Subject :　　　　　　　　　　　　　　　　　　　　　　Date :

سؤالات 5,3 با همفکری نوشته شده اند.

ا. تشکر بلند است ، تشکر مهربان!

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

$|\theta_1-\theta_2| \geq 2\delta$, $\Phi(\delta)\propto\delta^2 \to L2\text{-norm}$

$\hookrightarrow \theta_2=\theta_1+2\delta$

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

$\to$ so : $\min_{\hat\theta}\max_{\theta}\mathbb{E}[\|\theta-\hat\theta\|^2] \geq c\,\delta^2(1-2n\delta)$

$\rightsquigarrow \delta=\frac{1}{4n} \to \min_{\hat\theta}\max_{\theta}\mathbb{E}[\|\theta-\hat\theta\|^2] \geq \frac{c}{8n^2}$

$\rightsquigarrow C=\frac{c}{8} \to \min_{\hat\theta}\max_{\theta}\mathbb{E}[\|\theta-\hat\theta\|^2] \geq \frac{C}{n^2}$ ▨

۲. بهترین تخمین برای توزیع uniform را استفاده می‌کنیم : $\hat\theta = \max\{u_i\}-1 \to Y=\max\{u_i\}$

می‌دانیم u_iها مستقل هستند. داریم : $\mathbb{E}[\|\theta-\hat\theta\|^2] = \mathbb{E}[\|\theta+1-\max\{u_i\}\|^2]$

داریم $\to P\{\max\{u_i\}<x\} = \mathbb{P}\{\bigcap_{i=1}^{n} u_i<x\} = \prod_{i=1}^{n} P[u_i<x] = \prod_{i=1}^{n}\min\{1,x-\theta\}\,\mathbb{1}(x>\theta)$

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

so : $\mathbb{E}[(\theta+1-Y)^2] = n\int_{\theta}^{\theta+1}(\theta+1-x)^2(x-\theta)^{n-1}dx = n\int_0^1 \alpha^{n-1}(\alpha-1)^2 d\alpha =$

Subject :   Date :

$$= n\left[\frac{1}{n+2} + \frac{1}{n} - \frac{2}{n+1}\right] = \frac{2n}{n(n+1)(n+2)} < \frac{2}{n^2}$$

$\Longrightarrow$ so we have : $\sup_{\theta} \mathbb{E}\left[(\theta - \hat{\theta})^2\right] \leq \frac{c}{n^2}$ ▨

۳. همانطور که دیدیم، یک تخمین گر رایج را برای مسئله استفاده کردیم که برای توزیع های uniform رایج است و همچنین دیدیم که به کران minimax

ما در بخش اول مسئله رسیده است! پس احتمالا بهترین کران برای این مسئله همین باشد.

۲ f- انحراف و نامساوی داده

۱. $\longrightarrow D_f(P_{MX} T_{\hat{M}|X} \| P_M Q_X T_{\hat{M}|X}) = \mathbb{E}_{X,M,\hat{M}}\left[f\left(\frac{P_{MX}(m,x)\,T(\hat{m}|x)}{P_M(m)\,q_X(x)\,T(\hat{m}|x)}\right)\right]$

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

so we have $\Longrightarrow$ $D_f(P_{MX} T_{\hat{M}|X} \| P_M Q_X T_{\hat{M}|X}) = D_f(P_{MX} \| P_M Q_X)$ ▨

۲. $X \longrightarrow \boxed{T_{\hat{M}|X}} \longrightarrow \hat{M}$: $D_f(P_{MX} \| P_M Q_X) = D_f(P_{MX\hat{M}} \| P_M Q_{X\hat{M}}) \underbrace{\geq}_{\text{DP inequality}} D_f(P_{M\hat{M}} \| P_M Q_{\hat{M}})$

so we have : $D_f(P_{MX} \| P_M Q_X) \geq D_f(P_{M\hat{M}} \| P_M Q_{\hat{M}})$ ▨

Subject : Date :

we have: $P_{M\hat{M}} \rightsquigarrow M;\hat{M} \xrightarrow{\mathbb{1}[M\neq\hat{M}]} +$ $\rightsquigarrow Ber(P_E) : \mathbb{P}[M\neq\hat{M}] = P_E$.٣

$P_M Q_{\hat{M}} \rightsquigarrow$ $\rightsquigarrow Ber(1-\frac{1}{m})$: M is uniform $\Rightarrow \frac{1}{m}$ equality

Data Processing inquality: $D_f(P_{M\hat{M}} \| P_M Q_{\hat{M}}) \geq D_f(Ber(P_E) \| Ber(1-\frac{1}{m}))$

.٤

we have: $\forall Q_X \Rightarrow D_f(P_{MX} \| P_M Q_X) \geq D_f(P_{M\hat{M}} \| P_M Q_{\hat{M}}) \geq D_f(Ber(P_E) \| Ber(1-\frac{1}{m}))$

so we have $\Rightarrow \inf_{Q_X} D_f(P_{MX} \| P_M Q_X) \geq D_f(Ber(P_E) \| Ber(1-\frac{1}{m}))$ ✓

also we have:

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

$$= D_f(P_{X|M} \| Q_X | P_M) = D_f(P_{XM} \| Q_X P_M)$$

$$\Rightarrow \inf_{Q_X} \max_m D_f(P_{X|M=m} \| Q_X) \geq \inf_{Q_X} D_f(P_{MX} \| P_M Q_X)$$

so we proved $\Rightarrow \inf_{Q_X} \max_m D_f(P_{X|M=m} \| Q_X) \geq \inf_{Q_X} D_f(P_{MX} \| P_M Q_X) \geq D_f(Ber(P_E) \| Ber(1-\frac{1}{m}))$ ∎

MICRO

Subject : Date :

۳ نقطه در برابر حملات!

we have: $\ell : \Theta \times \hat{\Theta} \to \mathbb{R}$, $\theta_1 \leq \Theta$: $\ell(\theta_0, \alpha) + \ell(\theta_1, \alpha) \geq \Delta$, $\forall \alpha \in \hat{\Theta}$, $\theta_0 \leq \theta_1$

for fixed θ_0, θ_1, given estimator $\hat{\theta}$, define test $\tilde{\theta}$ =

$$\tilde{\theta} = \begin{cases} \theta_0 & \text{with probability: } \dfrac{\ell(\theta_1, \hat{\theta})}{\ell(\theta_0, \hat{\theta}) + \ell(\theta_1, \hat{\theta})} \\[2ex] \theta_1 & \text{with probability: } \dfrac{\ell(\theta_0, \hat{\theta})}{\ell(\theta_0, \hat{\theta}) + \ell(\theta_1, \hat{\theta})} \end{cases}$$

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

the same for θ_1 : $\mathbb{E}_{\theta_1}[\ell(\tilde{\theta}, \theta_1)] \leq \dfrac{\ell(\theta_0, \theta_1)}{\Delta}\, \mathbb{E}_{\theta_1}[\ell(\hat{\theta}, \theta_1)]$

$$E_\pi = \frac{1}{2}(\delta_{\theta_0} + \cancel{\mu}_{\theta_1})$$

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

we have: $\mathbb{P}[\hat{\theta} \neq \theta] \geq \dfrac{1}{2}(1 - TV(P_{\theta_0}, P_{\theta_1})) \xrightarrow[\text{for } \mathbb{E}_{\mu_{\theta_1}}]{\text{for some fixed } \theta_1} \dfrac{1}{2}(1 - TV(P_{\theta_0} \| \mathbb{E}_{\mu_{\theta_1}} P_{\theta_1}))$

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

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

Subject : Date :

.۲

we have: $d_{TV}(P\|Q) \leqslant \sqrt{\frac{1}{2} D_{KL}(P\|Q)}$

$$\Rightarrow d_{TV}^2(P\|Q) \leqslant \frac{1}{2} D_{KL}(P\|Q) \leqslant \frac{1}{2} D_{\chi^2}(P\|Q) \quad \overset{\text{just set}}{\longrightarrow} C \leqslant \frac{1}{2} \quad \blacksquare$$

.۳

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

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

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

so we have : $1 + D_{\chi^2}(\mathbb{E}_\mu[P_\theta], Q) = \mathbb{E}\left[\int \frac{P_\theta(x) P_{\theta'}(x)}{Q(x)} dx\right] \quad \blacksquare$

.٤

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

so: $X_i \sim N(\theta_i, 1)$, $\Theta = \mathbb{R}$, $\hat{\Theta} = \mathbb{R}$ $\rightarrow$ where Θ denotes the maximum

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

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

Subject : Date :

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

Let $\Theta_1 = [\theta_0 + \delta, \theta_0 + 2\delta]$, $\mu_{\theta_1} = \text{Unif}\{\Theta_1\}$, $\delta = \sqrt{2\Delta}$

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

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

$\rightsquigarrow \int \frac{P_{\theta_1} P_{\theta_2}}{P_{\theta_0}} dx = \mathbb{E}_{X\sim P_{\theta_0}}\left[\frac{P_{\theta_1}(x) P_{\theta_2}(x)}{P_{\theta_0}(x)^2}\right]$

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

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

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

so we have:

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

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

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

MICRO

Subject : Date :

θ_1, θ_2 : all zero except the max $\longrightarrow E_{\theta_1,\theta_2 \sim R}[\exp(\dots)] = \frac{1}{n} e^{10\delta^2} + \frac{n-1}{n} \times 1$

only non-zero with prob $\frac{1}{n}$

$$\Rightarrow E[\int_{2\Delta} \sim] - 1 = \frac{1}{n} E[e^{10\delta^2} - 1] = \frac{1}{n}[e^{2\Delta} - 1] \leq \frac{1}{n} e^{2\Delta}$$

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

$$\Rightarrow \frac{e^{2\Delta}}{n} = O(1) \Rightarrow \Delta = k \log(n)$$

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

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

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

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

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

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

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

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

MICRO

Subject : Date :

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

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

Chernoff - Rubin - Stein ε

$X_i \sim P_\theta$, $\theta \in [-a, a]$ .1

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

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

we know : Q : $\theta \sim \pi = Unif[-a, a]$, $X \sim P_\theta$

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

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

$$\rightarrow \chi^2(P_{\theta\hat x} \| Q_{\theta\hat x}) \geq \chi^2(P_{\theta\hat\theta} \| Q_{\theta\hat\theta}) \geq \chi^2(P_{\theta-\hat\theta} \| Q_{\theta-\hat\theta}) \geq \frac{(\mathbb{E}_P[\theta-\hat\theta]-\mathbb{E}_Q[\theta-\hat\theta])^2}{Var_Q(\hat\theta-\theta)}$$

by construction

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

$$\Rightarrow R_n^* \geq \sup_{\delta\neq 0} \frac{\delta^2}{\chi^2(P_{\theta\hat x} \| Q_{\theta\hat x})}$$

MICRO

Subject : Date :

$$\delta \to 0 : \quad \chi^2(P_\theta \| Q_\theta) = \chi^2\left(\pi \,\|\, \pi\left((\theta-\delta)\tfrac{a}{a+\delta}\right)\right) = (I(\pi) + o(1))\,\delta^2$$

$$\chi^2(P_{X|\hat\theta} \| Q_{X|\hat\theta}) = (I_X(\theta) + o(1))\,\delta^2 = (\underbrace{n I(\theta)}_{\text{additivity of fisher information}} + o(1))\,\delta^2$$

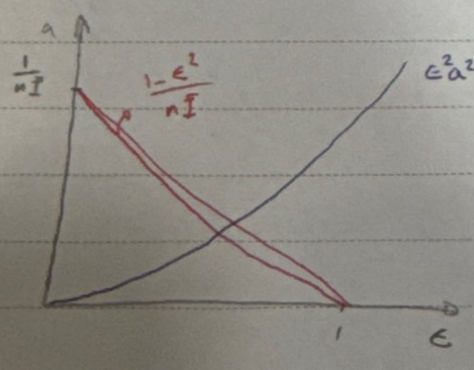
$$\Rightarrow R_\pi^* \geq \frac{1}{I(\pi) + E_\pi[n I(\theta)]} = \frac{1}{I(\pi) + n\bar{I}} , \quad I_\pi : \text{fisher information of } \pi$$

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

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

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

we also have:



according to figure : $\min_{0<\epsilon<0} \max\left(\epsilon^2 a^2, \frac{1-\epsilon^2}{n\bar{I}}\right)$; $\epsilon^2 a^2 = \frac{1-\epsilon^2}{n\bar{I}} \Rightarrow \epsilon^2 = \frac{1}{1+n\bar{I}a^2}$

$$\Rightarrow \min \max\left(\epsilon^2 a^2, \frac{1-\epsilon^2}{n\bar{I}}\right) = \frac{a^2}{1 + n\bar{I}a^2}$$

so we have: $R_\pi^*(\theta) \geq R_{\bar\pi}^* \geq \dfrac{a^2}{1+a^2 n\bar{I}} = \min\limits_{0\le\epsilon\le1} \max\left\{\epsilon^2 a^2, \dfrac{1-\epsilon^2}{n\bar{I}}\right\}$ ▨

$$\frac{a^2}{1+n\bar{I}a^2} = \frac{1}{a^{-2}+n\bar{I}} = \frac{1}{(a^{-1}+\sqrt{n\bar{I}})^2 - 2\sqrt{n\bar{I}}\,a^{-1}} \geq \frac{1}{(a^{-1}+\sqrt{n\bar{I}})^2}$$ ۲.

$$\Rightarrow R_n^*(\theta) \geq \left(\frac{1}{a^{-1}+\sqrt{n\bar{I}}}\right)^2$$ ▨

۳. شیفت دادن θ به اندازه θ_0 عملاً هیچ چیز را به جز اندازه شیفت می‌دهد و تغییر خاصی رخ نمی‌دهد.

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

now Let $a = n^{-\frac{1}{4}}$, $\bar{I} = \dfrac{1}{2a}\displaystyle\int_{-a+\theta_0}^{a+\theta_0} I(\theta)\,d\theta = \dfrac{1}{2a}\int_{-a}^{a} I(t+\theta_0)\,dt$

استفاده از پیوستگی و تیلور $\Rightarrow$ $= \dfrac{1}{2a}\displaystyle\int_{-a}^{a}\left[I(\theta_0) + t\,I'(\theta_0) + \dots\right] dt$

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

$$\Rightarrow \inf_{\hat\theta} \sup_{\theta\in[\theta_0-n^{-\frac14},\, \theta_0+n^{-\frac14}]} \mathbb{E}_\theta\left[\|\hat\theta-\theta\|_2^2\right] \geq \frac{1}{(a^{-1}+\sqrt{n\bar{I}})^2} = \frac{1}{\left(n^{\frac14}+n^{\frac12}I(\theta)^{\frac12}+O(n^{\frac14})\right)} = \frac{1}{nI(\theta)(1+O(1))}$$

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

۵. تخمین توزیع

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

get the log $\rightarrow$ Log likelihood $= \sum_{i=1}^{n} \log(P_{X_i}) = \sum_{i=1}^{n} \sum_{j=1}^{k} \mathbb{1}\{X_i = j\} \log P_j = \sum_{j=1}^{k} \log P_j \overbrace{\sum_{i=1}^{n} \mathbb{1}\{X_i = j\}}^{N_j}$

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

$$\frac{\partial}{\partial P_j} = 0 \implies \frac{N_j}{P_j} = \frac{N_k}{1 - \sum_{j \neq k} P_j} = \frac{N_k}{P_k} \implies \forall i: \frac{N_i}{P_i} = C \rightarrow P_i = \frac{N_i}{C}$$

$$\sum P_i = 1 \implies C = \sum N_i = n \xrightarrow{\text{so}} \boxed{P_i = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}\{X_j = i\}}$$ MLE

Now we have:

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

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

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

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

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

Subject : Date :

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

$$= \sum P_i \left\{ nP_i \geq \sum_{j=1}^{n} \mathbb{1}\{X_j = i\}, nP_i - 1 < \sum_{j=1}^{n-1} \mathbb{1}\{X_j = i\} \right\}$$

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

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

$$= \underline{\sum P_i (1 - P_i) \binom{n-1}{\lfloor nP_i \rfloor} P_i^{\lfloor nP_i \rfloor} (1 - P_i)^{n - \lfloor nP_i \rfloor - 1}}$$

this is symmetric $\rightarrow$ sup $\rightarrow$ should also be symmetric : $P_i = \frac{1}{k}$

$$\Rightarrow \sup_{P \in M_k} \mathbb{E}_P\{ d_{TV}(P, \hat{P}^{n,\mathrm{LE}}) \} = \sum_{i=1}^{k} \frac{1}{k} \left(1 - \frac{1}{k}\right) \binom{n-1}{\lfloor \frac{n}{k} \rfloor} \left(\frac{1}{k}\right)^{\lfloor \frac{n}{k} \rfloor} \left(1 - \frac{1}{k}\right)^{n - \lfloor \frac{n}{k} \rfloor - 1}$$

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

$\rightarrow$ for Large n : $\lfloor \frac{n}{k} \rfloor \approx \frac{n}{k}$

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

Subject :   Date :

$$\Rightarrow \sup_{P \in \mathcal{M}_k} \mathbb{E}[d_{TV}(P, \hat{P}^{MLE})] \asymp \sqrt{\frac{n}{(n-\frac{n}{k})(\frac{n}{k})}} \cdot \frac{(\frac{1}{k})^{\lfloor \frac{n}{k} \rfloor}}{(\frac{n}{k})^{\frac{n}{k}}} \cdot \frac{(1-\frac{1}{k})^{n-\lfloor \frac{n}{k} \rfloor}}{(n-\frac{n}{k})^{n-\lfloor \frac{n}{k} \rfloor - 1}} \times n^{n-1}$$

$$\asymp \sqrt{\frac{k}{n(1-\frac{1}{k})}} \cdot (1-\frac{1}{k}) = \sqrt{\frac{k}{n-1}}$$

$$\rightsquigarrow \sup_{P \in \mathcal{M}_k} \mathbb{E}[d_{TV}(P, \hat{P}^{MLE})] \asymp \min(1, \sqrt{\frac{k-1}{n}})$$

for an upper bound :

$$\mathbb{E}_p[d_{TV}(P, \hat{P}^{MLE})] = \mathbb{E}_p[\frac{1}{2} \sum_{i=1}^{k} |\frac{1}{n} \sum_{j=1}^{n} \mathbb{1}\{x_j = i\} - P_i|] = \frac{1}{2} \sum_{i=1}^{k} \mathbb{E}_p[|\frac{1}{n} \sum \mathbb{1}\{x_j = i\} - P_i|]$$

$$= \frac{1}{2} \sum_{i=1}^{k} \mathbb{E}[\sqrt{(\frac{1}{n} \sum \mathbb{1}\{x_j = i\} - P_i)^2}] \leq \frac{1}{2} \sum_{i=1}^{k} \sqrt{Var(\frac{1}{n} \sum_{j=1}^{n} \mathbb{1}\{x_j = i\})}$$

$$= \frac{1}{2} \sum_{i=1}^{k} \sqrt{\frac{1}{n^2} \times \sum_{j=1}^{n} P_i (1-P_i)}$$

$$\xrightarrow{\sup @ P_i = \frac{1}{k}} \sup_{P \in \mathcal{P}_k} d_{TV}(P, \hat{P}^{MLE}) \leq \frac{1}{2} \sum_{i=1}^{k} \sqrt{\frac{1}{n} \times \frac{1}{k} (1-\frac{1}{k})} =$$

$$\Rightarrow \sup_{P \in \mathcal{M}_k} \mathbb{E}[d_{TV}(P, \hat{P}^{MLE})] \leq \frac{1}{2} \sqrt{\frac{k-1}{n}}$$

Subject : Date :

$v \in \{\pm 1\}^{\lfloor \frac{k}{2} \rfloor} = V_{\lfloor \frac{k}{2} \rfloor}$, $P_v(2i) = \frac{1 - \varepsilon v_i}{k}$, $P(2i-1) = \frac{1 + \varepsilon v_i}{k}$ ۲

$\rightsquigarrow$ if k is odd : $P(k) = \frac{1}{k} \rightsquigarrow v_k = 0$

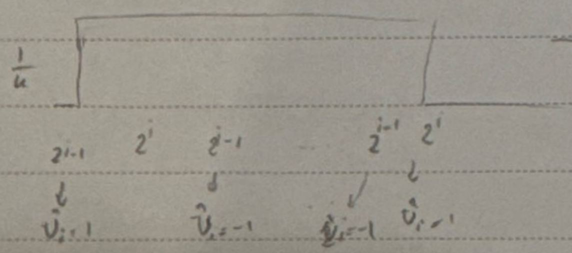
$R(\theta) \geq \delta \sum_{j=1}^{\lfloor \frac{k}{2} \rfloor} [1 - d_{TV}(P_{+j}, P_{-j})]$, $P_{\pm j}(a) = \frac{1}{2^{\lfloor \frac{k}{2} \rfloor - 1}} \sum_{v: v_j = \pm 1} P_v(\cdot)$

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

$\ell(\theta, \theta_v) = d_{TV}(P_\theta, P_{\theta v})$: let $P_i = \frac{1 + \eta_i}{k} \rightsquigarrow \eta_i \geq -1$

$\Rightarrow d_{TV}(P_{\theta_{\lfloor k \rfloor}}, P_{\theta_v}) = \sum_{i=1}^{k} \left| P_i - \frac{1 - ((-1)^i v_{\lfloor \frac{i}{2} \rfloor}}{k} \right| = \sum_{i=1}^{k} \left| \frac{\eta_i + \varepsilon (-1)^i v_{\lfloor \frac{i}{2} \rfloor}}{k} \right| = \frac{1}{k} \sum | \eta_i - \varepsilon (-1)^i v_{\lfloor \frac{i}{2} \rfloor} |$

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



$\rightsquigarrow$ $\hat{v}_i = -\text{sign}(\eta_{2i})$

$|\eta_{2i} + \varepsilon v_i| + |\eta_{2i-1} - \varepsilon v_i| = x_i$

$v_i = 1$: case 4 : $x_i \geq 0$ $(v_i = \hat{v}_i)$; case 3, 2 : $x_i \geq \varepsilon$; case 1 : $x_i \geq 2\varepsilon \geq \varepsilon$

$x_i \geq \varepsilon\, \mathbb{1}(v_i \neq \hat{v}_i)$ $\Longleftarrow$ $v_i = -1$: case 4 : $x_i \geq 2\varepsilon \geq \varepsilon$; cas 2, 3 : $x_i \geq \varepsilon$; case 1 : $x_i \geq 0$ $(v_i = \hat{v}_i)$

MICRO

Subject : Date :

so : $d_{TV}(P_\theta, P_{\theta_v}) = \frac{1}{k}\left[\sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} x_i + |q_k|\,\mathbb{1}\{k \text{ is odd}\}\right]$

$$\geq \frac{1}{k}\sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} x_i \geq \frac{\varepsilon}{k}\sum_{i=1}^{\lfloor \frac{k}{2} \rfloor} \mathbb{1}\{v_i \neq \hat{v}_i\}$$

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

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

$\rightarrow d^2_{TV}(P^{\otimes n}_{v,+j}, P^{\otimes n}_{v,-j}) \leq \frac{1}{2} D_{KL}(P^{\otimes n}_{v,+j}, P^{\otimes n}_{v,-j}) = \frac{n}{2} D_{KL}(P_{v,+j}, P_{v,-j})$

$$= \frac{n}{2}\left[\frac{1-\varepsilon}{k}\ln\left(\frac{1-\varepsilon}{1+\varepsilon}\right) + \frac{1+\varepsilon}{k}\ln\left(\frac{1-\varepsilon}{1+\varepsilon}\right)\right] = \frac{n\varepsilon}{k}\ln\left(\frac{1+\varepsilon}{1-\varepsilon}\right)$$

$$\leq \underbrace{\frac{n}{k} \times \frac{4\varepsilon^2}{1-\varepsilon^2}}_{\chi^2 \geq KL} \xrightarrow{\text{large } n} \approx \frac{2n\varepsilon^2}{k}$$

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

$\Rightarrow \frac{1}{4} = \frac{4\varepsilon^2 n}{k(1-\varepsilon^2)} \Rightarrow 16\varepsilon^2 n = k(1-\varepsilon^2) \xrightarrow{\text{use}} \frac{k}{16n} \geq \frac{k-1}{16n}$

so : $R(\theta) \geq \frac{1}{4}\varepsilon_{\frac{1}{2}} \geq \frac{1}{2}\sqrt{\frac{k-1}{16n}} \Rightarrow R(\theta) \geq c\sqrt{\frac{k-1}{n}}$ ☒

Subject : Date :

$$\inf_{\hat{P}} \sup_{P \in M_k} \mathbb{E}_P[d_{TV}(P,\hat{P})] \geq \Phi(S)\left[1 - \frac{I(V;X) + \log 2}{\log N}\right]$$.۳

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

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

$\rightarrow d_{TV}(P,P') = \frac{4\varepsilon}{k} \overbrace{d_H(v,v')}^{\geq k} \leq 2\delta \Rightarrow \delta \leq 4\varepsilon$

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

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

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

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

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

MICRO