Nonparametric estimation in the sequential lifetime model under random censorship

Abstract

In many applications one observes lifetimes X_1, X_2 in sequential order. For example, X_1 could be the incubation period of a disease and X_2 the duration until death. The sum of these lifetimes is often censored from the right by a random variable C. So the model consists of the twodimensional vector $(X_1, X_2) \sim F$ and the independent variable $C \sim G$. The observable variables are:

$$\begin{aligned} Z_1 &= \min(X_1, C) \\ Z_2 &= \min(X_2, (C - X_1)\, 1_{\{X_1 \leq C\}}) \\ \delta &= 1 + 1_{\{X_1 \leq C\}} + 1_{\{X_2 \leq C - X_1\}} \end{aligned}$$

The model and the results achieved in this thesis can easily be extended to the k-dimensional case where there are lifetimes $X_1, \ldots, X_k$ in a series.
In this multivariate censorship model one is not only interested in the joint distribution F of the lifetimes but more generally in the integral

$$\int \varphi(x, y) F d(x, y)$$

where φ is an arbitrary F-integrable function.
In this thesis a nonparametric estimator S_n for this target is developed by using an appropriate identifying equation. Some desirable properties for S_n are shown and the distributional structure is analysed. Under weak integrability assumptions on φ a linearisation is derived which leads to asymptotic normality of S_n. For the proof the theory of U-statistics and the Hájek-projection is used. Afterwards a related process indexed by a class of φ's is considered. It is shown that the estimator $S_n = S_n(\varphi)$ converges in distribution to a centered gaussian process if the index set is a Vapnik-Červonenkis-class . For this it is proved that the leading term of the linearisation forms a Donsker-class. For the remainder-term neglibility uniformly in φ is shown by using U-processes. Finally a simulation study is included to make the quality of the estimator for F for small and moderate sample sizes visible.