# Intro to Econometrics: Recitation 4

Gustavo Pereira

September 30, 2019

# Roadmap for today

- Review:
  1. Statistical problem
  2. Bayes rules, expected posterior loss
- PS3

# Review: Statistical Problem

- Components of a decision problem:
  - Statistical model: $\{P_\theta\}_{\theta \in \Theta}$
  - Action space $\mathcal{A}$
  - Loss function $\mathcal{L} : \mathcal{A} \times \Theta \to \mathbf{R}$
  - Decision rules: $d : \mathfrak{X} \to \mathcal{A}$
- Risk function: expected loss from decision $d$ when parameter is $\theta$:

$$R(d(\cdot), \theta) = \int_{\mathfrak{X}} \mathcal{L}(d(x), \theta) f_\theta(x) dx$$

# Review: Admissibility

- Decision rule $d_1$ is *dominated* by $d_2$ iff, for all $\theta \in \Theta$,

$$d_1 \precsim_\theta d_2$$

and $d_1 \prec_{\theta_0} d_2$ for at least one $\theta_0$

  - What does it mean for a rule to be *not dominated* by another rule?
- A rule $d$ that is not dominated by any other rule is called *admissible*
  - Expand the definition of admissible
- It is generally hard to find admissible rules.

# Review: priors, posteriors, etc...

- Suppose model is $\{f_\theta\}_{\theta \in \Theta}$, i.e., data has a density for all possible parameters
- Suppose also $\Theta \subseteq \mathbf{R}^k$, and pdf $\pi(\theta)$ summarizes some prior belief about $\theta$
  - With this, we're interpreting the parameter $\theta$ as a *random variable*
  - Before the prior was introduced, $\theta$ was merely an index
- With this structure, we can define the induced joint density of data and parameters,

$$f(x, \theta; \pi) = f_\theta(x)\pi(\theta)$$

  - Does this integrate to one?

# Review: priors, posteriors, etc...

- Given induced joint density,

$$f(x|\theta;\pi) = \frac{f_\theta(x)\pi(\theta)}{\pi(\theta)} = f_\theta(x)$$

- What about the marginal of data?
  - Recover it by integrating $\theta$ out:

$$f(x;\pi) = \int_{\theta\in\Theta} f_\theta(x)\pi(\theta)d\theta$$

- Conditional density of parameter given data?

$$f(\theta|x;\pi) = \frac{f(x,\theta;\pi)}{f(x;\pi)} = \frac{f_\theta(x)\pi(\theta)}{\int_{\theta\in\Theta} f_\theta(x)\pi(\theta)d\theta}$$

  - This is called *posterior density* in Bayesian jargon

# Review: Bayes rules

- Let's go back to the statistical decision problem
- Let $d(\cdot)$ be a decision rule, and $\pi$ a prior density over $\Theta$
- Bayes risk of $d(\cdot)$ given $\pi$ is

$$\begin{aligned} r(d(\cdot),\pi) &= \int_{\Theta} R(d(\cdot),\theta)\pi(\theta)d\theta \\ &= \int_{\Theta}\int_{\mathfrak{X}} \mathcal{L}(d(x),\theta)f(x,\theta;\pi)dx\,d\theta \end{aligned}$$

- A *Bayes decision rule* $d^*$ is one that minimizes Bayes risk given a prior $\pi$.

$$d^*_\pi(\cdot) = \arg\min_{d(\cdot)} r(d(\cdot),\pi)$$

- Important feature: *under mild assumptions, Bayes rules are admissible*

# Review: finding Bayes rules

- Rewrite the Bayes risk using Fubini's theorem

$$
\begin{aligned}
r(d(\cdot),\pi) &= \int_{\mathfrak{X}} \left[ \int_{\Theta} \mathcal{L}(d(x),\theta) f(\theta|x;\pi) d\theta \right] f(x;\pi) dx \\
&= \int_{\mathfrak{X}} \psi(d(x),x) f(x;\pi) dx
\end{aligned}
$$

where

$$
\psi(a,x) = \int_{\Theta} \mathcal{L}(a,\theta) f(\theta|x;\pi) d\theta
$$

- Let $d^*(x) = \arg\min_{a\in\mathcal{A}} \psi(a,x)$
  - Immediate consequence: for any decision rule $d(\cdot)$,

$$
\psi(d^*(x),x) \leq \psi(d(x),x)
$$

  - Important: optimization in space $\mathcal{A}$ is easier than in the space of all $d : \mathfrak{X} \to \mathcal{A}$!