Hiding in plain sight: latent grouping factors in linear models

Dr. Chris Franck
Department of Statistics
Virginia Tech, USA

Standard linear modeling approaches make potentially simplistic assumptions regarding the structure of categorical effects that may obfuscate more complex relationships governing data. For example, recent work focused on the two-way unreplicated layout has shown that hidden groupings among the levels of one categorical predictor frequently interact with the ungrouped factor. In this talk, I extend the notion of a ``latent grouping factor'' to linear models in general. The proposed work allows researchers to determine whether an apparent grouping of the levels of a categorical predictor reveals a plausible hidden structure given the observed data. Specifically, I offer a Bayesian model selection-based approach to reveal latent group-based heteroscedasticity, regression effects, and/or interactions. Failure to account for such structures can produce misleading conclusions. Since the presence of latent group structures is frequently unknown *a* priori to the researcher, I use fractional Bayes factor methods and mixture *g*-priors to overcome lack of prior information.