Survival Model Selection with Missing Data and Correlated Covariates

Sydeaka Watson, PhD
The University of Chicago, Biostatistics Laboratory
sydeaka@uchicago.edu

In this talk, I will describe an algorithm used to develop a survival prediction equation for pulmonary arterial hypertension patients awaiting lung transplantation. The transplant registry dataset featured censored survival times, missing covariate data, and a large number of highly correlated candidate predictor variables. We used a novel combination of existing methods to select a subset of the candidate variables, which could be used to predict survival probabilities for each patient. In our approach, we repeatedly applied penalized weighted least squares regression in bootstrap resamples of multiply imputed data and selected a parsimonious model that satisfied internal validation criteria of clinical interest. Simulation studies under various degrees of predictor variable missingness, survival time censoring, effect size, and proportion of variables unrelated to survival have shown that this method accurately recovers the true list of Cox regression predictor variables.

Biography

Sydeaka Watson was raised in New Orleans, Louisiana. She earned mathematics degrees from University of New Orleans (B.S., 2002) and Michigan State University (M.S., 2003). Sydeaka received her Ph.D. in statistics from Baylor University in 2011. In her dissertation, she developed a Bayesian Poisson regression model for interval censored count data (immune response counts) collected in HIV vaccine studies conducted at Los Alamos National Laboratory in New Mexico. She is currently a Research Associate (Assistant Professor) in the Department of Health Studies at The University of Chicago. She serves as a biostatistical collaborator and/or joint principal investigator on a number of biomedical research studies. As a member of the Internal Scientific Advisory Panel (ISAP) at the UChicago, she evaluates study designs and statistical methods in clinical and translational research protocols for which internal funding support is requested. Sydeaka currently chairs the American Statistical Association (ASA) Committee on Minorities in Statistics. For her work with an HIV vaccine research team at Los Alamos National Laboratory, the Center for HIV/AIDS Vaccine Immunology (CHAVI) recognized her as their Young Investigator of the Month (July 2011). Her primary research interest is in the area of survival model selection in the context of missing data and correlated candidate predictors.