University of California, Irvine
Statistics Seminar

Prediction and Inference under Competing Risks in High Dimension - An EHR Demonstration Project for Prostate Cancer

Ronghui Xu
Professor, Family Medicine and Public Health
University of California, San Diego

Thursday, October 4, 2018
4 p.m., 6011 Bren Hall
(Bldg. #314 on campus map)

Our work was motivated by the analysis project using the linked SEER-Medicare database to predict cancer versus non-cancer mortality in men of age 65 years or older who were diagnosed with prostate cancer. We consider existing R package implementations that are computationally feasible for such data sets with up to 100,000 human subjects and over 20,000 claim codes. We carried out simulation studies to compare lasso for cause-specific hazards regression and boosting for Fine-Gray sub-distribution regression models with different approaches to choosing the penalty parameters, with the goal of prediction accuracy of the cumulative incidence rates at 2 and 5 years from baseline. Separately, we also developed methods for constructing confidence intervals of the regression effects of predictors (clinical variables and claim codes) under the Fine-Gray model in high dimensions, using the one-step estimator along the line of van de Geer et al. (2014) and Zhang and Zhang (2014). The extension though, is non-trivial under competing risks and Cox type model formulation. The results of the SEER-Medicare database analysis will be presented.

For directions/parking information, please visit https://uci.edu/visit/maps.php and http://www.ics.uci.edu/about/visit/index.php