September 26, 2009

MUSINGS
(In no particular order)

- The fundamental statistical entity is the (joint, conditional) distribution or likelihood. Summaries are computed from it; it is the "evidence."
- Inferences should attend to all, relevant uncertainties, both "sampling" (for a specific model) and "non-sampling" (model uncertainty).
  - There is a fuzzy boundary between sampling and non-sampling uncertainties
- There is no free lunch in statistics, but there may be a reduced-price lunch
  - Building in a little bias can dramatically reduce variance, thus reducing MSE
- The best of a breed may still be a dog!
- Most of statistics, indeed science, relies on answering, "as compared to what?" or relying on, "under the following assumptions."
- Consider what properties you want of a procedure. For example, do you want location invariance (slide things; get the same answer) or scale invariance (doesn't matter if we measure in pounds or kilograms) or not. Make sure that those properties are satisfied and then try to optimize or at least "do well."
- Don't let statistical optimality dictate what your estimate. "Optimal" statistical weights may not address your goal. Least squares may fall on its face, ...
- Don't let natural parameters dictate what you estimate. Example, the odds ratio is the "natural parameter" in comparing two binomials, but it may not be the appropriate science or policy metric. If you want a risk difference (and have a design that allows you to get it!), figure out a way to estimate it.
  - Summarize in the right framework; absolute risk vs relative risk vs odds ratio. We can get SEs (better still, CIs) without delta theorems so don't let complexity stop you from doing the right thing.
- The Sampling Plan (probability space, sigma-algebra, probability measure) is fundamental: How did I get to see (and not see!) my data? Addressing these issues underly all of "causal analysis."
- Strike an effective trade-off between robustness and efficiency; robustness of efficiency; robustness of validity.
  - The, generally small, gain in efficiency produced by departing from working independence may not be worth the risks; but protection from violations of MCAR may make it worth it.
- Don't let the linear model lull you into incorrect statistical constructs or computing for non-linear models.
  - Population level is average of individual level predictions, not the prediction of individual-level averages. Prediction of the average $\neq$ average of the predictions and the latter is correct.
  - Marginal, conditional and transitional models are not always the same shape; Gaussian/Gaussian/Gaussian is deceptively simple
  - In the linear model, residual variance is estimated; in other basic models (e.g., binomial, Poisson), the residual variance is linked to the mean. Need to use "overdispersed" models to get "honest" uncertainties.

- Beware of multiple (and simple) regression coefficients. Predictions are safe (and evaluable); individual slopes are dangerous. X_1 and X_2 with $X_2 = X_1^2$ example.
  - Slopes depend on context. For example, they depend on measurement scale and on centering of higher order terms; true even for the intercept. Predictions are parameterization-free.
- Always indicate conditioning. For example, write $[Y \mid X] = ...$, not $Y =$.
- Look at your data, at residuals, ...
- The original variables are not sacred.
  - (D, R, T) example and examples with $V(Y|X) \propto X^p$
  - Transforms of Y and X
- Know the units of all variables and the coding. Otherwise, you can't interpret slopes.
- Consider reporting regression results as "per an interquartile range change in the regressor." These are invariant under the units for predictors and avoids deceptively large or small slopes.
- Come to the data with identified questions and a strategy. For true explorations, label them as such and, possibly, adjust for multiplicity. See the "Green Book" (Mosteller & Tukey)
- Measurement error issues.
- Don't use stepwise regression.
- Focus on estimates and confidence intervals, not on testing.
- Consider smooth departures from H_0. Do you really want to leap from $\hat{\beta} = 0$ to $\hat{\beta} = 7.0$ when the P-value goes from 0.11 to 0.09?
- Almost all H_0 are known to be false a priori. It's magnitude and uncertainty of key estimates that matter.
- Procedure evaluation
  - Performance under assumptions
  - Performance under violations of assumptions
  - Performance compared to another procedure. Needs to be a fair game. For example, only compare power for tests with the same size.
  - Resistance and robustness
- Goodness of fit
  - Comparing models with AIC, etc.
  - Best of breed may still be a dog
  - Look at residuals; evaluate predictions
  - Mean squared error, diagnostic accuracy
  - Cross-validation
  - Diagnostics and Robust estimation approaches
- Unbiasedness and its discontents, let's go for MSE
- Non-linearity
- Augmenting the model: powers, splines, loess, wavelets
- The perils of threshold hunting, maximally selected chi-square

- Confounding, mediation, interaction, threshold effects, etc.
- How to design. Relative to an analysis. Be Bayes. Use analysis programs to design
- Know the assumptions behind methods and which are the most fragile