

The Use of Misspecified or Simplified Models

Shaohua (Roy) Wu

(Abstract)

When we have some data, a mathematical model, and some initial parameter guesses, it is very interesting and important to determine which parameters we should estimate from the available data. If our data aren't very informative, should we simplify the model to remove some parameters or should we fix some parameters at their initial guesses and estimate only a subset of the model parameters? Will estimating a reduced number of parameters result in better predictions? Once we have estimated some parameters based on the misspecified or simplified model, the model predictions and parameter estimates will be biased. We need to determine how to do hypothesis tests, how to construct reliable confidence intervals and how to estimate the goodness of the approximation. We know these models are misspecified or simplified, and may not pass a statistical test for validity, but they may be useful nonetheless.

Our previous research proved that misspecified or simplified models are better than the "true" model (in sense of mean square prediction error) when the data are less-than-ideal. In the current research, we are trying to solve the problems described above. We will do theoretical analysis, Monte Carlo simulations, and some practical case studies. This is a multi-disciplinary research project involving statistical analysis, estimability/sensitivity analysis, experimental design, and model/subset selection.