MODEL SELECTION IN MIXTURE MODELING

Emilie Shireman

Dr. Douglas Steinley, Dissertation Supervisor

ABSTRACT

In the psychological sciences, mixture modeling (also referred to as latent class or latent profile analysis) is very commonly used to find sub-populations within a sample. However, the process by which researchers select a model (i.e., how many sub-populations and how many covariance parameters) is not standardized. Furthermore, many techniques that researchers use to select a model are ad hoc and have varied statistical theoretical support. This dissertation systematically examines three commonly used but not formally tested model selection heuristics for mixture modeling: (1) using several fit indices to collectively select a model, (2) using the difference in fit to differentiate “weak” versus “strong” evidence of one solution over another, and (3) examining the difficulty in convergence to indicate that a model is over-specified.