Journal of Statistical Planning and Inference 149 (2014) 33–45

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Maximum likelihood estimation for left-censored survival times in an additive hazard model

Alexander Kremer [a], Rafael Weißbach [a,*], Friedrich Liese [b]

[a] Faculty for Economic and Social Sciences, University of Rostock, Ulmenstraße 69, 18057 Rostock, Germany
[b] Institute of Mathematics, University of Rostock, Ulmenstraße 69, 18057 Rostock, Germany

ARTICLE INFO

Article history:
Received 26 April 2013
Received in revised form
29 October 2013
Accepted 26 February 2014
Available online 5 March 2014

Keywords:
Left censoring
Parametric maximum likelihood
Asymptotic normality
Time-dependent covariate
Additive hazard

ABSTRACT

Motivated by an application from finance, we study randomly left-censored data with time-dependent covariates in a parametric additive hazard model. As the log-likelihood is concave in the parameter, we provide a short and direct proof of the asymptotic normality for the maximal likelihood estimator by applying a result for convex processes from Hjort and Pollard (1993). The technique also yields a new proof for right-censored data. Monte Carlo simulations confirm the nominal level of the asymptotic confidence intervals for finite samples, but also provide evidence for the importance of a proper variance estimator. In the application, we estimate the hazard of credit rating transition, where left-censored observations result from infrequent monitoring of rating histories. Calendar time as time-dependent covariates shows that the hazard varies markedly between years.

1. Introduction

We consider parametric modeling of left-censored durations. Being a special case of double censoring (see for example Fernandez et al., 1999, 2002, and the papers cited therein), so far parametric left censoring has been applied in Lynn (2001). Our aim is to contribute a detailed proof for the asymptotic normality of the maximum likelihood estimator (MLE). One attempt to establish the asymptotic normality of the MLE would be a linearization of the score equation. To this end, *every* sequence of MLEs maximizing the likelihood must converge to the true parameter value. Wald was the first to establish conditions for this type of consistency (Wald consistency). In contrast, in Cramer's concept of consistency *at least one* sequence of solutions for the score equation needs to converge (Cramer consistency). The latter concept needs fewer technical effort, as compared to the Wald consistency. However, Cramer consistency is not sufficient to prove asymptotic normality (see Lehmann, 1998, Chapter 6, for a detailed discussion).

We restrict our analysis to parametric left censoring, as in Gomez et al. (1992) or Lynn (2001). In our model, the log-likelihood is a concave function of the parameters and allows for Theorem 2.2 of Hjort and Pollard (1993) to be applied. The approach circumvents the problem of establishing Wald consistency in a first step and, additionally, allows for not necessarily identically distributed duration times. In particular, we include time-dependent covariates. In order to confirm the conditions of the theorems, we use the boundedness of the covariates, the concavity of the log-likelihood which follows by the concavity of the log-density, and the finiteness of the expected second and third derivatives of the density and the log-density. Asymptotic normality of the MLE is required for constructing asymptotic confidence intervals. In order to

* Corresponding author. Tel.: +49 381 4984428; fax: +49 381 4984401.
E-mail address: rafael.weissbach@uni-rostock.de (R. Weißbach).

http://dx.doi.org/10.1016/j.jspi.2014.02.013
0378-3758

calculate the standard error, the Fisher information must be studied being defined as minus the expectation as the second derivative of the density. As under weak regularity conditions, the Fisher information coincides with the expectation of the squared first derivative, observed versions of both are studied. We prove that both the observed Fisher informations consistently estimate the Fisher information.

Monte Carlo simulations are carried out to verify that the nominal level of the confidence interval is attained. We compare the finite sample properties of all four Fisher informations mentioned above. The simulations show that even for medium size samples, the nominal level of the confidence interval is attained with high accuracy in three of four cases. With respect to the application, the probability of default (PD) is an important parameter in the framework of credit ratings. It is to be estimated with the help of internal default data (see Basel Committee on Banking Supervision, 2004, paragraph 461ff). As a consequence of the model in Weißbach and Mollenhauer (2011) transitions to an adjacent class are exponentially distributed, but, see Mählmann (2006), can only be observed left-censored. We include the influence of calendar time on the transition hazard, using time-dependent dummies. In this particular model, we prove that the Fisher information is positive definite, which has to be assumed in the general left censoring framework.

2. Censoring model

Instead of an event time X, or more precisely an age-at-event, we observe in the left censoring model only

$$(T,\Delta)=(\max(X,C),I_{[C,\infty)}(X)),$$

where C is an external unobservable influence and $I_{[C,\infty)}$ is the indicator function of $[C,\infty)$. Hence observations are in $\mathbb{R}\times\{0,1\}$. As dominating measure on this sample space we use $\mu\otimes\kappa$, where μ is the Lebesgue measure on the Borel sets of $\mathbb{R}$, κ is the counting measure on $\{0,1\}$, and $\mu\otimes\kappa$ is the symbol for the product measure. Moreover we use the abbreviation $dt{:=}\mu(dt)$ and the convention $0^0=0$. The proof of the following lemma is standard.

Lemma 1. *Suppose X and C are independent with c.d.f. F and G and (Lebesgue) densities f and g. If Q denotes the distribution of (T,Δ) then*

$$h(t,\delta):=\frac{dQ}{d(\mu\otimes\kappa)}(t,\delta)=f(t)^{\delta}F(t)^{1-\delta}G(t)^{\delta}g(t)^{1-\delta}. \tag{1}$$

We assume that $X_1,\ldots,X_n$ is a sample of independent, but not necessarily identically distributed durations from the parametrized families $(F_{i,\theta})_{\theta\in\Theta}$, $i=1,\ldots,n$. Denote by $(f_{i,\theta})_{\theta\in\Theta}$, $\Theta\subseteq\mathbb{R}$, the associated families of densities. Moreover, the censoring times $C_1,\ldots,C_n$ are assumed to be independently distributed and independent of $X_1,\ldots,X_n$. Assume that the censoring is non-informative, which means that its c.d.f. G_i do not depend on θ. To compute the log-likelihood, adapt f and F in (1) as $f_{i,\theta}$ and $F_{i,\theta}$, respectively. The resulting distribution and density depend on θ and will be denoted by $Q_{i,\theta}$, and $h_{i,\theta}(t,\delta)$. The log-likelihood is

$$\begin{aligned}\ln L_n(\theta,T_1,\ldots,T_n,\Delta_1,\ldots,\Delta_n)&=\sum_{i=1}^{n}\ln h_{i,\theta}(T_i,\Delta_i)\\&=\sum_{i=1}^{n}(\Delta_i\ln f_{i,\theta}(T_i)+(1-\Delta_i)(\ln(F_{i,\theta}(T_i))\\&\quad+\sum_{i=1}^{n}(\Delta_iG_i(T_i)+(1-\Delta_i)(\ln g_i(T_i)).\end{aligned}\tag{2}$$

Note that for estimating θ, it suffices to consider the first summands of the log-likelihood. In order to guarantee the assumptions needed to prove asymptotic normality it is necessary to formulate conditions on the event distribution. Recall that the density $f(t)$ of the univariate X and its hazard rate $\lambda(t)$ are related by

$$f(t)=I_{[0,\infty)}(t)\lambda(t)\exp\left\{-\int_0^t\lambda(s)\,ds\right\} \tag{3}$$

and that for a given function λ, the right-hand term in (3) defines a probability density if and only if

$$\lambda(t)\ge 0,\qquad\int_0^\infty\lambda(s)\,ds=\infty. \tag{4}$$

A flexible and tractable model for the event times is the additive hazard model with b time-dependent fixed regressors $z_{i,j}(t)$, see Andersen et al. (1993, Formular 7.1.4)

$$\lambda_{i,\theta}(t)=\sum_{j=0}^{b}\theta_j z_{i,j}(t),\quad z_{i,0}(t)\equiv 1, \tag{5}$$