

Model Averaging Review from Math/Stats View

Michael Messner

USEPA

Office of Ground Water and Drinking Water

Disclaimers

- This will be Mike's view (not that of EPA)
- My Job Title = Mathematical Statistician, so the “Math/Stats View” seems fitting, **but**
 - My perspective → statistical decision analysis
 - My experience → data are usually dirtier (more complex) than these
 - more explanatory variables / confounders
 - the usual errors (precision, bias) are also accompanied by blunders, so need to look for and deal with outliers
 - My blunder → not considering a sufficiently full set of models in analyzing dose-response data.
 - My expectation → model averaging with sufficient set of competing models will suggest a good plan for obtaining new data (value-of-information analysis)

While Waiting, found...

- BMA has a home page:

<http://www.research.att.com/~volinsky/bma.html>

- An R package:

<http://cran.r-project.org/src/contrib/Descriptions/BMA.html>

- **Nice, clear paper:** Variable selection and Bayesian model averaging in case-control studies, by Valerie Viallefont; Adrian E. Raftery, and Sylvia Richardson, in *Statistics in Medicine* (20:3215-3230), 2001.
- **And a tutorial:** <http://www.stat.colostate.edu/~jah/papers/statsci.pdf>
(Turns out I already had this one filed away.)

But Didn't Find

- Not much on Frequentist treatment. (Is bootstrap Frequentist and does it work? What about neural nets?)
 - Not sure I'd care, even if there had been a large body of stuff.
- Not much of anything on constructing informative priors.
 - Yes, some of the software allows prior weights to be entered, but...how does one select the prior weights?

Idea

- Consider the first dose-response experiment (doses 0, 21, 60 and # Subjects ~ 50, 50, 50) and imagine the results aren't in yet.
 - What is the universe (or reasonable shopping list) of models?
 - Include models whose complexity couldn't possibly be revealed by the experiment?
 - Include models that allow $\Pr\{\text{response} \mid \text{dose} = 0\} = 0$? (Reject if any responses at dose 0, but keep if no responses.)
 - Include $\Pr\{\text{response} \mid \text{dose}\} = \text{constant}$?
 - Assign probability mass to each type of model?
 - Are some models more plausible/likely than others?
 - Discount for complexity?
 - For each model, do we need to consider parameters?
 - Any constraints (monotonicity)?
 - Informative priors?

Shoot-Out

- Simulate data for fake contaminants using at least 3 different model forms.
 - Decent amount of data, but not enough to be certain about the model.
 - Six dose levels, 40 subjects each?
- Tell analysts there are 3 different model forms, but give them a bigger set (10?).
- Let the analysts have at it with different methods (classic, simple Bayes, model averaging, other)
- As for estimates of risk at 3 different dose levels for each data set
 - Near zero
 - Mid-range
 - High, but outside range of data
- See which methods do best in terms of predicting which of the fake contaminants is most risky at each dose level... or have some other measure of performance so we can have a winner.