Statistical Theory Problem Set 1

Bohao Tang

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

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:

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

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

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

$= E\{l_i(S(x))\,|\,\theta(i) = k\}$ --- because of the conditional independence

$= l(k, S(x))$ --- by definition

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 un practicle;

Let $l_i(a_1) = \begin{cases} 6 & \text{probability } \frac{1}{2} \\ -3 & \cdots\cdots \frac{1}{2} \end{cases}$

$l_i(a_2) = 5 l_i(a_1) - 5$ (determined by $l_i(a_1)$) $= \begin{cases} 25 & \cdots \frac{1}{2} \\ -20 & \cdots \frac{1}{2} \end{cases}$

$X_i = \begin{cases} +1 & \cdots \frac{4}{9} \\ -1 & \cdots \frac{5}{9} \end{cases}$ and $X_i \perp l_i(a_1)$, so $X_i \perp l_i(a_2)$

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

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

$$E[l_i(a_1) \mid \theta(i)=1] = 6 \cdot P[l_i(a_1)=6 \mid X_i \cdot l_i(a_1) \le 0] - 3 P(l_i(a_1)=-3 \mid X_i \cdot l_i(a_1) \le 0)$$

$$= 6 \cdot \frac{\frac{1}{2} \cdot \frac{5}{9}}{\frac{1}{2} \cdot \frac{4}{9} + \frac{1}{2} \cdot \frac{5}{9}} - 3 \cdot \frac{4}{9} = 2.$$

$$E[l_i(a_1) \mid \theta(i)=2] = 6 \cdot \frac{4}{9} - 3 \cdot \frac{5}{9} = 1$$

and $E[l_i(a_2) \mid \theta(i)] = 5 E[l_i(a_1) \mid \theta(i)] - 5$

So they satisfies the condition in this question

But:

$$E\{l[\theta(i), S(X_i)] \mid \theta(i)=1\} = 2 \cdot \frac{\frac{1}{2} \cdot \frac{4}{9}}{\frac{1}{2}} + 5 \cdot \frac{\frac{1}{2} \cdot \frac{5}{9}}{\frac{1}{2}} = \frac{33}{9} = \frac{11}{3}$$

$$E\{l_i[S(X_i)] \mid \theta(i)=1\} = -3 \cdot \frac{\frac{1}{2} \cdot \frac{4}{9}}{\frac{1}{2}} + 25 \cdot \frac{\frac{1}{2} \cdot \frac{5}{9}}{\frac{1}{2}} = \frac{113}{9}$$

so $E[l[\theta(i), S(X_i)] \mid \theta(i)] \neq E[l_i\, S(X_i) \mid \theta(i)]$

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

$E[L_i(S(X_i)) \mid \theta(i)=1] = 2.172854$ / 2.18006754

$E[L_i(S(X_i)) \mid \theta(i)=2] = \cancel{0.997985}$ $\cancel{0.023}$ 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: $E|X-a| - E|X-med(X)|$

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

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

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

$$= a\,P(X \le a) - a\,P(X > a) + 2\int_{a}^{0} x f(x)\,dx$$

1° if $a>0$, then $E|X-a| - E|X|$

$$= a\,P(X \le a) - a\,P(X>a) - 2\int_{0}^{a} x f(x)\,dx$$

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

$$= a\,P(X \le a) - a\,P(X>a) - 2a\left(P(X \le a) - \tfrac{1}{2}\right)$$

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

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

$$> a\,P(X \le a) - a\,P(X>a) + 2a\left(P(X>a) - \tfrac{1}{2}\right)$$

$$= a\left(P(X \le a) - 1\right) = (-a)\left[1 - P(X \le a)\right]$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

So by dominent convergent theorem (use it in $\mathbb{R}^2$, $C|x|e^{-\frac{n}{\sigma_0^2}x^2}\cdot N(0,\tau_0^2)$ 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 $\le E[|Z|]$

So by theorem of (Ferguson 1967)

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