

# Semiparametric Bayesian Classification with Longitudinal Markers

Rolando de la Cruz-Mesía

Departamento de Estadística, Pontificia Universidad Católica de Chile,  
Casilla 306, Correo 22, Santiago, CHILE.

*email:* rolando@mat.puc.cl

## Abstract

We discuss Bayesian statistical methods for the classification of observations into two or more groups based on longitudinal observations. Inference is based on a semi-parametric hierarchical model for each group. Specifically, we propose a species sampling model prior for the distribution of the random effects. The unknown random effects distributions are allowed to vary across groups, but are modeled dependent by using a design vector to select different features of a single underlying random probability measure. The resulting model can be interpreted as an extension of the dependent Dirichlet process (DDP) model.

Relevant posterior distributions are summarized using Markov chain Monte Carlo methods. The method is illustrated with data from a study involving 173 pregnant women. The main objective in this study is to predict normal versus abnormal pregnancy outcomes from data available at early stages of pregnancy.

**KEYWORDS:** Dependent nonparametric model; Discriminant analysis; Longitudinal data; MCMC sampling; Nonlinear hierarchical models; Nonparametric modeling; Random Effects Models; Species Sampling Models.