نظریه اطلاعات، آمار و یادگیری

دکتر یاسایی

مهندسی برق

برنا خدابنده ۴۰۰۱۰۹۸۹۸

تمرین شماره 1

تاریخ: 1403/01/14

Problem 9)

$$P_i \to X_i \leftarrow q_i, \quad X_i \to \boxed{HT} \to \begin{cases} H_0 : (Z=0) \\ H_1 : (Z=1)\end{cases}$$

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

$$P_{err} = P(\hat{H} \neq H) = \frac{1}{2}\left(1 - d_{TV}\left(\prod_{i=1}^{n} P_i, \prod_{i=1}^{n} q_i\right)\right) \quad : \quad H : \{H_i\}_{i=1}^{n} \quad : H_i : \begin{cases} 0 : X_i \sim P_i \\ 1 : X_i \sim q_i \end{cases}$$

$$P(\hat{H} \neq H) = P(\{H_i\}_{i=1}^{n} \neq \{\hat{H}_i\}_{i=1}^{n}) = P\left(\bigcup_{i=1}^{n} H_i \neq \hat{H}_i\right) \ge \sum_{i=1}^{n} P(H_i \neq \hat{H}_i) = \sum_{i=1}^{n} \frac{1}{2}(1 - d_{TV}(P_i, q_i))$$

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

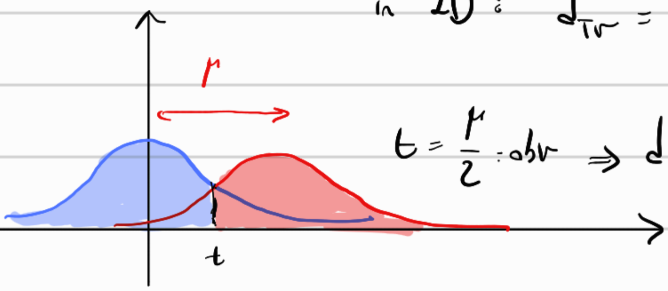
2) $d_{TV}(P_x, q_x) = d_{TV}(P_y, q_y) : Y = g(X)$: obvious from a HT perspective.

$$d_{TV}(P_x, q_x) = \frac{1}{2}\sum_{x} |P(x) - q(x)| \underset{\substack{\{ \\ g \,:\, \text{one-to-one} \,\Rightarrow\, P(X=x) = P(g(X) = g(x))}}{=} \frac{1}{2}\sum_{x} |P(g(x)) - q(g(x))| = \frac{1}{2}\sum_{y = g(x)} |P(y) - q(y)| = \boxed{d_{TV}(P_y, q_y)}$$

3) $d_{TV}(P_0, P_1) = d_{TV}(P_0 \otimes q, P_1 \otimes q)$: obvious from HT.

$$d_{TV}(P_0 \otimes q, P_1 \otimes q) = \frac{1}{2}\sum_{x,y} |P_0(x)q(y) - P_1(x)q(y)| = \frac{1}{2}\sum_{x}\sum_{y} q(y)|P_0(x) - P_1(x)| = \frac{1}{2}\underbrace{\sum_{x}|P_0(x) - P_1(x)|}_{d_{TV}(P_0,P_1)} \cdot \underbrace{\sum_{y} q(y)}_{1} = \boxed{d_{TV}(P_0, P_1)}$$

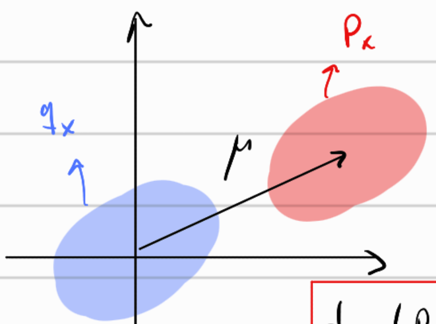
4) $d_{TV}(\mathcal{N}(0,C), \mathcal{N}(\mu, C)) = 1 - 2\Phi\left(\frac{1}{2}\|C^{-\frac{1}{2}}\mu\|_2\right)$



in 1D : $d_{TV} = \sum_{x} [P(x) - q(x)]_{+} = \int_{t}^{\infty} (dP_p - dP_q)\, dx = \int_{t}^{\infty} \frac{\left(e^{-\frac{(x-\mu)^2}{2C}} - e^{-\frac{x^2}{2C}}\right)}{\sqrt{2\pi}} dx$

$$t = \frac{\mu}{2} : \text{obv} \Rightarrow d_{TV} = \frac{1}{\sqrt{2\pi}}\int_{t}^{\infty}\left(e^{-\frac{(x-\mu)^2}{2C}} - e^{-\frac{x^2}{2C}}\right)dx = 1 - \frac{2}{\sqrt{2\pi}}\int_{t}^{\infty} e^{-\frac{x^2}{2C}} dx$$

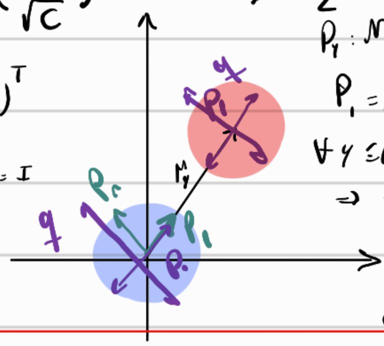
$$\Rightarrow d_{TV}(\mathcal{N}(0,C), \mathcal{N}(\mu, C)) = 1 - 2\Phi\left(\frac{t}{\sqrt{C}}\right) = 1 - 2\Phi\left(\frac{1}{2}C^{-\frac{1}{2}}\mu\right) \quad : \text{in 1D}$$



$C = E(XX^T) = C^{\frac{1}{2}}(C^{\frac{1}{2}})^T$

$Y = C^{-\frac{1}{2}}X : C_y = E(YY^T) = I$

$\mu_y = C^{-\frac{1}{2}}\mu$



$P_y : \mathcal{N}(0, I)$, $q_y : \mathcal{N}(\mu_y, I)$

$P_1 = \hat{\mu}$: decomposition :

$\forall y \in \mathbb{R}^n : X = \alpha P_1 + \beta P_2$) rotation.

$\Rightarrow P_y = \mathcal{N}(0,1) \otimes \mathcal{N}(0, I^{n-1})$

$\to Q_y = \mathcal{N}(\|\mu_y\|, 1) \otimes \mathcal{N}(0, I^{n-1})$

$\sim e^{-\frac{1}{2}\|y - \mu_y\|^2} = e^{-\frac{1}{2}\beta^2} \cdot e^{-\frac{1}{2}(\alpha - \|\mu_y\|)^2}$

$$\boxed{d_{TV}(P_x, q_x) = d_{TV}(P_y, q_y) = d_{TV}(P_y, Q_y) = d_{TV}(\mathcal{N}(0,1), \mathcal{N}(\|\mu_y\|, 1)) = 1 - 2\Phi\left(\frac{1}{2}\|\mu_y\|\right) = 1 - 2\Phi\left(\frac{1}{2}\|C^{-\frac{1}{2}}\mu\|\right)}$$

$$X \overset{\text{whiten}}{\rightsquigarrow} C^{-\frac{1}{2}}X = Y \overset{\text{rotation}}{\rightsquigarrow} \mathcal{N}(\|\mu_y\|, I) \otimes \mathcal{N}(0, I)$$

Problem 2)

T: sufficient statistic : $X \sim P_\theta$: $\theta \perp\!\!\!\perp X \mid T \Rightarrow T$ has all information of θ.

1. $T = t(X) \rightsquigarrow T$: sufficient iff $P_\theta(x) = g(t(x),\theta)\,h(x)$

(a) obviously : $\boxed{\text{if } P_\theta(x) = g(t(x),\theta)h(x)} = P(x|\theta)$, $T = t(X) \Rightarrow P_{T|X}(t|x) = \delta_{t,t(x)}$

$$\Rightarrow P_{X|T}(x|\tau) = \frac{P_{T|X}(\tau|x)P_\theta(x)}{\sum_x P_{T|X}(\tau|x)P_\theta(x)} = \frac{\delta_{\tau,t(x)}\, g(t(x),\theta)h(x)}{\sum_x g(t(x),\theta)h(x)\,\delta_{\tau,t(x)}} = \delta_{\tau,t(x)} \frac{g(\cancel{t},\theta)h(x)}{\sum_{x:t(x)=\tau} g(\cancel{t},\theta)h(x)}$$

$$= \delta_{\tau,t(x)} \frac{h(x)}{\sum_{x:t(x)=\tau} h(x)} = \frac{h(x)}{\sum_{x':t(x')=t(x)} h(x')} \Rightarrow \boxed{\text{independent of } \theta \to X|T \perp\!\!\!\perp \theta}$$

now the other way : $\boxed{\text{if } X|T \perp\!\!\!\perp \theta}$

$$P_{X|T,\theta}(x|\tau,\theta) = \frac{P_{T|X}(\tau|x)P_\theta(x)}{\sum P_{T|X}(\tau|x)P_\theta(x)} = \frac{\delta_{\tau,t(x)}P_\theta(x)}{\sum_{x'}\delta_{\tau,t(x')}P_\theta(x')} = \delta_{\tau,t(x)}\frac{P_\theta(x)}{\sum_{x':t(x')=\tau}P_\theta(x')} = \frac{P_\theta(x)}{\sum_{x':t(x')=t(x)}P_\theta(x')} = P_{X|T}(x|\tau) \ \forall x \in \mathcal{X}$$

$$\sum_{x':t(x)=t(x')} P_\theta(x') = g(\theta, t(x)) \Rightarrow P_{X|T,\theta}(x|\tau,\theta) = \frac{P_\theta(x)}{g(t(x),\theta)} = P_{X|T}(x|\tau) \quad \forall x \in \mathcal{X}$$

↳ function of $\theta \Rightarrow P_\theta(x)$ should cancel

$\Rightarrow \exists\, m : \mathcal{X} \to \mathbb{R}$ s.t. $\boxed{P_\theta(x) = g(t(x),\theta)\,m(x)}$

$$\longrightarrow \boxed{\text{if } T = t(X):\ X|T \perp\!\!\!\perp \theta \iff P_\theta(x) = g(t(x),\theta)h(x)}$$

b)

$$\begin{tikzcd} H_0: P_x \arrow[dr] & & & & \hat H_0 \\ & X \arrow[r] & \boxed{P_{H|X}} \arrow[r] & \hat H \arrow[ur] \arrow[dr] & \\ H_1: q_x \arrow[ur] & & & & \hat H_1 \end{tikzcd}$$

or :

$$\begin{tikzcd} H_0: P_x \arrow[dr] & & & & & \hat H_0 \\ & X \arrow[r, "\pi_{T|X}"] & T \arrow[r] & \boxed{P_{H|T}} \arrow[r] & \hat H \arrow[ur] \arrow[dr] & \\ H_1: q_x \arrow[ur] & & & & & \hat H_1 \end{tikzcd}$$

let : $P(x) = P_{\theta_1}(x)$, $Q(x) = P_{\theta_2}(x)$, $D_{KL}(P_x \| Q_x) = D_{KL}(P_{X|\theta_1} \| P_{X|\theta_2})$

which would be a measure of the importance of θ, for example: $\left.\begin{matrix}\theta_1 = \theta \\ \theta_2 = \theta + \Delta\end{matrix}\right\} \Rightarrow D_{KL}(P\|Q) \approx \frac{1}{2} I(\theta)\Delta^2$

$D_{KL}(P_{XT} \| Q_{XT}) = D_{KL}(P_X \| Q_X) + E_T[D_{KL}(P_{X|T} \| Q_{X|T})]$, $X|T \perp\!\!\!\perp \theta$ iff $P_{X|T} = Q_{X|T}\ \forall \theta_1, \theta_2$

$\Rightarrow D_{KL}(P_{X|T} \| Q_{X|T}) = 0$ iff T is sufficient $\rightsquigarrow D_{KL}(P_{XT} \| Q_{XT}) = D_{KL}(P_X \| Q_X)$ if T is SS.

$$D_{KL}(P_{XT} \| Q_{XT}) = D_{KL}(P_T \| Q_T) + E_x[\underbrace{\cancel{D_{KL}(\overset{\pi_{T|X}}{P_{T|X}} \| \overset{\pi_{T|X}}{Q_{T|X}})}}_{0}] = D_{KL}(P_T \| Q_T)$$

$$\Rightarrow D_{KL}(P_T \| Q_T) = D_{KL}(P_{XT} \| Q_{XT}) = D_{KL}(P_X \| Q_X) \rightarrow \text{iff } T \text{ is sufficent.}$$

$$\boxed{\Rightarrow D_{KL}(P_T \| Q_T) = D_{KL}(P_X \| Q_X) \text{ iff } T \text{ is a sufficient statistic of } X}$$

c)

$$T = \log\left(\frac{P(x)}{Q(x)}\right), \quad \text{for a HT}: \theta = Z = \begin{cases} 0 & X \sim P \\ 1 & X \sim Q \end{cases}$$

$$\Rightarrow P_\theta(x) = \begin{cases} P(x) & \theta = 0 \\ Q(x) & \theta = 1 \end{cases}$$

$$\Rightarrow P(x) = Q(x) \times \frac{P(x)}{Q(x)} = Q(x)\, e^{T} \Rightarrow \text{let } g(t,\theta) = \begin{cases} e^{t} & \theta = 0 \\ 1 & \theta = 1 \end{cases}$$

$$\Rightarrow P_\theta(x) = Q(x) \times g(t,\theta) = \underbrace{g(t(x),\theta)}_{g(t(x),\theta)} \times \underbrace{Q(x)}_{h(x)}$$

$$\boxed{\Rightarrow \exists\, h, g : P_\theta(x) = g(t(x),\theta)\, h(x) \overset{\text{part a}}{\rightsquigarrow} t(x) \text{ is sufficient}}$$

which makes intuitive sense since we calculate $\log\left(\frac{P(x)}{Q(x)}\right)$ for our ideal HT anyway.

Problem 3)

$H_0 : N(0,1) \rightarrow X \rightarrow \boxed{P_{Z|X}} \rightarrow Z \rightarrow Z=0 : \hat{H}_0 \quad : P(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$H_1 : N(\mu,1) \rightarrow X \rightarrow \boxed{P_{Z|X}} \rightarrow Z \rightarrow Z=1 : \hat{H}_1 \quad : Q(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}$

$$R(N(0,1), N(\mu,1)) = \{(x,y) : \exists P_{Z|X} : x = P(Z=0), y = Q(Z=0)\}$$

we know that the shape is symmetric around $(\frac{1}{2},\frac{1}{2})$ and convex.
finding the lower bound should be enough for the shape.
we will use the supp Hyperplanes.

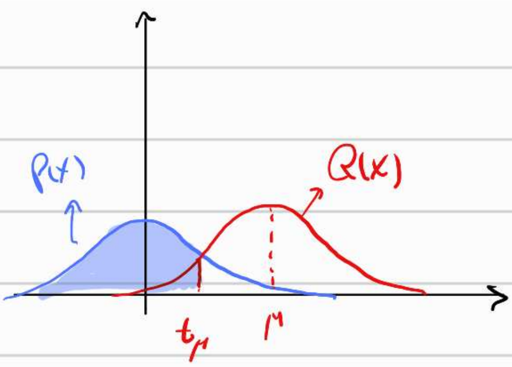
$$\Rightarrow f(\gamma) = \sup_{(x,y)\in R(P,Q)} x - \gamma y = \sup_{P_{Z|X}} \{P(Z=0) - \gamma Q(Z=0)\} = \sup_{P_{Z|X}} \{\sum_x P_{Z|X}(0|x)[P(x) - \gamma Q(x)]\}$$

<u>ideal test</u>

<u>deterministic</u>

$$\sum_x [P(x) - \gamma Q(x)]_+ = \int_{\frac{P(x)}{Q(x)} > \gamma} \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{x^2}{2}} - \gamma e^{-\frac{(x-\mu)^2}{2}}\right) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t_\gamma} \left(e^{-\frac{x^2}{2}} - \gamma e^{-\frac{(x-\mu)^2}{2}}\right) dx$$

$$= 1 - \Phi(t_\gamma) - \gamma \Phi(\mu - t_\gamma) = f(\gamma) = x_0 - \gamma y_0$$



$t_\gamma : e^{-\frac{t_\gamma^2}{2}} = \gamma e^{-\frac{(t_\gamma - \mu)^2}{2}} \Rightarrow -\ln(\gamma) + \frac{\mu^2}{2} = \mu t_\gamma$

$$\Rightarrow \boxed{t_\gamma = \frac{\mu}{2} - \frac{1}{\mu}\ln(\gamma)}$$

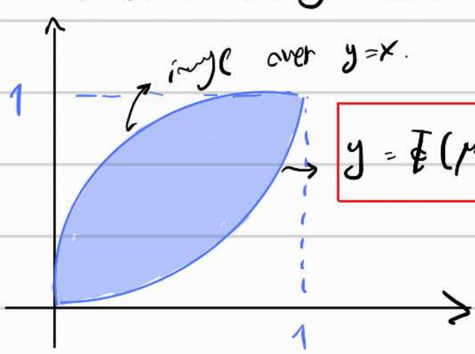
$$\Rightarrow x_0 = \gamma y_0 + 1 - \Phi(t_\gamma) - \gamma \Phi(\mu - t_\gamma) : \frac{\partial}{\partial \gamma} = 0 = y_0 + \frac{\partial t_\gamma}{\partial \gamma} \cancel{\frac{1}{\sqrt{2\pi}} e^{-\frac{t_\gamma^2}{2}}} - \Phi(\mu - t_\gamma) - \gamma \frac{\partial t_\gamma}{\partial \gamma} \cancel{\frac{1}{\sqrt{2\pi}} e^{-\frac{(t_\gamma-\mu)^2}{2}}}$$

$\Rightarrow y_0 = \Phi(\mu - t_\gamma) \Rightarrow x_0 = 1 - \Phi(t_\gamma)$, from γ : t_γ can range from $-\infty \rightarrow \infty$

$$\Rightarrow \boxed{R(P,Q) = \{(x,y) : x = 1 - \Phi(t), y = \Phi(\mu - t) : t \in \mathbb{R}, \text{ or } \begin{matrix} x \rightarrow y \\ y \rightarrow x \end{matrix} \text{ of this}\}}$$

$x = P(X<t)$, $y = Q(X<t)$, looks like a circle, but not exactly.

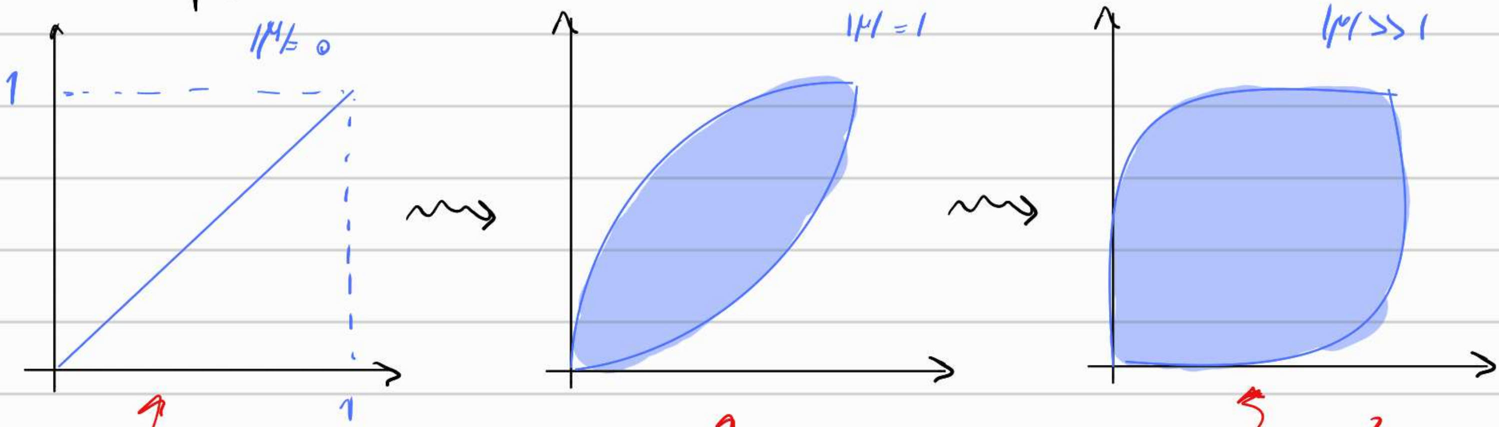
$$\boxed{\Phi(a) = \int_a^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx}$$



$\boxed{y = \Phi(\mu - \Phi^{-1}(1-x))}$ ~ to be exact.

$$\boxed{\Phi(\mu - \Phi^{-1}(1-x)) \leq y \leq 1 - \Phi(\mu - \Phi^{-1}(x))}$$

for example:



no clue, random guess is the best option

reasonable values. $\beta_{0.7}(P,Q) \approx 0.2$

almost $[0,1]^2$ can differentiate with perfect accuracy.

$$\boxed{\beta_\alpha(P,Q) = \bar{\Phi}\left(\mu - \bar{\Phi}^{-1}(1-\alpha)\right)}$$: weird expression indeed.

(b) if we have n samples: $X^1, X^2, \dots, X^n \sim P^{\otimes n}$ or $Q^{\otimes n}$

we showed that: $t = \log\left(\frac{P(x)}{Q(x)}\right)$ is a sufficient statistic for HT.

$$t = \log\left(\frac{P^{\otimes n}(x)}{Q^{\otimes n}(x)}\right) = \log\left(\prod_{i=1}^n \frac{P(x_i)}{Q(x_i)}\right) = \sum_{i=1}^n \log\left(\frac{P(x_i)}{Q(x_i)}\right) = \sum_{i=1}^n \log(P(x_i)) - \log(Q(x_i)) = \sum_{i=1}^n -\frac{x_i^2}{2} + \frac{(x_i-\mu)^2}{2}$$

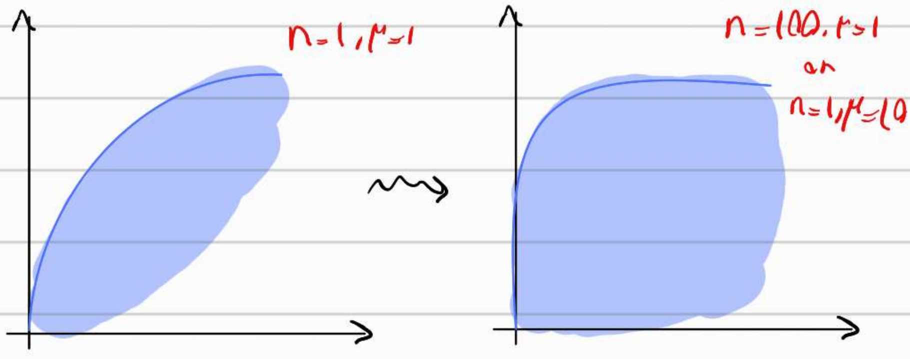
$$= \sum_{i=1}^n \frac{\mu^2 - 2x_i\mu}{2} = \frac{1}{2}n\mu^2 - \mu\left(\sum x_i\right) = n\mu\left[\frac{\mu}{2} - \frac{1}{n}\sum_{i=1}^n x_i\right] = t$$: is sufficient $\rightsquigarrow$ $\boxed{\frac{1}{n}\sum_{i=1}^n x_i \text{ is sufficient}}$

the two are equivalent for HT, thus they produce the same region. $\rightsquigarrow$ new HT: $\frac{\sum x_i}{n} \sim \bar{P}$ or $\bar{Q}$

$R(P^{\otimes n}, Q^{\otimes n}) = R(\bar{P}, \bar{Q})$. X_i are iid gaussians with $\sigma = 1$ $\rightsquigarrow$ $\begin{cases} \bar{Q} = \mathcal{N}(\mu, \frac{1}{n}) \\ \bar{P} = \mathcal{N}(0, \frac{1}{n}) \end{cases}$

equivalently, we can choose $t' = \frac{1}{\sqrt{n}}\sum_{i=1}^n x_i$, as a statistic $\rightarrow$ $P' = \mathcal{N}(\sqrt{n}\mu, 1)$, $Q' = \mathcal{N}(0,1)$

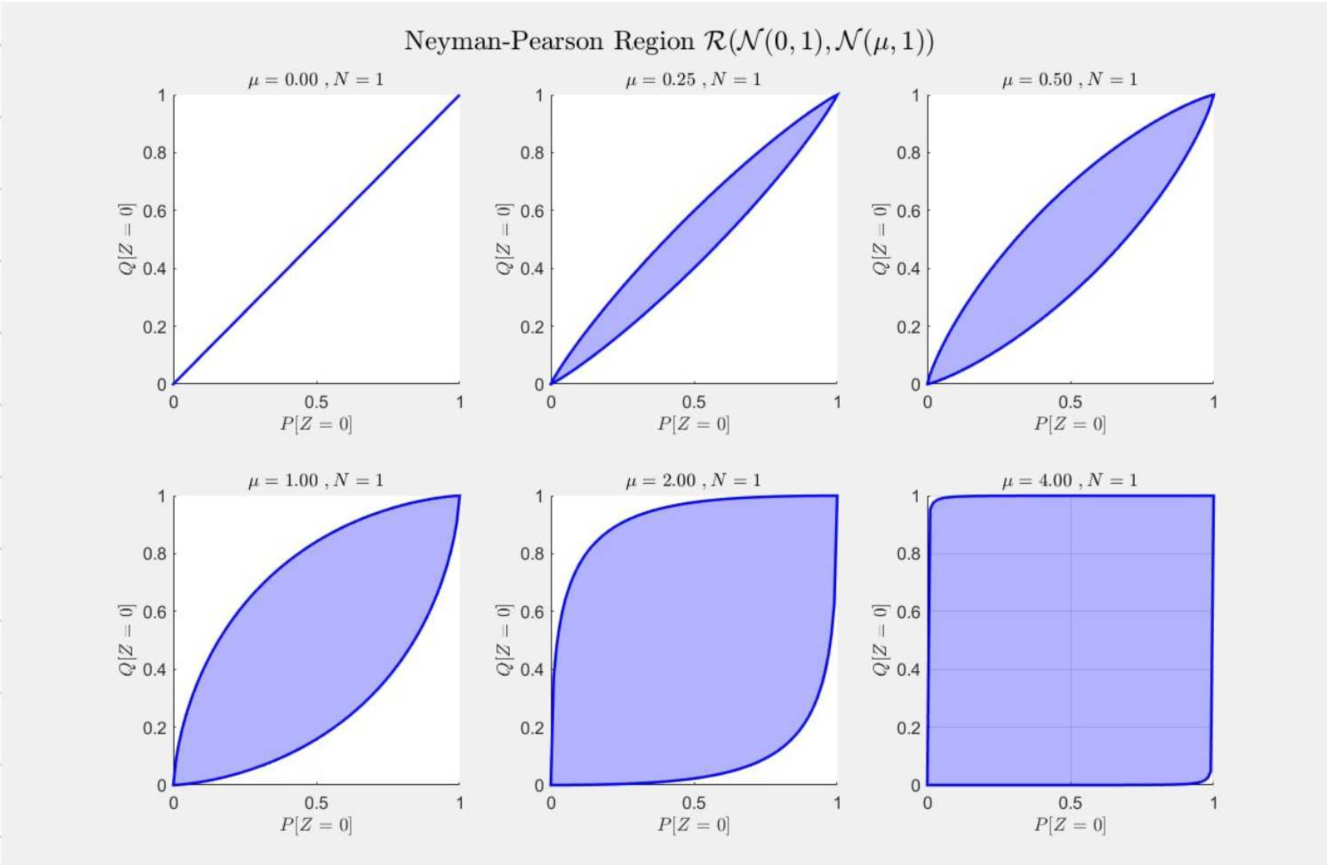
$$\rightarrow R(P^{\otimes n}, Q^{\otimes n}) = R(\bar{P}, \bar{Q}) = R(P', Q') = \boxed{R(\mathcal{N}(0,1), \mathcal{N}(\sqrt{n}\mu, 1)) = R(P^{\otimes n}, Q^{\otimes n})}$$



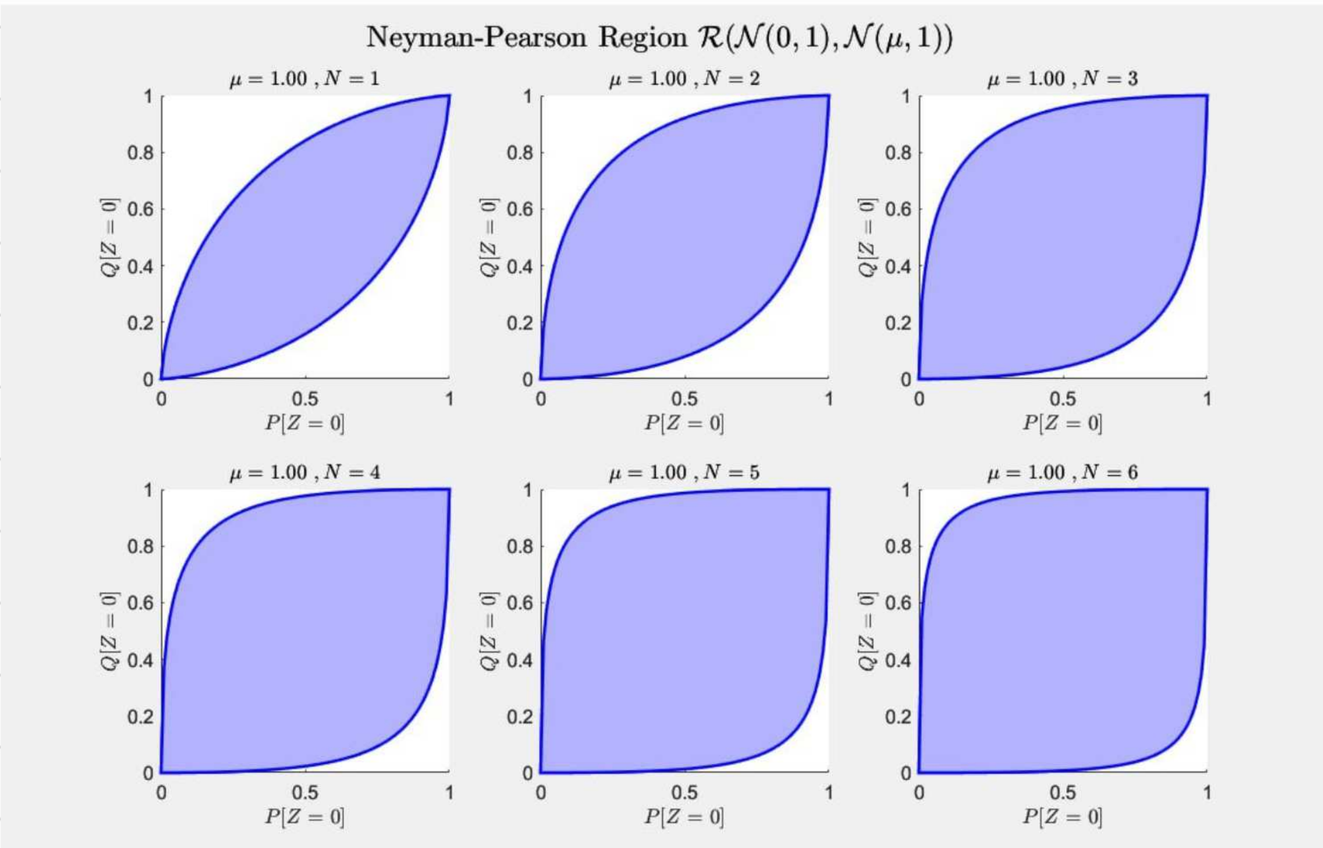
intuitively, increasing samples increases our accuracy.

Using matlab:

changing μ:



changing N:



Problem 4)

$\alpha = P(Z=0)$, $\beta_\alpha(P,Q) = \min\limits_{\substack{P(Z=0)\geq\alpha \\ T(\cdot|x)}} Q(Z=0)$

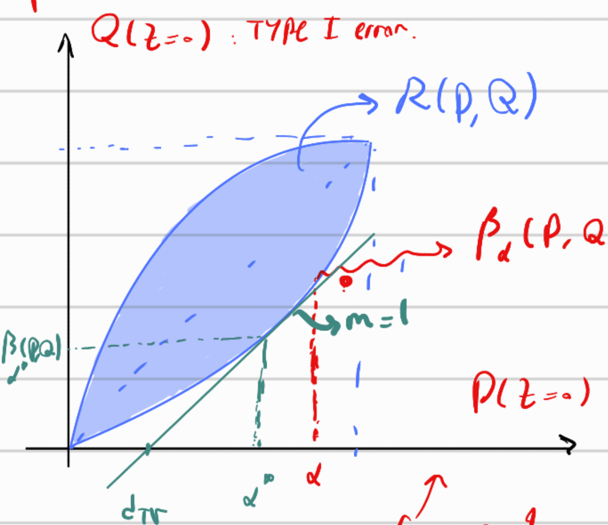
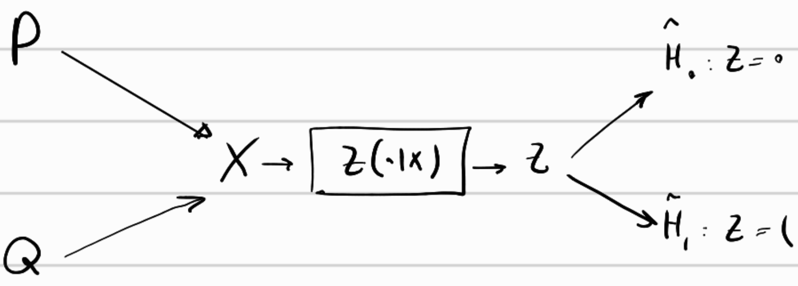


Figure 1.



a)

$$\sup_{Z(\cdot|x)} \{P(Z=0) - Q(Z=0)\} = \sup_{Z(\cdot|x)} \left\{\sum_x P(x) Z(0|x) - \sum_x Q(x) Z(0|x)\right\}$$

$$= \sup_{Z(\cdot|x)} \sum_x Z(0|x)\left[P(x) - Q(x)\right] = \sum_x \left[P(x) - Q(x)\right]_+ = d_{TV}(P,Q) = \sup_{x,y\in R(P,Q)} \{x-y\}$$

$$= \sup_{0<x<1} \sup_{y:(x,y)\in R(P,Q)} \{x-y\} = \sup_{0<x<1} \left\{x - \inf_{y:(x,y)\in R(P,Q)} y\right\} = \sup_{0<x<1} \{x - \beta_x(P,Q)\}$$

$$\Rightarrow \boxed{d_{TV}(P,Q) = \sup_{0<\alpha<1} \{\alpha - \beta_\alpha(P,Q)\} = \sup_{A\in\mathcal{F}} \{P(A) - Q(A)\}}$$

as obvious from figure 1, $\boxed{d_{TV} \text{ is the } x \text{ intersection}}$ of the lowest intersection of $R(P,Q)$ with lines with unit slope.

b)

$$P_E = \inf_{P_{Z|X}} \pi_0 \pi_{1|0} + \pi_1 \pi_{0|1} = \inf_{Z(\cdot|x)} \pi_0 P(Z=1) + \pi_1 Q(Z=0) = \inf_{Z(\cdot|x)} \pi_0 - \pi_0 P(Z=0) + \pi_1 Q(Z=0)$$

$$= \pi_0 - \sup_{Z(\cdot|x)} \{\pi_0 P(Z=0) - \pi_1 Q(Z=0)\} = \pi_0 - \sup_{Z(\cdot|x)} \left\{\pi_0 \sum_x P(x) Z(0|x) - \pi_1 \sum_x Q(x) Z(0|x)\right\}$$

1 when positive, 0 otherwise.

$$= \pi_0 - \sup_{Z(\cdot|x)} \left\{\sum_x Z(0|x)\left[\pi_0 P(x) - \pi_1 Q(x)\right]\right\} \overset{\text{I}}{=} \pi_0 - \sum_x \left[\pi_0 P(x) - \pi_1 Q(x)\right]_+$$

I $\Rightarrow$ $\boxed{\text{ideal test:}}$ $\begin{cases} \dfrac{P(x)}{Q(x)} \geq \dfrac{\pi_1}{\pi_0} \longrightarrow H_0 \\ \dfrac{P(x)}{Q(x)} < \dfrac{\pi_1}{\pi_0} \longrightarrow H_1 \end{cases}$ which is deterministic! / $P_{E,min} = \frac{1}{2}\left(1 - \sum |\pi_0 P(x) - \pi_1 Q(x)|\right)$

$$P_{E,min} = \pi_0 - \sum_x \left[\pi_0 P(x) - \pi_1 Q(x)\right]_+ = \pi_0 - \frac{1}{2}\sum_x |\pi_0 P(x) - \pi_1 Q(x)| + \cancel{\pi_0 P(x)} - \pi_1 \cancel{Q(x)} = \frac{\pi_0}{2} + \frac{\pi_1}{2} + \frac{1}{2}\sum |$$

c)

let $P_{Z|X}$ be the ideal test: $Z(a|x) = \mathbb{1}\left(\log\left(\frac{P(x)}{Q(x)}\right) > T\right)$

we showed that $T = \log\left(\frac{P(x)}{Q(x)}\right)$ is a measurable statistic.

$\hookrightarrow D_{KL}(P_X \| Q_X) = D_{KL}(P_T \| Q_T)$

$\Rightarrow P(Z=0) = P(T > T_\alpha) = \alpha, \quad Q(Z=0) = Q(T > \tilde{T}_\alpha), \quad T_\alpha = \log(\gamma_\alpha)$

we know that: supp-hyperplane $\Rightarrow \sup_{P_{Z|X}} P(\underbrace{Z=0}_{\alpha}) - \gamma Q(\overbrace{Z=0}^{\beta}) \Rightarrow \frac{1}{\gamma_\alpha} = \frac{d\beta_\alpha}{d\alpha}$

$\Rightarrow \boxed{T_\alpha = -\log\left(\frac{d\beta_\alpha}{d\alpha}\right)}$

Radon-Nikodym derivative

$D_{KL}(P_T \| Q_T) = E_T\left[\log\left(\frac{dP_T(t)}{dQ_T(t)}\right)\right] \qquad \frac{dP_T}{dQ_T} = \frac{d\mathbb{P}_{P_T}}{d\mathbb{P}_{Q_T}}$

Radon-Nikodym derivative

$P(T > T_\alpha) = \alpha \Rightarrow \frac{d\mathbb{P}_{P_T}}{dT} = \frac{d\alpha}{dT}, \quad \frac{d\mathbb{P}_{Q_T}}{dT} = \frac{d\beta_\alpha}{dT} \Rightarrow \frac{dP_T}{dQ_T} = \frac{d\alpha}{d\beta_T} = \gamma_\alpha$

$\Rightarrow D_{KL}(P_X \| Q_X) = D_{KL}(P_T \| Q_T) = E_{T \sim \frac{d\alpha}{dT}}[\log(\gamma_\alpha)] = \int_{-\infty}^{\infty} \frac{d\alpha}{\cancel{dT}} \log(\gamma_\alpha) \cancel{dT}$

$= \int_0^1 \log(\gamma_\alpha)\, d\alpha = -\int_0^1 T_\alpha\, d\alpha = \boxed{-\int_0^1 \log\left(\frac{d\beta_\alpha}{d\alpha}\right) d\alpha = D_{KL}(P_X \| Q_X)}$

basically, a change of measure

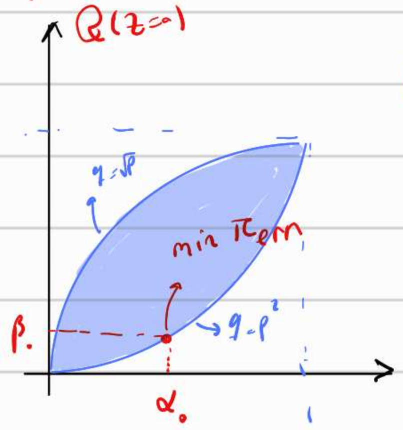
algebraically:

$\alpha = P(Z=0) = \sum_x \mathbb{1}\left(\log\frac{P(x)}{Q(x)} > T_\alpha\right) P(x) \Rightarrow \frac{d\alpha}{dT} = \sum_x \delta\left(\log\frac{P(x)}{Q(x)} - T\right) P(x)$

$-\int_0^1 \log\left(\frac{d\beta}{d\alpha}\right) d\alpha = \int_0^1 T_\alpha\, d\alpha = \int_T T_\alpha \frac{d\alpha}{dT_\alpha} dT_\alpha = \int_{-\infty}^{\infty} T_\alpha \sum_x \delta\left(\log\left(\frac{P(x)}{Q(x)}\right) - T\right) P(x)\, dT_\alpha$

$= \sum_x P(x) \underbrace{\int_{-\infty}^{\infty} T_\alpha \delta\left(\log\left(\frac{P(x)}{Q(x)}\right) - T\right) dT}_{\log\left(\frac{P(x)}{Q(x)}\right)} = \sum P(x) \log\left(\frac{P(x)}{Q(x)}\right) = \boxed{D_{KL}(P\|Q) = -\int_0^1 \log\left(\frac{d\beta}{d\alpha}\right) d\alpha}$

Problem 5)



$$\beta_\alpha(P,Q) = \inf_{P(Z=0)\geq 1-\alpha} Q(Z=0) = \alpha^2 , \quad \min_{P_{Z|X}} \overbrace{\pi_0\pi_{1|0}+\pi_1\pi_{0|1}}^{\pi_{err}}$$

from Neyman-person: $\log\left(\frac{P(x)}{Q(x)}\right)$: log-likelihood ratio.

$$\Rightarrow \begin{cases} \frac{P(x)}{Q(x)} > \gamma \rightsquigarrow H_0 \\ \quad " \quad < \gamma \rightsquigarrow H_1 \\ \quad " \quad = \gamma \Rightarrow \text{toss coin! with } \tau \end{cases}$$

Neyman-Pearson:

a)

$$\alpha = P(Z=0) = P\left(\frac{P(x)}{Q(x)} > \gamma\right) + \tau P\left(\frac{P(x)}{Q(x)} = \gamma\right) \quad : \exists \gamma, \tau$$

$$\beta_\alpha = Q(Z=0) = Q\left(\frac{P(x)}{Q(x)} > \gamma\right) + \tau Q\left(\frac{P(x)}{Q(x)} = \gamma\right) = \alpha^2$$

: $\alpha = \alpha(\gamma,\tau)$, $\beta = \beta(\gamma,\tau)$: keeping γ constant, changing τ:

$\frac{\Delta\alpha}{\Delta\tau} = cte$, $\frac{\Delta\beta}{\Delta\tau} = cte$, $\frac{d\beta}{d\alpha} = 2\alpha \neq cte \Rightarrow$ ✗ unless $\Delta\alpha = 0$, $\Delta\beta = 0$

$\frac{d^2\beta_\alpha}{d\alpha^2} > 0 \;\; \forall\alpha \Rightarrow$ no need for chance move $\boxed{P\left(\frac{P(x)}{Q(x)} = \gamma\right) = 0}$

↳ contact point is unique. $\rightarrow \boxed{\tau \text{ can be anything} \Rightarrow \tau = 0}$

b)

α_0, β_0 : $\underset{P_{Z|X}}{\operatorname{argmin}}\; \pi_1\pi_{0|1} + \pi_0\pi_{1|0} \xrightarrow{\text{Problem 4}} P^*_{Z|X} = \begin{cases} \frac{P(x)}{Q(x)} \geq \frac{\pi_1}{\pi_0} \rightarrow H_0 \\ \quad " \quad \leq \frac{\pi_1}{\pi_0} \rightarrow H_1 \\ \quad " \quad = \quad " \rightarrow \text{random} \end{cases}$

→ in log form: $\boxed{\tau = \log\left(\frac{\pi_1}{\pi_0}\right)}$

$\Rightarrow \boxed{\gamma = \frac{\pi_1}{\pi_0}}$, $\sup_{(x,y)\in R(P,Q)} x - \gamma y \Rightarrow x - \gamma y$ is adjacent to $R(P,Q)$

$$\Rightarrow \frac{1}{\gamma} = \left.\frac{d\beta_\alpha}{d\alpha}\right|_{\alpha_0} = 2\alpha_0 \Rightarrow \alpha_0 = \frac{1}{2\gamma} \Rightarrow \beta_{\alpha_0} = \frac{1}{4\gamma^2} \Rightarrow \boxed{\alpha = \frac{\pi_0}{2\pi_1}, \; \beta = \frac{\pi_0^2}{4\pi_1^2}}$$

c)

$\alpha, \beta \leq 1 \Rightarrow \alpha \leq 1$ suffices $\Rightarrow \frac{\pi_0}{2\pi_1} \leq 1 \Rightarrow \boxed{\pi_0 \leq 2\pi_1}$

Problem 6)

(1

we saw that: $P_{opt}(\hat{H} \neq H) = \frac{1}{2}(1 - d_{TV}(P_n, Q_n)) = \inf_{P_{Z|X}} \frac{1}{2}P(Z=1) + \frac{1}{2}Q(Z=0)$

α: type I error, β: type II error $\Rightarrow \alpha + \beta \geq \inf_{P_{Z|X}} \alpha + \beta = 1 - d_{TV}(P_n, Q_n)$

$\Rightarrow \alpha + \beta \geq 1 - d_{TV}(P_n, Q_n)$

$\sup_{A_n \in F_n} P_n(A_n) - Q_n(A_n) = \sup_{A_n} \sum_x (P(x) - Q(x)) A_n(x) = \sum_x [P(x) - Q(x)]_+ = d_{TV}(P_n, Q_n)$

$\Rightarrow \forall A_n \in F_n : P_n(A_n) - Q_n(A_n) \leq \sup_{A_n \in F_n} \dots = d_{TV}(P_n, Q_n)$

$\Rightarrow \alpha + \beta \geq 1 - d_{TV}(P_n, Q_n) \geq 1 - P(A_n) + Q(A_n) = P(A_n^c) + Q(A_n)$

$\Rightarrow$ $\boxed{\text{for all test, } \forall A_n \in F_n : \alpha + \beta \geq P_n(A_n^c) + Q_n(A_n)}$

$\mathbb{P} \triangleleft \mathbb{Q} \Rightarrow Q_n(A_n) \xrightarrow{n \to \infty} 0 \Rightarrow P_n(A_n) \xrightarrow{n \to \infty} 0 \Rightarrow P_n(A_n^c) \xrightarrow{n \to \infty} 1$

thus : $P_n(A_n^c) + Q_n(A_n) \xrightarrow{n \to \infty} 1$, similiar scenerio for $\mathbb{Q} \triangleleft \mathbb{P}$

thus if any of the probabilities are zero, $P_n(A_n) + Q_n(A_n) \neq 0$

thus we know: $\boxed{\forall A_n \in F_n : P_n(A_n^c) + Q_n(A_n) > 0 \Rightarrow \alpha + \beta > 0}$

(2

$\langle f, g \rangle = \mathbb{E}_{X \sim Q_n}[f(x) g(x)] : f, g : S_n \to \mathbb{R}$

$L_n = \frac{dP_n}{dQ_n}$, $\|L_n\|^2 = \langle L_n, L_n \rangle = \mathbb{E}_{X \sim Q_n}[(\frac{dP_n}{dQ_n}(x))^2] = \int (\frac{dP_n}{dQ_n}(\omega))^2 dQ_n(\omega)$

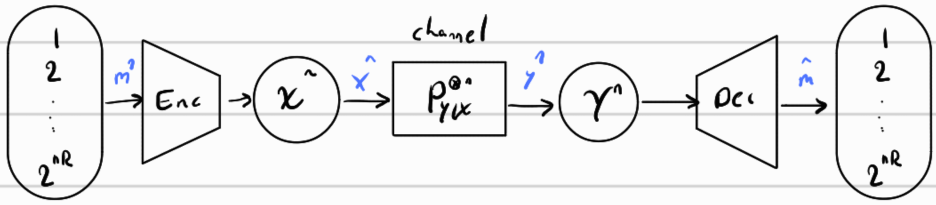
$\Rightarrow \|L_n\|^2 = \int \frac{dP_n}{dQ_n}(\omega) dP_n(\omega) < \infty \Rightarrow \|L_n\|^2 = \mathbb{E}_{X \sim P_n}[\frac{dP_n}{dQ_n}(X)] \geq t P_n[\frac{dP_n}{dQ_n}(X) > t] : \forall t > 0$

$\Rightarrow \exists c < \infty : P_n[\frac{dP_n}{dQ_n}(X) > t] \leq \frac{c}{t} \Rightarrow \lim_{t \to \infty} P_n[\frac{dP_n}{dQ_n}(X) > t] = 0$

assuming $\mathbb{P} \not\triangleleft \mathbb{Q} \Rightarrow \exists A_n \in F_n, \alpha > 0 : P_n(A_n) \geq \alpha,\ Q_n(A_n) \longrightarrow 0 \Rightarrow$

$\frac{P_n}{Q_n}(A_n) \longrightarrow \infty$, $P_n(A_n) \geq \alpha > 0 \Rightarrow \lim_{t \to \infty} P_n[\frac{dP_n}{dQ_n}(X) > t] > 0 \Rightarrow$ $\boxed{\times \Rightarrow \mathbb{P} \triangleleft \mathbb{Q}}$

Problem 7) $P_{\hat m|y}$



$$\mathbb{P}[\hat M = \hat m] = \frac{P^{\otimes n}_{Y|X}(y^n|\hat X(\hat m))}{\sum_m P^{\otimes n}_{Y|X}(y^n|\hat x(m))} \quad , \quad \mathbb{E}_{Enc}[\mathbb{P}[\hat M = M]] = \mathbb{E}_{Enc}\Big[\sum_{m,y^n} P_M(m)\, P^{\otimes n}_{Y|X}(y^n|\hat x(m))\, P_{\hat m|Y}(m|y^n)\Big]$$

a) $P_M(m) = \frac{1}{|M|}$ $\forall m$, codes are symmetric $\Rightarrow \mathbb{E}[f(\hat x(m))] = \mathbb{E}[f(\hat x(1))]$

$$\Rightarrow \mathbb{E}_{Enc}[\mathbb{P}(\hat M = M)] = \mathbb{E}_{Enc}\Big[\sum_{m,y^n} P_M(m)\, P^{\otimes n}_{Y|X}(y^n|\hat x(m))\, P_{\hat m|Y}(m|y^n)\Big] = \mathbb{E}_{Enc}\Big[\frac{1}{|M|}\sum_{m,y^n} P^{\otimes n}_{Y|X}(y^n|\hat x(m))\, P_{\hat m|Y}(m|y^n)\Big]$$

$$= \mathbb{E}_{Enc}\Big[\frac{1}{|M|}\underbrace{\sum_{m,y^n} \overbrace{P^{\otimes n}_{Y|X}(y^n|\hat x(1))\, P_{\hat m|Y}(1|y^n)}^{\neq f(m)}}_{|M|\times\cdots}\Big] = \boxed{\mathbb{E}_{Enc}\Big[\sum_{y^n} P^{\otimes n}_{Y|X}(y^n|\hat X(1))\, P_{\hat m|Y^n}(1|y^n)\Big] = \mathbb{E}_{Enc}[\mathbb{P}[\hat M = M]]}$$

b)

$$\mathbb{E}[\mathbb{P}(\hat M = M)] = \mathbb{E}_{Enc}\Big[\sum_{y^n} P^{\otimes n}_{Y|X}(y^n|\hat x(1))\, P_{\hat M|y^n}(1|y^n)\Big] = \mathbb{E}_{Enc}\Big[\sum_{y^n} P^{\otimes n}_{Y|X}(y^n|\hat x(1))\,\frac{P^{\otimes n}_{Y|X}(y^n|\hat x(1))}{\sum_m P^{\otimes n}_{Y|X}(y^n|\hat x(m))}\Big]$$

$$= \sum_{y^n}\mathbb{E}_{Enc}\Big[\frac{P^{\otimes n}_{Y|X}(y^n|\hat X(1))^2}{P^{\otimes n}_{Y|X}(y|\hat x(1)) + \sum_{m=2}^{2^{nR}} P^{\otimes n}_{Y|X}(y|\hat x(m))}\Big] = \sum_{y^n}\mathbb{E}_{X(1)}\big[\mathbb{E}_{X(2)\dots X(2^{nR})}[\dots]\big] \geq$$ using yensen on convex $\frac{1}{x_1+\dots x_n}$

$$\geq \sum_{y^n}\mathbb{E}_{X(1)}\Big[\frac{P^{\otimes n}_{Y|X}(y^n|\hat x(1))^2}{P^{\otimes n}_{Y|X}(y|\hat x(1)) + \mathbb{E}_{X(2)\dots X(2^{nR})}\Big[\sum_{m=2}^{2^{nR}} P^{\otimes n}_{Y|X}(y|\hat x(m))\Big]}\Big] = \sum_{y^n}\mathbb{E}_{X(1)}\Big[\frac{P^{\otimes n}_{Y|X}(y^n|\hat x(1))^2}{P^{\otimes n}_{Y|X}(y|\hat x(1)) + \sum_{m=2}^{2^{nR}}\mathbb{E}_{X(m)}\big[P^{\otimes n}_{Y|X}(y|\hat x(m))\big]}\Big]$$

$$\mathbb{E}_{X(m)}[P^{\otimes n}_{Y|X}(y|\hat x(m))] = P^{\otimes n}_{Y}(y^n) \Rightarrow \mathbb{E}_{Enc}[\mathbb{P}(\hat m = M)] \geq \sum_{y^n}\mathbb{E}_{X(1)}\Big[\frac{P^{\otimes n}_{Y|X}(y^n|\hat x(1))^2}{P^{\otimes n}_{Y|X}(y^n|\hat x(1)) + (2^{nR}-1)P^{\otimes n}_{Y}(y^n)}\Big]$$

$$\hat X(1)\in\mathcal{X}^n \Rightarrow \mathbb{E}_{Enc}[\mathbb{P}(\hat m = M)] \geq \sum_{x^n,y^n}\overbrace{P^{\otimes n}_X(x^n)\, P^{\otimes n}_{Y|X}(y^n|x^n)}^{P^{\otimes n}_{Y,X}(y^n,x^n)}\,\frac{P^{\otimes n}(y^n|x^n)}{P^{\otimes n}_{Y|X}(y^n|x^n) + (2^{nR}-1)P^{\otimes n}_Y(y^n)} = \mathbb{E}_{x^n,y^n}\Big[\frac{P^{\otimes n}(y^n|x^n)}{P^{\otimes n}_{Y|X}(y^n|x^n) + (2^{nR}-1)P^{\otimes n}_Y(y^n)}\Big]$$

$\hat X(1)\to x^n$

Yensen here would give an obvious result. $\frac{P^{\otimes n}(y^n|x^n)}{P^{\otimes n}_{Y|X}(y^n|x^n) + (2^{nR}-1)P^{\otimes n}_Y(y^n)} = \Big(1 + (2^{nR}-1)\Big(\frac{P^{\otimes n}_Y(y^n)}{P^{\otimes n}_{Y|X}(y^n|x^n)}\Big)\Big)^{-1}$

$$\Rightarrow \boxed{\mathbb{E}_{Enc}[\mathbb{P}(\hat m = M)]} \geq \mathbb{E}_{x^n,y^n}\Big[\frac{1}{1 + (2^{nR}-1)\cdot 2^{-\ell_2(\dots)}}\Big] \geq \boxed{\mathbb{E}_{x^n,y^n}\Big[\frac{1}{1 + 2^{nR}\cdot 2^{-i(x^n;y^n)}}\Big] \ , \ i(x^n;y^n) = \log_2\Big(\frac{P^{\otimes n}_{Y|X}(y^n|x^n)}{P^{\otimes n}_Y(y^n)}\Big)}$$

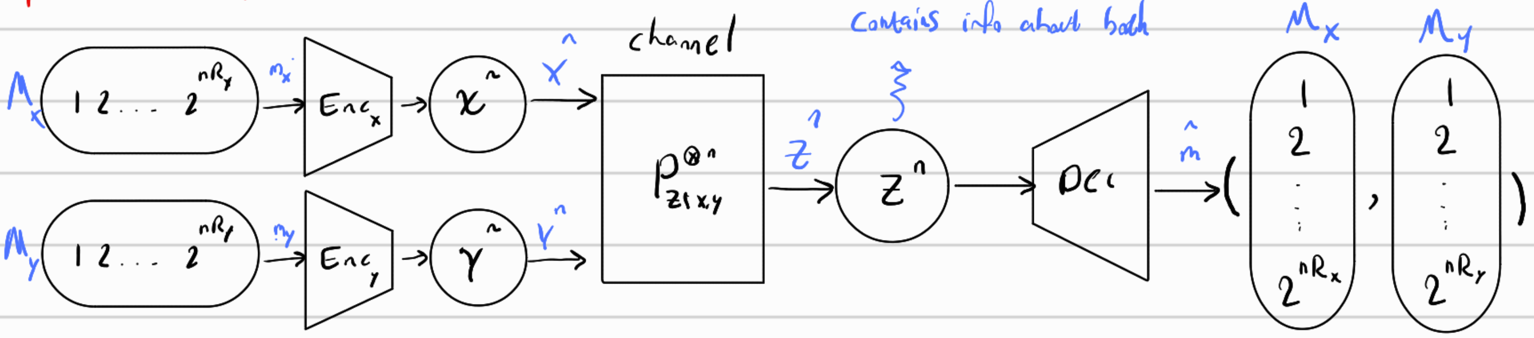
a bit tighter

c) $i(x^n;y^n) = \log_2\Big(\frac{P^{\otimes n}_{Y|X}(y^n|x^n)}{P^{\otimes n}_Y(y^n)}\Big) = \log_2\Big(\frac{\prod P_{Y|X}(y_i|x_i)}{\prod P_Y(y_i)}\Big) = \sum_{i=1}^n \log_2\Big(\frac{P_{Y|X}(y_i|x_i)}{P_Y(y_i)}\Big) \xrightarrow{\text{large } n} n\,\mathbb{E}_{X,Y}\Big[\log_2\Big(\frac{P_{Y|X}(y|x)}{P_Y(y)}\Big)\Big] = n\,I(X;Y)$

$$\Rightarrow \text{large } n:\ \mathbb{E}_{Enc}[\mathbb{P}(\hat m = M)] \geq \mathbb{E}_{x^n,y^n}\Big[\frac{1}{1 + 2^{nR - i(x^n;y^n)}}\Big] \overset{\text{large } n}{\approx} \mathbb{E}_{x^n,y^n}\Big[\frac{1}{1 + 2^{n(R - I(X;Y))}}\Big] = \boxed{\frac{1}{1 + 2^{n(R - I(X;Y))}}}\ \begin{cases} R > I(X;Y) \to 0 \\ \boxed{R < I(X;Y) \to 1}\end{cases}$$

$\Rightarrow$ $\boxed{R < I(X;Y) \Rightarrow \mathbb{P}(\hat M = M) \to 1}$ $\Rightarrow$ less symbols, perfect recovery is possible.

Problem 8) multi-Access : (n, R_x, R_y)



(a) to find bounds for possible rates, we will use a random relation as the last question. as in X, Y will both be random mapings from $M_x \to \hat{X}$ and $M_y \to \hat{Y}$.

$$\Rightarrow \boxed{\forall m_x \in [1 : 2^{nR_x}] : X(m_x) \in_R \mathcal{X}^n, \ \forall m_y \in [1 : 2^{nR_y}] : Y(m_y) \in_R \mathcal{Y}^n}$$

(b) to find similiar bounds, we will follow a similiar process as last question.

we use a variational decoder inspired by the last question to derive bounds.

$$\hat{m} = (\hat{m}_x, \hat{m}_y) : \mathbb{P}[\hat{M} = (\hat{M}_x, \hat{M}_y) = (\hat{m}_x, \hat{m}_y)] = \frac{P^{\otimes n}_{Z|(X,Y)}[\hat{z}^n | X(\hat{m}_x), Y(\hat{m}_y)]}{\sum_{m_x, m_y} P^{\otimes n}_{Z|(X,Y)}[\hat{z}^n | X(\hat{m}_x), Y(\hat{m}_y)]}$$

$$\Rightarrow \mathbb{P}[\hat{M}_x = \hat{m}_x] = \sum_{m_y} \mathbb{P}[(\hat{M}_x, \hat{M}_y) = (\hat{m}_x, m_y)]$$, similiar relation for $\mathbb{P}[\hat{M}_y = \hat{m}_y]$

similiarly : $$\mathbb{P}[\hat{M} = M] = \sum_{m, \hat{z}^n} P_M(m) P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | X(m_x), Y(m_y)) P_{\hat{m}|\hat{z}^n}(m | \hat{z}^n) , \ P_M(m) = P_{M_x}(m_x) P_{M_y}(m_y)$$

we will follow a similiar process, acknowledging similiarities with the last problem

$$Enc = \{X(m) | m \in M_x\} \otimes \{Y(m) | m \in M_y\} , \quad \mathbb{E}_{X(1), \dots, X(2^{nR_x}), Y(1), \dots, Y(2^{nR_y})} \longrightarrow \mathbb{E}_{Enc}$$

$$\mathbb{E}_{Enc}[\mathbb{P}(\hat{M} = M)] = \mathbb{E}_{Enc}\left[\sum_{m, \hat{z}^n} P_M(m) P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | \hat{X}(m_x), \hat{Y}(m_y)) P_{\hat{m}|\hat{z}^n}(m | \hat{z}^n)\right]$$

by symmetry $\xrightarrow[\text{similiarly}]{}$ $$\mathbb{E}_{Enc}[\mathbb{P}(\hat{M} = M)] = \mathbb{E}_{Enc}\left[\sum_{\hat{z}^n} P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | \hat{X}(1), \hat{Y}(1)) P_{\hat{m}|\hat{z}^n}(m | \hat{z}^n)\right] \overset{\text{plugging in } P_{\hat{m}|\hat{z}^n}}{=}$$

$$\sum_{\hat{z}^n} \mathbb{E}_{Enc}\left[\frac{P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | \hat{X}(1), \hat{Y}(1))^2}{\sum_{m_x, m_y} P_{Z|(X,Y)}(\hat{z}^n | \hat{X}(m_x), \hat{Y}(m_y))}\right]$$, decompose into parts. $\hat{X}(1), \hat{Y}(1)$

$$= \sum_{\hat{z}^n} \mathbb{E}_{Enc}\left[\frac{P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | \hat{X}(1), \hat{Y}(1))^2}{P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | \hat{X}(1), \hat{Y}(1)) + \sum_{m_x} P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | \hat{X}(m_x), \hat{Y}(1)) + \sum_{m_y} P^{\otimes n}_{Z|(X,Y)}(\hat{z}^n | \hat{X}(1), \hat{Y}(m_y)) + \sum_{m_x=2, m_y=2}^{2^{nR_x}, 2^{nR_y}} P_{Z|(X,Y)}(\hat{z}^n | \hat{X}(m_x), \hat{Y}(m_y))}\right]$$

using Jensen, similiarly : $\mathbb{E}_{Enc \setminus X(1), Y(1)}\left[\mathbb{P}^{\otimes n}_{Z|(X,Y)}(z^n|(x^n(m_x), y^n(1)))\right] = P^{\otimes n}_{Z|Y}(\hat{z}|\hat{y}(1))$

$\mathbb{E}_{Enc \setminus X(1), Y(1)}\left[\mathbb{P}^{\otimes n}_{Z|(X,Y)}(z^n|(x^n(1), y^n(m_y)))\right] = P^{\otimes n}_{Z|X}(\hat{z}|\hat{x}(1))$, $\mathbb{E}_{Enc \setminus X(1), Y(1)}\left[\mathbb{P}^{\otimes n}_{Z|(X,Y)}(z^n|(x^n(m_x), y^n(m_y)))\right] = P^{\otimes n}_{Z}(\hat{z})$

taking expectation to denominator :

$$\mathbb{E}_{Enc}[\mathbb{P}(\hat{m}=m)] \geq \sum_{z^n} \mathbb{E}_{X(1),Y(1)}\left[\frac{P^{\otimes n}_{Z|(X,Y)}(z^n|x^n(1), y^n(1))^2}{P^{\otimes n}_{Z|(X,Y)}(z^n|x^n(1), y^n(1)) + (2^{nR_x}-1)P^{\otimes n}_{Z|Y}(z^n|Y(1)) + (2^{nR_y}-1)P^{\otimes n}_{Z|X}(z^n|X(1)) + (2^{n(R_x+R_y)} - 2^{nR_x} - 2^{nR_y} - 1)P^{\otimes n}_{Z}(z^n)}\right]$$

$$\geq \sum_{z^n} \mathbb{E}_{X(1),Y(1)}\left[\frac{P^{\otimes n}_{Z|(X,Y)}(z^n|x^n(1), y^n(1))^2}{P^{\otimes n}_{Z|(X,Y)}(z^n|x^n(1), y^n(1)) + 2^{nR_x}P^{\otimes n}_{Z|Y}(z^n|Y(1)) + 2^{nR_y}P^{\otimes n}_{Z|X}(z^n|X(1)) + 2^{n(R_x+R_y)}P^{\otimes n}_{Z}(z^n)}\right]$$

$$= \sum_{z^n, x^n, y^n} \underbrace{P^{\otimes n}_X(x^n) P^{\otimes n}_Y(y^n) P^{\otimes n}_{Z|(X,Y)}(z^n|(x^n, y^n))}_{P^{\otimes n}_{Z,Y,X}(z^n, y^n, x^n)} \frac{P^{\otimes n}_{Z|(X,Y)}(z^n|(x^n, y^n))}{P^{\otimes n}_{Z|(X,Y)}(z^n|x^n, y^n) + 2^{nR_x}P^{\otimes n}_{Z|Y}(z^n|y^n) + 2^{nR_y}P^{\otimes n}_{Z|X}(z^n|x^n) + 2^{n(R_x+R_y)}P^{\otimes n}_{Z}(z^n)}$$

$$= \mathbb{E}_{X^n, Y^n, Z^n}\left[\frac{1}{1 + 2^{nR_x} \cdot 2^{-i(X;Z|Y)} + 2^{nR_y} \cdot 2^{-i(Y;Z|X)} + 2^{n(R_x+R_y)} \cdot 2^{-i(X,Y;Z)}}\right]$$

where : $i(X;Z|Y) = \log_2\left(\frac{P^{\otimes n}_{Z|Y}(z^n|y^n)}{P^{\otimes n}_{Z|(X,Y)}(z^n|(x^n, y^n))}\right)$, $i(X,Y;Z) = \log_2\left(\frac{P^{\otimes n}_{Z}(z^n)}{P^{\otimes n}_{Z|(X,Y)}(z^n|x^n, y^n)}\right)$

similiarly, for large n : $i(X;Z|Y) \rightsquigarrow n\, I(X;Z|Y)$, $i(X,Y;Z) \rightsquigarrow n\, I(X,Y;Z)$

by law of large numbers, thus the expected value drops!

$$\Rightarrow \mathbb{E}_{Enc}[\mathbb{P}(\hat{m}=m)] \geq \left(1 + 2^{n(R_x - I(X;Z|Y))} + 2^{n(R_y - I(Y;Z|X))} + 2^{n(R_x+R_y - I(X,Y;Z))}\right)^{-1}$$

if either of these exponents are positive $\rightarrow$ denominator goes to infinity $\rightarrow 0$

so we are guaranteed to have perfect codes if all are negative

$$\Rightarrow \boxed{R_x < I(X;Z|Y), \quad R_y < I(Y;Z|X), \quad R_x + R_y < I(X,Y;Z)}$$

which makes sense, as it is normal coding, viewing from perspective of either message, with one effectively concatenating m_x, m_y.

$Z|Y$: not caring about Y, what rate is R_x allowed?