STAT 530: Decision Theory Continued

Feb 3, 2009

Start with loss function, $L(\theta, a)$, recall definitions:

- $R(\theta, \delta)$ is **risk function** of rule (estimator) $\delta()$:
  - expected value of $L(\theta, \delta(Y))$, w.r.t. sampling $(Y|\theta)$
- $r(\pi, \delta)$ is **Bayes risk** of of $\delta()$:
  - average risk w.r.t. $\theta \sim \pi$
  - think of π as *nature's prior*.
- $\rho_\pi(a; y)$ is **posterior risk** of action (estimate) a:
  - average of $L(\theta, a)$ w.r.t. posterior on $(\theta | Y = y)$
  - here think of π as *investigator's prior*

Big Fact #1: Choice of loss function influences choice of which Bayesian estimator

General Bayesian principle: 'the' Bayesian estimator is:

$$\delta_B(y) \equiv \text{argmin}_a E\{L(\theta, a)|Y = y\}$$

Examples:

- $L(\theta, a) = (a - \theta)^2$ implies $\delta_B(y) = E(\theta|Y = y)$
- $L(\theta, a) = |a - \theta|$ implies $\delta_B(y) = \text{Median}(\theta|Y = y)$

Big Fact #2: Sense in which Bayes estimators are best possible

The Bayes risk based on nature's prior π_N is minimized - amongst all possible estimation procedures $\delta()$ - by the Bayes estimator which uses π_N as the investigator's prior.

Proof by changing the order of integration.

Big Fact # 3: There is no point in considering non-Bayesian estimators

Recall: $\delta()$ **inadmissible** if there exists $\delta^*()$ such that

$$R(\theta, \delta) \leq R(\theta, \delta^*)$$

for all θ, with strict inequality for at least one θ.

And an estimator is **admissible** if it is not inadmissible.

It is *almost* the case that:

$\delta()$ is admissible $\leftrightarrow$ $\delta()$ is a Bayes estimator w.r.t. some prior.

Say $\delta()$ is Bayes estimator w.r.t. π, show $\delta()$ admissible

Look at simple case $\theta \in \{\theta_1, \theta_2, \ldots\}$

Say $\delta()$ is admissible, show $\delta()$ is Bayes w.r.t. some prior π

Proof by picture in very simple case $\theta \in \{\theta_1, \theta_2\}$:

Visualize any $\delta()$ as point $(R_1, R_2) = R(\theta_1, \delta), R(\theta_2, \delta)$ in 'risk set'.

- risk set necessarily convex (think of randomized rules)
- 'lower boundary' of the set corresponds to admissible rules
- at any boundary point, convex set 'supported' by a line
- express line as $\pi_1 R_1 + \pi_2 R_2 = c$
- Voila! rule minimizes Bayes risk w.r.t. this π

More precisely, admissible estimators are Bayes estimators plus some 'limits' of Bayes estimators

E.g. $Y \sim N(\theta, 1)$, $L(\theta, a) = (a - \theta)^2$.

$\delta(Y) = Y$ is

- admissible
- not the posterior mean arising from any proper prior
- the $\tau^2 \to \infty$ limit of $E(\theta | Y = y)$ via the prior $\theta \sim N(0, \tau^2)$.

But things get crazy...

$Y \sim N_p(\theta, I_p)$, $L(\theta, a) = \|a - \theta\|^2$

Consider $\delta(Y) = Y$

- $p = 1$: admissible
- $p = 2$: admissible
- $p >= 3$: **inadmissible**
  - James-Stein (1960)
  - shocking!

James-Stein. . .

Showed that $\delta_{JS}(Y) = (1 - (p-2)/\|Y\|^2)Y$ beats $\delta(Y) = Y$

Later, someone showed $\delta_{JS}()$ is beaten by
$\delta_{JS}^{+}(Y) = \max\{0, 1 - (p-2)/\|Y\|^2)\}Y$.
Later someone showed δ_{JS}^{+} inadmissible,
without actually finding an estimator that beats it!

Isn't math great?