(4) Bayes rules, admissibility and complete class theorems

Theorem (1) If δ is a Bayes rule wrt prior π and it is unique up to equivalence (two Bayes rules have the same risk functions) then δ is admissible.

Proof: Suppose $R(\cdot, \delta_0) \le R(\cdot, \delta)$ then

$r(\delta_0) \le r(\delta)$ which proves $r(\delta_0) = r(\delta)$ and

δ_0 is a Bayes rule. This implies $R(\cdot, \delta_0)$

$= R(\cdot, \delta)$ and so δ is admissible.

Theorem (2) If $R(\cdot, \delta)$ is continuous $\forall \delta \in \mathcal{D}$ and $\pi(\Theta) > 0$ $\forall \Theta \in \Theta$ then a Bayes rule is admissible.

Proof: Suppose δ_* is a Bayes rule and

$R(\cdot, \delta) \le R(\cdot, \delta_*)$ with $R(\Theta_0, \delta) < R(\theta_0, \delta_*)$.

Then $r(\delta) = \int_{B_\varepsilon^c(\theta_0)} R(\theta, \delta)\, \pi(d\theta) + \int_{B_\varepsilon(\theta_0)} R(\theta, \delta)\, \pi(d\theta)$

$< \int_{B_\varepsilon^c(\theta_0)} R(\theta, \delta_*)\, \pi(d\theta) + \int_{B_\varepsilon(\theta_0)} R(\theta, \delta_*)\, \pi(d\theta)$

$= r(\delta^*)$ ⊗. Therefore δ_* is admissible.

eg quadratic loss

$R(\theta,d) = (\Psi(\theta) - E_\theta(d))' A(\;) + tr(A\, Var_\theta(d))$

$\therefore$ so if Ψ, $E_\theta(d)$, $Var_\theta(d)$ are continuous in θ $\forall d \in D$ then the Bayes rule $E_{\pi(\cdot|x)}(\Psi(\theta))$ is admissible.

- with models of exponential form these conditions are satisfied.

eg $\Theta = \{\theta_1, \ldots, \theta_k\}$ and every fn on Θ is continuous.

$\Psi = \{\psi_1, \ldots, \psi_k\}$, $L(\theta, \psi) = L^*(\Psi(\theta), \theta)$

- so in very general circumstances a Bayes rule is admissible but when is an admissible δ a Bayes rule wrt some prior π?

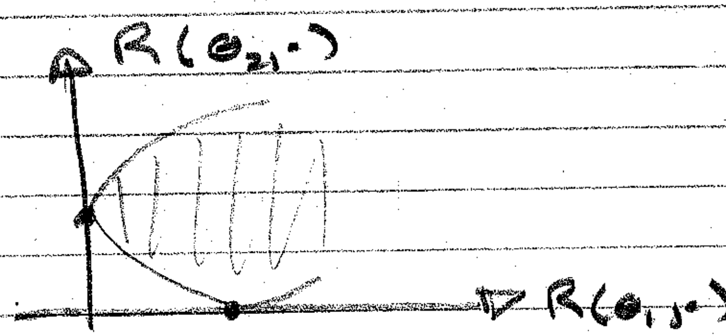
Def The risk set for a dec. problem is given by $\mathcal{R} = \{R(\cdot;\delta) : \delta \in \mathcal{D}\}$.

- so $\mathcal{R}$ a set of functions $\Theta : \to [0,\infty]$

- if $R(\cdot,\delta_1), R(\cdot,\delta_2) \in \mathcal{R}$ and $\alpha \in [0,1]$ then $\alpha R(\cdot,\delta_1) + (1-\alpha) R(\cdot,\delta_2) = R(\cdot, \alpha\delta_1 + (1-\alpha)\delta_2)$ so $\mathcal{R}$ is convex

- when $\Theta = \{\theta_1, \ldots, \theta_k\}$ then identify $R(\cdot,\delta)$ with the k-tuple $(R(\theta_1,\delta), \ldots, R(\theta_k,\delta))$

eg $k=2$



Def For $f : \Theta \to [0,\infty]$ define $Q_f = \{g : g \le f\}$ the lower quadrant of f.

- Q_f is closed (under pointwise convergence)

Def f is a lower boundary point of convex C if $Q_f \cap \bar{C} = \{f\}$ and let $\lambda(C)$ = set of lower bdry pts of C and say C is closed from below if $\lambda(C) \subseteq C$.

- if C is convex then $\bar{C}$ is convex and $\lambda(C) \subseteq \bar{C}$ so C is closed from below if it is closed.

Theorem (3) If $R(\cdot,\delta_0) \in \lambda(\mathcal{R})$ then δ_0 is admissible and if δ_0 is admissible then $R(\cdot,\delta_0) \in \lambda(\bar{\mathcal{R}})$

Proof: Suppose $R(\cdot,\delta) \leq R(\cdot,\delta_0)$ then $R(\cdot,\delta) \in Q_{R(\cdot,\delta_0)} \cap \bar{\mathcal{R}}$ which implies $R(\cdot,\delta) = R(\cdot,\delta_0)$ and so δ_0 is admissible.

If $R(\cdot,\delta_0) \notin \lambda(\mathcal{R})$, then $\exists\ \delta$ s.t. $R(\cdot,\delta) \in Q_{R(\cdot,\delta_0)} \cap \bar{\mathcal{R}}$ and $R(\cdot,\delta) \neq R(\cdot,\delta_0)$ which implies δ_0 is not admissible.

- note - when Θ is finite then

$$r_\pi(\delta) = E_\pi(R(\theta,\delta)) = \sum_{i=1}^{k} R(\theta_i,\delta)\,\pi(\{\theta_i\})$$

is a linear functional on $\mathcal{R}$

$R(\theta_1,\cdot)$ - $\{R(\cdot,\delta) : E_\pi(R(\theta,\delta)) = c\}$ is a hyperplane

- δ is a Bayes rule wrt π iff $R(\cdot,\delta)$

$R(\theta_2,\cdot)$ is in hyperplane given by $c = \inf_{\delta \in \mathcal{D}} r_\pi(\delta)$

Theorem (14) Suppose $\Theta = \{\theta_1, \dots, \theta_k\}$, then

(i) $\lambda(R) \neq \phi$

(ii) if R is closed from below then a Bayes rule exists for any prior with support Θ.

Proof: (i) Let π be a prior with support Θ, and let $s_n \in \mathcal{D}$ be s.t. $r_\pi(s_n) \to \inf_s r_\pi(s)$.

Then $(R(\theta_1, s_n), \dots, R(\theta_k, s_n))$ is a bdd sequence (otherwise $\max_\theta R(\theta, s_n) \to \infty$ and $r_\pi(s) \to \infty$) and so has a convergent subsequence, say $R(\cdot, s_{n_i}) \to f$. Then $f \in \bar{R}$ and so $f \in Q_f \wedge \bar{R}$ and $\int_\Theta f(\theta)\pi(d\theta) = \inf_s r_\pi(s)$. (have to show it is unique)

Suppose $g \in (Q_f \wedge \bar{R}) \setminus \{f\}$. Then $g \in \bar{R}$ implies $\exists\ s_n^* \in \mathcal{D}$ s.t. $R(\cdot, s_n^*) \to g$ and so $\int_\Theta R(\theta, s_n^*)\pi(d\theta) \to \int_\Theta g(\theta)\pi(d\theta) < \inf_s r_\pi(s)$ since $g \in Q_f \setminus \{f\}$ and support π is Θ. This implies $r_\pi(s_n^*) < \inf_s r_\pi(s)$ $\forall$ n large enough $\otimes$

Therefore $Q_f \wedge \bar{R} = \{f\}$; $f \in \lambda(R)$ and so $\lambda(R) \neq \phi$.

(ii) Using $\bar{\pi}$ in (i) we have $\exists$ unique f s.t.

$Q_f \wedge \bar{R} = \{f\} \in \lambda(R) \subseteq R$ and so $\exists\, s_*$ s.t.

$R(\cdot, s_*) = f$ and $r(s_*) = \int_{\Theta} R(\theta, s_*)\, \bar{\pi}(d\theta) = \inf_s r(s)$.

Theorem 5 If $\Theta = \{\theta_1, \ldots, \theta_k\}$ and δ is admissible, then δ is a Bayes rule wrt some π.

Proof: Put $f = R(\cdot, \delta)$ and thus $Q_f \cap R = \{f\}$ (since δ is admissible). Since Q_f° and R are disjoint convex sets, by the Separating Hyperplane Thm, $\exists\ \underset{\sim}{p} \in \mathbb{R}^k \setminus \{\underset{\sim}{0}\}$ s.t. $\underset{\sim}{p}'\underset{\sim}{x} \le \underset{\sim}{p}'\underset{\sim}{y}$ $\forall\ \underset{\sim}{x} \in Q_f^\circ$, $\underset{\sim}{y} \in R$. If $p_i < 0$ then, letting $x_i \to -\infty$ would imply that inevitably $\underset{\sim}{p}'\underset{\sim}{x} > \underset{\sim}{p}'\underset{\sim}{y}$ for a $y \in R$. So we can assume that $p_i \geq 0$ and wlog $\underset{\sim}{p}'\underset{\sim}{1} = 1$. Putting π s.t. $\pi(\{\theta_i\}) = p_i$

$$r_\pi(\delta) = \sum_{i=1}^{k} p_i R(\theta_i, \delta) \le \underset{\sim}{p}'\underset{\sim}{r} \quad \forall\ \underset{\sim}{r} \in R$$

(by cty of $\underset{\sim}{p}'\underset{\sim}{x}$) and so δ is a Bayes rule wrt π.

note - Theorem 2 Bayes rule (general) ($\pi(\theta) > 0\ \forall\theta$) is admissible.

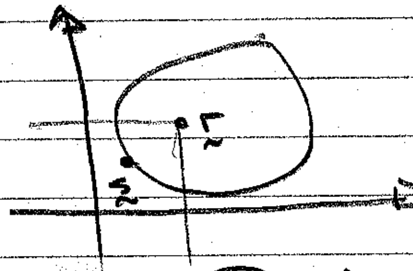
Theorem 5 admissible rule is Bayes (finite Θ)

Theorem (6) If $\Theta = \{\theta_1, \dots, \theta_k\}$ and $\mathcal{R}$ is closed from below then $\mathcal{D}_\lambda = \{\delta : R(\cdot, \delta) \in \lambda(\mathcal{R})\}$ is a minimal complete class.

Proof: Suppose $\delta \notin \mathcal{D}_\lambda$ and let $\underset{\sim}{r} = (R(\theta_1, \delta), \dots, R(\theta_k, \delta))'$

Then $\underset{\sim}{r} \in \mathcal{R} \setminus \lambda(\mathcal{R})$ and $Q_{\underset{\sim}{r}} \cap \bar{\mathcal{R}} \neq \phi$ is convex, closed (and so closed from below) and $\lambda(Q_{\underset{\sim}{r}} \cap \bar{\mathcal{R}}) \neq \phi$

(same proof as Theorem (4) (ii))



Let $\underset{\sim}{s} \in \lambda(Q_{\underset{\sim}{r}} \cap \bar{\mathcal{R}})$ then

$\{\underset{\sim}{s}\} = Q_{\underset{\sim}{s}} \cap (Q_{\underset{\sim}{r}} \cap \bar{\mathcal{R}}) = Q_{\underset{\sim}{s}} \cap Q_{\underset{\sim}{r}} \cap \bar{\mathcal{R}} = Q_{\underset{\sim}{s}} \cap \bar{\mathcal{R}}$

since $\underset{\sim}{s} \in Q_{\underset{\sim}{r}}$. Therefore, $\underset{\sim}{s} \in \lambda(\mathcal{R})$ and $\exists\, \delta'$ st

$\underset{\sim}{s} = (R(\theta_1, \delta'), \dots, R(\theta_k, \delta'))'$ and $\delta' < \delta$.

so $\mathcal{D}_\lambda$ is complete. Since every element of $\mathcal{D}_\lambda$ is admissible (Theorem (3)) there can't be a proper complete subclass.

Theorem 7 (Complete Class Theorem)
If $\Theta = \{\theta_1, \ldots, \theta_k\}$ and R is closed from below, then the class of Bayes rules is complete and A is a minimal complete subclass.

Proof: Each element of A is Bayes (Theorem 5), and $A = \mathcal{D}_r$ since $\mathcal{D}_r$ is a minimal complete subclass (Theorem 6 and Theorem II.2.7).

So A is a complete class contained in the class of Bayes rules this implies that the class of Bayes rules is complete.

- the only question remaining for the $\Theta = \{\theta_1, \ldots, \theta_k\}$ case: when is R closed from below?

eg - suppose $\mathcal{X} = \{x_1, \ldots, x_m\}$ so $(\delta(x_i, \{\cdot\}))$ specifies δ and $(d(x_1), \ldots, d(x_m)) \in \mathcal{A}^m$ specifies $d \in D$

- let ν be a prob. measure on $\mathcal{A}^m$ then
$\delta_\nu(x, A) = \nu(\{d : d(x) \in A\})$ is a dec. fn

- also if δ is dec fn, then for $B \subseteq \mathcal{A}^m$
$\nu_\delta(B) = (\delta(x_1, \cdot) \times \cdots \times \delta(x_m, \cdot))(B)$ is a prob. measure

on $\mathcal{I}^m$ s.t. $\delta_{r_\delta}(x_i, A) = r_\delta(\{d : d(x_i) \in A\})$
$= \delta(x_i, A)$

- so two equivalent ways of expressing randomized rules.

- this can be generalized Wald and Wolfowitz (1951) Annals of Mathematics 53, 3, 581-596.

- if $A \subseteq \mathbb{R}^k$ the <u>convex hull</u> of A is the smallest convex set containing A, denoted $C(A)$

- if $A = \{a_1, \ldots, a_l\}$ then

$$C(A) = \{\sum_{i=1}^{l} \lambda_i a_i : \lambda_i \geq 0, \lambda_1 + \cdots + \lambda_l = 1\}$$

- if A is compact, then $C(A)$ is compact

<u>Theorem</u> (8) If $\mathcal{X}$ and Θ are finite then $\mathcal{R}$ is closed from below

Proof: Let $\mathcal{R}_D = \{R(\cdot, d) : d \in D\}$. Since $\mathcal{R}_D$ is finite it is compact. For any $\lambda_i \geq 0$, $\lambda_1 + \cdots + \lambda_m = 1$, then $\sum_{i=1}^{m} \lambda_i R(\cdot, I_{\{d(a_i)\}}) = R(\cdot, \sum_{i=1}^{m} \lambda_i I_{\{d(a_i)\}})$ and putting $r(\{d(a_i)\}) = \sum_{d(a_j) = d(a_i)} \lambda_j$ then $\delta_r = \sum_{i=1}^{m} \lambda_i I_{\{d(a_i)\}}$

Therefore, $C(\mathcal{R}_D) = \mathcal{R}$ (by equivalence) and so $\mathcal{R}$ is compact and thus closed and closed from below.