

Day 1 (Tuesday, May 29) Instructor: Teddy Seidenfeld. Theme: *Some information-based techniques in data-mining*. Measures of statistical information play a central role in statistical inference. We examine several information-based concepts in statistical inference and their applications in data-mining. Topics to be covered include data-reduction, prediction with sparse data, and techniques with undirected graphical models. We will examine information-constraints operating in numerical algorithms, such as in the EM algorithm.

Day 2 (Wednesday, May 30) Instructor: Richard Scheines. Theme: *Graphical Causal Models and TETRAD*. Tetrad is an integrated suite of free software for specifying, estimating, and searching for graphical causal models. It contains over 20 algorithms for searching for a variety of model classes, e.g., path analytic models, Bayes nets, factor analytic and structural equation models, time series models general latent variable models of conditional independence structures, and Markov blankets. In this workshop, we will briefly introduce graphical causal models, show how to build, simulate data from, and estimate such models in Tetrad, explain model search, teach how to use a variety of the search algorithms in Tetrad on real and simulated data, and go through a few case studies in which Tetrad is applied to real data. We will devote particular attention to comparing provably reliable causal search algorithms to multiple regression, which is provably unreliable even on population probabilities.

Day 3 (Thursday, May 31) Instructor: Teddy Seidenfeld. Theme: *Some information-based techniques in statistical decision-making*. Statistical measures of information also play an important role when gauging the value of new information for a decision maker. What is the value of new information? We will examine the value of information both for asymptotic and short-run decision problems, involving cases with both single and multiple decision makers. Topics to be covered include the role of information in experimental design, and in multi-agent problems, such as in games, where there are issues of strategic interaction.

Day 4 (Friday, June 1) Instructor: Kevin Zollman. Theme: *Strategic emergence of information transfer*. When there is information in a system there are a number of important issues about how that information comes to be transmitted by economic actors. In the case of simple two-agent interactions, information might be strategically released, strategically modified, or withheld altogether. There are a number of (game theoretic) analyses that outline the conditions under which one should expect information to be transmitted. As one moves from a two-agent situation to a multi-agent situation, things become even more complex. Even when everyone desires that information be transmitted, forming optimal institutions for the transmission of that information can be difficult. This day will focus on both the issue of strategic information transmission and the strategic formation of institutions for the transmission of information.

Day 5 (half-day: Saturday morning, June 2) Instructor: Teddy Seidenfeld. Theme: *Information and some non-parametric Graphical and Bayesian methods*. We illustrate inference methods for large data sets using a combination of ideas developed in previous lectures, including variational and hierarchical models.