Problem 1.

$$U_\theta \sim \mathrm{Unif}[\theta, \theta+1] \quad , \quad \ell(\theta, \hat{\theta}) = \mathbb{E}_{U \sim P_\theta}\left[(\theta - \hat{\theta}(U))\right]$$

(1

$$\min_{\hat{\theta}} \max_{\theta} \mathbb{E}_{U^n \sim P_\theta^{\otimes n}}\left[\|\theta - \hat{\theta}(U^n)\|^2\right] \geqslant \Phi(\delta)\,\mathbb{P}[\hat{I} \neq I] \geqslant \frac{1}{2}\Phi(\delta)\left[1 - d_{TV}(P_{\theta_1}^{\otimes n}, P_{\theta_2}^{\otimes n})\right]$$

$|\theta_1 - \theta_2| \geqslant 2\delta$, $\Phi(\delta) \propto \delta^2$ (as we are using ℓ_2-norm)

↳ let $\theta_2 = \theta_1 + 2\delta$

$TV(P_{\theta_1}^{\otimes n}, P_{\theta_2}^{\otimes n})$, $D_{KL}(P_{\theta_1}, P_{\theta_2}) \to \infty \Rightarrow$ not useful

$$TV(P_{\theta_1}^{\otimes n}, P_{\theta_2}^{\otimes n}) \leqslant \sum_{i=1}^{n} TV(P_{\theta_1}, P_{\theta_2}) = n\,TV(P_{\theta_1}, P_{\theta_1+2\delta}) = n \times \frac{1}{2} \times 4\delta = 2n\delta$$

$$\Rightarrow \min_{\hat{\theta}} \max_{\theta} \mathbb{E}\left[\|\theta - \hat{\theta}\|^2\right] \geqslant c\,\delta^2(1 - 2n\delta) \xrightarrow{\delta = \frac{1}{4n}} (1-2n\delta) = \frac{1}{2},\ \delta^2 = \frac{1}{4n^2}$$

$$\Rightarrow \min_{\hat{\theta}} \max_{\theta} \mathbb{E}\left[\|\theta - \hat{\theta}\|^2\right] \geqslant \frac{c}{4n^2} \times \frac{1}{2} \overset{C = \frac{c}{8}}{\Longrightarrow} \boxed{\min_{\hat{\theta}} \max_{\theta} \mathbb{E}\left[(\theta - \hat{\theta})^2\right] \geqslant \frac{C}{n^2}}$$

(2

simple estimation : $\hat{\theta} = \max\{U_i\}_{i \in [N]} - 1$

$$\mathbb{E}\left[(\theta + 1 - \max\{U_i\})^2\right] = \mathbb{E}\left[(\theta + 1 - Y)^2\right] . \quad Y = \max\{U_i\}_{i \in [N]}$$

$$\mathbb{P}[Y < x] = \mathbb{P}[\max\{U_i\} < x] = \mathbb{P}\left[\bigcap_{i=1}^{n} U_i < x\right] = \prod_{i=1}^{n} \mathbb{P}[U_i < x] = \prod_{i=1}^{n} \min\{1, x - \theta\}\mathbb{1}\{x > \theta\}$$

$$= \mathbb{1}\{x > \theta\} \min\{1, x - \theta\}^n \Rightarrow f_Y(x) = n(x-\theta)^{n-1}\mathbb{1}\{\theta < x < \theta + 1\}$$

$$\mathbb{E}\left[(\theta + 1 - Y)^2\right] = n\int_{\theta}^{\theta+1} (\theta + 1 - x)^2 (x - \theta)^{n-1}\,dx = n\int_{0}^{1} \alpha^{n-1}(\alpha - 1)^2\,d\alpha = n\left[\frac{1}{n+2} + \frac{1}{n} - \frac{2}{n+1}\right]$$

$$= \frac{2n}{n(n+1)(n+2)} \leqslant \frac{2}{n^2} \Rightarrow \boxed{\sup_{\theta} \mathbb{E}\left[(\theta - \hat{\theta})^2\right] \leqslant \frac{C}{n^2}}$$

(3 این باند حالت تعمیمی های ما را برای $U[a,b]$ پیدا می‌کردیم. و همچنین می‌بینیم که به

باند minimax رسیده پس احتمالا بهترین باند برای چنین مسئله ای است.

Problem 2)

(1

$$D_f(P_{MX}T_{\hat{M}|X} \| P_M Q_X T_{\hat{M}|X}) = \mathbb{E}_{X,M,\cancel{\hat{M}}}\left[f\left(\frac{P_{MX}(x,m)\cancel{t(\hat{m}|x)}}{P_M(m)q_X(x)\cancel{t(\hat{m}|x)}}\right)\right] = \mathbb{E}_{m,x}\left[f\left(\frac{dP_{MX}}{dP_M dQ_X}\right)\right]$$

$$= \boxed{D_f(P_{MX}\|P_M Q_X) = D_f(P_{MX}T_{\hat{M}|X}\|P_M Q_X T_{\hat{M}|X})}$$

(2

$$X \xrightarrow{T_{\hat{M}|X}} \hat{M} \quad : D_f(P_{MX}\|P_M Q_X) = D_f(P_{MX\hat{M}}\|P_M Q_{X\hat{M}}) \geq D_f(P_{M\hat{M}}\|P_M Q_{\hat{M}})$$

also obvious from DP inequality : $\boxed{D_f(P_{MX}\|P_M Q_X) \geq D_f(P_{M\hat{M}}\|P_M Q_{\hat{M}})}$

(3

$D_f(P_{M\hat{M}}\|P_M Q_{\hat{M}})$ → nonetheless, same channel

$P_{M\hat{M}} \rightsquigarrow M,\hat{M} \xrightarrow{\mathbb{1}[M\neq\hat{M}]} t \rightsquigarrow \mathrm{Ber}(P_E)$: $\mathbb{P}[M\neq M] = P_E$

$P_M Q_{\hat{M}} \rightsquigarrow M,\hat{M} \xrightarrow{\mathbb{1}[M\neq\hat{M}]} t \rightsquigarrow \mathrm{Ber}(1-\frac{1}{m})$: M is uniform $\Rightarrow \frac{1}{m}$: equality

DP : $\boxed{D_f(P_{M\hat{M}}\|P_M Q_{\hat{M}}) \geq D_f(\mathrm{Ber}(P_E)\|\mathrm{Ber}(1-\frac{1}{m}))}$

(4

$$\forall Q_X : D_f(P_{MX}\|P_M Q_X) \geq D_f(P_{M\hat{M}}\|P_M Q_{\hat{M}}) \geq D_f(\mathrm{Ber}(P_E)\|\mathrm{Ber}(1-\tfrac{1}{m}))$$

$$\Rightarrow \inf_{Q_X} D_f(P_{MX}\|P_M Q_X) \geq D_f(\mathrm{Ber}(P_E)\|\mathrm{Ber}(1-\tfrac{1}{m}))$$ ~ independent of Q_X

$$\forall Q_X : \max_m D_f(P_{X|M=m}\|Q_X) \geq \frac{1}{m}\sum_{i=1}^{m} D_f(P_{X|M=i}\|Q_X) = \mathbb{E}_M[D_f(P_{X|M}\|Q_X)]$$

$$= D_f(P_{X|M}\|Q_X|P_M) = D_f(P_{XM}\|Q_X P_M) \rightsquigarrow \inf_{Q_X}\max_m \cdots \geq \inf_{Q_X} D_f(P_{MX}\|P_M Q_X)$$

$\Rightarrow$

$$\boxed{\inf_{Q_X}\max_m D_f(P_{X|M=m}\|Q_X) \geq \inf_{Q_X} D_f(P_{MX}\|P_M Q_X) \geq D_f(\mathrm{Ber}(P_E)\|\mathrm{Ber}(1-\tfrac{1}{m}))}$$

Problem 3)

$\ell : \Theta \times \hat{\Theta} \to \mathbb{R}$, $\Theta_1 \subseteq \Theta$: $\ell(\theta_0, a) + \ell(\theta_1, a) \geq \Delta \quad \forall a \in \hat{\Theta}$, $\theta_1 \in \Theta_1$

(1

for fixed θ_0, θ_1,

given estimator $\hat{\theta}$, define test $\tilde{\theta} = \begin{cases} \theta_0 & \text{probability } \dfrac{\ell(\theta_1, \hat{\theta})}{\ell(\theta_1, \hat{\theta}) + \ell(\theta_0, \hat{\theta})} \\ \theta_1 & \quad \text{''} \quad \dfrac{\ell(\theta_0, \hat{\theta})}{\ell(\theta_1, \hat{\theta}) + \ell(\theta_0, \hat{\theta})} \end{cases}$

$$\mathbb{E}_{\theta_0}[\ell(\tilde{\theta}, \theta_0)] = \ell(\theta_0, \theta_1)\, \mathbb{E}_{\theta_0}\left[\frac{\ell(\theta_0, \hat{\theta})}{\underbrace{\ell(\theta_1, \hat{\theta}) + \ell(\theta_0, \hat{\theta})}_{\geq \Delta}}\right] \leq \frac{\ell(\theta_0, \theta_1)}{\Delta} \mathbb{E}_{\theta_0}[\ell(\theta_0, \hat{\theta})]$$

$$\mathbb{E}_{\theta_1}[\ell(\tilde{\theta}, \theta_1)] \leq \frac{\ell(\theta_0, \theta_1)}{\Delta} \mathbb{E}_{\theta_1}[\ell(\hat{\theta}, \theta_1)]$$

the same holds for θ_1 , $\mathbb{E}_{\pi}$: $\pi = \frac{1}{2}(\delta_{\theta_0} + \mu_{\theta_1})$

$$\frac{\ell(\theta_0, \theta_1)}{\Delta} \mathbb{E}[\ell(\theta, \hat{\theta})] \geq \mathbb{E}[\ell(\tilde{\theta}, \theta)] \geq \mathbb{P}[\hat{\theta} \neq \theta]\, \ell(\theta_1, \theta_0)$$

$$\mathbb{P}[\hat{\theta} \neq \theta] \geq \frac{1}{2}(1 - TV(P_{\theta_0}, P_{\theta_1})) \quad = \frac{1}{2}(1 - TV(P_{\theta_0} \| \mathbb{E}_{\mu_{\theta_1}} P_{\theta_1}))$$

for same fixed θ_1, now taking $\mathbb{E}_{\mu_{\theta_1}}$

$$\Rightarrow \frac{\ell(\cancel{\theta_0}, \theta_1)}{\Delta} \mathbb{E}[\ell(\theta, \hat{\theta})] \geq \ell(\cancel{\theta_0}, \theta_1) \frac{1}{2}(1 - TV(P_{\theta_0} \| \mathbb{E}_{\mu_{\theta_1}} P_{\theta_1}))$$

$$\Rightarrow \boxed{\mathbb{E}[\ell(\theta, \hat{\theta})] \geq \frac{\Delta}{2}\left[1 - TV(P_{\theta_0} \| \mathbb{E}_{\mu_{\theta_1}} P_{\theta_1})\right]}$$

(2

$$d_{TV}(P \| Q) \leq \sqrt{\tfrac{1}{2} D_{KL}(P \| Q)} \quad \& \quad \boxed{d_{TV}^2(P \| Q) \leq \frac{1}{2} D_{KL}(P \| Q) \leq \frac{1}{2} D_{\chi^2}(P \| Q)} \Rightarrow c \leq \frac{1}{2}$$

(3

$$\mathbb{E}_{\theta, \theta' \sim \mu}\left[\int \frac{P_\theta(x) P_{\theta'}(x)}{Q(x)} dx\right] = \mathbb{E}_{\theta, \theta' \sim \mu} \mathbb{E}_{x \sim Q}\left[\frac{P_\theta(x) P_{\theta'}(x)}{Q(x)^2}\right]$$

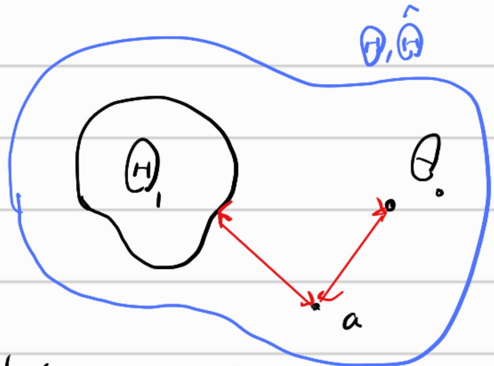
$$= \mathbb{E}_{x \sim Q} \mathbb{E}_{\theta, \theta' \sim \mu}\left[\frac{P_\theta(x) P_{\theta'}(x)}{Q^2(x)}\right] = \mathbb{E}_{x \sim Q}\left[\frac{\mathbb{E}_\mu[P_\theta]^2(x)}{Q(x)^2}\right] \quad : \quad P_\mu := \mathbb{E}_\mu[P_\theta]$$

$$\Rightarrow \mathbb{E}_{X\sim Q}\left[\frac{P_\mu^2(x)}{Q(x)^2}\right] = \mathbb{E}_{X\sim Q}\left[\frac{P_\mu^2(x) - Q(x)^2}{Q(x)^2}\right] + 1 = D_{\chi^2}(P_\mu \| Q) + 1$$

$$\Rightarrow \boxed{1 + D_{\chi^2}(\mathbb{E}_\mu[P_\theta], Q) = \mathbb{E}\left[\int \frac{P_\theta(x) P_{\theta'}(x)}{Q(x)} dx\right]}$$

$X_i = \theta_i + Z_i$, $Z_i \sim N(0,1)$, $L(\theta, T) = (T - \theta_{max})^2$ (4

$X_i \sim N(\theta_i, 1)$, $\Theta = \mathbb{R}$, $\hat{\Theta} = \mathbb{R}$



where Θ denotes the maximum

we can consider that $\max_i X_i$ is a sufficient statistic but we won't make that assumption

$$\forall a, \theta_0 \in \mathbb{R}, \theta_1 \in \Theta_1 \subseteq \mathbb{R}: \|\theta_0 - a\|^2 + \|\theta_1 - a\|^2 \geq \frac{1}{2}\|\theta_0 - \theta_1\|^2$$

$$\Rightarrow \ell(\theta_1, a) + \ell(\theta_0, a) \geq \frac{1}{2}\inf_{\theta_0, \theta_1}\|\theta_0 - \theta_1\|^2 = \Delta \qquad \forall \theta_0, \theta_1, a$$

$L(T, \theta) = \ell(T, \theta_{max})$ $\rightsquigarrow$ does not really matter $L(\theta_0, a) + L(\theta_1, a) \geq \Delta$

let: $\Theta_1 = [\theta_0 + \delta, \theta_0 + 2\delta]$, $\mu_{\theta_1} = \mathrm{Unif}\{\theta_1\}$, $\delta = \sqrt{2\Delta}$

$$\min_T \max_\theta \mathbb{E}[L(T, \theta)] \geq \frac{\Delta}{2}\left[1 - TV(P_{X|\theta_{max}=\theta_0} \| \mathbb{E}_\mu P_{X|\theta_{max}=\theta_1})\right]$$

$$TV(P_{X|\theta_0} \| \mathbb{E}_\mu P_{X|\theta_\mu})^2 \leq \frac{1}{2} D_\chi^2(P_{X|\theta_0} \| \mathbb{E}_\mu P_{X|\theta_\mu}) = \frac{1}{2}\left[\mathbb{E}_{\theta_1, \theta_2 \sim \mu}\left[\int \frac{P_{\theta_1} P_{\theta_2}}{P_{\theta_0}} dx\right] - 1\right]$$

$$\int_{\mathcal{X}} \frac{P_{\theta_1} P_{\theta_2}}{P_{\theta_0}} dx = \mathbb{E}_{X\sim P_{\theta_0}}\left[\frac{P_{\theta_1}(x) P_{\theta_2}(x)}{P_{\theta_0}(x)^2}\right] = \mathbb{E}_{X\sim P_{\theta_0}}\left[\exp\left\{\sum_{i=1}^n (x_i - \theta_1^i)^2 + (x_i - \theta_2^i)^2 - 2(x_i - \theta_0^i)^2\right\}\right]$$

$(x_i - \theta_0^i)^2 + (\theta_0^i - \theta_1^i)^2 + 2(x_i - \theta_0^i)(\theta_0^i - \theta_1^i)$

$$= \mathbb{E}_{X\sim P_{\theta_0}}\left[\exp\left\{\sum_{i=1}^n (\theta_0^i - \theta_1^i)^2 + (\theta_2^i - \theta_0^i)^2 + (x_i - \theta_0^i)\overbrace{\{2(\theta_0^i - \theta_1^i) + 2(\theta_0^i - \theta_2^i)\}}^{t_i}\right\}\right]$$

$P_{\theta_0} = N(\theta_0, I)$: $\mathbb{E}[e^{tX_i}] = e^{\mu t + \frac{\sigma^2 t^2}{2}}$, $X_i - \theta_0^i \sim N(0, I)$

$$= \exp\left\{ \sum_{i=1}^{n} (\theta_0^i - \theta_1^i)^2 + (\theta_2^i - \theta_0^i)^2 + \frac{1}{2}\left[2(\theta_0^i - \theta_1^i) + 2(\theta_0^i - \theta_2^i)\right]^2 \right\}$$

$$= \exp\left\{ \sum_{i=1}^{n} \underbrace{3(\theta_0^i - \theta_1^i)^2 + 3(\theta_0^i - \theta_2^i)^2 + 4(\theta_0^i - \theta_1^i)(\theta_0^i - \theta_2^i)}_{S_i} \right\}, \quad \inf |\theta_0 - \theta_1| = \delta$$

$$= \exp\left\{ 3\|\theta_0 - \theta_1^i\|_2^2 + 3\|\theta_0 - \theta_2^i\|_2^2 + 4\langle \theta_0 - \theta_1, \theta_0 - \theta_2 \rangle \right\}$$

θ_1, θ_2 : all zero except the one

$$\xrightarrow[\text{only one term with prob } \frac{1}{n}]{} \mathbb{E}_{\theta_1, \theta_2 \sim \mathcal{R}}\left[\exp(\cdots)\right] = \frac{1}{n} e^{10\delta^2} + \frac{n-1}{n} \times 1$$

$$\Rightarrow \mathbb{E}\left[\int \cdots\right] - 1 = \frac{1}{n}\left[e^{10\delta^2} - 1\right] = \frac{1}{n}\left[e^{20\Delta} - 1\right] \leq \frac{e^{20\Delta}}{n}$$

$$\Rightarrow \chi^2 \leq \frac{e^{20\Delta}}{n} \rightsquigarrow TV^2 \leq \frac{1}{2}\frac{e^{20\Delta}}{n}, \quad TV = o(1) = \frac{1}{2}$$

$$\Rightarrow \frac{e^{20\Delta}}{n} = o(1) \Rightarrow \underline{\Delta = K\log(n)}$$

$$\Rightarrow R \geq \frac{\Delta}{2}(1 - TV) \rightsquigarrow \boxed{R \geq c\log(n)}$$

$T(X) = \max_{i \in [n]} X_i$ (5

$$\forall t > 0: \exp\{t \mathbb{E}[T]\} \leq \mathbb{E}[\exp\{tT\}] = \mathbb{E}[\max_{i \in [n]} e^{t(\theta_i + Z_i)}]$$

$$\leq \mathbb{E}\left[\sum_{i=1}^{n} e^{t\theta_i + tZ_i}\right] = \sum_{i=1}^{n} e^{t\theta_i} \mathbb{E}[e^{tX_i}] = \sum_{i=1}^{n} e^{t\theta_i} e^{t^2/2}$$

$$\Rightarrow \mathbb{E}[T] \leq \frac{t}{2} + \frac{1}{t} \log\left[\sum e^{t\theta_i}\right] \leq \frac{t}{2} + \frac{1}{t} \log(n e^{t\theta_{\max}}) = \theta_{\max} + \frac{t}{2} + \frac{\log(n)}{t}$$

$$t = \sqrt{2\log n} \Rightarrow \mathbb{E}[T] \leq \theta_{\max} + \sqrt{2\log n}$$

$$\mathbb{E}[T] = \mathbb{E}[\max_{i \in [n]} X_i] \geq \max_{i \in n} \mathbb{E}[X_i] = \max_{i \in n} \theta_i = \theta_{\max}$$

$$\mathrm{Var}(T) = \mathrm{Var}(\max_{i \in [n]} X_i) \leq \mathrm{Var}(X) = 1$$ (equality for $n=1$, $\mathrm{Var}(T)$ decreases with n)

$$\mathbb{E}[(T - \theta_{\max})^2] = \mathrm{Var}(T) + (\mathbb{E}[T] - \theta_{\max})^2 \leq (\sqrt{2\log n})^2 + 1$$

✓ $o(\log(n))$

$$\Rightarrow \boxed{\sup_{\theta} \mathbb{E}[(T - \theta_{\max})^2] \leq 2\log n}$$

Problem 4)

(1

$X_i \sim P_\theta,\ \theta \in [-a,a]$

$$R_n^*(\Theta) \triangleq \inf_{\hat\theta}\sup_{\theta\in[-a,a]} \mathbb{E}_\theta[\|\hat\theta-\theta\|_2^2] \geq \sup_{\theta\in[-a,a]}\inf_{\hat\theta} \mathbb{E}_\theta[\|\theta-\hat\theta\|_2^2]$$

$$\geq \inf_{\hat\theta} \mathbb{E}_\pi[\|\theta-\hat\theta\|_2^2] = R_\pi \quad : \text{bayesian risk}$$

$\theta \sim \pi'$: $(\theta-\delta)\frac{a}{a+\delta} \sim \theta-\delta \sim \mathrm{Unif}[-a,a] = \pi$

Q: $\theta \sim \pi = \mathrm{Unif}[-a,a],\ X \sim P_\theta$

P: $\theta \sim \pi' = \mathrm{Unif}[-a,a+2\delta],\ X \sim P_{(\theta-\delta)\frac{a}{a+\delta}}$

$$\forall \hat\theta : \mathbb{E}_\pi[\|\theta-\hat\theta\|_2^2] \geq \mathrm{Var}_Q(\theta-\hat\theta)$$

$$\chi^2(P_{\theta X}\|Q_{\theta X}) \geq \chi^2(P_{\theta\hat\theta}\|Q_{\theta\hat\theta}) \geq \chi^2(P_{\theta-\hat\theta}\|Q_{\theta-\hat\theta}) \geq \frac{(\mathbb{E}_P[\theta-\hat\theta]-\mathbb{E}_Q[\theta-\hat\theta])^2}{\mathrm{Var}_Q(\hat\theta-\theta)}$$

by construction

$$P_X = Q_X \Rightarrow \mathbb{E}_P[\hat\theta(X)] = \mathbb{E}_Q[\hat\theta(X)],\ \mathbb{E}_P[\theta] = \mathbb{E}_Q[\theta]+\delta$$

$$R_\pi^* \geq \sup_{\delta\neq 0} \frac{\delta^2}{\chi^2(P_{\theta X}\|Q_{\theta X})}$$

$$\chi^2(P_{\theta X}\|Q_{\theta X}) = \chi^2(P_\theta\|Q_\theta) + \mathbb{E}_Q\left[\chi^2(P_{X|\theta}\|Q_{X|\theta})\underbrace{\cancel{\left(\frac{dP_\theta}{dQ_\theta}\right)^2}}_{\approx 1+o(1)}\right]$$

$\delta \to 0$: $\chi^2(P_\theta\|Q_\theta) = \chi^2(\pi\|\pi((\cdot-\delta)\frac{a}{a+\delta})) = (I(\pi)+o(1))\delta^2$

$$\chi^2(P_{X|\theta}\|Q_{X|\theta}) = (I_F(\theta)+o(1))\delta^2 = (nI(\theta)+o(1))\delta^2$$

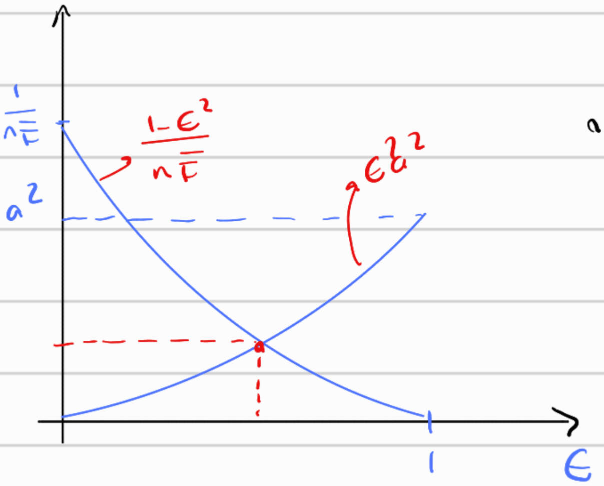
additivity of fisher info

$$R_\pi^* \geq \frac{1}{I(\pi)+\mathbb{E}_\pi[nI(\theta)]} = \frac{1}{I(\pi)+n\bar I}, \quad I_\pi : \text{fisher information of } \pi$$

$$\pi' = \mathrm{Unif}[-a,a+2\delta],\ \pi = \mathrm{Unif}[-a,a] \Rightarrow D_{\chi^2}(P_\theta\|Q_\theta) = \int_{-a}^{a} \frac{\left(\frac{1}{2a}-\frac{1}{2a+2\delta}\right)^2}{\left(\frac{1}{2a}\right)}dx$$

$$= 2a \times 2a \times \left(\frac{2\delta}{2a(2a+2\delta)}\right)^2 \approx \frac{\delta^2}{a^2} \Rightarrow I_\pi = a^{-2}$$

$$\Rightarrow R_\pi^* \geq \frac{1}{a^{-2}+n\bar I} = \frac{a^2}{1+n\bar I a^2}$$



$$\min_{a\le\epsilon\le 1}\max\left(\epsilon^2a^2,\frac{1-\epsilon^2}{n\bar I}\right):$$

$$\epsilon^2a^2=\frac{1-\epsilon^2}{n\bar I}\Rightarrow n\bar I a^2\epsilon^2=1-\epsilon^2\Rightarrow\epsilon^2=\frac{1}{1+n\bar I a^2}$$

$$\Rightarrow\min_{0\le\epsilon\le1}\max\left(\epsilon^2a^2,\frac{1-\epsilon^2}{n\bar I}\right)=\epsilon_c^2a^2=\frac{a^2}{1+n\bar I a^2}$$

$$\rightarrow\boxed{R_n^*(\theta)\geq R_R^*\geq\frac{a^2}{1+n\bar I a^2}=\min_{a\le\epsilon\le1}\max\left\{\epsilon^2a^2,\frac{1-\epsilon^2}{n\bar I}\right\}}$$

(2

$$\frac{a^2}{1+n\bar I a^2}=\frac{1}{a^{-2}+n\bar I}=\frac{1}{(a^{-1}+\sqrt{n\bar I})^2-2\sqrt{n\bar I}\,a^{-1}}\geq\frac{1}{(a^{-1}+\sqrt{n\bar I})^2}\Rightarrow\boxed{R_n^*(\theta)\geq\frac{1}{(a^{-1}+\sqrt{n\bar I})^2}}$$

(3

shifting θ by θ_0 does not change any hands, as it is equivalent to finding $(\theta-\theta_0)$ assuming $X\sim P_{\theta-\theta_0}$ and then estimating θ by using $(\theta-\theta_0)+\theta_0$. ⇝ just shift everything.

$$\Rightarrow\inf_{\hat\theta}\sup_{\theta\in[a+\theta_0,a+\theta_0]}\mathbb{E}_\theta\left[\|\theta-\hat\theta\|_2^2\right]=R_n^*(\theta)\geq\frac{1}{(a^{-1}+\sqrt{n\bar I})^2},\quad\bar I=\frac{1}{2a}\int_{\theta_0-a}^{\theta_0+a}I(\theta)\,d\theta$$

now let: $a=n^{-1/4}$, $\bar I=\frac{1}{2a}\int_{-a+\theta_0}^{a+\theta_0}I(\theta)d\theta=\frac{1}{2a}\int_{-a}^{a}I(t+\theta_0)dt=\frac{1}{2a}\int_{-a}^{a}[I(\theta_0)+tI'(\theta_0)+\dots]dt$

$$=I(\theta_0)+O(a)=I(\theta_0)+O(n^{-1/4})$$

$$\Rightarrow\inf_{\hat\theta}\sup_{\theta\in[\theta_0-n^{-1/4},\theta_0+n^{-1/4}]}\mathbb{E}_\theta[\|\hat\theta-\theta\|_2^2]\geq\frac{1}{(a^{-1}+\sqrt{n\bar I})^2}=\frac{1}{(n^{1/4}+n^{1/2}I(\theta_0)^{1/2}+O(n^{1/4}))^2}=\frac{1}{nI(\theta_0)(1+O(1))}$$

$$\Rightarrow\boxed{R_n^*(\theta)\geq\frac{1+O(1)}{nI(\theta_0)}}$$

Problem 5)

$X_1,\dots,X_n \sim P : P \in \Delta([k]) \rightsquigarrow X_i \in [k]$

MLE: $\mathbb{P}\{X_1,\dots,X_n \sim P\} = \prod_{i=1}^{n} P_{X_i}$ = likelihood = $L(P)$ (1

$$\log L(P) = \sum_{i=1}^{n} \log(P_{X_i}) = \sum_{i=1}^{n}\sum_{j=1}^{k} \mathbb{1}\{X_i=j\}\log P_j = \sum_{j=1}^{k} \log P_j \overbrace{\sum_{i=1}^{n} \mathbb{1}\{X_i=j\}}^{N_j}$$

$$= \sum_{j=1}^{K} N_j \log P_j = N_K \log\Big(1-\sum_{j\neq K} P_j\Big) + \sum_{j\neq K} N_j \log(P_j)$$

$$\frac{\partial}{\partial P_j} = 0 \rightsquigarrow \frac{N_j}{P_j} = \frac{N_K}{1-\sum_{j\neq K}^{k} P_j} = \frac{N_K}{P_K} \rightsquigarrow \forall i: \frac{N_i}{P_i} = c \Rightarrow P_i = \frac{N_i}{c}$$

$$\sum P_i = 0 \rightsquigarrow c = \sum_{i=1}^{k} N_i = n \Rightarrow \boxed{P_i^{MLE} = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}\{X_j = i\}} \rightsquigarrow \text{as expected}$$

$$d_{TV}(P,\hat{P}^{MLE}) = \sum_{i:P_i>\hat{P}_i} P_i - \hat{P}_i = \sum_{i=1}^{n} \mathbb{1}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\}\Big\}\Big(P_i - \frac{1}{n}\sum_{m=1}^{n}\mathbb{1}\{X_m=i\}\Big)$$

$$= \sum_{i=1}^{n} P_i \mathbb{1}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\}\Big\} - \frac{1}{n}\sum_{i=1}^{n}\sum_{m=1}^{n} \mathbb{1}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\}, X_m = i\Big\}$$

$$\Rightarrow \mathbb{E}_P[d_{TV}(P,\hat{P}^{MLE})] = \sum_{i=1}^{n} P_i \mathbb{P}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\}\Big\} - \frac{1}{n}\sum_{i=1}^{n}\sum_{m=1}^{n} \underbrace{\mathbb{P}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\}, X_m=i\Big\}}_{\text{all equal}}$$

$$= \sum_{i=1}^{n}\Big[P_i \mathbb{P}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\}\Big\} - \mathbb{P}\Big\{P_i \geq \frac{1}{n}\sum\{X_j=i\}, X_n=i\Big\}\Big]$$

$$= \sum_{i=1}^{k} P_i\Big[\mathbb{P}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\}\Big\} - \mathbb{P}\Big\{P_i \geq \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j=i\} \,\Big|\, X_n=i\Big\}\Big]$$

$$= \sum_{i=1}^{n} P_i\Big[\mathbb{P}\Big\{nP_i \geq \sum_{j=1}^{n}\mathbb{1}\{X_j=i\}\Big\} - \mathbb{P}\Big\{nP_i - 1 \geq \sum_{j=1}^{n-1}\mathbb{1}\{X_j=i\}\Big\}\Big]$$

$$= \sum_{i=1}^{k} P_i \mathbb{P}\Big\{nP_i \geq \sum_{j=1}^{n}\mathbb{1}\{X_j=i\}, nP_i - 1 < \sum_{j=1}^{n-1}\mathbb{1}\{X_j=i\}\Big\}$$

$$= \sum P_i\Big\{\mathbb{P}\{X_n=i\}\underbrace{\cancel{\mathbb{P}\Big\{nP_i-1 \geq \sum_{j=1}^{n-1}\mathbb{1}\{X_j=i\}, nP_i-1<\sum_{j=1}^{n-1}\mathbb{1}\{X_j=i\}\Big\}}}_{=0} + \mathbb{P}\{X_n \neq i\}\mathbb{P}\Big\{nP_i \geq \sum_{i=1}^{n-1}\cdots, nP_i - 1 < \sum\Big\}\Big\}$$

$$= \sum p_i \mathbb{P}\{X_n \neq i\} \mathbb{P}\left\{ \sum_{j=1}^{n-1} \mathbb{1}\{X_j = i\} = \lfloor n p_i \rfloor \right\}$$

$$= \sum_{i=1}^{K} p_i (1-p_i) \binom{n-1}{\lfloor n p_i \rfloor} p_i^{\lfloor n p_i \rfloor} (1-p_i)^{n - \lfloor n p_i \rfloor - 1}$$ ⇝ something! (symmetric)

$\sup_{p \in M_K}$ ⇝ should also be symmetric ⇝ let: $p_i = \frac{1}{K}$

$$\Rightarrow \sup_{p \in M_K} \mathbb{E}_p[d_{TV}(p, \hat{p}^{MLE})] = \sum_{i=1}^{K} \frac{1}{K}\left(1-\frac{1}{K}\right)\binom{n-1}{\lfloor \frac{n}{K} \rfloor}\left(\frac{1}{K}\right)^{\lfloor \frac{n}{K} \rfloor}\left(1-\frac{1}{K}\right)^{n - \lfloor \frac{n}{K} \rfloor - 1}$$

$$= \left(1-\frac{1}{K}\right)^{n - \lfloor \frac{n}{K} \rfloor} \left(\frac{1}{K}\right)^{\lfloor \frac{n}{K} \rfloor} \binom{n-1}{\lfloor \frac{n}{K} \rfloor}$$

for large n: by stirling approximation, $\lfloor \frac{n}{K} \rfloor \approx \frac{n}{K}$

$$\binom{n-1}{\lfloor \frac{n}{K} \rfloor} \approx \frac{\sqrt{2\pi(n-1)} \left(\frac{n-1}{\cancel{e}}\right)^{n-1}}{\sqrt{2\pi(n - \frac{n}{K} - 1)} \left(\frac{n - \lfloor \frac{n}{K} \rfloor - 1}{\cancel{e}}\right)^{n - \lfloor \frac{n}{K} \rfloor - 1} \cdot \sqrt{2\pi\left(\frac{n}{K}\right)} \left(\frac{\frac{n}{K}}{\cancel{e}}\right)^{\lfloor \frac{n}{K} \rfloor}}$$

$$\Rightarrow \sup_{p \in M_K} \mathbb{E}[d_{TV}(p, \hat{p}^{MLE})] \approx \sqrt{\frac{n}{(n - \frac{n}{K})(\frac{n}{K})}} \cdot \frac{\left(\cancel{\frac{1}{K}}\right)^{\lfloor \frac{n}{K} \rfloor}}{\left(\cancel{\frac{n}{K}}\right)^{n/K}} \cdot \frac{\left(1 - \cancel{\frac{1}{K}}\right)^{n - \lfloor \frac{n}{K} \rfloor}}{\left(\cancel{n - \frac{n}{K}}\right)^{n - \lfloor \frac{n}{K} \rfloor - 1}} \cdot \cancel{n}^{n-1}$$

$$\approx \sqrt{\frac{K}{n(1 - 1/K)}} \cdot \left(1 - \frac{1}{K}\right) = \sqrt{\frac{K}{n-1}} \Rightarrow \boxed{\sup_{p \in M_K} \mathbb{E}[d_{TV}(p, \hat{p}^{MLE})] \asymp \min\left(1, \sqrt{\frac{K-1}{n}}\right)}$$

for an upper bound:

$$\mathbb{E}_p[d_{TV}(p, \hat{p}^{MLE})] = \mathbb{E}_p\left[\frac{1}{2}\sum_{i=1}^{K}\left|\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{X_j = i\} - p_i\right|\right] = \frac{1}{2}\sum_{i=1}^{K}\mathbb{E}_p\left[\left|\frac{1}{n}\sum \cdots - p_i\right|\right]$$

$$= \frac{1}{2}\sum_{i=1}^{K}\mathbb{E}\left[\sqrt{\left(\frac{1}{n}\sum \cdots - p_i\right)^2}\right] \leq \frac{1}{2}\sum_{i=1}^{K}\sqrt{\mathrm{Var}\Big(\frac{1}{n}\sum_{j=1}^{n}\underbrace{\mathbb{1}\{X_j = i\}}_{\mathrm{Ber}(p_i)}\Big)} = \frac{1}{2}\sum_{i=1}^{K}\sqrt{\frac{1}{n^2} \times \sum_{j=1}^{n} p_i(1-p_i)}$$

$$= \frac{1}{2}\sum_{i=1}^{K}\sqrt{\frac{1}{n}p_i(1-p_i)} \xrightarrow{\sup @ p_i = \frac{1}{K}} \sup_{p \in P_K} d_{TV}(p, \hat{p}^{MLE}) \leq \frac{1}{2}\sum_{i=1}^{K}\sqrt{\frac{1}{n}\cdot\frac{1}{K}\left(1-\frac{1}{K}\right)} = \frac{1}{2}K \times \cdots$$

$$\Rightarrow \boxed{\sup_{p \in M_K} \mathbb{E}[d_{TV}(p, \hat{p}^{MLE})] \leq \frac{1}{2}\sqrt{\frac{K-1}{n}}}$$ ⇝ later $\asymp$ by minimax bounds.

$v \in \{\pm 1\}^{\lfloor k/2 \rfloor} = \mathcal{V}$, $P_v(2i) = \frac{1-\epsilon v_i}{k}$, $P(2i-1) = \frac{1+\epsilon v_i}{k}$, if k is odd $P(k) = \frac{1}{k}$ ~ $v_k = 0$ ~ ?

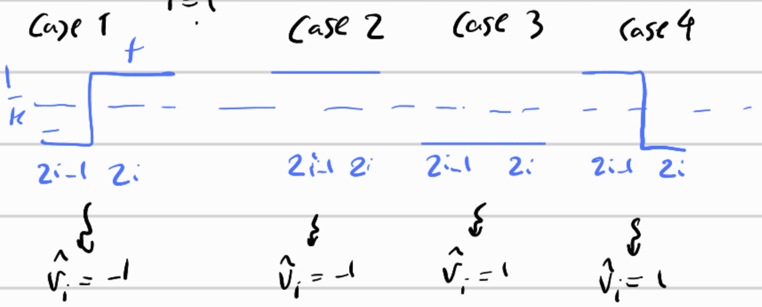
$$R(\theta) \geq \delta \sum_{j=1}^{\lfloor k/2\rfloor} [1 - d_{TV}(P_{+j}, P_{-j})], \quad P_{\pm j}(\cdot) = \frac{1}{2^{\lfloor k/2\rfloor -1}} \sum_{v: v_j = \pm 1} P_v(\cdot)$$

$\forall v \in \mathcal{V}: \ell(\theta, \theta_v) \geq \delta\, d_H(\hat v(\theta), v)$, $\hat v(\theta)$: maps θ to v

$\ell(\theta,\theta_v) = d_{TV}(P_\theta, P_{\theta_v})$. let $p_i = \frac{1+\eta_i}{k}$ ~ $\eta_i \geq -1$

$$\Rightarrow d_{TV}(P_\theta,P_{\theta_v}) = \sum_{i=1}^k \left|p_i - \frac{1-\epsilon(-1)^i v_{\lceil i/2\rceil}}{k}\right| = \sum_{i=1}^k \left|\frac{\eta_i + \epsilon(-1)^i v_{\lceil i/2 \rceil}}{k}\right| = \frac1k \sum_{i=1}^k |\eta_i + \epsilon(-1)^i v_{\lceil i/2\rceil}|$$

$$= \frac1k \left[\sum_{i=1}^{\lfloor k/2\rfloor} |\eta_{2i} + \epsilon v_i| + |\eta_{2i-1} - \epsilon v_i| + \overbrace{|\eta_k| \mathbb{1}\{k \text{ is odd}\}}^{\geq 0}\right]$$



~ $\hat v_i = -\text{Sign}(\eta_{2i})$

$|\eta_{2i}+\epsilon v_i| + |\eta_{2i-1} - \epsilon v_i| = \mathcal{X}_i$

$v_i = 1$; $\begin{cases} \text{Case 4: } \mathcal X_i \geq 0 \quad (v_i = \hat v_i) \\ \text{Case 3,2: } \mathcal X_i \geq \epsilon \\ \text{Case 1: } \mathcal X_i \geq 2\epsilon \geq \epsilon \end{cases}$

$v_i = -1$; $\begin{cases} \text{case 4: } \mathcal X_i \geq 2\epsilon \geq \epsilon \\ \text{case 2,3: } \mathcal X_i \geq \epsilon \\ \text{case 1: } \mathcal X_i \geq 0 \quad (v_i = \hat v_i) \end{cases}$ $\Longrightarrow \mathcal X_i \geq \epsilon\, \mathbb 1\{v_i \neq \hat v_i\}$

$$d_{TV}(P_\theta,P_{\theta_v}) = \frac1k\left[\sum_{i=1}^{\lfloor k/2\rfloor} \mathcal X_i + |\eta_k|\mathbb 1\{k \text{ is odd}\}\right] \geq \frac1k \sum_{i=1}^{\lfloor k/2\rfloor}\mathcal X_i \geq \frac{\epsilon}{k}\sum_{i=1}^{\lfloor k/2\rfloor} \mathbb 1\{v_i\neq\hat v_i\}$$

$\Rightarrow d_{TV}(P_\theta,P_{\theta_v}) \geq \frac{\epsilon}{k} d_H(\hat v(\theta), v)$ ~ $\delta = \frac{\epsilon}{k}$

$$R(\theta) \geq \delta \sum_{i=1}^{\lfloor k/2\rfloor}(1 - d_{TV}(P_{+j},P_{-j})) \geq \frac{\epsilon}{\cancel{k}}\cdot\frac{\cancel{k}}{2}\left(1 - \max_{v,j} d_{TV}(P_{v,+j}, P_{v,-j})\right)$$

$$d_{TV}^2(P^{\otimes n}_{v,+j}, P^{\otimes n}_{v,-j}) \leq \frac12 D_{KL}(P^{\otimes n}_{v,+j}, P^{\otimes n}_{v,-j}) = \frac n2 D_{KL}(P_{v,+j},P_{v,-j}) = \frac n2\left[\frac{1-\epsilon}{k}\ln\left(\frac{1-\epsilon}{1+\epsilon}\right)\right.$$

$$\left. + \frac{1+\epsilon}{k}\ln\left(\frac{1-\epsilon}{1+\epsilon}\right)\right] = \frac{n\epsilon}{k}\ln\left(\frac{1+\epsilon}{1-\epsilon}\right) \leq \underbrace{\frac nk\cdot\frac{4\epsilon^2}{1-\epsilon^2}}_{\chi^2 \geq KL}$$

, large n ~ small ϵ $\Rightarrow \approx \frac{2n\epsilon^2}{k}$

$$\Rightarrow R(\theta) \geq \frac{\varepsilon}{2}\left[1 - \sqrt{\frac{n}{k} \cdot \frac{4\varepsilon^2}{1-\varepsilon^2}}\right] \quad , \text{let: } d_{TV} = \frac{1}{2}$$

$$\Rightarrow \frac{1}{4} = \frac{4\varepsilon^2 n}{k(1-\varepsilon^2)} \Rightarrow 16\varepsilon^2 n = k(1-\varepsilon^2) \to \varepsilon^2 = \frac{k}{k+16n} \rightsquigarrow \frac{k}{16n} \leq \frac{k-1}{16n}$$

$$\Rightarrow R(\theta) \geq \frac{1}{4}\varepsilon_{1/2} \geq \frac{1}{2}\sqrt{\frac{k-1}{16n}} \Rightarrow \boxed{R(\theta) \gtrsim c\sqrt{\frac{k-1}{n}}}$$

(3

$$\inf_{\hat{p}} \sup_{p \in \mathcal{M}_k} E_p\left[d_{TV}(p,\hat{p})\right] \geq \Phi(\delta)\left[1 - \frac{I(V;X^n) + \log 2}{\log N}\right]$$

we need $\approx 2^k$ distributions that are close to each other.

same v as before: $N = 2^{\lfloor \frac{k}{2} \rfloor} \Rightarrow \log N = \lfloor \frac{k}{2} \rfloor$

$$d_{TV}(p,p') = \frac{4\varepsilon}{k} \underbrace{d_H(v,v')}_{\geq k} \leq 2\delta \Rightarrow \delta \leq 4\varepsilon$$

$$I(V;X^n) \leq \max_{v,v'} D_{KL}(P_{X^n|V=v} \| P_{X^n|V=v'}) = n \max_{v,v'} D_{KL}(P_v \| P_{v'})$$

$$\leq n\lfloor \tfrac{k}{2} \rfloor \cdot \left[\frac{2\varepsilon}{k}\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right)\right] = n\varepsilon \log\left(\frac{1+\varepsilon}{1-\varepsilon}\right)$$

$$\Rightarrow R \geq 4\varepsilon\left[1 - \underbrace{\frac{n\varepsilon\log(\frac{1+\varepsilon}{1-\varepsilon}) + 1}{k/2}}_{1/2}\right] \rightsquigarrow R \geq 2\varepsilon_{1/2}$$

$$\varepsilon_{1/2} : \frac{k}{2} - 1 = n\varepsilon\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right) \geq \underbrace{\frac{4\varepsilon^2 n}{1-\varepsilon^2}}_{\chi^2 \geq KL} \Rightarrow \varepsilon_{1/2}^2 \leq \frac{1}{4}\frac{k-1}{n} \quad (k \geq 2)$$

$$\Rightarrow \boxed{R \geq c\sqrt{\frac{k-1}{n}}}$$