SCORE TESTS FOR DEPENDENT CENSORING WITH SURVIVAL DATA

MÉRIEM SAÏD, NADIA GHAZZALI AND LOUIS-PAUL RIVEST
Université Laval

In a standard survival data analysis, the observed time is the minimum of the survival time T and a censoring time independent of T. In this paper, we consider models featuring two censoring times U and V. The distribution of U is possibly related with that of T, while the second censoring time V is independent of both T and U. These models involve an Archimedean copula to incorporate a possible dependency between T and U. Score tests for dependent censoring are derived when a parametric model, e.g. Weibull or exponential, is assumed for T. One is fully parametric; it assumes that the marginal distributions of T, U, and V are either exponential or Weibull. The other is semiparametric; no assumptions are made on neither U nor V. The relative efficiencies of these two tests are compared with that of tests involving uncensored data. A numerical example is presented.

1. Introduction

A standard assumption underlying most statistical models for the analysis of lifetime data is the independence between survival and censoring times. This assumption is acceptable for *administrative* censoring associated with the termination of a study. However it is questionable in follow-up studies with self-selected removal where patients leave for causes possibly associated with the study variable. For such withdrawals, Frangakis and Rubin (2001) use the term *dropout* censoring. It is convenient to distinguish these two kinds of censoring. Administrative censoring is referred to as being "independent" while dropout censoring is called "dependent." The goal of this paper is to develop score tests to ascertain whether dropout censoring is truly dependent.

In a typical study the data set consists of the minimum (X) of survival, dependent censoring and independent censoring times T, U and V, respectively, is observed, with an indicator function δ taking value 1 if $X = T$, 0 if $X = U$ and -1 if $X = V$. Tsiatis (1975) showed that the independence between T, U and V is needed for their marginal distributions to be identifiable. This suggests that it might be difficult, not to say impossible, to test the independence between T and U without assuming a parametric model for at least one of their two marginal distributions. Nonparametric tests of independence might be feasible using additional follow-up of patients that either dropped out (see Lee and Wolfe, 1995, 1998, Lee, 1996) or "died" (see Lin, Robins and Wei, 1996) before the study termination. When only $\{(X_i, \delta_i)\}$ is observed, Zheng and Klein (1995) acknowledge that the level of dependency between T and U is not estimable from the data. It has to be