Monday 22 July 2024

13:30-15:00 Invited Session 2 (Main Room)

Recent advances in survival analysis with complex data structures (Chair: Giorgos Bakoyannis)

Bayesian semiparametric modeling of spatially-referenced multistate current status data

Dipankar Bandyopadhyay (Virginia Commonwealth University, USA)

Assessment of multistate disease progression is commonplace in biomedical research, such as, in periodontal disease (PD). However, the presence of current status (a severe form of interval-censoring) endpoints, where a single snapshot of study subjects transitioning through a sequence of well-defined disease states at random inspection times, complicates the inferential framework. In addition, these endpoints can be clustered (tooth-level event status within subjects), and spatially associated, where a group of proximally located teeth may experience similar PD status compared to distally-located teeth. Failure to adjust for the aforementioned complexities may lead to biased and imprecise inference. Motivated by a clinical study recording multistate event time progression of PD, we propose a Bayesian semiparametric accelerated failure time model with a Wishart proposal for accommodating (spatial) random effects, and flexible errors that follows a Dirichlet process mixture of Gaussians. For elegant clinical interpretation, the systematic component of the event times is modeled using a monotone single index model, whose (unknown) link function is estimated via a novel integrated basis expansion, with basis coefficients enjoying constrained Gaussian process priors. In addition to promising parameter identifiability, we present scalable computing via a combination of elliptical slice sampling, fast circulant embedding techniques, and smoothing of hard constraints, leading to straightforward computation of the parameter estimates, state occupation, and transition probabilities from posterior estimates. Using synthetic data, we study the finite sample properties of our Bayesian estimates, and its performance under model misspecification. We also illustrate our methodology via application to a real dataset recording PD status of Type-2 diabetic African-Americans living in coastal South Carolina.