THE JOY OF PSEUDO-VALUES

Per Kragh Andersen, Section of Biostatistics, University of Copenhagen, Denmark.

Survival analysis is characterized by the need to deal with incomplete observation of the outcome variable, most frequently caused by right-censoring, and several – now standard – inference procedures have been developed to deal with this. Examples include the Kaplan-Meier estimator for the survival function and partial likelihood for estimating regression coefficients in the proportional hazards (Cox) model. During the last decades, methods based on pseudo-values have been studied. Here, the idea is to apply a transformation of the incompletely observed survival data and, thereby, to create a more simple data set for which 'standard' techniques (i.e., for complete data) may be applied, e.g., methods using generalized estimating equations.

An advantage of this approach is that it applies quite generally to (marginal) parameters for which no or few other regression methods are directly available (including average time spent in a state of a multi-state model). Another advantage is that it allows the use of a number of graphical techniques, otherwise unavailable in survival analysis. Disadvantages include that the method is not fully efficient and that it, in its simplest form, assumes covariate-independent censoring (though generalizations to deal with this have been developed).

We will review the development in the field since the idea was put forward by Andersen, Klein and Rosthøj in a 2003 Biometrika paper. Focus will be on graphical methods but the theoretical properties of the approach will also be touched upon.