BAYESIAN SUBCOHORT SELECTION FOR LONGITUDINAL COVARIATE MEASUREMENTS IN FOLLOW-UP STUDIES

Jaakko Reinikainen

Department of Mathematics and Statistics, University of Jyväskylä, Finland, jaakko.o.reinikainen@jyu.fi

Juha Karvanen

Department of Mathematics and Statistics, University of Jyväskylä, Finland

A fundamental question in study design is how to answer the research question as precisely as possible with a limited budget. We consider planning longitudinal data collection in follow-up studies where covariates are time-varying. We assume that the entire cohort cannot be selected for longitudinal measurements due to financial limitations and study how a subset of the cohort should be selected optimally in order to obtain precise estimates of covariate effects in a survival model. In our approach, the study will be designed sequentially utilizing the data collected in previous longitudinal measurements as prior information. We propose using a Bayesian optimality criterion in the subcohort selections, which is compared to simple random sampling with simulated and real follow-up data. This study extends our previous results, where optimal subcohort selection was studied with only one re-measurement and one covariate, to more realistic cases where several covariates and measurement points are allowed. Our results show that the optimal subcohort includes individuals with a high risk of an event and, on the other hand, with extreme covariate values. The results support the conclusion that the precision of the estimates can be clearly improved by optimal design.