

The University of Chicago

Department of Statistics

Seminar Series

SHAM KAKADE

Toyota Technical Institute

“Calibration via Regression”

MONDAY, April 17, 2006 at 4:00 PM
133 Eckhart Hall, 5734 S. University Avenue

Refreshments following the seminar in Eckhart 110.

ABSTRACT

In a sequential prediction setting where the learner makes a prediction of (say a binary) outcome at each timestep, the concept of calibration entails having the empirical (conditional) frequencies of the outcome match the claimed predicted probabilities — in essence, a calibrated algorithm is unbiased in a certain sense. The concept of calibration has had significant impact in the areas of game theory and economics, and more, recently, the notion of calibration is being used in machine learning to improve prediction algorithms.

As an example, the most minimal calibration condition would be to have our overall predictions be unbiased, i.e. to have the overall empirical frequency of the event match our average predicted frequency. More generally, if we have some binary covariates $\{x_i\}$ at each time step, then on the subsequence in which the i -th covariate is 1 (when $x_i = 1$) we might like our predictions to be unbiased.

In the worst case (where no probabilistic assumptions are made about the sequence), it has been shown that calibration is possible using rather sophisticated machinery, namely Blackwell’s approachability theorem. Here, we show how merely an online regression algorithm can be used for calibration, in the worst case. In essence, we show how the ‘normal equations’ (used in least squares regression) can be satisfied, even when we make no probabilistic assumptions.

This is joint work with Dean Foster.