Survival Analysis:
Analysis of Right Censored Time to Event Data

Scott S. Emerson, M.D., Ph.D.
Professor of Biostatistics
University of Washington

May 1, 2004

Abstract

The analysis of data measuring time to event is often complicated by incomplete observations: Some subjects have not yet had an event at the time of data analysis. A wide variety of statistical methods have been developed for use in this setting of "right censored data", including parametric and semiparametric regression models, as well as a broad array of nonparametric methods. In this short course, I review the scientific and statistical issues involved in the analysis of right censored data and provide an overview of the spectrum of analytic methods available. I then focus on the nonparametric analysis of censored time to event data in the presence of nonproportional hazards. In particular, I consider a class of weighted logrank statistics, and contrast the use of those statistics with the behavior of statistics based on the weighted difference of survival curves. I then explore the effect that a varying censoring distribution has on the statistical operating characteristics of these methods when applied in sequential sampling.

Contact:

Scott S. Emerson, M.D., Ph.D.
Department of Biostatistics Box 357232
University of Washington
Seattle, Washington 98195
(206) 543-1044
semerson@u.washington.edu

Course Outline

Session 1: Overview of the Censored Data Setting

In the first session, I provide an overview of the types of scientific problems where right censored data typically arises and discuss the probability models that are used to quantify the scientific hypotheses. I consider mechanisms of censoring and the effect that such censoring has on the choice of distributional parameters used as the basis of inference and how the computation of those parameters must be modified in the censored data setting. In particular, I consider the special problems posed by competing risks and time-varying covariates.

Session 2: Nonparametric Analysis of Time to Event Data in the Presence of Nonproportional Hazards

In the second session, I focus on a class of weighted logrank statistics often used when nonproportional hazards are anticipated. I also consider a class of statistics based on the weighted difference between survival curves. I explore the operating characteristics of both classes of tests with respect to a variety of scientific and statistical criteria. In particular, I consider the sensitivity of these statistics to the censoring distribution, which can affect both the scientific interpretation of the parameters tested and the statistical power of the resulting analyses.

Session 3: Nonparametric Analysis of Time to Event Data under Sequential Sampling

In the clinical trial setting and some industrial settings, it is common to collect data under a sequential sampling plan in order to minimize study costs—scientific, economic, and ethical. Such sequential sampling causes special problems in the setting of censored survival data, because the censoring distribution is generally not constant across the various interim analyses of the data. I present approaches for investigating the effect that sequential sampling can have on a study's operating characteristics, and propose modifications to the standard survival analyses which mitigate some of the issues which arise in this setting. We have found that such modifications can have relatively minor effects on the efficiency of the analytic techniques, while allowing great improvements in the ability of the sequential sampling plan to reproduce the statistical behavior of a fixed sample test.