Statistical Theory Problem Set 4

Bohao Tang

(i) Suppose ten explanations are described by $\theta_1 \cdots \theta_{10}$

Then $E\left[l(\hat{\theta},\theta) \mid \theta=\theta_i\right] = P\left(\hat{\theta}(X) \neq \theta_i \mid \theta=\theta_i\right)$

$$= 1 - P\left(\hat{\theta}(X)=\theta_i \mid \theta=\theta_i\right)$$

$$= 1 - \int_{\hat{\theta}(X)=\theta_i} \prod_{t=1}^{n} f_i(X_t)\, dX.$$

(ii) We have the likelihood $P(X \mid \theta=\theta_i) = \prod_{t=1}^{n} f_i(X_t)$

and prior $P(\theta=\theta_i) = \pi_i$

So the posterior $P(\theta \mid X) \sim \sum_{i=1}^{10}\left[1_{\{\theta=\theta_i\}}\, \pi_i \prod_{t=1}^{n} f_i(X_t)\right]$

The Bayes estimator is then derived by maximize $\sum_{i=1}^{10} \pi_i 1_{\{\theta=\theta_i\}} \prod_{t=1}^{n} f_i(X_t)$

which is equal as finding i_0 with biggest $\pi_i \prod_{t=1}^{n} f_i(X_t)$

So $\hat{\theta}_{Bayes}(X) = \underset{k\in\{1,\cdots,10\}}{\arg\max}\ \pi_k \prod_{t=1}^{n} f_k(X_t)$

(iii) If complete class theorem holds, then all admissible estimators $\hat{\theta}$ are Bayes estimator, so the estimator have form of

$$\underset{k\in\{1,\cdots,10\}}{\arg\max}\ \pi_k \prod_{t=1}^{n} f_k(X_t) \qquad \pi_k \in [0,1],\ \sum \pi_k = 1$$

If we constrain the ~~est~~ decision space to let Bayes estimator be unique, then form ① are all the ~~others~~ admissible estimators since unique Bayes estimator ~~one~~ is admissible.

(iv) MLE of θ is admissible in every situation.

Here we have MLE is the Bayes estimator of prior $\pi_k = \frac{1}{10}$ $\forall k$.

we prove if $\pi_k > 0$ $\forall k$, then its Bayes estimator $\hat{\theta}_\pi$ is admissible.

Proof: Denote risk $E[l(\hat{\theta}, \theta) \mid \theta = \theta_i]$ be $R(\hat{\theta}, \theta_i)$

then if $\hat{\theta}_\pi$ is not admissible then there exists $\check{\theta}$ such that

$$R(\check{\theta}, \theta_i) \leq R(\hat{\theta}_\pi, \theta_i) \quad \forall i$$

and $\exists i_0$ $R(\check{\theta}, \theta_{i_0}) < R(\hat{\theta}_\pi, \theta_{i_0})$

$\Rightarrow$ ~~$\sum_{i=1}^{10} R(\check{\theta}, \theta_i) < \sum_{i=1}^{10} R(\hat{\theta}_\pi, \theta_i)$~~ $\sum_{i \neq i_0} \pi_i R(\check{\theta}, \theta_i) \leq \sum \pi_i R(\hat{\theta}_\pi, \theta_i)$; $\pi_{i_0} R(\check{\theta}, \theta_{i_0}) < \pi_{i_0} R(\hat{\theta}_\pi, \theta_{i_0})$

$$\Rightarrow \sum_{i=1}^{10} \pi_i R(\check{\theta}, \theta_i) < \sum_{i=1}^{10} \pi_i R(\hat{\theta}_\pi, \theta_i) \Rightarrow E_\pi R(\check{\theta}, \theta_i) < E_\pi R(\hat{\theta}_\pi, \theta_i)$$

which is a contradiction to $\hat{\theta}_\pi$ being Bayes estimator for π.

(V) Since the posterior is $p(\theta \mid x) = \dfrac{\sum_{i=1}^{10} \pi_i 1_{\{\theta = \theta_i\}} \prod_{t=1}^{n} f_i(x_t)}{\sum_{i=1}^{10} \pi_i \prod_{t=1}^{n} f_i(x_t)}$

and the likelihood $p(x \mid \theta) = \sum_{i=1}^{10} 1_{\{\theta = \theta_i\}} \prod_{t=1}^{n} f_i(x_t)$

Whether we use Bayes or MLE, $\hat{\theta}$ is far from being sufficient.

Maybe I also want to keep

1° the first 3 θ_i with biggest likelihood.

2° the likelihood ratio between θ_i of biggest and second biggest likelihood.