

Nonparametric estimators for interval censoring problems*

Piet Groeneboom
Delft University of Technology

Abstract

We study weighted least squares estimators for the distribution function of observations which are only visible via interval censoring, i.e., in the situation where one only has information about an interval to which the variable of interest belongs and where one cannot observe it directly. The least squares estimators are shown to be closely related to nonparametric maximum likelihood estimators (NPMLE's) and to coincide with these in certain cases. New algorithms for computing the estimators are presented and it is shown that they converge from any starting point (in contrast with the EM-algorithm in this situation). Finally, the estimation of non-smooth and smooth functionals of the model is considered; for the latter case, we discuss \sqrt{n} -consistency and efficiency of the NPMLE.

1 Introduction

An extensive statistical theory exists for treating right censored data. Much less is known about more general types of censorship. This paper considers estimators for data subject to interval censoring. In this situation one only has information about an interval to which the observation of interest belongs; so only indirect information about the observation of interest is available.

Most of the time the interval will be a time interval, but the following interesting spatial version of this situation was brought to our attention by professor Dietz. In examinations of skin tissue, possibly affected by skin cancer, successive (roughly) circular incisions are made to determine the region of affected tissue; in this case one tries to estimate the smallest “safe” radius determining the region on which the operation should take place. On the one hand one tries to minimize the number of incisions, but on the

*AMS 1991 subject classifications. 60F17, 62E20, 62G05, 62G20, 45A05.

Key words and phrases. nonparametric maximum likelihood, empirical processes, asymptotic distributions, asymptotic efficiency, Fredholm integral equations.