Methods for interval-censored data and testing for stochastic dominance

Angel G. Angelov

Department of statistics, Umeå University

Abstract

This thesis includes four papers: the first three of them are concerned with methods for interval-censored data, while the forth paper is devoted to testing for stochastic dominance.

In many studies, the variable of interest is observed to lie within an interval instead of being observed exactly, i.e., each observation is an interval and not a single value. This type of data is known as interval-censored. It may arise in questionnaire-based studies when the respondent gives an answer in the form of an interval without having pre-specified ranges. Such data are called self-selected interval data. In this context, the assumption of noninformative censoring is not fulfilled, and therefore the existing methods for interval-censored data are not necessarily applicable.

A problem of interest is to estimate the underlying distribution function. There are two main approaches to this problem: (i) parametric estimation, which assumes a particular functional form of the distribution, and (ii) nonparametric estimation, which does not rely on any distributional assumptions. In Paper A, a nonparametric maximum likelihood estimator for self-selected interval data is proposed and its consistency is shown. Paper B suggests a parametric maximum likelihood estimator. The consistency and asymptotic normality of the estimator are proven.

Another interesting problem is to infer whether two samples arise from identical distributions. In Paper C, nonparametric two-sample tests suitable for self-selected interval data are suggested and their properties are investigated through simulations.

Paper D concerns testing for stochastic dominance with uncensored data. The paper explores a testing problem which involves four hypotheses, that is, based on observations of two random variables X and Y, one wants to discriminate between four possibilities: identical survival functions, stochastic dominance of X over Y, stochastic dominance of Y over X, or crossing survival functions. Permutation-based tests suitable for two independent samples and for paired samples are proposed. The tests are applied to data from an experiment concerning the individual's willingness to pay for a given environmental improvement.

For a copy of the thesis, please find it here:

http://umu.diva-portal.org/smash/record.jsf?pid=diva2:1241348