Causal inference in survival analysis using pseudo-observations

Per Kragh Andersen
Biostatistics, University of Copenhagen,
Denmark

Causal inference for non-censored response variables, such as binary or quantitative outcomes, is often based on one of the following two approaches: (1) inverse probability of treatment assignment weights ('propensity score'), (2) direct standardization ('g-formula'). To do causal inference in survival analysis one needs to address right-censoring and, often, special techniques are required for that purpose.

We will show how censoring can be dealt with 'once and for all' by the means of so-called pseudo-observations when doing causal inference in survival analysis. Suppose that the target is the average causal effect expressed by the mean, $E(f(T))$, of some transformation $f(\cdot)$ of the survival time, T. Examples include $f(T) = I(T > t)$ and $f(T) = \min(T, \tau)$ leading to the average causal effect for the t-year survival probability $S(t) = E(I(T > t))$ and for the τ-restricted mean life time $E(\min(T; \tau))$, respectively. Without censoring, causal inference for such parameters could proceed as for other completely observed responses. In the presence of right-censoring we will replace the incompletely observed random variable $f(T_i)$ by its pseudo-observation obtained as follows:

Suppose that θ is a consistent estimator of $E(f(T))$ which may be calculated based on a right-censored sample of n independent subjects and that θ_i is the same estimator applied to the sample (of size $n - 1$) obtained by eliminating subject i from the total data set. The ith pseudo-observation is then

$$n\theta - (n-1)\theta_i.$$

We will then show how 'standard' causal inference techniques, such as (1) or (2) above, may be applied to the right-censored survival data. The same idea applies to competing risks settings.

The methods will be illustrated via a study of patients with acute myeloid leukemia who received either myeloablative or non-myeloablative conditioning before allogeneic hematopoetic cell transplantation. We will estimate the average causal effect of the conditioning regime on outcomes such as the 3-year leukemia-free survival probability and the 3-year risk of acute graft-versus-host disease.