Y}\left[\log \frac{P_Y}{q_Y}\right]}_{D_{kl}(P_Y \| q_Y)} + \mathbb{E}_{P_Y}\left[D_{kl}\left(\frac{P_{X|Y}}{q_{X|Y}}\right)\right]$$

$$P_X^{\theta} P_{T|X} = P_T^{\theta} P_{X|T}$$

T بسنده ای کافی از X است با پارامتر θ اگر

let $\theta = X, T = Y \longrightarrow P_{X|Y} = P_{Y|X}$

Data Processing Inequality $\longrightarrow D(P_X \| q_X) = D(P_Y \| q_Y) + D(P_{X|Y} \| q_{X|Y} | P_Y)$

$\forall \theta_0, \theta_1 : P_{X|Y} = q_{X|Y}$ در این صورت $X \perp\!\!\!\perp \theta | Y$ اگر T بسنده کافی از X است آنگاه در این صورت

$\forall \theta_0, \theta_1 \Rightarrow P_{X|Y} = q_{X|Y}$

if $P_{X|Y} = q_{X|Y} \Rightarrow X \perp\!\!\!\perp \theta | Y$

$\left(\forall \theta_0, \theta_1 \Leftrightarrow P_{X|Y} = q_{X|Y} \right)$

حال چون دو طرف اثبات شده داریم.

(ii)

طبق فرض اگر X، T بسنده کافی باشد آنگاه در این صورت $P_{X|T} = q_{X|T}$ در نتیجه

$$D(P_X \| q_X) = D(P_T \| q_T) + \underbrace{D(P_{X|T} \| q_{X|T} | P_T)}_{0} \longrightarrow D(P_X \| q_X) = D(P_T \| q_T)$$

T بسنده کافی است $\leftarrow X \perp\!\!\!\perp \theta | T \leftarrow P_{X|T} = q_{X|T} \leftarrow D(P_{X|T} \| q_{X|T}) = 0$

(III)

$$P_\theta(x) = \mathbb{1}(\theta = 0) P_X + \mathbb{1}(\theta = 1) q_X = q(x)\left(\frac{P(x)}{q(x)} \mathbb{1}(\theta = 0) + \mathbb{1}(\theta = 1)\right)$$

$$= \underbrace{q(x)}_{h(x)} \exp\left(\underbrace{\left(\log \frac{P(x)}{q(x)}\right) \mathbb{1}(\theta = 0) + \mathbb{1}(\theta = 1) \times \log 1}_{g(T(x); \theta)}\right)$$

در نتیجه g و h ثابت شدند در نتیجه $\log \frac{P(x)}{q(x)} = T(x)$ بسنده ای کافی برای X است

سوال 3: (I) می‌دانیم که برای $R=(P(Z=0),Q(Z=0))$ که

$$P_{Z|X}(0|x)=\begin{cases}0 & \log\frac{P(x)}{q(x)}\geq\gamma\\ 1 & \log\frac{P(x)}{q(x)}<\gamma\end{cases}$$

نسبت درستنمایی:

$$\Lambda(x)=\frac{P(x)}{q(x)}=\frac{1/\sqrt{2\pi}}{1/\sqrt{2\pi}}\times\frac{\exp\left(\frac{-x^2}{2}\right)}{\exp\left(\frac{-(x-\mu)^2}{2}\right)}=\exp\left(\frac{+\mu^2-2x\mu}{2}\right)$$

$$\log\Lambda(x)=\frac{\mu^2}{2}-\mu x\geq\log\gamma(x)\longrightarrow x\leq\frac{\mu}{2}-\frac{\log\gamma}{\mu}$$

حال ناحیه Neyman-Pearson بصورت مقابل است

$$R(P,Q)=\left(P(\log\Lambda(x)\geq\log\gamma),\ Q(\log\Lambda(x)\leq\log\gamma)\right)=\left(P\left(x\leq\frac{\mu}{2}-\frac{\log\gamma}{\mu}\right),\ Q\left(x\leq\frac{\mu}{2}-\frac{\log\gamma}{\mu}\right)\right)$$

$$P=N(0,1)\longrightarrow P\Big(x\leq\underbrace{\frac{\mu}{2}-\frac{\log\gamma}{\mu}}_{\lambda}\Big)=\int_{-\infty}^{\lambda}dx=1-\Phi(\lambda)$$

$$Q=N(\mu,1)\longrightarrow Q\left(x\leq\frac{\mu}{2}-\frac{\log\gamma}{\mu}\right)=\int_{-\infty}^{\lambda}dx=\Phi(\mu-\lambda)$$

تساوی به همراه که در صفحات بعدی آمده است. ← (*)

(II)

$$\Lambda(x)=\frac{\prod_{i=1}^{n}P(x_i)}{\prod_{i=1}^{n}q(x_i)}=\frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n}{\left(\frac{1}{\sqrt{2\pi}}\right)^n}\times\frac{\prod_{i=1}^{n}\exp\left(\frac{-x_i^2}{2}\right)}{\prod_{i=1}^{n}\exp\left(\frac{-(x_i-\mu)^2}{2}\right)}=\prod_{i=1}^{n}\exp\left(-\frac{1}{2}(2x_i\mu-\mu^2)\right)=\exp\left(\frac{n\mu^2}{2}-\mu\bar{X}\right)$$

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

$$\Rightarrow\log\Lambda(x)=n\mu\left(\frac{\mu}{2}-\bar{X}\right)$$

از قسمت قبل می‌دانیم که $T(X)=\log\frac{P^{n}(X)}{Q^{n}(X)}$ یک سنجه آماری است، در نتیجه:

$$R(P^{\otimes n},Q^{\otimes n})=R(P_{\bar{X}},Q_{\bar{X}})\Rightarrow \begin{matrix}P_{\bar{X}}\sim N(0,\frac{1}{n}) & \equiv P_{\bar{X}}\left(x\leq\frac{\mu}{2}-\frac{\log\gamma}{n\mu}\right)\Rightarrow\\ Q_{\bar{X}}\sim N(\mu,\frac{1}{n}) & \equiv Q_{\bar{X}}\left(x\leq\frac{\mu}{2}-\frac{\log\gamma}{n\mu}\right)\end{matrix}$$

برای نرمالایز کردن $\bar{X}$ را بصورت $\tilde{X}=\sqrt{n}\bar{X}$ تعریف می‌کنیم:

$$\left.\begin{matrix}P_{\tilde{X}}\sim N(0,1)\\ Q_{\tilde{X}}\sim N(\sqrt{n}\mu,1)\end{matrix}\right\}\quad R(P_{\bar{X}},Q_{\bar{X}})=R(P_{\tilde{X}},Q_{\tilde{X}})$$

T برای $\tilde{X}$ نیز بسنده کافی است.

(*) تساوی به همراه که در صفحات بعدی آمده است

سوال ۴:

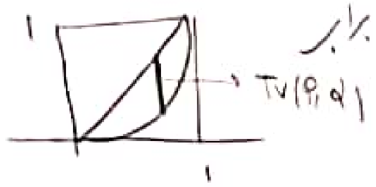
(I) ابتدا می‌دانیم که TV تعریف $d_{TV}(P,Q)=\sup_{E}\left[P(E)-Q(E)\right]$ نیز ترتیب کرد و همچنین داریم $P(Z=0)\ge\alpha$

$\beta_\alpha(P,Q)=\inf_{\substack{P(Z=0)\ge\alpha \\ Z(\cdot|x)}} Q(Z=0)$ برای $0\le\alpha\le1$ است. در نتیجه برای α ثابت شده داریم: (E را فرض می‌کنیم که $Z=0$ است! $Z=1$)

$$d_{TV}(P,Q)=\sup_{Z(\cdot|x)}\left[P(Z=0)-Q(Z=0)\right]\le\sup\left[P(Z=0)\right]-\inf\left[Q(Z=0)\right]=\alpha-\beta_\alpha(P,Q)$$

$\blacksquare$ Q.E.D.

$E: Z=0$



برای پیدا کردن TV کافی است $\sup\ \alpha-\beta_\alpha(P,Q)$ را بیابیم یعنی برابر نصف طول ماکسیمم خط عمودی بین منحنی α و $\beta_\alpha(P,Q)$ است

(II)

$$\begin{cases} H_0 & X\sim P,\ \pi_0 \\ H_1 & X\sim Q,\ \pi_1 \end{cases}\Rightarrow P_E=\inf_{P_{Z|X}}\pi_0\pi_{1|0}+\pi_1\pi_{0|1}=\pi_0\inf_{P_{Z|X}}\left[P(Z=1)+\frac{\pi_1}{\pi_0}Q(Z=0)\right]$$

$\pi_{1|0}$
$\pi_{0|1}$

$$=\pi_0\inf_{P_{Z|X}}\left[P(Z=1)-\underbrace{\frac{\pi_1}{\pi_0}}_{\gamma}Q(Z=1)\right]+\pi_1$$

$$\xrightarrow{(*)} P(Z=1)-\gamma Q(Z=1)=\sum_{x\in\mathcal{X}}P(x)Z(1|x)-\gamma q(x)Z(1|x)=\sum_{x\in\mathcal{X}}Z(1|x)\left(P(x)-\gamma q(x)\right)$$

طبق قضیه ML $Z(1|x)=\mathbb{1}\left[P(x)\ge\gamma q(x)\right]$

$$\underset{P_{Z|X}}{\arg\min}\ P_E=\begin{cases}\mathbb{1}\left[P(x)\ge\gamma q(x)\right] & Z=1\\ \mathbb{1}\left[P(x)\le\gamma q(x)\right] & Z=0\end{cases}\qquad \gamma=\frac{\pi_1}{\pi_0}$$

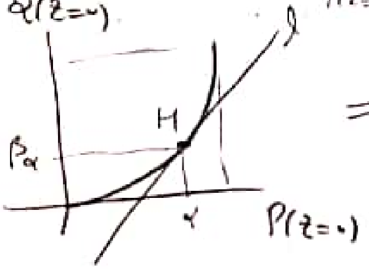
در نتیجه $Z(0|x)=\mathbb{1}\left[P(x)\le\gamma q(x)\right]$ ~ در واقع تست بهینه طبق قضیه ML است:

(III) می‌دانیم که $P_{Z|X}$ برای β_α به ازای $P_{Z|X}=\mathbb{1}\left(\frac{P(x)}{Q(x)}>\gamma\right)$ رخ می‌دهد که $\gamma>\frac{\pi_1}{\pi_0}$. چون که $T(x)=\log\frac{P(x)}{Q(x)}$ است

- اگر T فرض کنیم X است در نتیجه داریم (*). $D_{KL}(P_X\|Q_X)=D_{KL}(P_T\|Q_T)$

$P(Z=0)=\alpha$

$Q(Z=0)=\beta_\alpha(P,Q)=\min_{P(Z=0)\ge\alpha}Q(Z=0)$



حال خطوط $P(Z=0)-\gamma Q(Z=0)$ را در نظر بگیرید ماکسیمم آن برابر

$$\Rightarrow\sup_{P_{Z|X}}P(Z=0)-\gamma Q(Z=0)$$

جایی رخ می‌دهد که شیب در آن برابر $\frac{1}{\gamma}$ است.

$$\Rightarrow\frac{1}{\gamma}=\frac{d}{d\alpha}\beta_\alpha(P,Q)=\frac{dQ_X}{dP_X}=\frac{dP_T}{dQ_T}$$

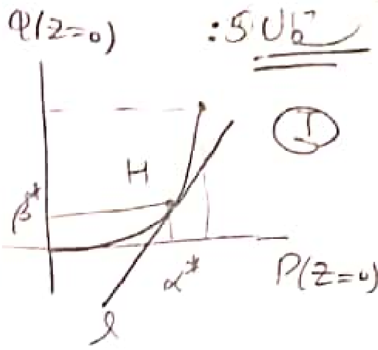
$$\xrightarrow{(*)} D_{KL}(P_T\|Q_T)=D_{KL}(P_X\|Q_X)=\underset{T\sim\log\left(\frac{dP}{dQ}\right)=\tau}{\mathbb{E}}\left[\log\frac{dP_T}{dQ_T}\right]=\int_{-\infty}^{\infty}\frac{d\alpha}{d\tau}\log\left(\frac{dP}{dQ}\right)d\tau=\int_0^1\log\frac{dP}{dQ}\,d\alpha$$

$$\hookrightarrow D_{KL}(P_X\|Q_X)=-\int_0^1\left(\log\frac{dQ}{dP}\right)d\alpha=-\int_0^1\log\left(\frac{d}{d\alpha}\beta_\alpha(P,Q)\right)d\alpha$$

سوال ۵:

I

$$\begin{cases} H_0 & X\sim P \\ H_1 & X\sim Q \end{cases} \qquad \underset{P(Z=0)\geq\alpha}{\arg\min}\; Q(Z=0) = \beta_\alpha(P,Q)$$



$$\underset{P_{Z|X}}{\arg\min}\; P_e = \underset{P_{Z|X}}{\arg\min}\; \pi_0\pi_{1|0} + \pi_1\pi_{0|1} = \underset{P_{Z|X}}{\arg\min}\; \pi_{1|0} + \overbrace{\frac{\pi_1}{\pi_0}}^{\gamma}\pi_{0|1}$$

$$= \sum_{x\in\mathcal{X}} P(z|x)\left(P(x) - \gamma q(x)\right)$$

حال . $P(z|x) = \mathbb{1}\left(\frac{p(x)}{q(x)} > \gamma\right)$ اگر بسازیم فعلا $\underset{P_{Z|X}}{\arg\min}\; P_e$

طبق قضیه ML ، بهترین حالت حالتی است که

$P(Z=0) - \gamma Q(Z=0) = C$ در این صورت

$$\beta = \frac{\pi_0}{\pi_1}\alpha + \frac{\pi_0}{\pi_1}C$$

or $\qquad \beta^* = \min_{\alpha>\alpha^*}\beta$

در نتیجه $\boxed{C = \tau = \frac{\pi_1}{\pi_0}}$ Q.E.D. ∎

II حال داریم که $\frac{d}{d\alpha}\beta_\alpha = 2\alpha$ چون که منحنی β_α محدب است

$$\alpha^* - \gamma\beta^*_\alpha = C \longrightarrow 1 = \gamma\frac{d\beta_\alpha}{d\alpha}$$

$$\boxed{\beta^* = \alpha^{*2} = \frac{\pi_0^2}{4\pi_1^2} \longleftarrow 2\alpha^* = \frac{\pi_0}{\pi_1} = \gamma}$$

∎ Q.E.D.

III داریم $0<\alpha<1$ است و همچنین $\tau = 2\alpha^* = \frac{\pi_0}{\pi_1}$ در نتیجه: $0 < \frac{\pi_0}{\pi_1} < 2$ خواهد بود.

سوال ۶:

$\mathbb{P}=\{P_n\}_{n=1}^{\infty}$
$\mathbb{Q}=\{Q_n\}_{n=1}^{\infty}$
دنباله‌ها

$$\mathbb{P}\lhd\mathbb{Q} \Leftrightarrow \forall A_n\in\mathcal{X}_n : \text{if } \lim_{n\to\infty} Q_n(A_n)\to 0 \quad \text{then: } \lim_{n\to\infty} P_n(A_n)\to 0$$

I) فرض می‌کنیم که
$$\begin{cases} H_0 : X\sim P_n \\ H_1 : X\sim Q_n \end{cases}$$

حال فرض کنید A_n بخش‌هایی از $\mathcal{X}_n$ باشد که به H_0 تعلق دارند. حال مشخص است که احتمال خطای نوع دوم برابر $Q_n(A_n)\to 0$ $\lim_{n\to\infty}$ است. در این صورت طبق $\mathbb{P}\lhd\mathbb{Q}$ نتیجه می‌گیریم که $\lim_{n\to\infty} P_n(A_n)=0$ است. در حالی که ۱

برای اینکه بگوییم با $n\to\infty$، مجموع احتمال خطای اول و دوم صفر می‌شود؛ نیاز داریم که $Q_n(A_n)\to 0$ و $P_n(A_n)\to 1$. اما با $Q_n(A_n)$ به علت به هم پیوستگی $\mathbb{P}$ و $\mathbb{Q}$ نتیجه می‌شود که $P_n(A_n)\to 0$ که این مورد مطلوب نیست. در نتیجه مجموع احتمال خطای نوع یک و دو صفر نمی‌شود.

$$P_e = P_{e_1}+P_{e_2} = Q_n(A_n)+\left[1-P_n(A_n)\right] \to 1 \neq 0 \qquad \blacksquare \text{ Q.E.D.}$$

II) یک دنباله از مجموعه‌های A_n را در نظر می‌گیریم که $\lim_{n\to\infty} Q_n(A_n)=0$.

$$P_n(A_n)=\underset{X\sim P_n}{\mathbb{E}}\left[\mathbb{1}(A_n)\right] \overset{\text{change of measure}}{=} \underset{X\sim Q_n}{\mathbb{E}}\left[\mathbb{1}(A_n)\underbrace{\frac{dP_n}{dQ_n}}_{L_n}\right] \leq \|L_n\|\sqrt{\underset{X\sim Q_n}{\mathbb{E}}\left[\mathbb{1}\{A_n\}\right]} = \sqrt{Q_n(A_n)}\times\|L_n\|$$

حال اگر در نامساوی فوق $Q_n(A_n)\to 0$ برود، به دلیل آنکه $\|L_n\|<\infty$، نتیجه می‌گیریم که $\lim_{n\to\infty} P_n(A_n)\to 0$، و در نتیجه $\mathbb{P}\lhd\mathbb{Q}$ ∎
Q.E.D.

سوال ۷:

سوال 7 (I) فرض می‌کنیم که m پیام ارسالی باشد که $m \in \{1,2,\dots,N\}$ ، $N=2^{nR}$. کد $X^n:\{1,\dots,N\}\to\mathcal{X}^n$
و دیکودر $D:\mathcal{Y}^n\to\{1,\dots,N\}$ ، مشاهده y^n ، و $\hat{m}$ خروجی است. همچنین n بار استفاده داشتیم.

$$M \to \boxed{X^n:\{1,\dots,N\}\to\mathcal{X}^n} \xrightarrow{X^n(m)} \boxed{P_{Y|X}} \xrightarrow{Y^n} \boxed{D:\mathcal{Y}^n\to\{1,\dots,N\}} \to \hat{M}$$

حال برای کانال $P_{Y|X}^{\otimes n}$ احتمال $P_{\hat{m}}$ این گونه است که به شرط y^n از ارسال $X^n(m)$ بین پیام‌های ممکن اشتراک بین $X^n(1), X^n(2), \dots, X^n(N)$.

$$\mathbb{P}[\hat{M}=\hat{m}] = \frac{P_{Y|X}^{\otimes n}[y^n \mid X^n(\hat{m})]}{\sum_{m=1}^{N} P_{Y|X}^{\otimes n}(y^n \mid X^n(m))}$$

حال فرض می‌کنیم که m توزیع یکنواختی روی مجموعه $\{1,\dots,N\}$ داشته باشد
امید ریاضی اینکه m ارسال شود و m نیز دریافت شود را تحت همه
کتاب کدهای $X^n(i)$ ها بررسی کنیم.

$$\mathbb{E}_{X^n(1)\dots X^n(N)}\left[\mathbb{P}[\hat{M}=m]\right] = \mathbb{E}_{X^n(1)\dots X^n(N)}\left[\sum_{m,y^n} \underbrace{P_M(m)}_{1/N} P_{Y|X}^{\otimes n}(y^n \mid X^n(m))\, P_{\hat{M}|Y^n}(m \mid y^n)\right]$$

حال جمع روی m تقارن دارد، پس می‌توانیم این شکل را فقط برای $X^n(1)$ حساب کنیم، داریم:

$$\mathbb{E}_{X^n(1)\dots X^n(N)}\left[\mathbb{P}[\hat{M}=m]\right] = \underbrace{N \times \frac{1}{N}}_{1}\ \mathbb{E}_{X^n(1)\dots X^n(N)}\left[\sum_{y^n} P_{Y|X}^{\otimes n}(y^n \mid X^n(1))\, P_{\hat{M}|Y^n}(1 \mid y^n)\right]$$

Q.E.D.

(II) حال می‌دانیم که $P_{\hat{M}|Y^n}(1 \mid y^n) = \dfrac{\frac{1}{N} P_{Y|X}^{\otimes n}(y^n \mid X^n(1))}{\frac{1}{N}\sum_{m=1}^{N} P_{Y|X}^{\otimes n}(y^n \mid X^n(m))}$ که نتیجه می‌دهد:

$$\mathbb{E}_{X^n(1)\dots X^n(N)}\left[\mathbb{P}[\hat{M}=m]\right] = \sum_{y^n} \mathbb{E}_{X^n(1)\dots X^n(N)}\left[P_{Y|X}^{\otimes n}(y^n \mid X^n(1)) \times \frac{P_{Y|X}^{\otimes n}(y^n \mid X^n(1))}{\sum_{m=1}^{N} P_{Y|X}^{\otimes n}(y^n \mid X^n(m))}\right]$$

حال چون تابع $f(x) = \frac{k}{x+c}$ محدب است می‌دانیم که $\mathbb{E} f(x) \geq f(\mathbb{E}x)$. در نتیجه:

$$\mathbb{E}_{X^n(1)\dots X^n(N)}\left[\mathbb{P}[\hat{M}=m]\right] \geq \mathbb{E}_{X^n(1)}\left[\sum_{y^n} P_{Y|X}^{\otimes n}(y^n \mid X^n(1)) \times \frac{P_{Y|X}^{\otimes n}(y^n \mid X^n(1))}{P_{Y|X}^{\otimes n}(y^n \mid X^n(1)) + \mathbb{E}_{X^n(2)\dots X^n(N)}\left[\sum_{m=2}^{N} P_{Y|X}^{\otimes n}(y^n \mid X^n(m))\right]}\right]$$

به علت تقارن

$$= \mathbb{E}_{X^n(1)}\left[\sum_{y^n} P_{Y|X}^{\otimes n}(y^n \mid X^n(1)) \frac{P_{Y|X}^{\otimes n}(y^n \mid X^n(1))}{P_{Y|X}^{\otimes n}(y^n \mid X^n(1)) + (N-1)\underbrace{\mathbb{E}_{X^n(2)\dots X^n(N)}\left[P_{Y|X}^{\otimes n}(y^n \mid X^n(m))\right]}_{P_Y^{\otimes n}(y^n)}}\right]$$

$$\frac{1}{x+(N-1)a} \geq \frac{1}{x+Na}$$

$$\geq \mathbb{E}_{X^n(1)}\left(\sum_{y^n} P_{Y|X}^{\otimes n}(y^n \mid X^n(1)) \times \frac{1}{1 + N\,\frac{P_Y^{\otimes n}(y^n)}{P_{Y|X}^{\otimes n}(y^n \mid X^n)}}\right)$$

$$N = 2^{nR}$$

$$a^{-\log_a x} = \frac{1}{x}$$

$$= \sum_{y^n, X^n(1)} \underbrace{P_{Y|X}^{\otimes n}(y^n \mid X^n(1))\, P_X^{\otimes n}(X^n)}_{P(X^n, Y^n)} \frac{1}{1 + 2^{nR} \cdot 2^{-\log_2\left(\frac{P_{Y|X}^{\otimes n}(y^n \mid X^n)}{P_Y^{\otimes n}(y^n)}\right)}}$$

و در نهایت به این رسیدیم که

$$\mathbb{E}_{X^n(1),\dots,X^n(M)}\Big[\mathbb{P}[\hat{M}=M]\Big] \geq \mathbb{E}_{X^n,Y^n}\left[\frac{1}{1+2^{nR}\cdot 2^{-\log_2\left(\frac{P^{\otimes n}_{Y|X}(y^n|x^n)}{P^{\otimes n}_{Y}(y^n)}\right)}}\right]$$

■ Q.E.D.

(III) حال می‌دانیم که

$$R < I(X;Y) \quad , \quad I(X^n;Y^n) = \mathbb{E}_{X^n,Y^n}\left[\frac{P^{\otimes n}_{Y|X}(y^n|x^n)}{P^{\otimes n}_{Y}(y^n)}\right]$$

$f(x)=\frac{k}{x+c}$ محدب

$$\mathbb{E}_{X^n(1),\dots,X^n(M)}\Big[\mathbb{P}[\hat{m}=m]\Big] \geq \mathbb{E}_{X^n,Y^n}\left[\frac{1}{1+2^{nR}\cdot 2^{-\log_2\left(\frac{P^{\otimes n}_{Y|X}(y^n|x^n)}{P^{\otimes n}_{Y}(y^n)}\right)}}\right] \geq \frac{1}{1+2^{nR}\,\mathbb{E}_{X^n,Y^n}\left[2^{-\log_2(-)}\right]}$$

$f(x)=2^{-x}$ محدب

$$\geq \frac{1}{1+2^{nR}\underbrace{2^{-\mathbb{E}_{X^n,Y^n}\left[+\log_2\left(\frac{P_{Y|X}(y^n|x^n)}{P^{\otimes n}_{Y}(y^n)}\right)\right]}}_{2^{I(X^n;Y^n)}}} = \frac{1}{1+2^{nR-nI(X;Y)}}$$

از طرف دیگر می‌دانستیم که $I(X^n;Y^n) = nI(X;Y)$ (چون که X ها مستقلند). حال اگر n به سمت ∞ میل دهیم:

$$\lim_{n\to\infty}\mathbb{E}_{X^n(1),\dots,X^n(2^{nR})}\Big[\mathbb{P}[\hat{m}=m]\Big] \geq \lim_{n\to\infty}\frac{1}{1+2^{nR-nI(X;Y)}} = \frac{1}{1+\lim_{n\to\infty}2^{nR-nI(X;Y)}} = \frac{1}{1+0} = 1$$

$R<I(X;Y)$

حال که $\mathbb{E}(\mathbb{P}(\hat{m}=m))$ بین 1 و 0 است نتیجه شد با این احتمال فقط با فرض $R < I(X;Y)$ تحت تمام انتخاب‌های $X^n(1),\dots,X^n(2^{nR})$ و با افزایش n به سمت 1 میل می‌کند

$$\lim_{n\to\infty}\mathbb{E}_{X^n(1),\dots,X^n(2^{nR})}\Big[\mathbb{P}[\hat{M}=M]\Big] = 1$$

■ Q.E.D.

ادامه سوال 7.

سوال 8: این بار $M=(M_x,M_y)$ و Z^n تعریف شده و کانال n بار استفاده شده است.

$M_x \in \{1,\dots,N_x\},\ N_x = 2^{nR_x}$

$M_y \in \{1,\dots,N_y\},\ N_y = 2^{nR_y}$

$$M=(M_x,M_y) \longrightarrow \boxed{\begin{array}{l} X^n:\{1,\dots,N_x\}\to \mathcal{X}^n \\ Y^n:\{1,\dots,N_y\}\to \mathcal{Y}^n \end{array}} \longrightarrow \boxed{P_{Z|X,Y}} \longrightarrow \boxed{D: Z^n \to \{1,\dots,N_x\}\times\{1,\dots,N_y\}} \longrightarrow \hat{M} = (\hat{M}_x,\hat{M}_y)$$

$$\mathbb{P}(\hat{m}=\hat{m}) = \mathbb{P}(\hat{M}_x=\hat{m}_x, \hat{M}_y=\hat{m}_y) = \frac{P_{Z|X,Y}^{\otimes n}(z^n|X^n(m_x),Y^n(m_y))}{\sum_{m_x=1}^{N_x}\sum_{m_y=1}^{N_y} P_{Z|X,Y}^{\otimes n}(z^n|X^n(m_x),Y^n(m_y))}$$

حال روی تمامی کدهای $X^n_{(i)}$ و $Y^n_{(j)}$ امید ریاضی را می‌گیریم

$$\mathbb{E}_{X^n_{(1)}\dots X^n_{(N_x)}}\left[\mathbb{E}_{Y^n_{(1)}\dots Y^n_{(N_y)}}\left[\sum_{m_x,m_y,z^n} \underbrace{P_{M_xM_y}(m_x,m_y)}_{\frac{1}{N_xN_y}} P_{Z|X,Y}^{\otimes n}(z^n|X^n(m_x),Y^n(m_y))\, P_{\hat{M}_x\hat{M}_y|Z^n}(m_x,m_y|z^n)\right]\right]$$

به سبب تقارن

m_y, m_x

$$= N_xN_y\cdot\frac{1}{N_xN_y}\ \mathbb{E}_{X^n_{(1)}\dots X^n_{(N_x)}}\left[\mathbb{E}_{Y^n_{(1)}\dots Y^n_{(N_y)}}\left[\sum_{z^n} P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)})\, P_{\hat{M}_x\hat{M}_y|Z^n}(1,1|z^n)\right]\right]$$

حال چون که

$$P_{\hat{M}_x\hat{M}_y|Z^n}(1,1|z^n) = \frac{\cancel{\frac{1}{N_xN_y}}\,P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)})}{\cancel{\frac{1}{N_xN_y}}\sum_{m_x=1}^{N_x}\sum_{m_y=1}^{N_y} P_{Z|X,Y}^{\otimes n}(z^n|X^n(m),Y^n(m))}$$

در نتیجه:

$$\mathbb{E}_{X^n_{(1)}\dots}\ \mathbb{E}_{Y^n_{(1)}\dots}\ \mathbb{P}(\hat{M}=m) = \mathbb{E}_{X^n_{(1)}\dots}\ \mathbb{E}_{Y^n_{(1)}\dots}\sum_{z^n} P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)}) \frac{P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)})}{\sum_{m_x=1}^{N_x}\sum_{m_y=1}^{N_y} P_{Z|X,Y}^{\otimes n}(z^n|X^n(m),Y^n(m))}$$

حال چون تابع $f(x)=\frac{k}{x+c}$ محدب است داریم

$$\mathbb{E}_{X^n_{(1)}\dots}\ \mathbb{E}_{Y^n_{(1)}\dots}\ \mathbb{P}(\hat{M}=m) \geq \mathbb{E}_{X^n_{(1)},Y^n_{(1)}}\left[\sum_{z^n} P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)}) \cdot \frac{P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)})}{\mathbb{E}_{X^n_{(2)}\dots X^n_{(N_x)}}\ \mathbb{E}_{Y^n_{(2)}\dots Y^n_{(N_y)}}(\dots)}\right]$$

$$\& \geq \mathbb{E}_{X^n_{(1)},Y^n_{(1)}}\left[\sum_{z^n} P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)}) \cdot \frac{P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)})}{\mathbb{E}_{X^n_{(2)}\dots,\,Y^n_{(2)}\dots}\left[\begin{array}{l} P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n_{(1)}) + \sum_{m_x=2}^{N_x} P_{Z|X,Y}^{\otimes n}(z^n|X^n(m_x),Y^n_{(1)}) \\ + \sum_{m_y=2}^{N_y} P_{Z|X,Y}^{\otimes n}(z^n|X^n_{(1)},Y^n(m_y)) + \sum_{m_x,m_y=2,2}^{N_x,N_y} P_{Z|X,Y}^{\otimes n}(z^n|X^n(m_x),Y^n(m_y)) \end{array}\right]}\right]$$

حال از Jensen داریم:

$$\mathbb{E}_{X^n_{(1)}\cdots}\ \mathbb{E}_{Y^n_{(1)}\cdots}\ \mathbb{P}(\hat{M}=M) \geq \mathbb{E}_{X^n_{(1)},Y^n_{(1)}}\left[\sum_{z^n} P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)}) \times \frac{P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})}{P^{\otimes n}_{Z|X,Y}(z|X^n_{(1)},Y^n_{(1)}) + (N_x-1)\,P^{\otimes n}_{Z|Y}(z^n|Y^n_{(1)}) + (N_y-1)\,P^{\otimes n}_{Z|X}(z^n|X^n_{(1)}) + (N_x-1)(N_y-1)P_Z'(z^n)}\right.$$

$$\geq \sum_{z^n} \mathbb{E}_{X^n_{(1)},Y^n_{(1)}} P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)}) \times \frac{1}{1 + N_x \dfrac{P^{\otimes n}_{Z|Y}(z^n|Y^n_{(1)})}{P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})} + N_y \dfrac{P^{\otimes n}_{Z|X}(z^n|X^n_{(1)})}{P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})} + N_x N_y \dfrac{\mathbb{P}^{\otimes n}_Z(z^n)}{P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})}}$$

$$\geq \sum_{z^n} \underbrace{P^{\otimes n}_X(X^n_{(1)})\,P^{\otimes n}_Y(Y^n_{(1)})\,P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})}_{P^{\otimes n}_{X,Y,Z}(X^n,Y^n,Z^n)} \times \frac{1}{1 + 2^{nR_x}\,2^{-\log_2\left(\frac{P^{\otimes n}_{Z|Y}}{P^{\otimes n}_{Z|X,Y}}\right)} + 2^{nR_y}\,2^{-\log_2\left(\frac{P^{\otimes n}_{Z|X}}{P^{\otimes n}_{Z|X,Y}}\right)} + 2^{n(R_x+R_y)}\cdot 2^{-\log_2\left(\frac{P^{\otimes n}_Z}{P^{\otimes n}_{Z|X,Y}}\right)}}$$

$$= \mathbb{E}_{X^n,Y^n,Z^n}\left(\cdots\right)$$

حال جنسن چون $\frac{k}{x+c}$ محدب است (داریم):

$$\geq \frac{1}{1 + \mathbb{E}_{X^n,Y^n,Z^n}\left[2^{nR_x}\cdot 2^{-\log\left(\frac{P^{\otimes n}_{Z|Y}}{P^{\otimes n}_{Z|X,Y}}\right)} + 2^{nR_y}\cdot 2^{-\log_2\left(\frac{P^{\otimes n}_{Z|X}}{P^{\otimes n}_{Z|X,Y}}\right)} + 2^{nR_x+nR_y}\cdot 2^{-\log_2\left(\frac{P^{\otimes n}_Z}{P^{\otimes n}_{Z|X,Y}}\right)}\right]}$$

حال بین تابع

2^{-x} محدب است داریم

$$\geq \frac{1}{1 + 2^{nR_x - \mathbb{E}_{X^n,Y^n,Z^n}\log_2\left(\frac{P^{\otimes n}_{Z|Y}(z^n|Y^n_{(1)})}{P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})}\right)} + 2^{nR_y - \mathbb{E}_{X^n,Y^n,Z^n}\log_2\left(\frac{\mathbb{P}^{\otimes n}_{Z|X}(z^n|X^n_{(1)})}{P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})}\right)} + 2^{nR_x+nR_y - \mathbb{E}_{X,Y,Z}\log_2\left(\frac{P^{\otimes n}_Z(z^n)}{P^{\otimes n}_{Z|X,Y}(z^n|X^n_{(1)},Y^n_{(1)})}\right)}}$$

$$\geq \frac{1}{1 + 2^{nR_x - I(X^n;Z^n|Y^n)} + 2^{nR_y - I(Z^n|X^n;Y^n)} + 2^{nR_x+nR_y - I(Z^n;X^n,Y^n)}}$$

بعلت iid بودن $X^n;Y^n,Z^n$

$$= \frac{1}{1 + 2^{nR_x - nI(X;Z^n|Y^n)} + 2^{nR_y - nI(Z|X;Y)} + 2^{nR_x+nR_y - nI(Z;X,Y)}}$$

حال مشخص است که اگر سه شرط

$$\begin{cases} R_X \leq I(X;Z|Y) \\ R_Y \leq I(Y;Z|X) \\ R_X + R_Y \leq I(X,Y;Z) \end{cases}$$

تواماً برقرار باشند در این صورت

$$\lim_{n\to\infty} \mathbb{E}_{X^n(1),\ldots X^n(N_X)} \left[\mathbb{E}_{Y^n(1),\ldots Y^n(N_Y)} \left[\mathbb{P}(\hat{m}=m) \right] \right] \geq \lim_{n\to\infty} \frac{1}{1+2^{nR_X - nI(X;Z|Y)} + 2^{nR_Y - nI(Z|X;Y)} + 2^{nR_X + nR_Y - nI(Z;X,Y)}}$$

$$= \frac{1}{1+\lim_{n\to\infty} 2^{\overbrace{nR_X - nI(X;Z|Y)}^{\leq 0}} + 2^{nR_Y - nI(Z|X;Y)} + 2^{nR_X+nR_Y - nI(Z;X,Y)}}$$

$$= \frac{1}{1+0} = 1$$

حکم اثبات شد، چرا که

$$\mathbb{E}_{X^n(1),\ldots,X^n(2^{nR_X})} \left[\mathbb{E}_{Y^n(1),\ldots,Y^n(2^{nR_Y})} \left[\mathbb{P}(\hat{m}=m) \right] \right] = \boxed{1}$$

Q.E.D.