University of California, Irvine
Statistics Seminar

Dynamic Risk Prediction Triggered by Intermediate Events Using Survival Tree Ensembles

Yifei Sun
Assistant Professor, Biostatistics
Columbia University

4 p.m., Thursday, Oct. 6, 2022
6011 Donald Bren Hall

With the availability of massive amounts of data from electronic health records and registry databases, incorporating time-varying patient information to improve risk prediction has attracted great attention. To exploit the growing amount of predictor information over time, we develop a unified framework for landmark prediction using survival tree ensembles, where an updated prediction can be performed when new information becomes available. Compared to conventional landmark prediction with fixed landmark times, our methods allow the landmark times to be subject-specific and triggered by an intermediate clinical event. Moreover, the nonparametric approach circumvents the thorny issue of model incompatibility at different landmark times. In our framework, both the longitudinal predictors and the event time outcome are subject to right censoring, and thus existing tree-based approaches cannot be directly applied. To tackle the analytical challenges, we propose a risk-set-based ensemble procedure by averaging martingale estimating equations from individual trees. Extensive simulation studies are conducted to evaluate the performance of our methods. The methods are applied to the Cystic Fibrosis Patient Registry (CFFPR) data to perform dynamic prediction of lung disease in cystic fibrosis patients and to identify important prognosis factors.