

UCLA

BIOSTATISTICS SEMINAR

SPRING 2017

Some model-building tools for Gaussian processes, using an approximate form of the restricted likelihood

Jim Hodges, PhD

Professor

Division of Biostatistics, School of Public Health

University of Minnesota, Minneapolis

Wednesday, May 17, 2017

3:30pm - 4:30pm, CHS 33-105

Refreshments served at 3:00 PM in room 51-254 CHS

ABSTRACT:

Gaussian processes (GPs) are widely used in statistical modeling, often as the distribution of a random effect in a linear mixed model. The GP's unknowns are commonly estimated by maximizing the restricted likelihood or doing a Bayesian analysis, which are closely related. However, it is unclear how a GP's variance and range parameters and the error variance are fit to features in the data, largely because the restricted likelihood does not have a closed form.

To understand how these parameters are fit to the data, we need a simple, interpretable form of the restricted likelihood. We use the spectral approximation to obtain an approximate restricted likelihood with a simple form, which is identical to the likelihood arising from a gamma-errors generalized linear model (GLM) with the identity link. We use this GLM to make conjectures about how GP parameters are fit to data and then investigate those conjectures by introducing features into simulated data, e.g., outliers and mean shifts, and observing how those features affect parameter estimates.

The last part of the talk describes briefly how to use this representation to derive diagnostic tools intended to identify potential covariates and to examine whether and how the data support their inclusion and what effect their inclusion will have on the fit of the GP part of the mixed linear model.