Dealing with Incomplete Covariates in Survival Models: A Bayesian Approach

Nicole Erler*[1]

[1]Department of Biostatistics, Erasmus University Medical Center [Rotterdam] – Netherlands

Abstract

Missing values are a challenge regularly encountered in the analysis of real-world data. Multiple imputation via a fully conditional specification (FCS) is popular to address the missing data problem. The separation of imputation and analysis and the specification of a set of univariate full-conditional models are attractive features that allow re-using the imputed data and facilitate a straightforward specification of imputation models for mixed-type variables. In time-to-event analyses, settings involving non-linear associations or when analysing multi-level data, however, important assumptions of the FCS approach are likely violated, leading to biased results. A fully Bayesian analysis, in which the parameters of interest are estimated jointly with the missing values, is an attractive alternative in such settings. The joint distribution of response variable(s), incomplete variables and parameters can be conveniently split into a sequence of (univariate) conditional distributions, allowing the choice of appropriate distributions for mixed-type variables while not requiring the inclusion of the response into a linear predictor. This makes the Bayesian approach suitable for highly complex substantive models, such as, for instance, multivariate joint models of longitudinal and survival data. Moreover, any complex association structures specified in the substantive model(s) are automatically taken into account during imputation, ensuring compatibility between all sub-models involved.

*Speaker