STAT6201: Theoretical Statistics I

Satvik Saha

Homework 6

1. (a) Suppose that δ_0 is Bayes for some prior π on Θ, and is an equalizer, i.e. has constant risk $\overline{R}(\delta_0) = R(g(\cdot), \delta_0)$ on Θ. Then, for any other estimator δ of $g(\theta)$, we have

$$\begin{aligned}
\overline{R}(\delta_0) &= R(g(\cdot), \delta_0) \\
&= \mathbb{E}_{\theta\sim\pi}[R(g(\theta)), \delta_0] && \text{(Constant risk)} \\
&\leq \mathbb{E}_{\theta\sim\pi}[R(g(\theta), \delta)] && (\delta_0 \text{ is Bayes}) \\
&\leq \sup_{\theta\in\Theta} R(g(\theta), \delta) \\
&= \overline{R}(\delta).
\end{aligned}$$

   This shows that δ_0 is minimax for Θ.

   (b) Suppose that δ_0 is extended Bayes, and is an equalizer on Θ. Let δ be some other estimator of $g(\theta)$. Fix $\epsilon > 0$, and let π be a prior such that

$$\mathbb{E}_{\theta\sim\pi}[R(g(\theta), \delta_0)] \leq \inf_{\eta} \mathbb{E}_{\theta\sim\pi}[R(g(\theta), \eta)] + \epsilon.$$

   Then,

$$\begin{aligned}
\overline{R}(\delta_0) &= R(g(\cdot), \delta_0) \\
&= \mathbb{E}_{\theta\sim\pi}[R(g(\theta)), \delta_0] && \text{(Constant risk)} \\
&\leq \mathbb{E}_{\theta\sim\pi}[R(g(\theta), \delta)] + \epsilon && (\delta_0 \text{ is extended Bayes}) \\
&\leq \sup_{\theta\in\Theta} R(g(\theta), \delta) + \epsilon \\
&= \overline{R}(\delta) + \epsilon.
\end{aligned}$$

   Since $\epsilon > 0$ was arbitrary, we must have $\overline{R}(\delta_0) \leq \overline{R}(\delta)$, whence δ_0 is minimax for Θ_0.

2. Let $X \sim \text{Binomial}(n, p)$ for some $n \geq 1$, where $p \in (0, 1)$. Set $\delta_0(X) = X/n$. Then, using $\text{var}(X) = np(1-p)$,

$$R(p, \delta_0) = \mathbb{E}_X[\ell(p, \delta_0(X))] = \mathbb{E}_X\left[\frac{(p - X/n)^2}{p(1-p)}\right] = \frac{1}{n}.$$

   Thus, δ_0 is an equalizer. Next, consider a prior π such that $\theta \sim \text{Beta}(1,1) \sim \text{Uniform}[0,1]$, then $\theta \mid X \sim \text{Beta}(1 + X, 1 + n - X)$. To find the corresponding Bayes estimator, we minimize the posterior expected losses

$$\begin{aligned}
\mathbb{E}_{p|X=x}[\ell(p, \delta(x))] &= \mathbb{E}_{p|X=x}\left[\frac{(p - \delta(x))^2}{p(1-p)}\right] \\
&= \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} \int_0^1 (p - \delta(x))^2 p^{x-1}(1-p)^{n-x-1}\, dx.
\end{aligned}$$

When $x = 0$, this is ∞ unless $\delta(x) = 0$; similarly, when $x = n$, this is ∞ unless $\delta(x) = 1$. Otherwise, this is just

$$\left(\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} \cdot \frac{\Gamma(x)\Gamma(n-x)}{\Gamma(n)}\right) \mathbb{E}_{q\sim\text{Beta}(x,n-x)}[(q-\delta(x))^2],$$

which we know is minimized when

$$\delta(x) = \mathbb{E}_{q\sim\text{Beta}(x,n-x)}[q] = \frac{x}{n}.$$

This shows that we must have $\delta_\pi(x) = x/n$.

With this, δ_0 is Bayes for π. By the result in the previous problem, it is minimax on $p \in (0,1)$. Furthermore, the corresponding Bayes risk is just $1/n$ which is equal to the minimax risk, and our Bayes estimator is unique, which makes δ_0 unique minimax.

3. Let $\Theta_0 \subseteq \Theta$, and let δ be minimax for $g(\theta)$ under Θ_0. Further suppose that

$$\sup_{\theta\in\Theta} R(g(\theta),\delta) = \sup_{\theta\in\Theta_0} R(g(\theta),\delta).$$

Then, for any estimator η for $g(\theta)$,

$$\begin{aligned}\sup_{\theta\in\Theta} R(g(\theta),\delta) &= \sup_{\theta\in\Theta_0} R(g(\theta),\delta) \\ &\leq \sup_{\theta\in\Theta_0} R(g(\theta),\eta) && (\delta \text{ minimax under } \Theta_0) \\ &\leq \sup_{\theta\in\Theta} R(g(\theta),\eta), && (\text{Supremums})\end{aligned}$$

whence δ is minimax under Θ.

4. Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, where $\mu \in \mathbb{R}$ and $\sigma^2 \in (0, \infty)$.

(a) Suppose that $\sigma^2 \in (0, M]$ for some constant $M > 0$. Note that

$$R(\mu, \bar{X}) = \mathbb{E}[(\mu - \bar{X})^2] = \frac{\sigma^2}{n},$$

so $\overline{R}(\mu, \bar{X}) = M/n$. Using the fact that $\bar{X}$ is minimax for fixed σ^2, for any other estimator δ,

$$\sup_{\mu,\sigma^2} R(\mu,\delta) = \sup_{\sigma^2\in(0,M]} \sup_{\mu\in\mathbb{R}} R(\mu,\delta) \geq \sup_{\sigma^2\in(0,M]} \frac{\sigma^2}{n} = \frac{M}{n}.$$

Now, $\bar{X}$ attains the lower bound M/n on the supremum risk over $(\mu, \sigma^2) \in \mathbb{R} \times (0, M]$, hence must be minimax.

Remark: For fixed σ^2, one can show that $\bar{X}$ has supremum risk σ^2/n, which is the limit of Bayes risks of the (least favorable) sequence of priors $N(0, m)$ as $m \to \infty$, whence $\bar{X}$ is minimax.

(b) Since $\bar{X}$ is minimax for fixed σ^2, note that when $\sigma^2 \in (0, \infty)$, we have

$$\sup_{\mu,\sigma^2} R(\mu,\delta) \geq \sup_{\sigma^2\in(0,\infty)} \frac{\sigma^2}{n} = \infty.$$

Thus, estimators such as $\bar{X}$ are minimax.

(c) Note that with the loss

$$\ell((\mu, \sigma^2), \delta) = \frac{(\delta - \mu)^2}{\sigma^2},$$

we have

$$R(\mu, \bar{X}) = \mathbb{E}\left[\frac{(\mu - \bar{X})^2}{\sigma^2}\right] = \frac{1}{n}.$$

Thus, $\bar{X}$ is an equalizer. Again, for any other estimator δ, since $\bar{X}$ is minimax for fixed σ^2,

$$\begin{aligned}
\sup_{\mu,\sigma^2} R(\mu, \delta) &= \sup_{\sigma^2 \in (0,\infty)} \sup_{\mu \in \mathbb{R}} R(\mu, \delta) \\
&= \sup_{\sigma^2 \in (0,\infty)} \sup_{\mu \in \mathbb{R}} \frac{\mathbb{E}[(\mu - \delta(X))^2]}{\sigma^2} \\
&= \sup_{\sigma^2 \in (0,\infty)} \left(\frac{1}{\sigma^2} \sup_{\mu \in \mathbb{R}} \mathbb{E}[(\mu - \delta(X))^2]\right) \\
&\geq \sup_{\sigma^2 \in (0,\infty)} \left(\frac{1}{\sigma^2} \sup_{\mu \in \mathbb{R}} \mathbb{E}[(\mu - \bar{X})^2]\right) \\
&= \frac{1}{n}.
\end{aligned}$$

Thus, $\bar{X}$ attains the lower bound $1/n$ on the supremum risk over $(\mu, \sigma^2) \in \mathbb{R} \times (0, \infty)$, hence must be minimax.

5. Let $X \mid \theta \sim N(\theta, 1)$ and $\theta \sim \pi \in \Gamma$, where Γ is a class of priors on $\mathbb{R}$ with

$$\mathbb{E}_\pi[\theta] = 0, \qquad \mathbb{E}_\pi[\theta^2] = 1.$$

(a) For $\delta_{a,b}(X) = aX + b$, observe that

$$\begin{aligned}
R(\pi, \delta) &= \mathbb{E}[(\theta - (aX + b))^2] \\
&= \mathbb{E}_\theta[\mathbb{E}_{X|\theta}[a^2(\theta - X)^2 + ((1-a)\theta - b)^2 + 2a\cancelto{0}{(\theta - X)}((1-a)\theta - b)]] \\
&= a^2 + \mathbb{E}_\theta[(1-a)^2\theta^2 + b^2 - \cancelto{0}{2b(1-a)\theta}] \\
&= a^2 + (1-a)^2 + b^2.
\end{aligned}$$

This is clearly minimized when $a = 1/2, b = 0$, with a value of $1/2$.
Remark: The map $t \mapsto \sum_i (t - x_i)^2$ is minimized at $\bar{x}$.

(b) Set $\pi_0 \sim N(0, 1)$ and $\delta_0(X) := \delta_{1/2,0}(X) = X/2$. Then, observe that for $\pi \in \Gamma$, our previous calculations give

$$\inf_\delta R(\pi, \delta) \leq \inf_{\delta_{a,b}} R(\pi, \delta_{a,b}) = \frac{1}{2},$$

so $\sup_{\pi \in \Gamma} \inf_\delta R(\pi, \delta) \leq 1/2$. On the other hand, δ_0 is Bayes for $\pi_0 \in \Gamma$, with Bayes risk $1/2$, so

$$\frac{1}{2} = R(\pi_0) = \inf_\delta R(\pi_0, \delta) \leq \sup_{\pi \in \Gamma} \inf_\delta R(\pi, \delta) \leq \frac{1}{2}.$$

This forces $\sup_{\pi \in \Gamma} \inf_\delta R(\pi, \delta) = 1/2$.

6. Let $X_1, \dots, X_n \overset{iid}{\sim} F \in \mathcal{P}$, where $\mathcal{P}$ is the class of probability distributions supported on $[0,1]$. Set $\mu_F = \mathbb{E}_{X\sim F}[X]$ and $\sigma_F^2 = \mathrm{var}_{X\sim F}[X]$.

   Now, let
$$\delta_0(X) = \frac{\sqrt{n}\bar{X} + \frac{1}{2}}{\sqrt{n}+1}.$$
   We will show that δ_0 is minimax for μ_F under the squared error loss.

   (a) Note that
$$\begin{aligned}
R(\mu_F, a\bar{X}+b) &= \mathbb{E}[(\mu_F - (a\bar{X}+b))^2] \\
&= \mathbb{E}[a^2(\mu_F - \bar{X})^2 + ((1-a)\mu_F - b)^2 + 2a\cancelto{0}{(\mu_F - \bar{X})}((1-a)\mu_F - b)] \\
&= \frac{1}{n}a^2\sigma_F^2 + ((1-a)\mu_F - b)^2,
\end{aligned}$$
   so for δ_0,
$$\begin{aligned}
R(\mu_F, \delta_0) &= \frac{\sigma_F^2}{(\sqrt{n}+1)^2} + \left(\frac{\mu_F - \frac{1}{2}}{\sqrt{n}+1}\right)^2 \\
&= \frac{1}{(\sqrt{n}+1)^2}\mathbb{E}_{X\sim F}\left[\left(X - \frac{1}{2}\right)^2\right] \\
&\le \frac{1}{4(\sqrt{n}+1)^2}. && (|X - 1/2| \le 1/2)
\end{aligned}$$
   Furthermore, this upper bound is attained when $F \sim \mathrm{Bernoulli}(\frac{1}{2})$; note that in that case, $\mu_F = 1/2$ and $\sigma_F^2 = 1/4$. Thus,
$$\sup_{F\in\mathcal{P}} R(\mu_F, \delta_0) = \frac{1}{4(\sqrt{n}+1)^2}.$$

   (b) Let $\mathcal{P}_0 \subseteq \mathcal{P}$ be the class of distributions where the above supremum is attained. For this, we demand $\mathbb{E}[(X-1/2)^2] = 1/4$, which is possible only when $|X - 1/2| = 1/2$ almost surely. This forces $X \in \{0,1\}$ almost surely, i.e. $X \sim \mathrm{Bernoulli}(p)$ for some $p \in [0,1]$. Thus,
$$\mathcal{P}_0 = \{\mathrm{Bernoulli}(p) \colon p \in [0,1]\}.$$
   With this, let π be a prior on $\mathcal{P}$ taking probability 1 on $\mathcal{P}_0$, with $p \sim \mathrm{Beta}(\sqrt{n}/2, \sqrt{n}/2)$. Then, we have the posterior $p \mid X \sim \mathrm{Beta}(n\bar{X} + \sqrt{n}/2, n - n\bar{X} + \sqrt{n}/2)$, whence the Bayes estimator is the posterior expectation
$$\delta_\pi(X) = \frac{n\bar{X} + \sqrt{n}/2}{n + \sqrt{n}} = \delta_0(X).$$
   The Bayes risk for this estimator is precisely the supremum risk in (a), since $R(\mu_F, \delta_0)$ is constant on $F \in \mathcal{P}_0$.

   (c) We have shown that δ_0 is minimax for μ_F under $\mathcal{P}_0$, since it is the Bayes estimator under π and an equalizer. Since $\sup_{F\in\mathcal{P}_0} R(\mu_F, \delta_0) = \sup_{F\in\mathcal{P}} R(\mu_F, \delta_0)$, we must have δ_0 minimax for μ_F under $\mathcal{P}$ via the result in Problem 3.