Statistical Inference I

Final exam: 2012/1/15(Tues)

Q1
Q2
Q3
Q4

YOUR NAME______________________

NOTE1: Please write down the derivation of your answer very clearly for all questions. The score will be reduced when you only write answer. Also, the score will be reduced if the derivation is not clear. The score will be added even when your answer is incorrect but the derivation is correct.

1. Bayes optimality, Admissibility and Mimimaxity

Assume i.i.d. samples $X = (X_1,..., X_n) \sim Bin(n, p)$. Consider estimation of p under the square loss $L(a, p) =| p - a |^2$.

1) Show that $\overline{X}$ is the unique UMVUE of p (with proof).

2) Find the Bayes estimator under the prior $p \sim Beta(\alpha, \beta)$ (with derivation)

3) Find the mininimax estimator of p (with proof).

4) Calculate the risk of a randomized estimator

$$\delta(X) = \begin{cases} \overline{X} & \text{with probability } n/(n+1) \\ 1/2 & \text{with probability } 1/(n+1) \end{cases}$$

5) Is $\delta(X)$ mimimax? Is $\delta(X)$ admissible?

6) Draw the graph of risks for the 3 estimators ($\overline{X}$, minimax estimator, and $\delta(X)$) when n=4.

7) Find the asymptotic MSEs of $\delta(X)$ and $\overline{X}$ ($amse_{\delta(X)}(p)$; $amse_{\overline{X}}(p)$). Which is asymptotically more efficient?

2. MLE

Let $X_1,..., X_n \overset{iid}{\sim} f_\theta(x)$, where

$$f_\theta(x) = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{ -\frac{(x-\theta)^2}{2\sigma^2} \right\} I(x \geq \theta)$$

is a truncated normal distribution, truncated at unknown value $\theta \in R$.

1) Find the MLE $\hat{\theta}$.

2) Show $\hat{\theta}_n \xrightarrow{P} \theta$.

3. Asymptotics for MLE

Let $X_1,\ldots,X_n \overset{iid}{\sim} f_\theta(x)$, where $\theta \in \Theta$. Also, let $s_n(\theta)=\frac{\partial}{\partial\theta}\log \ell(\theta)$ and $\hat{\theta}_n$ be the MLE with $\hat{\theta}_n \xrightarrow{P} \theta$. We assume usual regularity conditions (A1)-(A7).

i) Write $\sqrt{n}(\hat{\theta}_n-\theta)$ in terms of derivatives of $s_n(\theta)$ and θ_n^*, a midpoint between θ and $\hat{\theta}_n$.

ii) Derive the asymptotics for $-\frac{1}{n}\frac{\partial}{\partial\theta}s_n(\theta)$.

iii) Derive the asymptotics for $\frac{1}{\sqrt{n}}s_n(\theta)$.

iv) Derive the asymptotics for $\frac{1}{2n}\frac{\partial^2}{\partial\theta^2}s_n(\theta_n^*)(\hat{\theta}_n-\theta)$.

v) Derive the asymptotics for $\sqrt{n}(\hat{\theta}_n-\theta)$.

4. Simultaneous estimation

Based on data $X \sim N_p(\theta, I_p)$, consider simultaneous estimation of $\theta=(\theta_1,\ldots,\theta_p)'$ under the loss $\|a-\theta\|^2$.

i) Show that X is UMVUE of θ. In UMVUE, what “V” stand for?

ii) Derive the Bayes estimator under a prior $\theta_1,\ldots,\theta_p \overset{iid}{\sim} N(\xi,\tau^2)$

iii) Show that X is minimax.

iv) Applying a Theorem in the textbook, calculate the risk of the James-Stein estimator $\delta(X)=X-\frac{p-2}{\|X\|^2}X$.

v) Show that $\delta(X)$ is also minimax.

iv) Calculate the Bayes risk of $\delta(X)$ under a prior $\theta_1,\ldots,\theta_p \overset{iid}{\sim} N(0,\tau^2)$

Answers: This is a simplified answer. In the exam, you need to write more detailed calculations.

<u>Answer 1</u>

1) $\overline{X}$ is CSS and unbiased. Hence, it is UMVUE.

2) $\delta^{Bayes}(X) = \dfrac{n}{n+\alpha+\beta}\overline{X} + \dfrac{\alpha+\beta}{n+\alpha+\beta}\dfrac{\alpha}{\alpha+\beta}$

3)

$$R_{\delta^{Bayes}}(p) = E(\delta^{Bayes}(X) - p)^2 = E\left\{\frac{n}{n+\alpha+\beta}(\overline{X}-p) + \frac{\alpha+\beta}{n+\alpha+\beta}\left(\frac{\alpha}{\alpha+\beta} - p\right)\right\}^2$$

$$= \left(\frac{n}{n+\alpha+\beta}\right)^2 E(\overline{X}-p)^2 + \left(\frac{\alpha+\beta}{n+\alpha+\beta}\right)^2\left(\frac{\alpha}{\alpha+\beta} - p\right)^2$$

$$= \frac{1}{(n+\alpha+\beta)^2}[np(1-p) + \{\alpha - (\alpha+\beta)p\}^2]$$

$$R_{\delta^{Bayes}}(p) = E(\delta^{Bayes}(X) - p)^2 = E\left\{\frac{n}{n+\alpha+\beta}(\overline{X}-p) + \frac{\alpha+\beta}{n+\alpha+\beta}\left(\frac{\alpha}{\alpha+\beta} - p\right)\right\}^2$$

$$= \left(\frac{n}{n+\alpha+\beta}\right)^2 E(\overline{X}-p)^2 + \left(\frac{\alpha+\beta}{n+\alpha+\beta}\right)^2\left(\frac{\alpha}{\alpha+\beta} - p\right)^2$$

$$= \frac{1}{(n+\alpha+\beta)^2}[np(1-p) + \{\alpha - (\alpha+\beta)p\}^2]$$

For $R_{\delta^{Bayes}}(p)$ to be constant risk, we solve

$$\frac{d}{dp}R_{\delta^{Bayes}}(p) \propto n - 2np - 2(\alpha+\beta)\{\alpha - (\alpha+\beta)p\}$$

$$= n - 2\alpha(\alpha+\beta) - 2np + 2(\alpha+\beta)^2 p = 0 \qquad \forall p$$

Thus, for $\alpha = \beta = \sqrt{n}/2$,

$$\delta^{\min}(X) = \frac{n}{n+\sqrt{n}}\overline{X} + \frac{\sqrt{n}}{n+\sqrt{n}}\frac{1}{2}$$

is the Bayes estimator with constant risk. Hence it is minimax.

4)

$$R_{\delta}(p) = E(\delta(X) - p)^2 = \frac{n}{n+1}E(\overline{X}-p)^2 + \frac{1}{n+1}(1/2 - p)^2$$

$$= \frac{p(1-p)}{n+1} + \frac{(1/2-p)^2}{n+1} = \frac{1}{4(n+1)}$$

5) $\delta(X)$ is not mimimax. $\delta(X)$ is not admissible

6) Omit

7) $amse_{\delta(X)}(p) = 1/4$, $amse_{\overline{X}}(p) = p(1-p)$. $\overline{X}$ is asymptotically more efficient.

Answer 2

1) Maximizing the likelihood function

$$l(\theta)=\left(\frac{2}{\pi\sigma^2}\right)^{n/2}\exp\left\{-\frac{\sum_{i=1}^n (X_i-\theta)^2}{2\sigma^2}\right\}\exp\left\{-\frac{n(\overline{X}-\theta)^2}{2\sigma^2}\right\}I(X_{(1)}\geq\theta).$$

is equivalent to mimimizing

$$\phi(\theta)=(\overline{X}-\theta)^2 \quad \text{subject to} \quad X_{(1)}\geq\theta.$$

Since $\phi(\theta)$ is convex with minimum at $\theta=\overline{X}$, the MLE is is $\hat{\theta}=X_{(1)}$.

2) Note that for any $\varepsilon>0$, $\Pr(X>\varepsilon+\theta)<1$. Hence,

$$\Pr(|X_{(1)}-\theta|>\varepsilon)=\Pr(X_{(1)}>\varepsilon+\theta)=\Pr(X>\varepsilon+\theta)^n\to 0 \quad (n\to\infty).$$

Answer 3

i) Since $\hat{\theta}_n$ is MLE, $s_n(\hat{\theta}_n)=0$. Using a Taylor expansion,

$$0=s_n(\theta)+\frac{\partial}{\partial\theta}s_n(\theta)(\hat{\theta}_n-\theta)+\frac{1}{2}\frac{\partial^2}{\partial\theta^2}s_n(\theta_n^*)(\hat{\theta}_n-\theta)^2.$$

Multiplying $1/\sqrt{n}$ to both sides,

$$\begin{aligned}0&=\frac{1}{\sqrt{n}}s_n(\theta)+\frac{1}{n}\frac{\partial}{\partial\theta}s_n(\theta)\sqrt{n}(\hat{\theta}_n-\theta)+\frac{1}{2n}\frac{\partial^2}{\partial\theta^2}s_n(\theta_n^*)(\hat{\theta}_n-\theta)\sqrt{n}(\hat{\theta}_n-\theta)\\&=\frac{1}{\sqrt{n}}s_n(\theta)+\left\{\frac{1}{n}\frac{\partial}{\partial\theta}s_n(\theta)+\frac{1}{2n}\frac{\partial^2}{\partial\theta^2}s_n(\theta_n^*)(\hat{\theta}_n-\theta)\right\}\sqrt{n}(\hat{\theta}_n-\theta)\end{aligned}.$$

Hence,

$$\sqrt{n}(\hat{\theta}_n-\theta)=\frac{\frac{1}{\sqrt{n}}s_n(\theta)}{-\frac{1}{n}\frac{\partial}{\partial\theta}s_n(\theta)-\frac{1}{2n}\frac{\partial^2}{\partial\theta^2}s_n(\theta_n^*)(\hat{\theta}_n-\theta)}.$$

ii) $-\frac{1}{n}\frac{\partial}{\partial\theta}s_n(\theta)=-\frac{1}{n}\sum_{i=1}^n\frac{\partial^2}{\partial\theta^2}\log f_\theta(X_i)\to_p I(\theta)$

iii) $\frac{1}{\sqrt{n}}s_n(\theta)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{\partial}{\partial\theta}\log f_\theta(X_i)\to_d N(0,I(\theta))$

iv) $R_n=\frac{1}{2n}\frac{\partial^2}{\partial\theta^2}s_n(\theta_n^*)(\hat{\theta}_n-\theta)\xrightarrow{P}0$.

$\because$ $\Pr(|R_n|>\varepsilon)=\Pr(|R_n|>\varepsilon,|\hat{\theta}_n-\theta|>\delta)+\Pr(|R_n|>\varepsilon,|\hat{\theta}_n-\theta|\leq\delta).$

$\Pr(|R_n| > \varepsilon, |\hat{\theta}_n - \theta| > \delta) \le \Pr(|\hat{\theta}_n - \theta| > \delta) \to 0$

By condition (A7), there exists an integrable function $M(x)$ such that

$\sup_{|\theta - t| \le \delta} \left| \frac{\partial^3}{\partial \theta^3} \log f_\theta(x) \right|_{\theta = t} \le M(x)$. If $|\hat{\theta}_n - \theta| \le \delta$,

$$\left| \frac{1}{n} \frac{\partial^2}{\partial \theta^2} s_n(\theta_n^*) \right| = \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^3}{\partial \theta^3} \log f_\theta(X_i) \right| \le \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\partial^3}{\partial \theta^3} \log f_\theta(X_i) \right| \le \frac{1}{n} \sum_{i=1}^{n} M(X_i).$$

Hence, if $|\hat{\theta}_n - \theta| \le \delta$, $|R_n| \le \frac{\delta}{2} \frac{1}{n} \sum_{i=1}^{n} M(X_i) \xrightarrow{P} \frac{\delta}{2} E[M(X_1)]$. By taking δ arbitrary small, $\Pr(|R_n| > \varepsilon, |\hat{\theta}_n - \theta| \le \delta)$. become arbitrary small. Therefore, $\Pr(|R_n| > \varepsilon, |\hat{\theta}_n - \theta| \le \delta) \to 0$.

v) $\sqrt{n}(\hat{\theta}_n - \theta) = \dfrac{\frac{1}{\sqrt{n}} s_n(\theta)}{I(\theta) + o_P(1)} \to_d N(0, I(\theta)^{-1})$.

Answer 4

i) Let $\tilde{\delta}(X)$ be an unbiased estimator, i.e., $E[\tilde{\delta}_i(X)] = \theta_i$. Since X_i is UMVUE, $E|\tilde{\delta}_i(X) - \theta_i|^2 \ge E|X_i - \theta_i|^2$. Hence,

$$E\|\tilde{\delta}(X) - \theta\|^2 = \sum_{i=1}^{p} E|\tilde{\delta}_i(X) - \theta_i|^2 \ge \sum_{i=1}^{p} E|X_i - \theta_i|^2 = E\|X - \theta\|^2.$$

As shown in the preceding inequality, $\mathrm{trVar}\{\tilde{\delta}(X)\} \ge \mathrm{trVar}(X)$. Hence, "V" stands for "trace of variance".

ii) $\delta_i^{Bayes}(X) = E[\theta_i | X] = E[\theta_i | X_i] = \dfrac{1}{\tau^2 + 1}\xi + \dfrac{\tau^2}{\tau^2 + 1} X_i$ (Example 2.25)

Hence, $\delta^{Bayes}(X) = \dfrac{\tau^2}{\tau^2 + 1} \xi \mathbf{1}_n + \dfrac{\tau^2}{\tau^2 + 1} X$.

iii) $X = \lim_{\tau^2 \to \infty} \delta^{Bayes}(X)$:

The limit of Bayes estimator with constant risk is mimimax (Theorem 4.12).

$$
\begin{aligned}
R_{\delta^{Bayes}}(\theta) &= \sum_{i=1}^{p} E_\theta\left(\frac{1}{\tau^2+1}\xi + \frac{\tau^2}{\tau^2+1}X_i - \theta_i\right)^2 \\
&= \sum_{i=1}^{p}\left[\left(\frac{1}{\tau^2+1}\right)^2(\xi-\theta_i)^2 + \left(\frac{\tau^2}{\tau^2+1}\right)^2 E_\theta(X_i-\theta_i)^2\right] \\
&= \left(\frac{1}{\tau^2+1}\right)^2\sum_{i=1}^{p}(\xi-\theta_i)^2 + p\left(\frac{\tau^2}{\tau^2+1}\right)^2 \\
&\xrightarrow[\tau^2\to\infty]{} p \quad \text{(constant risk)}
\end{aligned}
$$

iv) $R_\delta(\theta) = p - (p-2)^2 E_\theta[\| X \|^{-2}]$.

v) Note that $\lim_{\theta\to\infty} E_\theta[\| X \|^{-2}] = 0$. Hence, $\sup_\theta R_\delta(\theta) = p = R_X(\theta)$.

vi) Since $X \mid \theta \sim N_p(\theta, I_p)$ and $\theta_1, ..., \theta_p \overset{iid}{\sim} N(0,\tau^2)$, we have

$E[X] = EE[X \mid \theta] = E(\theta) = 0$, $Var[X] = EVar[X \mid \theta] + VarE[X \mid \theta] = (1+\tau^2)I_p$.

It follows that $X \sim N(0,(1+\tau^2)I_p)$, and hence, $\dfrac{\| X \|^2}{1+\tau^2} \sim \chi_p^2$.

This means that $E[\| X \|^{-2}] = \dfrac{1}{(p-2)(1+\tau^2)}$. Therefore,

$$
r_\delta(\Pi) = \int_\Theta R_\delta(\theta)d\theta = E[R_\delta(\theta)] = p - (p-2)^2 E[\| X \|^{-2}] = p - \frac{p-2}{1+\tau^2}.
$$