The University of Chicago
Department of Statistics

Seminar Series

HANNES LEEB

Department of Statistics
Yale University

Conditional Predictive Inference Post Model Selection

MONDAY, February 2, 2009 at 4:00 PM
133 Eckhart Hall, 5734 S. University Avenue

Refreshments following the seminar in Eckhart 110.

ABSTRACT

We give a finite-sample analysis of predictive inference procedures after model selection in regression with random design. The analysis is focused on a statistically challenging scenario where the number of potentially important explanatory variables can be infinite, where no regularity conditions are imposed on unknown parameters, where the number of explanatory variables in a 'good' model can be of the same order as sample size, and where the number of candidate models can be of larger order than sample size. The performance of inference procedures is evaluated conditional on the training sample. Under weak conditions on only the number of candidate models and on their complexity, and uniformly over all data-generating processes under consideration, we show that a certain prediction interval is approximately valid and short with high probability in finite samples, in the sense that its actual coverage probability is close to the nominal one, and in the sense that its length is close to the length of an infeasible interval that is constructed by actually knowing the 'best' candidate model. Similar results are shown to hold for predictive inference procedures other than prediction intervals like, e.g., tests of whether a future response will lie above or below a given threshold.

For further information and about building access for persons with disabilities, please contact Kelly Macias at 773.834.5169 or send email (kmacias@galton.uchicago.edu). If you wish to subscribe to our email list, please visit the following web site: https://lists.uchicago.edu/web/info/statseminars.