MODEL EVALUATION AND VARIABLE SELECTION  
FOR INTERVAL–CENSORED DATA

Tyler Cook

Dr. (Tony) Jianguo Sun, Dissertation Supervisor

ABSTRACT

Survival analysis is a popular area of statistics dealing with time-to-event data. A special characteristic of survival data is the presence of censoring. Censoring occurs when the survival time is only partially known. This dissertation focuses on the analysis of interval-censored data, where the failure time is only known to belong to some interval of observation times.

One problem researchers face when analyzing survival data is how to handle the censoring distribution. It is often assumed that the observation process generating the censoring is independent of the event time of interest, but this assumption is clearly not always realistic. Researchers have a choice between using methods designed for informative or noninformative censoring. Part of this dissertation investigates the effectiveness of different methods developed for the analysis of informative interval-censored data under both types of censoring. Extensive simulation studies indicate that the methods produce unbiased results in the presence of both informative and noninformative censoring. The results of these simulation studies can provide guidelines for deciding between models when facing a practical problem where one is unsure about the dependence of the censoring distribution.

Another important problem seen in survival analysis is variable selection. Several methods using penalized likelihood procedures have been proposed for right-censored data to perform parameter estimation and variable selection simultaneously. We propose an imputation approach for variable selection of interval-censored data that utilizes these penalized likelihood procedures. Comprehensive simulation studies illustrate the effectiveness of this new approach. Also, this method is attractive due to how easy it is to implement, since it can take advantage of existing software for variable selection of right-censored data.