A FRAILTY MODEL APPROACH FOR REGRESSION ANALYSIS OF BIVARIATE INTERVAL-CENSORED SURVIVAL DATA

Chi-Chung Wen\textsuperscript{1} and Yi-Hau Chen\textsuperscript{2}

\textsuperscript{1} Tamkang University and \textsuperscript{2} Academia Sinica

Abstract: Owing to the fact that general semiparametric inference procedures are still underdeveloped for multivariate interval-censored event time data, we propose semiparametric maximum likelihood estimation for the gamma-frailty Cox model under mixed-case interval censoring. We establish the consistency of the semiparametric maximum likelihood estimator (SPMLE) for the model parameters, including the regression coefficients and the cumulative hazard functions in the Cox model, and the variance of the gamma frailty. The SPMLEs of the cumulative hazard functions are shown to have a $n^{1/3}$-rate of convergence, while those of the regression coefficients and the frailty variance have a $n^{1/2}$-rate of convergence; here $n$ denotes the number of study units. The asymptotic normality for the regression coefficients and the frailty variance is also established, with the asymptotic variance given by the inverse of the efficient Fisher information matrix. A profile-likelihood approach is proposed for estimating the asymptotic variance. Based on the self-consistency equations and the contraction principle, we propose a very stable and efficient computation algorithm. Simulation results reveal that the large sample theories work quite well in finite samples. We analyze a dataset from an AIDS clinical trial by the proposed methods to assess the effects of the baseline CD4 cell counts on the times to CMV shedding in blood and urine.

Key words and phrases: Correlated data, interval censoring, proportional hazards, self-consistency.

1. Introduction

Data on survival or event time are often subject to censoring due to limitation in the observational process. For example, right censoring occurs when time to the event is beyond the end of observation, while interval censoring occurs when the observation is only made at several examination times, and hence we can only know the event time lies in some interval bracketed by two examination times. The incomplete nature of censored event time data complicates the subsequent statistical analysis, including the event time regression analysis where the covariate effects on the event time are to be assessed. In particular, interval-censored data generally create more difficulties