Title: Regression Spline Mixed Models for Testing Meaningful Landmarks in Time Series Data

Author: Karen Nielsen

Date, Time: December 15th, 1pm

Location: 340 West Hall

Abstract:

Time series data, particularly biological, are becoming increasingly common as we explore the relationship between biology and behavior. Event-Related Potentials (ERPs), which are brain responses to time-locked stimuli measured using Electroencephalography (EEG), are one example of such data. The goal in ERP research is to make inferences about neural circuits and mechanisms used when responding to stimuli. We can draw from the substantive expectations researchers have about the anticipated shape of individuals' waveforms, particularly in local regions of interest, and translate these verbalized assumptions into mathematical basis sets. These assumptions allow us to derive properties of the representative waveform and implement them as parameters of a statistical model. We then test hypotheses on landmark parameters of the basis sets via multilevel modeling, which allows us to account for temporal patterns, patterns across channels, individual differences, and differences across experimental conditions.

The process of translating verbalized assumptions into basis sets is nontrivial. We begin with common basis sets. For example, polynomials are common in some areas of psychology and Fourier transforms are often used in engineering and physics. Basis sets that are more specific to the research hypotheses can also be parameterized and tested, and these are likely to be most useful for testing specific functional forms and for testing competing models. Thus, the properties of basis set transformations can be leveraged in powerful ways, especially when they can give model parameters more natural, domain-relevant interpretations.

Basis sets can be fit to the data in many ways. We focus on how Regression Spline Mixed Models (RSMM) can combine problem-specific basis sets with a hierarchical random effects framework. These methods can address multiple levels and sources of variability, and therefore can be valuable to researchers working with time series data. Having the ability to work at multiple levels with theoretically meaningful parameters is useful for recognizing outliers, learning about variance and statistical significance, testing specific predictions about waveforms, and inspiring further analyses or future studies.