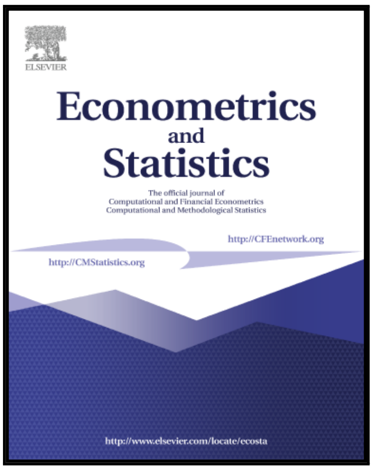
Accepted Manuscript



The copula-graphic estimator in censored nonparametric location-scale regression models

Aleksandar Sujica, Ingrid Van Keilegom

PII: S2452-3062(17)30063-1
DOI: 10.1016/j.ecosta.2017.07.002
Reference: ECOSTA 70

To appear in: *Econometrics and Statistics*

Received date: 16 November 2016
Revised date: 4 July 2017
Accepted date: 18 July 2017

Please cite this article as: Aleksandar Sujica, Ingrid Van Keilegom, The copula-graphic estimator in censored nonparametric location-scale regression models, *Econometrics and Statistics* (2017), doi: 10.1016/j.ecosta.2017.07.002

The copula-graphic estimator in censored nonparametric location-scale regression models

Aleksandar Sujica[a], Ingrid Van Keilegom [1 b]

[a]Institute of Statistics, Biostatistics and Actuarial Sciences, Université catholique de Louvain, Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium.

[b]ORSTAT, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium. E-mail address : ingrid.vankeilegom@kuleuven.be

Abstract

A common assumption when working with randomly right censored data, is the independence between the variable of interest Y (the survival time) and the censoring variable C. This assumption, which is not testable, is however unrealistic in certain situations. Let us assume that for a given covariate X, the dependence between the variables Y and C is described via a known copula. Additionally assume that Y is the response variable of a heteroscedastic regression model $Y = m(X) + \sigma(X)\varepsilon$, where the error term ε is independent of the explanatory variable X, and the functions m and σ are 'smooth'. An estimator of the conditional distribution of Y given X under this model is then proposed, and the asymptotic normality of this estimator is shown. The small sample performance of the estimator is also studied, and the advantages/drawbacks of this estimator with respect to competing estimators are discussed.

Keywords: Asymptotic normality, asymptotic representation, copula, dependent censoring, kernel estimator, nonparametric regression, right censoring.

1. Introduction

Consider the following nonparametric location-scale model

$$Y = m(X) + \sigma(X)\varepsilon , \tag{1}$$

where the error ε is assumed to be independent of a one dimensional covariate X. The function $m(\cdot)$ is a conditional location functional and $\sigma(\cdot)$ is a conditional scale functional representing possible heteroscedasticity. We assume that Y is a possible (given) transformation of a survival time and is subject to random right censoring, i.e. instead of observing Y we only observe (T, Δ), where $T = \min(Y, C)$, $\Delta = I(Y \leq C)$ and C represents the censoring time. Let (T_i, X_i, Δ_i), $i = 1, \ldots, n$ be n independent vectors having the same distribution as (T, X, Δ).

The motivation for considering model (1) comes from the fact that the model offers important advantages with respect to the completely nonparametric model when one is interested in the estimation of the conditional distribution $F(\cdot|x) = P(Y \leq \cdot|X = x)$ of Y given $X = x$. [1] showed how advantage can be taken from model (1) to estimate this conditional distribution. The advantages are especially apparent in the right tail of the distribution. In this region the completely nonparametric competitor proposed by [2] (see also [3], [4], [5], [6], [7], among others) suffers from inconsistency problems especially when censoring is heavy. This phenomenon is similar to what happens in the right tail of the Kaplan-Meier estimator in the absence of covariates. Under model (1), [1] showed that the right tail of the distribution $F(\cdot|x)$ can be well estimated for all values of X, provided there is a region of X where censoring is light. This is because under model (1) the conditional distribution $F(y|x)$ can be written as

$$F(y|x) = F_e\left(\frac{y - m(x)}{\sigma(x)}\right), \tag{2}$$

[1]This research was supported by the European Research Council (2016-2021, Horizon 2020 / ERC grant agreement No. 694409), and by IAP Research Network P7/06 of the Belgian State.