Week 3
Spring 2009

Lecture 5. Bayes estimation, minimaxity and Admissibility (cont.).

Admissibility

Conditions on priors and admissibility: conditions on the prior measure which guarantees that the corresponding generalized Bayes procedure is admissible.

Define

$$J_x(h) = \int h(\theta)\varphi(x-\theta)\,d\theta.$$

Let $S_1 = \{x \in \mathbb{R}^p : \|x\| \le 1\}$.

Assumption

Growth Condition:

$$\int_{S_1^c} \frac{g(\theta)}{\|\theta\|^2 \log^2(\|\theta\|)} < \infty$$

Asymptotic flatness condition:

$$\int_{S_1^c} J_x\left\{\left\|\frac{\nabla g}{g} - \frac{J_x(\nabla g)}{J_x(g)}\right\|^2 g\right\} dx < \infty$$

It can be shown that $\int_{S_1^c} \|\nabla g\|^2 / g dx < \infty$ implies the flatness condition.

Theorem. Let G be a prior satisfying two conditions above. Then δ_G is admissible.

For the normal mean estimation problem with squared error loss,

Blyth's Method. Let δ be an estimator. Let $\{G_j\}$ be a sequence of finite prior measures such that:(i) $r(G_j, \delta) - r(G_j, \delta_{G_j}) \to 0$ as $j \to \infty$; (ii) $\inf_j \{G_j(S_1)\} > 0$. Then δ is an admissible estimator.

<u>hint</u>: Let $\delta'' = (\delta' + \delta)/2$. If $R(\theta, \delta') \le R(\theta, \delta)$ for all θ and with strict inequality for some θ, then $R(\theta, \delta'') < R(\theta, \delta)$ for all θ. For all j,

$$\begin{aligned} r(G_j, \delta_{G_j}) &\le r(G_j, \delta'') \le \int_{S_1} R(\theta, \delta'') G_j(d\theta) + \int_{S_1^c} R(\theta, \delta) G_j(d\theta) \\ &= r(G_j, \delta) + \int_{S_1} \left[R(\theta, \delta'') - R(\theta, \delta)\right] G_j(d\theta) \le r(G_j, \delta) - \varepsilon \end{aligned}$$

where $\varepsilon = \int_S \left[R(\theta, \delta'') - R(\theta, \delta)\right] G_j(d\theta)$. Contradiction!

Example. Let $X \sim N(\theta, 1)$. Let $g_j(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2j}\right)$, and $g_j = dG_j/d\mu$ where μ is the Lebesgue measure. It is easy to show

$$r(G_j, X) = \sqrt{j},\ r(G_j, \delta_{G_j}) = \sqrt{j}\frac{j}{j+1}$$

then

$$r(G_j, X) = \frac{\sqrt{j}}{j+1} \to 0.$$

Proposition.

$$r(G,\delta) - r(G,\delta_G) = \int \|\delta_G - \delta\|^2 g^*(x)\, dx$$

<u>*Proof*</u> :

$$\begin{aligned} r(G,\delta) - r(G,\delta_G) &= E_X E\left(\|\theta - \delta_G + \delta_G - \delta\|^2 - \|\theta - \delta_G\|^2 |X\right) \\ &= E_X E\left(\|\delta_G - \delta\|^2 |X\right) = \int \|\delta_G - \delta\|^2 g^*(x)\, dx \end{aligned}$$

<u>*Proof of the theorem*</u>: Please read page 374 of Stein (1961).
Define $g_j = h_j^2 g$ where

$$h_j = \begin{cases} 1 & \|\theta\| \le 1 \\ 1 - \frac{\log(\|\theta\|)}{\log j} & 1 \le \|\theta\| \le j \\ 0 & \|\theta\| > j \end{cases}, \ j = 2, 3, \ldots.$$

It is easy to see

$$\delta_{G_j} = \frac{\int \theta h_j^2 g(\theta)\varphi(x-\theta)\, d\theta}{\int h_j^2 g(\theta)\varphi(x-\theta)\, d\theta} \to \delta_G \text{ a.s.}$$

and

$$g_j^*(x) = \int h_j^2 g(\theta)\varphi(x-\theta)\, d\theta \le g^*(x).$$

Write

$$\begin{aligned} \delta_G(x) &= x + \frac{\nabla g^*}{g^*} = x + \frac{J_x(\nabla g)}{J_x(g)}, \\ \delta_{G_j}(x) &= x + \frac{\nabla g^*}{g^*} = x + \frac{J_x\left(h_j^2 \nabla g + g \nabla h_j^2\right)}{J_x\left(h_j^2 g\right)} \end{aligned}$$

where the second equality for each equation follows from integration by parts.
Hence

$$\begin{aligned} & r(G_j, \delta_G) - r(G_j, \delta_{G_j}) \\ = & \int_{S_1^c} \left\|\delta_G - \delta_{G_j}\right\|^2 g_j^*(x)\, dx + \int_{S_1^c} \left\|\delta_G - \delta_{G_j}\right\|^2 g_j^*(x)\, dx \ (\leftarrow \text{ apply DCT, since } g_j^*(x) \le g^*(x) \text{ finite.}) \\ \le & \ 2\int_{S_1^c} \left\| \frac{J_x\left(g \nabla h_j^2\right)}{J_x(g_j)} \right\|^2 g_j^*(x)\, dx + 2\int_{S_1^c} \left\| \frac{J_x(\nabla g)}{J_x(g)} - \frac{J_x\left(h_j^2 \nabla g\right)}{J_x(g_j)} \right\|^2 g_j^*(x)\, dx + o(1) \\ = & \ 2A_j + 2B_j + o(1) \end{aligned}$$

Show $A_j \to 0$ by DCT:

$$\begin{aligned}
A_j &= \int_{S_1^c} \left\| \frac{J_x\left(g\nabla h_j^2\right)}{J_x\left(g_j\right)} \right\|^2 g_j^*\left(x\right) dx \\
&= 4\int_{S_1^c} \left\| \frac{J_x\left(gh_j\nabla h_j\right)}{J_x\left(g_j\right)} \right\|^2 g_j^*\left(x\right) dx = 4\int_{S_1^c} \left\| \frac{J_x\left(g_j^{1/2}\cdot g_j^{1/2}\nabla h_j\right)}{J_x\left(g_j\right)} \right\|^2 g_j^*\left(x\right) dx \\
&\leq 4\int_{S_1^c} J_x\left(g\left\|\nabla h_j\right\|^2\right) dx \text{ (Cauchy-Schwartz inequality and } g_j \leq g) \\
&\leq 4\int_{S_1^c} \left\|\nabla h_j\left(\theta\right)\right\|^2 g\left(\theta\right) d\theta
\end{aligned}$$

and

$$\left\|\nabla h_j\left(\theta\right)\right\|^2 = \frac{1}{\left\|\theta\right\|^2 \log^2\left(j\right)} I_{[1,j]}\left(\left\|\theta\right\|\right) \leq \frac{1}{\left\|\theta\right\|^2 \log^2\left(\left\|\theta\right\| \vee 2\right)} I_{[1,j]}\left(\left\|\theta\right\|\right).$$

Show $B_j \to 0$ by DCT again:

$$\begin{aligned}
& \left\| \frac{J_x\left(\nabla g\right)}{J_x\left(g\right)} - \frac{J_x\left(h_j^2\nabla g\right)}{J_x\left(g_j\right)} \right\|^2 g_j^*\left(x\right) \text{ (Note that } h_j^2\nabla g \text{ is } g_j\frac{\nabla g}{g}) \\
&= \frac{\left\| J_x\left(g_j\frac{J_x(\nabla g)}{J_x(g)} - h_j^2\nabla g\right)\right\|^2}{J_x\left(g_j\right)} \\
&= \frac{\left\| J_x\left[g_j\left(\frac{J_x(\nabla g)}{J_x(g)} - \frac{\nabla g}{g}\right)\right]\right\|^2}{J_x\left(g_j\right)} \\
&\leq J_x\left(g_j\left\|\frac{J_x\left(\nabla g\right)}{J_x\left(g\right)} - \frac{\nabla g}{g}\right\|^2\right) \text{ (Cauchy-Schwartz inequality)} \\
&\leq J_x\left(g\left\|\frac{J_x\left(\nabla g\right)}{J_x\left(g\right)} - \frac{\nabla g}{g}\right\|^2\right).
\end{aligned}$$

Admissibility of $\delta_0 = X$ for p= 1, 2
Let $g\left(\theta\right) = 1$, then

$$\int_{S_1^c} \frac{g\left(\theta\right)}{\left\|\theta\right\|^2 \log^2\left(\left\|\theta\right\|\right)} = 2\pi\int_2^\infty \frac{1}{r\log^2 r} dr < \infty.$$

Homework problem (you pick one part to work on**)**. Let $X_i \sim \text{Poisson}\left(\lambda_i\right)$ be independent, $i = 1, 2, \ldots, p$. Denote $X = \left(X_1, \ldots, X_p\right)$ and $\lambda = \left(\lambda_1, \ldots, \lambda_p\right)$. Under the loss $L\left(\lambda, \delta\right) = \sum_{i=1}^p \left(\delta_i - \lambda_i\right)^2 / \lambda_i$, show that (i) for $p = 1$, X is an admissible estimator of λ using Blyth's method; (2) for $p \geq 2$, X is not an admissible estimator of λ.

Reference: (i) Clevenson and Zidek (1975), *Simultaneous Estimation of the Means of Independent Poisson Laws*, JASA.

(ii) Brown and Hwang (1982), *A unified admissibility proof*, Statistical Decision Theory and Related Topics, III. S. S. Gupta and J. O. Berger (eds.)

Lecture 6. Bayes estimation, minimaxity and Admissibility (cont.).

Superharmonic priors and minimaxity.

Definition. Let $h : \mathbb{R}^p \to \mathbb{R}$ be twice differentiable. We call h *superharmonic* if

$$\nabla^2 h(x) \leq 0 \text{ for all } x \in \mathbb{R}^p, \text{ where } \nabla^2 h(x) = \sum_{j=1}^{p} \frac{\partial^2}{\partial x_i^2} h(x)$$

and call h *harmonic* if $\nabla^2 h^*(x) = 0$ for all $x \in \mathbb{R}^p$. The operator ∇^2 is called Laplace operator (other notations include Δ and $\nabla \cdot \nabla$).

Let $X \sim N(\theta, I)$. Let G be any prior such that $g^*(x) < \infty$ for all x. Then the generalized Bayes estimator δ_G is

$$E(\theta | X = x) = x + \nabla(\log(g^*(x))).$$

By Stein's unbiased estimate of the risk, we have

$$\begin{aligned} R(\theta, \delta_G) - p &= E\left\{2\nabla \cdot \frac{\nabla g^*}{g^*} + \left\|\frac{\nabla g^*}{g^*}\right\|^2\right\} \\ &= E\left\{2\frac{\nabla^2 g^*}{g^*} - \left\|\frac{\nabla g^*}{g^*}\right\|^2\right\}. \end{aligned}$$

Let $h^* = \sqrt{g^*}$. We see

$$R(\theta, \delta_G) - p = 4E\frac{\nabla^2 h^*}{h^*} \text{ (or } 4E\frac{\nabla^2 \sqrt{g^*}}{\sqrt{g^*}}\text{)}.$$

since

$$\begin{aligned} E\left\{2\nabla \cdot \frac{\nabla g^*}{g^*} + \left\|\frac{\nabla g^*}{g^*}\right\|^2\right\} &= 4E\left\{\nabla \cdot \frac{\nabla h^*}{h^*(x)} + \left\|\frac{\nabla h^*}{h^*}\right\|^2\right\} \\ &= 4E\left\{\frac{\nabla^2 h^*}{h^*} - \frac{\|\nabla h^*\|^2}{(h^*)^2} + \left\|\frac{\nabla h^*}{h^*}\right\|^2\right\} \end{aligned}$$

Theorem. If $\nabla^2 \sqrt{g^*} \leq 0$ (or $\nabla^2 g^*$), then δ_G is minimax.

Since

$$R(\theta, \delta_G) - p = E\left\{2\frac{\nabla^2 g^*}{g^*} - \left\|\frac{\nabla g^*}{g^*}\right\|^2\right\}.$$

and

$$\nabla^2 g^*(x) = \int \nabla^2 g(\theta)\, \varphi(x - \theta)\, d\theta,$$

then δ_G is minimax if g is super-harmonic.

Calculation of the Harmonic Bayes procedure.

Let

$$g(\theta) = 1/\|\theta\|^{p-2}$$

This prior density is called the "harmonic prior" because it is harmonic at every point except $\theta = 0$. Note that it is a a valid generalized prior density when $p \geq 3$ in the sense that $g^*(x) < \infty$ for all x . And it is true that $g^*(x)$ is superharmonic.

Write

$$\begin{aligned} g(\theta) &= 1/\|\theta\|^{p-2} \propto \int_0^\infty \omega^{(p-4)/2} e^{-\omega\|\theta\|^2/2} d\omega \\ &= \int_0^1 \frac{1}{v^2} \left(\frac{v}{1-v}\right)^{p/2} e^{-\frac{v\|\theta\|^2}{2(1-v)}} dv \end{aligned}$$

then

$$g^*(x) \propto \int_0^1 v^{(p-4)/2} \exp\left(-\frac{\|x\|^2 v}{2}\right) dv,$$

and thus

$$\frac{\nabla g^*}{g^*} = \frac{\int_0^1 v^{p/2-1} \exp\left(-\frac{\|x\|^2 v}{2}\right) dv}{\int_0^1 v^{p/2-2} \exp\left(-\frac{\|x\|^2 v}{2}\right) dv} x = \frac{1}{\|x\|^2} \frac{\int_0^{\|x\|^2} y^{p/2-1} \exp(-y/2)\, dy}{\int_0^{\|x\|^2} y^{p/2-2} \exp(-y/2)\, dy} x$$

We have

$$\delta_G(x) = x + \nabla(\log(g^*(x))) = \left(1 - \frac{p-2}{\|x\|^2} \cdot \frac{1 - e^{-\|x\|^2/2} T_{p+2}(\|x\|^2)}{1 - e^{-\|x\|^2/2} T_p(\|x\|^2)}\right) x$$

where

$$T_p\left(\|x\|^2\right) = \begin{cases} \Sigma_{k=0}^{p/2-2} (s/2)^k / k! & p \geq 4 \text{ even} \\ \Sigma_{k=0}^{(p-1)/2-1} 2^k k! s^{2k+1} / (2k+1)! & p \geq 3 \text{ odd} \end{cases}.$$