2011 Seminars

06/29/2011

Misspecifying the shape of a random effects distribution: Why getting it wrong may not matter.

Charles E. McCulloch, PhD

Professor and Head, Division of Biostatistics, Dept. of Epidemiology and Biostatistics, University of California at San Francisco

Joint work with John M. Neuhaus

Abstract: Statistical models that include random effects are commonly used to analyze longitudinal and correlated data, often with strong and parametric assumptions about the random effects distribution. There is marked disagreement in the literature as to whether such parametric assumptions are important or innocuous. In the context of generalized linear mixed models used to analyze clustered or longitudinal data, we examine the impact of random effects distribution misspecification on a variety of inferences, including prediction, inference about covariate effects, prediction of random effects, and estimation of random effects variances. We describe examples, theoretical calculations, and simulations to elucidate situations in which the specification is and is not important. A key conclusion is the large degree of robustness of maximum likelihood for a wide variety of commonly encountered situations.

05/23/2011

New Findings from Terrorism Data: Dirichlet Process Random Effects Models for Latent Groups

George Casella, PhD

University of Florida

Abstract: Data obtained describing terrorist events are particularly difficult to analyze, due to the many problems associated with the both the data collection process, the inherent variability in the data itself, and the usually poor level of measurement coming from observing political actors that seek not to provide reliable data on their activities. Thus, there is a need for sophisticated modeling to obtain reasonable inferences from these data. Here we develop a logistic random effects specification using a Dirichlet process to model the random effects. We first look at how such a model can best be implemented, and then we use the model to analyze terrorism data. We see that the richer Dirichlet process random effects model, as compared to a normal random effects model, is able to remove more of the underlying variability from the data, uncovering latent information that would not otherwise have been revealed.