Subject:---------- Date:----------

سوال یک

الف

$\chi^2(P\|Q) = ?$

$$\chi^2(P\|Q) = \int \frac{(P-Q)^2}{Q}\,dx = \int \frac{(P-Q)^2}{Q^2}\,Q = \int \frac{P^2}{Q^2}\,Q - \cancel{\int 2P} + \cancel{\int Q}$$

$$\Rightarrow \chi^2(P\|Q) = E_Q\left(\left(\frac{P}{Q}\right)^2\right) - 1$$

$$\Rightarrow \chi^2(P_\pi\|Q) = E_Q\left[\left(\frac{P_\pi P_{\pi'}}{Q^2}\right)\right] - 1 = \int \frac{\int P_\theta \pi(\theta) \int P_{\theta'} \pi(\theta')}{Q^2}\,Q\,dx - 1$$

$$= \int \frac{E_{\theta\sim\pi}[P_\theta]\, E_{\theta'\sim\pi}[P_{\theta'}]}{Q^2}\,Q\,dx - 1 = \underbrace{\left(E_Q\left[E_{\theta\sim\pi}\frac{(\theta)}{Q}\, E_{\theta'\sim\pi}\frac{(\theta')}{Q}\right]\right)}_{I} - 1$$

$$I = E_{\theta\sim\pi} E_Q\left[\frac{\theta}{Q}\right] E_{\theta'\sim\pi} E_Q\left(\frac{\theta'}{Q}\right) = E_{\theta,\theta'\underset{iid}{\sim}\pi} E_Q\left(\frac{\theta\theta'}{Q^2}\right) = E_{\theta,\theta'\sim\pi}\left[G(\theta,\theta')\right]$$

$$\Rightarrow \chi^2(P_\pi\|Q) = E_{\theta,\theta'\underset{iid}{\sim}\pi}\left[G(\theta,\theta')\right] - 1 \quad ✓$$

ب

$$\lim_{n\to\infty} E_{Q_n}\left[\left(\frac{P_n}{Q_n}\right)^2\right] - 1 = o(1) \Rightarrow \lim_{n\to\infty} E_{Q_n}\left[\left(\frac{P_n}{Q_n}\right)^2\right] = 1$$

if $Q_n \underset{n\to\infty}{\to} c \overset{\text{implies}}{\Longrightarrow} P_n \underset{n\to\infty}{\to} c$. ⇐ یعنی

$$\Rightarrow \lim E_{Q_n}\left[\frac{P_n}{Q_n}\right] = o(1) \Rightarrow \lim_{n\to\infty}\left(E_{Q_n}\left[\frac{P_n}{Q_n}\right]\right)^2 = o(1)$$

$$\Rightarrow \lim_{n\to\infty} E_{Q_n}\left[\left(\frac{P_n}{Q_n}\right)^2\right] - E_{Q_n}\left[\left(\frac{P_n}{Q_n}\right)\right]^2 = o(1) - o(1) = 0 \Rightarrow$$

چون واریانس صفره

$$Var_{Q_n}\left(\frac{P_n}{Q_n}\right) = 0 \Rightarrow \frac{P_n}{Q_n} = \text{constant}$$

$\Rightarrow$ if $Q_n \to 0$ then $P_n \xrightarrow{n\to\infty} 0$

if $Q_n \to c$ then $P_n \to \cancel{k}c$ $\Rightarrow$ $\boxed{P \triangle Q}$

constant.

NAHAL

Scanned by CamScanner

Subject: Date:

سوال ۹

الف) تخمین‌گر زیر را در نظر بگیرید

$\hat{\theta}(x) = (x_1, \dots, 0)$

$$E_\theta[\|\hat{\theta}-\theta\|_2^2] = E_\theta[\underbrace{(\theta_1-x_1)^2}_{z_1^2} + \overbrace{\|\theta_{1:}\|^2}^{\sum_{i=2}^{n}(\theta_i-0)^2}] = E_\theta[z_i^2 + \|\theta_{1:}\|^2] =$$

$$E_\theta[z_i^2] + E(\|\theta_{1:}\|^2) =$$

$(\theta_1 - x_1) \sim N(0,1) \Rightarrow z_i \sim N(0,1) \Rightarrow E_\theta(z_i^2) \lesssim 1$

$$E[\|\theta_{1:}\|^2] \lesssim E_\theta[2(1-p^{-1/4}|\theta_1|)^2] \lesssim E_\theta[2^2] \lesssim 4$$

$$\Rightarrow E_\theta[\|(\hat{\theta}-\theta)^2\|^2] \lesssim 5$$

$$\Rightarrow \inf_{\hat{\theta}} \sup_{\theta\in\Theta} \|\hat{\theta}-\theta\|_2 \lesssim 1$$

ب)

$$\theta_{mle} = \arg\min_{\hat{\theta}} \|x-\theta\|_2$$

$$2\times(1-\cancel{p^{-1/4}p^{1/4}}) = 0$$

$$\sup_{\theta\in\Theta} E_\theta[\|\hat{\theta}_{mle}-\theta\|^2] \gtrsim \sqrt{p}$$

فرض کنید بردار اصل برابر ماکسیمم زمانی است که تنها یک مولفه باشد

$\theta_1 = p^{1/4} \xrightarrow{\text{imlty.}} \theta_i = 0 \quad i \neq 1$

$\theta = (p^{1/4}, \dots, 0)$

we know MLE estimator is equal to intrical estimator

$$\theta_{ML} = \frac{1}{p}\sum_{i=1}^{p} x_i \rightsquigarrow E_\theta[\|\theta_{ML}-\theta\|^2] = E_\theta[\|\tfrac{1}{p}\sum x_i - \theta\|^2]$$

$$= \frac{1}{p^2}E_\theta[\|\sum x_i - p\theta\|^2] \geq \frac{1}{p^2}\times|p^2\sqrt{p}| \geq \sqrt{p}$$

$(p^{1/4}, 0, 0, 0, 0, \dots, 0)$

نهال NAHAL

Scanned by CamScanner

Date:

Subject:

سوال 8

الف

$$\hat{\theta} = \underset{\theta \in R^p}{\operatorname{argmin}} \sum_i (\theta_i - y_i)^2 + \lambda \{\theta_i \neq 0\}$$

$\Rightarrow$ independence of each index $\Rightarrow$ $\hat{\theta}_i = \underset{\theta_i}{\operatorname{argmin}} \overset{①}{(\theta_i - y_i)^2} + \overset{②}{\lambda \{\theta_i \neq 0\}}$

$\Rightarrow$ $\hat{\theta}_i \neq 0 \rightarrow$ better $\hat{\theta}_i = \hat{y}_i$ so ① term is zero $\rightarrow$ loss is λ

$\hat{\theta}_i = 0 \rightarrow y_i^2$

$\Rightarrow$ $\hat{\theta}_i \neq 0 \rightarrow \hat{\theta}_i = \hat{y}_i \rightarrow$ loss is λ

$\hat{\theta}_i = 0 \rightarrow$ loss is y_i^2

if $y_i^2 > \lambda \Rightarrow \hat{\theta} = y_i$

$y_i^2 < \lambda \Rightarrow \hat{\theta} = 0$

$$\Rightarrow \hat{\theta}_i = \begin{cases} y_i & y_i^2 > \lambda \\ 0 & y_i^2 < \lambda \end{cases} \Rightarrow \tau = \sqrt{\lambda}$$

$$\Rightarrow \hat{\theta}_i = \begin{cases} y_i & |y_i| > \lambda \\ 0 & |y_i| < \lambda \end{cases}$$

$$\hat{\theta} = \underset{i}{\operatorname{argmin}} \, E\left[(\hat{\theta}_i - y_i)^2 + \lambda |\hat{\theta}_i|\right]$$

ب

$$\hat{\theta}_i = \underset{\theta_i}{\operatorname{argmin}} \, (\theta_i - y_i)^2 + \lambda |\theta_i|$$

$\Rightarrow$ if $\hat{\theta}_i > 0 \rightarrow (\theta_i - y_i)^2 + \lambda \theta_i \Rightarrow \frac{\partial}{\partial \theta_i} = 2(\theta_i - y_i) + \lambda = 0 \Rightarrow \hat{\theta} = y_i - \frac{\lambda}{2}$

if $\hat{\theta}_i < 0 \rightarrow$ " $- \lambda \theta_i \Rightarrow \hat{\theta}_i = y_i + \frac{\lambda}{2}$

$\Rightarrow$ $\hat{\theta}_i > 0 \rightarrow y_i - \lambda/2 \rightarrow y_i > \frac{\lambda}{2} \leftarrow \hat{\theta}_i > 0$

$\hat{\theta}_i < 0 \rightarrow y_i + \frac{\lambda}{2} < 0 \rightarrow y_i < -\frac{\lambda}{2}$

$\hat{\theta}_i = 0 \rightarrow -\frac{\lambda}{2} \leq y_i \leq \frac{\lambda}{2}$

NAHAL

Scanned by CamScanner

Subject: Date:

بنابراین

$$\hat{\theta}_i = \begin{cases} y_i - \frac{\lambda}{2} & y_i > \frac{\lambda}{2} \\ 0 & -\frac{\lambda}{2} \le y_i \le \frac{\lambda}{2} \\ y_i + \frac{\lambda}{2} & y_i < -\frac{\lambda}{2} \end{cases}$$

$$\tau = \frac{\lambda}{2}$$

ج

$$\underset{\theta}{\arg\min} \; \|y-\theta\|_\infty \le \tau \; \|\theta\|_0$$

تابع هزینه برابر تعداد درایه‌های ناصفر است مشروط در $\max_i |y_i-\theta_i|$

تا زمانی که $\max_i (y_i-\theta_i) < \tau$ باشد.

بنابراین ابتدا $\forall i : \hat{\theta}_i = 0$ می‌گذاریم. θ مشکل زمانی است که $|y_i| > \tau$ باشد. برای آن کافی است.

$\hat{\theta}_i$ را برابر y_i بگذاریم تا $y_i-\theta_i$ برابر صفر شود:

$$\Rightarrow \hat{\theta}_i = 0 \quad \text{except for } |y_i| > \tau$$

$$\Rightarrow \hat{\theta}_i = \begin{cases} 0 & |y_i| \le \tau \\ y_i & |y_i| > \tau \end{cases} = \hat{\theta}_i^{H_T}$$

د

در ایجاد می‌خواهیم تک درایه‌های $|\theta|$ را کمینه کنیم.

تا زمانی که $\max(y_i-\theta_i) < \tau$ است $\hat{\theta}_i = 0$ حال اگر $|y_i-\theta_i| \ge \tau \Rightarrow$

① $y_i - \theta_i \ge \tau \Rightarrow y_i \ge \tau \Rightarrow \theta_i \le \tau - y_i \Rightarrow$

② $\theta_i - y_i \ge \tau \Rightarrow \theta_i \ge y_i + \tau \Rightarrow$

$y_i - \theta_i > 0$

① $\rightarrow \theta_i > 0 \rightarrow \tau - y_i < 0 \Rightarrow y_i > \tau \rightsquigarrow \hat{\theta}_i = \tau - y$

② $\rightarrow \theta_i < 0 \rightarrow y_i + \tau < 0 \Rightarrow y_i < -\tau \rightsquigarrow \hat{\theta}_i = \tau + y$

نهال NAHAL

Scanned by CamScanner

Date:

Subject:

بنابراین

$$\hat{\theta} = \begin{cases} 0 & |y_i| \le \tau \\ y_i + \tau & y_i < -\tau \\ y_i - \tau & y_i > \tau \end{cases} = \hat{\theta}^{ST}$$

سوال هفتم

حوزه‌ی زیر را در نظر بگیرید:

$\theta \sim \text{Uniform}\{\tau e_j\}$ $\rightarrow$ $H_0 = \theta = 0$

e_k unit standards basis.

حال قرار دهید $\Delta = \frac{\tau^2}{2}$ تا شرط 28 رعایت شود.

چون گوسی‌اند

$$\chi^2(P_1, P_0) = E_{P_0}[e^{\theta^T \tilde{\theta}}] - 1 = \frac{\exp(\tau^2) - 1}{p}$$

$$TV(P_1, P_0) \le \sqrt{\frac{D(P_1 \| P_0)}{2}} \le \sqrt{\frac{\chi^2(P_1 \| P_0)}{2}}$$

حال می‌خواهیم که این مقدار به صفر برود.

$$\Rightarrow \exp(\tau^2) - p \Rightarrow$$

$$\tau = \sqrt{\log p}$$

اما از آنجا که می‌خواهیم مقدار χ^2 به صورت سری به صفر برود τ را دیگر می‌گیریم

$$\tau = \sqrt{\varepsilon} \sqrt{\log p}$$

$$\tau = \sqrt{\varepsilon \log p}$$

$$\Rightarrow \tau^2 = \varepsilon \log p$$

Two point method

$$\inf \sup E[(\hat{\theta} - \theta_{max})^2] \ge \frac{\tau^2}{2} \times \left(\frac{1}{2} - \cancel{\frac{1}{2} TV}^{\,0}\right) \ge \frac{\tau^2}{4}$$

$$\tau^2 = \varepsilon \log p \rightarrow \theta \; \Omega \ge \frac{\varepsilon \log p}{4} = c \log p$$

تخمین‌گر زیر را در نظر بگیرید:

$$\tau = x_{max} = \max_{i \in [p]} x_i$$

نهال
NAHAL

Scanned by CamScanner

Subject: Date:

$$E_\theta[(X_{max}-\theta)^2] \le E_\theta[\overbrace{|X_{max}-\theta|}^{z}{}^2] = E\|z\|_\infty^2]$$

$$e^{tE[z_{max}]} \le E\, e^{t z_{max}} \le E\left[\sum_i^p e^{tz_i}\right] = p\sum_{i=1}^p E[e^{tz_i}] = pe^{t^2/2}$$

$$\Rightarrow e^{tE[z_{max}]} \le pe^{t^2/2} \Rightarrow E[z_{max}] \le \frac{\log p}{t} + \frac{t}{2}$$

$$\frac{\partial}{\partial t}\left(\frac{\log p}{t} + \frac{t}{2}\right) \Rightarrow -\frac{\log p}{t^2} + \frac{1}{2} \overset{\text{بهینه سازی}}{=} 0 \Rightarrow t^2 = 2\log p$$

از ایده sanov استفاده کردیم. $\Leftarrow$ $\min_t$

$$t = \sqrt{2\log p} \rightarrow E[z_{max}] \le \sqrt{\frac{\log p}{2}} \times 2 = \sqrt{2\log p}$$

$$\Rightarrow E[|z_{max}|^s] \le E[|\max z|^s] \le c(\sqrt{\log p})^s \Rightarrow E[z^2] \le c\log p$$

سوال 2

$$P_\sigma(A)_{ij} = \begin{cases} \sigma_i=1,\ \sigma_j=1 \rightsquigarrow P(A_{ij}) \\ \sigma_i=-1,\ \sigma_j=-1 \rightsquigarrow P(A_{ij}) \\ \sigma_i=-1,\ \sigma_j=1 \rightsquigarrow f(A_{ij}) \\ \sigma_i=+1,\ \sigma_j=-1 \rightsquigarrow f(A_{ij}) \end{cases}$$

$$\Rightarrow P_\sigma(A) = \prod_{\substack{i,j \\ \sigma_i\sigma_j=1 \\ i<j}} P(A_{ij}) \prod_{\substack{i,j \\ i<j \\ \sigma_i\sigma_j=-1}} f(A_{ij}) = \prod\left[\frac{1+\sigma_i\sigma_j}{2}P(A_{ij}) + \frac{1-\sigma_i\sigma_j}{2}f(A_{ij})\right]$$

$\searrow$ 1+

$$= \prod_{ij}\left[\frac{1}{2}\left(P(A_{ij}) + f(A_{ij}) + \sigma_i\sigma_j\left(P(A_{ij}) - f(A_{ij})\right)\right)\right]$$

$$P_\sigma(A) = \prod_{\substack{i,j \\ i<j}} \frac{1}{2}\left(P(A_{ij}) + f(A_{ij}) + \sigma_i\sigma_j\left(P(A_{ij}) - f(A_{ij})\right)\right)$$

$$P_n(A) = \prod \frac{1}{2}\left(P(A_{ij}) + f(A_{ij})\right)$$

نهال NAHAL

Scanned by CamScanner

Subject: Date:

$$\frac{P_\sigma(A)\ P_{\hat\sigma}(A)}{P_0(A)} = \prod_{\substack{ij\\ i<j}} \frac{1}{2}\,\frac{\left(P(A_{ij})+f(A_{ij})+\sigma_i\sigma_j\left(P(A_{ij})-f(A_{ij})\right)\right)\times\left(P(A_{ij})+f(A_{ij})+\hat\sigma_i\hat\sigma_j\left(P(A_{ij})-f(A_{ij})\right)\right)}{P(A_{ij})+f(A_{ij})}$$

$\frac{1}{2}\times\frac{1}{1/2}$

بنابراین

$$\prod_{\substack{ij\\ i<j}}\left(1+\sigma_i\sigma_j\frac{\left(P(A_{ij})-f(A_{ij})\right)}{P(A_{ij})+f(A_{ij})}\right)\left(P(A_{ij})+f(A_{ij})+\hat\sigma_i\hat\sigma_j\left(P(A_{ij})-f(A_{ij})\right)\right) =$$

$$\underbrace{\prod_{\substack{ij\\ i<j}}\frac{1}{2}\left[P(A_{ij})+f(A_{ij})+\left(\sigma_i\sigma_j+\hat\sigma_i\hat\sigma_j\right)\left(P(A_{ij})-f(A_{ij})\right)+\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\frac{\left(P(A_{ij})-f(A_{ij})\right)^2}{P(A_{ij})+f(A_{ij})}\right]}_{\frac{1}{2}\beta}$$

$$\int\frac{P_\sigma P_{\hat\sigma}}{P_0} = \prod_{\substack{ij\\ i<j}}\int\frac{1}{2}\beta = ?$$

$$\frac{1}{2}\int\beta = \frac{1}{2}\Big(\cancel{\int P(A_{ij})+f(A_{ij})} + \sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\cancel{\int P(A_{ij})-f(A_{ij})} + \sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\cancel{\int\frac{\left(P(A_{ij})-f(A_{ij})\right)^2}{P(A_{ij})+f(A_{ij})}}\rightarrow 2\rho$$

$$\Rightarrow \frac{1}{2}\int\beta = \frac{1}{2}\left(\cancel{2}^{1} + 0 + \cancel{2}^{1}\rho\,\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\right)$$

$$\Rightarrow P_\sigma = \prod\left(1+\rho\,\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j\right) \Rightarrow$$

$$e^x = 1+x+\frac{x^2}{2}\cdots \Rightarrow 1+x \leq e^x$$

$$P_\sigma < \prod e^{\rho\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j} = e^{\rho\sum_{i<j}\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j}$$

$$\sum_{i,j}\sigma_i\sigma_j\hat\sigma_i\hat\sigma_j \leq \frac{1}{2}\sum\hat\sigma_i^2\hat\sigma_j^2 + \frac{1}{2}\sum\hat\sigma_i\hat\sigma_j + \frac{1}{2}\times 2\,\sigma_i\sigma_j\hat\sigma_j\hat\sigma_j =$$

$$\leq \frac{1}{2}\left(\hat\sigma_i\hat\sigma_j+\sigma_i\sigma_j\right)^2 \leq \frac{1}{2}\left(\langle\sigma,\hat\sigma\rangle^2\right)$$

$$\Rightarrow P_\sigma \leq e^{\rho/2\,\langle\sigma,\hat\sigma\rangle^2}$$

نهال NAHAL

Scanned by CamScanner

Subject: Date:

ب.

$$\rho = \frac{1}{2} \sum \frac{(P(x)-Q(x))^2}{P(x)+Q(x)} =$$

$P \to p$

$Q \to q$

$$= \frac{1}{2}\left[\frac{(p-q)^2}{(p+q)^2} + \frac{(1-p-(1-q))^2}{(1-p+1-q)^2}\right] = \frac{1}{2}\,\frac{\left(\frac{a}{n}-\frac{b}{n}\right)^2}{\left(\frac{a+b}{n}\right)} + \frac{1}{2}\,\frac{\left(\frac{b}{n}-\frac{a}{n}\right)^2}{2-\frac{a+b}{n}}$$

$$= \frac{(a-b)^2}{2n(a+b)} + \frac{(b-a)^2}{2n(2n-a-b)} = \frac{\tau}{n} + \frac{o(1)}{n} = \frac{\tau + o(1)}{n}$$

$\frac{o(1)}{n}$

2.

$P(\sigma_i \hat{\sigma}_i = 1) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}$

$P(\sigma_i \hat{\sigma}_i = -1) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}$

$\Rightarrow E[\sigma_i \hat{\sigma}_i] = \frac{1}{2} \times -1 + \frac{1}{2} \times 1 = 0$

$\mathrm{var}(\sigma_i \hat{\sigma}_i) = E[(\sigma_i \sigma_j)^2] - E^2 = \frac{1}{2} \times 1 + \frac{1}{2} \times 1 = 1$

$\Rightarrow Z^2 = \frac{1}{n} \langle \hat{\sigma}, \sigma \rangle^2 \xrightarrow{n \to \infty} Z = \frac{1}{\sqrt{n}} \langle \hat{\sigma}, \sigma \rangle \sim N(0,1)$

از همگرایی MGF

$\chi^2(P_1 \| P_0) = E[G(\theta, \hat{\theta})] - 1$

$$\Rightarrow \chi^2(P_1 \| P_0) \leq E[e^{G(\theta,\hat{\theta})}] - 1 \leq E_{\sigma,\hat{\sigma}}[e^{\frac{\rho}{2}\langle \sigma, \hat{\sigma} \rangle^2}] - 1$$

$$= E_{\sigma,\hat{\sigma}}\left[\exp\left(\frac{n\rho}{2} \frac{\langle \sigma, \hat{\sigma} \rangle^2}{\sqrt{n}^2}\right)\right] - 1 = E_Z\left[\exp\left(\frac{n\rho}{2} Z^2\right)\right] - 1$$

Z^2

$n \to \infty$

NAHAL نهال

Scanned by CamScanner

Date:

Subject:

ادامه ب

$$n \to \infty \to E_Z\left[\exp\left(\frac{n}{2}\times\left(\frac{\tau}{n}+\frac{O(1)}{n}\right)Z^2\right]-1$$

$$\xrightarrow{N(0,1)} \frac{1}{\sqrt{2\pi}}\underbrace{\int_{-\infty}^{\infty} e^{\left(\frac{\tau+O(1)}{2}\right)Z^2}\, e^{-Z^2/2}\, dz}_{e^{\tau Z^2}} - 1 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{(\tau+O(1)-1)Z^2/2}\, dz - 1$$

if $\tau + O(1) - 1 > 0 \to$ همگرایی ناچیز $\to$ $P_0, P_1 \to$ distinguishable.

$\tau \geq 1 \to$ همگرایی برقرار

if $\tau + O(1) - 1 < 0 \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-Z^2/2} < 0$ بازدارد.

$\Rightarrow n \to \infty \to \chi^2(P,Q) < \underset{\text{یک عدد}}{\underset{\downarrow}{A}}$ $\to$ به صفر می‌رود

$\to 0$

نمی‌توان فاصله دو توزیع را کم کرد. $\leftarrow$ قابل تمایز نیستند.

نهال NAHAL

Scanned by CamScanner

Date:

Subject

سوال ۴

قسمت ب

فرض اول ←

$$\varepsilon^2 a^2 \geq \frac{1-\varepsilon^2}{n\hat{I}} \Rightarrow \varepsilon^2 a^2 \hat{I}_n \geq 1-\varepsilon^2 \Rightarrow$$

$$\varepsilon^2 (a^2 I_n + 1) \geq 1 \Rightarrow$$

$$\varepsilon^2 a^2 \left(\hat{I}_n + \frac{1}{a^2}\right) \geq 1 \Rightarrow \varepsilon^2 \geq \left(a^2\left(\hat{I}_n + \frac{1}{a^2}\right)\right)^{-1} \geq \underbrace{(a^2 \hat{I}_n + 1)^{-1}}_{t^2}$$

$$\varepsilon^2 a^2 \geq \frac{1}{\left(I_n + \frac{1}{a^2}\right)} \Rightarrow (a^{-1} + \sqrt{I_n})^2 = a^{-2} + I_n + 2a^{-1}\sqrt{I_n}$$

$$\geq a^{-2} + I_n \Rightarrow$$

$$\left(\frac{1}{a^{-1}+\sqrt{I_n}}\right)^2 \leq \frac{1}{a^{-2}+I_n} \quad (1)$$

$$\rightarrow \varepsilon > t$$

$$(1) \Rightarrow \varepsilon^2 a^2 \geq \left(\frac{1}{a^{-1}+\sqrt{I_n}}\right)^2 \checkmark$$

یعنی $\varepsilon < t$

حال فرض کنید که

$$\frac{1-\varepsilon^2}{n\hat{I}} \geq \varepsilon^2 a^2 \Rightarrow 1 \geq \varepsilon^2 a^2 \left(\hat{I}_n + \frac{1}{a^2}\right) \Rightarrow$$

$$t^2 = (a^2 I_n + 1)^{-1}$$

$$\Rightarrow \min_{0<\varepsilon<t} \left(\frac{1-\varepsilon^2}{n\hat{I}}\right) \rightarrow \frac{1-t^2}{nI} \rightarrow$$

$$\Rightarrow \min \max \quad a^2 I_n$$

$$\Rightarrow \frac{1-(a^2 I_n+1)}{a^2 I_n + 1} \times \frac{1}{nI} = \left(\frac{1}{\hat{I}_n + \frac{1}{a^2}}\right)$$

$$\frac{1}{I_n + a^{-2}} \geq \left(\frac{1}{\sqrt{I_n} + a^{-1}}\right)^2 \checkmark$$

ثابت کرد که

اثبات

Subject: Date:

$H_0 \to P$
$H_1 \sim Q$

سوال 5

$$E[Z] = E[Z|H_0]P(H_0) + E[Z|H_1]P(H_1)$$

$$= \frac{1}{2}\left[E_P[P_\theta(H_0|x)] + E_Q[P_\theta(H_0|x)]\right]$$

$$= \frac{1}{2}E_P[P_\theta(H_0|x)] + \frac{1}{2} - \frac{1}{2}E_Q[P_\theta(H_0|x)]$$

$$\sup\{E[Z]\} = \frac{1}{2} + \frac{1}{2}\sup E_P[\overbrace{P_\theta(H_0|x)}^{f(x)}] - E_Q[P_\theta(H_0|x)]$$

$$\frac{1}{2} + \frac{1}{2}\sup[E_P[f(x)]] - E_Q[f(x)] \to$$

$$\text{apply } \frac{1}{2} \text{ and } +\frac{1}{2} \to \sup_{f'}\{E_P(f') - E_Q[f'^* - \tfrac{1}{2}]\} \longrightarrow \text{II}$$

conjugate of (f)

$f^* = \sup_x[xy - f(x)] = f(f(x)) - \frac{1}{2}$ باید تابع ای پیدا کنیم

$$G(f) = \begin{cases} f & 0 \le f \le \frac{1}{2} \\ +\infty & o.w \end{cases}$$

if $f > \frac{1}{2} \to$ then $\sup[E[f] - E_Q[G(f)]] + \frac{1}{2}$

should is $+\infty$ so it shand occurs in $0 \le f < \frac{1}{2}$

$G(f(x)) \to f(x) - \frac{1}{2}$ shoud be $f \Leftarrow$ between 0 and $\frac{1}{2}$

lets assume liner $f(x) = \begin{cases} ax+b & \text{where } 0 \le ax+b \le \frac{1}{2} \\ 0 & o.w \end{cases}$

~~for sup~~

نهال NAHAL

Scanned by CamScanner

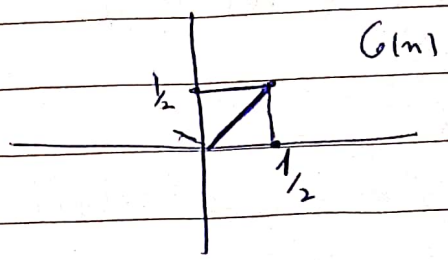
Subject: Date:

conjugate of Y

$$\sup_{Y} \{ E[Y] - E_a[f(y)] \}$$

$$f(y) = \sup_{x} (Yx - G(x)) = \sup_{0 \le x \le 1/2} \{ Yx - G(x) \}$$

$$G'(x) = 1 \rightsquigarrow 0 \le x \le 1/2$$
$$G'(x) = 0 \quad x > 1/2,\ x < 0 \quad \Rightarrow \begin{cases} Y = 1 \\ Y = 0 \end{cases}$$



$$\Rightarrow \begin{cases} x - G(x) = 0 & 0 \le x \le \frac{1}{2} \\ 0 - G(x) = & x > 1/2,\ x < 0 \rightsquigarrow -x \rightsquigarrow \end{cases}$$

$$\Rightarrow f(y) = \begin{cases} 0 & Y < 1 \\ & Y \ge 1 \end{cases}$$

$G'(x) = 0 \rightsquigarrow$ xY

$0 < x \le 1/2 \rightsquigarrow \frac{1}{2}x$

$Y \ge 1 \rightsquigarrow f(1) = 0 \Rightarrow \frac{1}{2}x \rightsquigarrow -\frac{1}{2}$ sub $\Rightarrow f(x) - \frac{1}{2}$

$$\Rightarrow a = \frac{1}{2},\ b = -\frac{1}{2}$$

$$\Rightarrow f(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}x - \frac{1}{2} & x \ge 1 \end{cases}$$

نهال NAHAL

Scanned by CamScanner

Subject: Date:

$$D_f(P\|Q) = E_Q\left[f\left(\frac{P(x)}{q(x)}\right)\right] \Rightarrow$$

$$\frac{P(x)}{q(x)} \geq 1 \rightarrow \frac{1}{2}\frac{P(x)}{q(x)} - \frac{1}{2}$$

$$\frac{P(x)}{q(x)} < 1 \rightarrow 0$$

$$\Rightarrow D_f(P\|Q) = \sum_{P(x) > q(x)} \left(\frac{1}{2}\frac{P(x)}{q(x)} - \frac{1}{2}\right) q(x) = \frac{1}{2}\sum_{P(x) \geq q(x)} [P(x) - q(x)] = \frac{1}{2} TV$$

$$\boxed{D_f = \frac{1}{2} TV(P\|Q)} + \frac{1}{2} \rightarrow \text{II}$$

$$D_f = \frac{1}{2}(TV(P\|Q) + 1) \leftarrow \text{II}$$

ب

$$E[\log P_\theta(z|x)] = \cancel{E_{x|H_0}} \log(P_\theta(z|x) \mid H_0)\, \cancel{P(H_0)}\, \frac{1}{2}$$

$$+ \cancel{E_{x|H_1}} \log P_\theta(z|x) \mid H_1)\, \cancel{P(H_1)}\, \frac{1}{2}$$

$$E_{x\sim P}[\underbrace{\log P_\theta(H_0|x)}_{f(x)}] + E_{x\sim Q}[\log \underbrace{P_\theta(H_1|x)}_{(1 - P_\theta(H_0|x)}]$$

$$\Rightarrow \theta = 1)$$

$$E_P[f(x)] + E_Q[\log(1 - e^{f(x)})] \leftarrow \log(1 - P(x)) = \log(1 - e^{\log P(x)})$$

$$\log P_\theta(H_0|x) = f(x)$$

$$P_\theta(H_0|x) \leq 1 \Rightarrow$$

$$\log P(H_0|x) \leq 0$$

NAHAL

Scanned by CamScanner

Subject: Date:

$$\Rightarrow \sup_{f:\, f\geq 1} E_p[f(x)] - E_q\left[\ln\left(\frac{1}{1-e^{f(x)}}\right)\right]$$

$$G(f(x)) = \ln\left(\frac{1}{1-e^{f(x)}}\right)$$

$$\Rightarrow L(\hat{\theta}) \leq \sup[f(x)] - E_q[G(f(x))]$$

$$f(y) = \sup_{x}\{yx - G(x)\} = \sup_{x}\{yx + \ln(1-e^{x})\}$$ ← تابع f

$$\cancel{y - (1-e^{x})x}$$

$$y + \frac{-e^{x}}{1-e^{x}} = 0 \Rightarrow y = \frac{e^{x}}{1-e^{x}} \Rightarrow e^{x} = y - e^{x}y \Rightarrow y = (y+1)e^{x}$$

$$\Rightarrow e^{x} = \frac{y}{y+1} \Rightarrow f(y) = y\ln\frac{y}{y+1} + \ln\left(\cancel{1-\frac{y}{y+1}}\ \frac{1}{y+1}\right) =$$

$$y\ln y - y\ln y+1 - \ln(y+1) = y\ln y - (y+1)\ln(y+1)$$

$$f(1) = -2\ln 2 = -2 \Longrightarrow +2$$ باید اضافه کنیم

Scanned by CamScanner

$$f(x) = x\log x - (x+1)\log(x+1) + 2$$

$$f'(x) = \log x - \log(x+1)$$

$$f''(x) = \frac{1}{x} - \frac{1}{x+1} \Rightarrow \underset{x>0}{\frac{x+1-x}{x(x+1)}} = \underset{x>0}{\frac{1}{x^2+x}} > 0$$

$f''(x) > 0$ محدب است.

$$\boxed{f(x) = x\log(x) - (x+1)\log(x+1) + 2}$$

$f(1) = 0$

تابع محدب است ✓

ج) باید حداقل L را کاهش دهیم تا تلاش کند که $D_f(P\|Q)$ کاهش یابد.

$$\min_{Q_\theta} \max_{P_{\theta'}} \left[2\log\left(P_{\theta'}(z|x)\right) \right] = \min_{Q_\theta} \max_{P_{\theta'}} \left\{ E_P\left[P_{\theta'}(H_0|x)\right] \right.$$

θ'

$$\left. - E_Q\left[\log \frac{1}{1 - P_{\theta'}(H_1|x)}\right] \right\}$$

θ

نهال
NAHAL

Scanned by CamScanner

Subject: Date:

سوال ۶

الف

$$L(f(x), Y) = \mathbb{1}\{f(x) \neq Y\} = \mathbb{1}\{f(x)Y < 0\}$$

$$\phi(f(x), y) = \phi(f(x)y) \Rightarrow R_\phi(f) = E[\phi(Y f(x))]$$

$$= \sum_{x, y=1} P(y, x)\, \phi f(x) + \sum_{\substack{x \\ y=-1}} P(y, x)\, \phi(-f(x))$$

$$\sum_x P(1, x)\, \phi f(x) + \sum_x \underbrace{P(x) - P(1, x)}_{P(x) - P(1,x) \to P(x)(1 - P(1|x))} \phi(-f(x))$$

$$\Rightarrow P(1, x) = P(1|x)\, P(x) = \eta_x(x)\, P(x)$$

$$\Rightarrow \sum_x P(x)\, \eta(x)\, \phi f(x) + \sum P(x)(1 - P(1|x))\, \phi(-f(x))$$

$$= \sum_x P(x) \left[\eta(x)\, \phi f(x) + (1 - \eta(x))\, \phi(-f) \right]$$

$$\Rightarrow E_x[\ell(f(x), \eta(x))]$$

$$\ell(f(x), \eta(x)) = \eta\, \phi(x) + (1 - \eta)\, \phi(-f(x))$$

Scanned by CamScanner

سوال ۶ب

$$R_{\phi^*} = \int \inf_{\alpha} \phi(\alpha)\, P_1(x)\, \pi + \phi(-\alpha)\, P_{-1}(x)\,(1-\pi)\, dx =$$

$$R_{\phi} = \int \inf_{\alpha} \left(\phi(\alpha) \frac{P_1(x)\,\pi}{P(x)} + \frac{\phi(-\alpha)\, P_1(x)\,(1-\pi)}{P(x)} \right) P(x)\, dx =$$

$$R_{\phi, Prior^*} = E_x \Big[\overbrace{\underbrace{\inf_{\alpha} P(Y=1)\,\phi(\alpha) + P(Y=-1)\,\phi(-\alpha)}_{\text{مستقل از } x}}^{F} \Big] = F\, \cancel{E_x[1]}$$

$\Rightarrow$ ~~$R_{\phi, Prior} = E_x \left[\sup_{\alpha} P(Y=1)\,\phi(\alpha) - P(Y=-1)\,\phi(-\alpha) \right]$~~

~~$R_{Prior} - R_{\phi^*} = \int$~~ $\ell^*_{\phi} = \inf_{\alpha} \ell_{\phi}(\alpha, \pi) \mapsto R_{\phi^*, Prior^*}$. تعریف relate ℓ^* با

$$R_{\phi, Prior^*} = E_x \left[\ell^*_{\phi}(\pi) \right]$$

$$R_{\phi, Prior} - R_{\phi^*} = \int \left(\sup_{\alpha} \ell^*_{\phi}(\pi) - R_{*i} \right) P(x)$$

$$R_{\phi, Prior^*} - R_{\phi} = \int \left[\sup_{\alpha} \ell^*_{\phi}(\pi) - \frac{\phi(\alpha)\, P_1(x)\,\pi}{P(x)} - \phi(-\alpha) \frac{P_{-1}(x)\,(1-\pi)}{P(x)} \right] P(x)\, dx$$

$$= \int \sup_{\alpha} \left[\ell^*_{\phi}(\pi) - \frac{\phi(\alpha) \frac{P_1(x)}{P(x)}\, \pi}{\frac{P_1(x)}{P_{-1}(x)}} - \phi(-\alpha) \frac{\cancel{\frac{P_{-1}(x)}{P(x)}}\,(1-\pi)}{\frac{P(x)}{P_{-1}(x)}} \right] P(x)\, dx$$

نهال NAHAL

Scanned by CamScanner

Date:

$$= \int \sup_{\alpha} \left[\ell_{\phi}^{*}(\pi) - \frac{\phi(\alpha)\, x\pi + \phi(-\alpha)\, 1-\pi}{x} \right] P(x)\, dx$$

$$\frac{P(x)}{x} = P_{-1}(x)$$

$$\frac{P(x)}{P_{1}(x)}$$

تعني $P(x)$ بـ $P_{1}(x)$

$$\Rightarrow \int \sup_{\alpha} \left[\ell_{\phi}^{*}(\pi) \frac{P(x)}{} - \phi(\alpha)\, \pi x + \phi(-\alpha)(1-\pi)\, P_{-1}(x) \right] dx$$

$$= \int \sup_{\alpha} \left[\ell_{\phi}^{*}(\pi) - \frac{\pi\, \phi(\alpha)\, x + \phi(-\alpha)(1-\pi)}{\pi x + (1-\pi)} \right] (\pi x + (1-\pi))\, P_{-1}(x)$$

$$= E_{P_{-1}x} \Big[\underbrace{\sup \left[\ell_{\phi}^{*}(\pi) - \frac{\pi\, \phi(\alpha)\, x + \phi(-\alpha)(1-\pi)}{\pi x + 1-\pi} \right] (\pi x + (1-\pi))}_{f(x)} \Big]$$

$$f(x) = \sup_{\alpha} \left[\ell_{\phi}^{*}(\pi) - \frac{\pi(\alpha)\, x + (1-\pi)\, \phi(-\alpha)}{\pi x + 1-\pi} \right] (\pi x + 1-\pi)$$

Scanned by CamScanner