Statistical Theory Problem Set 1

Bohao Tang

1. When condition 1 or 3 holds then $E[l_i(S(X_i))/\theta(i)] = E[l(\theta(i), S(X_i))|\theta(i)]$

Actually condition 1 is contained in condition 3, because condition 1 says that given $\theta(i)$, $l_i(a)$ is just a constant, and constant will be independent with every random variable X_i. So we just need to prove for condition 3.

Proof:

Here the probability space is for $(X_i, \theta(i), l_i(a_1), l_i(a_2))$, and the expectation will perform on the joint distribution of $X_i, \theta(i), l_i(a_1), l_i(a_2)$.

Then:

$$E\{l_i[S(X_i)]|\theta(i)\} = E\{E[l_i[S(X_i)] \mid X_i, \theta(i)] \mid \theta(i)\}$$

Consider

$$\begin{aligned} & E\{l_i[S(X_i)] \mid X_i = x, \theta(i) = k\} \\ &= E\{l_i[S(x)] \mid X_i = x, \theta(i) = k\} \\ &= E\{l_i(S(x)) \mid \theta(i) = k\} \quad \text{--- because of the conditional independence} \\ &= l(k, S(x)) \quad \text{--- by definition} \end{aligned}$$

So $E[l_i(S(X_i)) | X_i, \theta(i)] = l(\theta(i), S(X_i))$

So $E\{l_i[S(X_i)] | \theta(i)\} = E[l(\theta(i), S(X_i)) | \theta(i)]$

For the counter-example that condition 2 is not sufficient, it needs some computation and looked unpractical:

Let $l_i(a_1) = \begin{cases} 6 & \text{probility } \frac{1}{2} \\ -3 & \cdots\cdots \frac{1}{2} \end{cases}$ $\qquad l_i(a_2) = 5\, l_i(a_1) - 5$

$X_i = \begin{cases} +1 & \cdots \frac{5}{9} \\ -1 & \cdots \frac{4}{9} \end{cases}$ and $X_i \perp l_i(a_1)$

Set $\theta(i) = \mathbb{1}_{X_i \cdot l_i(a_1) > 0} + 1$

the $X_i \perp l_i(a_1)$ but given $\theta(i)$ we will know that if they are of the same sign, so the are not conditional independent.

Directly compute

$$E\left[l(\theta(i), S(X_i)) \mid \theta(i)=1\right] = 2 \cdot \frac{\frac{1}{2}\cdot\frac{4}{9}}{\frac{1}{2}} + 5 \cdot \frac{\frac{1}{2}\cdot\frac{5}{9}}{\frac{1}{2}} = \frac{33}{9}$$

$$E\{l[S(X_i)] \mid \theta(i)=1\}$$

$$= \cancel{3 \cdot \frac{5}{9}}$$

$$E\left[l[S(X_i)] \mid X_i \cdot l_i(a_1) < 0\right]$$

$$= -3 \cdot \frac{\frac{5}{9}\cdot\frac{1}{2}}{\frac{1}{2}} + 25 \cdot \frac{\frac{4}{9}\cdot\frac{1}{2}}{\frac{1}{2}}$$

$$= \frac{85}{9}$$

$$\frac{85}{9} \neq \frac{33}{9}$$

so 2. is not sufficient.

2. We have that

$$P(X_i^* > 0 \mid \theta(i)=1) = 0.94$$

$$P(X_i^* < 0 \mid \theta(i)=2) = 0.98$$

and $pr(X_i^* \mid \theta(i)) = N[\mu[\theta(i)], 1]$

so we have that. $E[X_i^* \mid \theta(i)=1] = 1.554774$

and $E[X_i^* \mid \theta(i)=2] = -2.053749$

Also, $P_r(L_i(a_1), L_i(a_2), X_i^* \mid \theta(i)=1) = N\left(\begin{pmatrix} 2 \\ 5 \\ 1.554774 \end{pmatrix}, \begin{pmatrix} 10 & 0 & \frac{7\sqrt{10}}{10} \\ 0 & 10 & \frac{7\sqrt{10}}{10} \\ \frac{7\sqrt{10}}{10} & \frac{7\sqrt{10}}{10} & 1 \end{pmatrix} \right)$

$P_r(L_i(a_1), L_i(a_2), X_i^* \mid \theta(i)=2) = N\left(\begin{pmatrix} 1 \\ 0 \\ -2.053749 \end{pmatrix}, \begin{pmatrix} 10 & 0 & \frac{7}{\sqrt{10}} \\ 0 & 10 & \frac{7}{\sqrt{10}} \\ \frac{7}{\sqrt{10}} & \frac{7}{\sqrt{10}} & 1 \end{pmatrix} \right)$

(we assume $L_i(a_1), L_i(a_2)$ are independent given $\theta(i)$)

Use simulation, we get that (1000 simulations) / Simulate 0.2 billion times

$E[L_i(S(X_i)) \mid \theta(i)=1] = 2.1721854$ / 2.18006754

$E[L_i(S(X_i)) \mid \theta(i)=2] = \cancel{0.991985}$ $\cancel{0.023}$ 0.008223 / 0.02008129

$E[L(\theta(i), S(X_i)) \mid \theta(i)=1] = 2.201$ / 2.18001206

$E[L(\theta(i), S(X_i)) \mid \theta(i)=2] = \cancel{0.918}$ 0.023 / 0.01999644

3: Without any loss, we can suppose median$(X)=0$

Because $E|X-a| = E|(X-med(X))-(a-med(X))|$

and 0 is the median of $X-med(X)$

if we proof $b=0$ is the argmin of $E|(X-med(X))-b|$

the $a=0+med(X)=med(X)$ is the argmin of $E|X-a|$

Suppose the density is $f(x)$

Then for any a: $E|X-a|-E|X-med(X)|$

$$= E|X-a|-E|X|$$

$$= \int_{-\infty}^{a}(a-x)f(x)dx + \int_{a}^{+\infty}(x-a)f(x)dx$$

$$- \int_{-\infty}^{0}-xf(x)dx - \int_{0}^{+\infty}xf(x)dx$$

$$= aP(X\leq a) - aP(X>a) + 2\int_{a}^{0}xf(x)dx$$

1° if $a>0$, then $E|X-a|-E|X|$

$$= aP(X\leq a)-aP(X>a)-2\int_{0}^{a}xf(x)dx$$

$$\xrightarrow{\text{continuous}} > aP(X\leq a)-aP(X>a)-2a\int_{0}^{a}f(x)dx$$

$$= aP(X\leq a)-aP(X>a)-2a(P(X\leq a)-\tfrac{1}{2})$$

$$= a[1-P(X>a)] > a(1-\tfrac{1}{2}) > 0$$

2° if $a<0$, then: $E|X-a|-E|X| = aP(X\leq a)-aP(X>a)+2\int_{a}^{0}xf(x)dx$

$$> aP(X\leq a)-aP(X>a)+2a(P(X>a)-\tfrac{1}{2})$$

$$= a(P(X\leq a)-1) = (-a)[1-P(X\leq a)]$$

$$> -a\cdot\tfrac{1}{2} > 0.$$

So $a=0$ is indeed the minimum point.

4: 1: $Pr(\theta|X) = \dfrac{Pr(X|\theta)Pr(\theta)}{Pr(X)} = \dfrac{Pr(X|\theta)Pr(\theta)}{\int_R Pr(X|\theta)Pr(\theta)\,d\theta}$

$$Pr(X|\theta)\cdot Pr(\theta) \sim e^{-\frac{(x-\theta)^2}{2\sigma_0^2}} \cdot e^{-\frac{(\theta-\mu_0)^2}{2\tau_0^2}}$$

$$= e^{-\frac{1}{2}\left(\frac{\theta^2}{\tau_0^2}+\frac{\theta^2}{\sigma_0^2} - \left(\frac{2x}{\sigma_0^2}+\frac{2\mu_0}{\tau_0^2}\right)\theta + \frac{x^2}{\sigma_0^2}+\frac{\mu_0^2}{\tau_0^2}\right)}$$

$$\sim e^{-\frac{1}{2}\left[\theta^2\cdot\left(\frac{1}{\tau_0^2}+\frac{1}{\sigma_0^2}\right) - 2\left(\frac{x}{\sigma_0^2}+\frac{\mu_0}{\tau_0^2}\right)\theta\right]}$$

$$\sim e^{-\frac{1}{2}\left(\frac{\sigma_0^2\tau_0^2}{\sigma_0^2+\tau_0^2}\right)^{-1}\cdot\left(\theta - \frac{X\tau_0^2+\mu_0\sigma_0^2}{\tau_0^2+\sigma_0^2}\right)^2}$$

$$\Rightarrow Pr(\theta|X) = N\left(\frac{X\tau_0^2+\mu_0\sigma_0^2}{\tau_0^2+\sigma_0^2}, \frac{\sigma_0^2\tau_0^2}{\sigma_0^2+\tau_0^2}\right)$$

2: $E\{E[l(\theta,S(X))|\theta]\} = E\{l[\theta,S(X)]\} = E\{E[l(\theta,S(X))|X]\}$

for $E[l(\theta,S(X))|X]$ we can minimize every $E[l(\theta,S(X))|X=x]$
$= E\{|\theta - S(x)|\ |X=x\}$

to get the minimizer of original function

Since give $X=x$, $\theta \sim N\left(\frac{x\tau_0^2+\mu_0\sigma_0^2}{\tau_0^2+\sigma_0^2}, \frac{\sigma_0^2\tau_0^2}{\sigma_0^2+\tau_0^2}\right)$, $S(x)$ should be choosen as the median of $Pr(\theta|x)$

So $\delta(X) = \dfrac{X\tau_0^2+\mu_0\sigma_0^2}{\tau_0^2+\sigma_0^2}$

5: Suppose samples are $X_1, X_2 \cdots X_n$, then

$$Pr(X|\theta)\cdot Pr(\theta) \sim \prod_{i=1}^{n} e^{-\frac{(X_i-\theta)^2}{2\sigma_0^2}} \cdot e^{-\frac{\theta^2}{2\tau_0^2}}$$

$$\sim e^{-\frac{1}{2}\left(\frac{\tau_0^2\sigma_0^2}{n\tau_0^2+\sigma_0^2}\right)^{-1}\left[\theta-\frac{(\Sigma X_i)\tau_0^2}{n\tau_0^2+\sigma_0^2}\right)^2}.$$

$$\Rightarrow Pr(\theta|X) = N\left(\frac{(\Sigma x_i)\tau_0^2}{n\tau_0^2+\sigma_0^2},\ \frac{\tau_0^2\sigma_0^2}{n\tau_0^2+\sigma_0^2}\right)$$

So the Bayes strategy for prior $N(0,\tau_0^2)$ is $S_{0,\tau_0} = \frac{(\Sigma X_i)\tau_0^2}{n\tau_0^2+\sigma_0^2}$

~~When~~ And consider the loss $L(\Pi_{0,\tau_0}, S_{0,\tau_0})$

$$= E_\theta E_{X|\theta}\left|\bar{X}\left(\frac{\tau_0^2}{\tau_0^2+\frac{\sigma_0^2}{n}}\right)-\theta\right| = E_\theta E_{X|\theta}[Y]$$

where $Y \sim N\left[\left(\frac{\tau_0^2}{\tau_0^2+\frac{\sigma_0^2}{n}}-1\right)\theta,\ \frac{\tau_0^4}{(\tau_0^2+\frac{\sigma_0^2}{n})^2}\cdot\frac{\sigma_0^2}{n}\right]$

$$\text{So } L(\Pi_{0,\tau_0}, S_{0,\tau_0}) = \int_{-\infty}^{+\infty} N(0,\tau_0^2)\,d\theta \int_{-\infty}^{+\infty} \underbrace{|x|\frac{1}{\sqrt{2\pi}\,\frac{\tau_0^2\sigma_0}{\sqrt{n}(\tau_0^2+\frac{\sigma_0^2}{n})^2}}\, e^{-\frac{1}{2}\frac{\left(x+\frac{\frac{\sigma_0^2}{n}}{\tau_0^2+\frac{\sigma_0^2}{n}}\theta\right)^2}{\cdots}}}_{f(\theta,\tau_0^2)}\,dx$$

$$0 \le f(\theta,\tau_0^2) \le \frac{|x|\sqrt{n}\left(1+\frac{\sigma_0^2}{n}\right)^2}{\sqrt{2\pi}\,\sigma_0}\, e^{-\frac{n}{2\sigma_0^2}\left[x+\frac{\sigma_0^2}{n}\Big/\left(\tau_0^2+\frac{\sigma_0^2}{n}\right)\theta\right]^2} \quad (\text{we suppose } \tau_0 \ge 1)$$

$$\le C|x|\, e^{-\frac{n}{2\sigma_0^2}\left(2x^2+2\left[\frac{\sigma_0^2}{n}\Big/\left(\tau_0^2+\frac{\sigma_0^2}{n}\right)\right]^2\theta^2\right)}$$

$$\le C|x|\, e^{-\frac{n}{\sigma_0^2}x^2} \quad \text{for some } C.$$

and $C|x|e^{-\frac{n}{\sigma_0^2}x^2}$ is integrable

So by dominent convergent theorem (use it in R^2, $C|x|e^{-\frac{n}{\sigma_0^2}x^2}\cdot N(0,\tau_0^2)$ dominent $N(0,\tau_0^2)f(\theta,\tau_0^2)$)

we have $$\lim_{\tau_0\to+\infty} L(\Pi_{0,\tau_0}; S_{0,\tau_0}) = \int_{-\infty}^{+\infty} N(0,\tau_0^2)\,d\theta \int_{-\infty}^{+\infty} \lim_{\tau_0\to+\infty} f(\theta,\tau_0^2)\,dx$$

$= E[|Z|]$ where $Z \sim N(0, \frac{\sigma_0^2}{n})$

Also, for strategy $S(X) = \bar{X}$

the loss is ~~$E[\theta|X]$~~ $= L(\theta, X) = E_{X|\theta}[|\bar{X} - \theta|]$

$= E[|Z|]$ $Z \sim N(0, \frac{\sigma_0^2}{n})$

is constant to θ and of course $\leq E[|Z|]$

So by theorem of (Ferguson 1967)

$S(X) = \bar{X}$ is minimax strategy.