Stat 643
Solution Key to Homework Assignment 9
Apr 26, 2007

Q72

Lets prove it.

Note that if there is no better rules than a decision rule d, we have to keep it in the complete class by definition.

Suppose $\mathcal{C}_1 \bigcap \mathcal{C}_2$ is not essentially complete, i.e. $\exists$ a decision rule $d \in \bar{\mathcal{C}_1} \bigcup \bar{\mathcal{C}_2}$ that no rules in $\mathcal{C}_1 \bigcap \mathcal{C}_2$ is at least as good as d.

Now, if there $\exists$ a decision rule d' that is also in $\bar{\mathcal{C}_1} \bigcup \bar{\mathcal{C}_2}$ and is better than d. We can repeat this until we find the "best" rule from d outside of the intersection, call it d^*, it is clear no other rule is better than d^*. But then d^* should belong to every complete class by definition hence belong to $\mathcal{C}_1 \bigcap \mathcal{C}_2$, we have a contradiction here.

If there is no better rule out there(in $\bar{\mathcal{C}_1} \bigcup \bar{\mathcal{C}_2}$) that is better than d, then d itself has to belong to every complete class and belong to $\mathcal{C}_1 \bigcap \mathcal{C}_2$, we have a contradiction again.

Q73

(a) It is straightforward to show the posterior dsn of $\theta|X$ is $Beta(1 + X, 1 + n - X)$, note the loss function is Weighted Square, the Bayes rule wrt to G which is $U(0, 1)$ is:

$$\delta(X) = \frac{E_{\theta|X} p w(p)}{E_{\theta|X} w(p)}$$

with $w(p) = p^{-1}(1 - p)^{-1}$. Simple integration will get you the result $\frac{X}{n}$.

(b) G is uniform/diffuse so that every open ball inside $(0, 1)$ has positive probability and obviously every neighborhood of any point in $(0, 1)$ intersect with $(0, 1)$, by Thm 152 and part (a), $\frac{X}{n}$ is admissible.

(c)

$$\begin{aligned} R(p, \frac{X}{n}) &= E_p p^{-1}(1-p)^{-1}(p - \frac{X}{n})^2 \\ &= p^{-1}(1-p)^{-1}\mathrm{Var}_p(\frac{X}{n}) \\ &= p^{-1}(1-p)^{-1}\frac{1}{n^2}np(1-p) \\ &= \frac{1}{n} \end{aligned}$$

Cheers, we get an equalizer!
Easy to see $R(G)$, the Bayes risk against G is $\frac{1}{n}$, by (a) and Thm 165, G is the least favorable prior.

(d) ...

Q74

(a) Let us say the prior is $G = \{p_0, p_1\}$ on $\Theta = \{0, 1\}$, then the posterior dsn is $\{\frac{p_0 f_0(X)}{p_0 f_0(X)+p_1 f_1(X)}, \frac{p_1 f_1(X)}{p_0 f_0(X)+p_1 f_1(X)}\}$. The Bayes rule is to minimize the posterior expected loss, i.e.

$$\delta(X) = I[f_0(X)p_0L(0,0) + f_1(X)p_1L(1,0) > f_0(X)p_0L(0,1) + f_1(X)p_1L(1,1)]$$

(b) $L(\theta, a) = I[\theta \neq a]$, then $L(0,0) = L(1,1) = 0$ and $L(0,1) = L(1,0) = 1$, then the result of Bayes rule in (a) become

$$\delta(X) = I[f_1(X)p_1 > f_0(X)p_0]$$

This says we can base our decision on the ratio of $\frac{f_1(X)}{f_0(X)}$ in simple Vs simple hypothesis testing.

Q75

(a)

$$\begin{aligned} R(p, \delta) &= p|p - 3/4| + (1-p)|p - 1/4| \\ &= \begin{cases} 1/4 - p/2 & p < 1/4 \\ -2p^2 + 2p - 1/4 & 1/4 \leq p \leq 3/4 \\ p/2 - 1/4 & p > 3/4 \end{cases} \end{aligned}$$

For $p \geq 0, 1/4 - p/2 \leq 1/4$, at the other end $p \leq 1, p/2 - 1/4 \leq 1/4$. In the middle, $-2p^2 + 2p - 1/4$ is quadratic function, achieving its maximum $1/4$ when $p = 1/2$. So indeed $R(p, \delta) \leq 1/4$.

(b) Let us say the prior G assign probability $\{p_0, p_1, p_2\}$ to points $\{0, 1/2, 1\}$. Then the posterior dsn on these points are $\{\frac{p_0}{p_0+p_1/2}, \frac{p_1/2}{p_0+p_1/2}, 0\}$ when $X = 0$ and $\{0, \frac{p_1/2}{p_2+p_1/2}, \frac{p_2}{p_2+p_1/2}\}$ when $X = 1$. To make the $1/4$ and $3/4$ their posterior median respectively, we can make those non-zero probability equal, i.e. let $p_0 = p_2 = 1/3, p_1 = 2/3$. Then against this prior, δ is Bayes.

(c) $R(p, \delta)$ is no equalizer. Happily we put the "better" Corollary 163' into use. Easy to check $R(G) = 1/4$(The losses and risks are all $1/4$ for the three θ points), then by Corollary 163' and Thm 165, δ is minimax and G is the least favorable prior.

Q76

(a) The least favorable prior is putting $\{0.5, 0.5\}$ mass on $\{5, 6\}$ points. The corresponding Bayes rule ϕ is $I[X > 5.5]$. The risk of this is:

$$R(G) = 0.5R(5, 1) + 0.5R(6, 0) = \Phi(-.5)$$

For any $\theta \in (-\infty, 5] \bigcup [6, \infty)$, $R(\theta, \phi) = \Phi(-|5.5 - \theta|) \leq \Phi(-.5)$, then by Corollary 163', ϕ is indeed minimax, and by Thm 165, we confirm that our guess G is the least favorable prior.

(b) Don't we talked about it in class?

Q77

(a)

$$\begin{aligned} R(\lambda, X) &= E_\lambda \lambda^{-1}(\lambda - X)^2 \\ &= \lambda^{-1}\text{Var}_\lambda(X) \\ &= \lambda^{-1}\lambda \\ &= 1 \end{aligned}$$

Another equalizer.

(b) First, since we have a "uniform" prior, the posterior density is just the likelihood function itself. the Bayes rule wrt to G, Lebesque is:

$$\delta(X) = \frac{E_{\lambda|X}\lambda w(\lambda)}{E_{\lambda|X} w(\lambda)}$$

with $w(\lambda) = \lambda^{-1}$. The numerator is

$$\int_0^\infty \lambda^{-1} X \frac{\lambda^X}{X!} e^{-\lambda} \, \mathrm{d}\lambda = \int_0^\infty \frac{\lambda^{X-1}}{(X-1)!} e^{-\lambda} \, \mathrm{d}\lambda = 1$$

and the denominator is

$$\int_0^\infty \lambda^{-1} \frac{\lambda^X}{X!} e^{-\lambda} \, \mathrm{d}\lambda = \frac{1}{X} \int_0^\infty \frac{\lambda^{X-1}}{(X-1)!} e^{-\lambda} \, \mathrm{d}\lambda = \frac{1}{X}$$

So the result is no surprise, X.

(c) $\lambda | X \propto \lambda^X e^{-\lambda} \lambda^{\alpha-1} e^{-\beta\lambda} = \lambda^{X+\alpha-1} e^{-(\beta+1)\lambda}$, it is a Beta$(X+\alpha, \beta+1)$ distribution, the Bayes estimator is the posterior mean, $\frac{X+\alpha}{\beta+1}$.

(d) δ is equalizer and is Bayes, then by Corollary 163', it is minimax.

Q80

$X_1, \ldots, X_n$ i.i.d. Unif$(0, \theta)$ with $\theta > 0$

(a) Show that $\frac{\mathrm{d}}{\mathrm{d}\theta} \int x f_\theta \, \mathrm{d}x \neq \int x \frac{\mathrm{d}}{\mathrm{d}\theta} f_\theta \, \mathrm{d}x$, where f_θ is the density of $X_{(n)}$

Pf.

$$\begin{aligned} F_\theta(x) = P(X_{(n)} \leq x) = (\frac{x}{\theta})^n \\ \Longrightarrow f_\theta(x) = n\theta^{-n} x^{n-1} I_{(0,\theta)}(x) \end{aligned}$$

So

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\theta} \int x f_\theta \, \mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}\theta} (\frac{n}{\theta^n} \int_0^\theta x^n \, \mathrm{d}x) \\ &= \frac{\mathrm{d}}{\mathrm{d}\theta} \frac{n\theta}{n+1} \\ &= \frac{n}{n+1} \end{aligned}$$

While

$$\int x \frac{\mathrm{d}}{\mathrm{d}\theta} f_\theta \, \mathrm{d}x = -\frac{n^2}{\theta^{n+1}} \int_0^\theta x^n \, \mathrm{d}x = -\frac{n^2}{n+1}$$

(b) Show that the Fisher information inequality does not hold for the UMVUE of θ. The UMVUE of θ is $\frac{(n+1)X_{(n)}}{n}$ with variance $\frac{\theta^2}{n(n+2)}$. On the other hand, the FI is $I(\theta) = n\theta^{-2}$. $[I(\theta)] - 1 = \frac{\theta^2}{n} > \frac{\theta^2}{n(n+2)}$.