Statistics and Probability Letters 80 (2010) 911–915

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Identifiability conditions for covariate effects model on survival times under informative censoring

A. Adam Ding *

Department of Math, Northeastern University, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 7 March 2009
Received in revised form 31 August 2009
Accepted 28 January 2010
Available online 10 February 2010

Keywords:
Archimedean copula
Artificial censoring
Dependent censoring

ABSTRACT

We provide the identifiability conditions for the covariate effects modeling of bivariate survival data when one survival time is dependently censored by the other survival time. The covariate effects are specified through three components of the copula decomposition. Many commonly used copula families are shown to satisfy the identifiable condition. A condition for causal interpretation is also provided.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

In clinical trials, the time to disease X is often subject to censoring by another survival time Y. When X and Y are dependent, the censoring is informative. Under informative censoring (also called dependent censoring), the marginal distribution of X is not identifiable. Therefore it is impossible to estimate the effect of a covariate Z on X without additional assumptions. Robins (1995a) studied the identifiability issues for this informative censoring setting with a binary covariate Z. Since it is not easy to separate out the covariate effect on X directly from the joint covariate effects model on X and Y, Robins (1995a) eventually focused on the simple case that covariate Z does not affect the dependence structure between X and Y. For this simple case, the covariate effect of Z on X is easy to interpret and conditions were imposed for causal interpretation of this effect. Robins (1995b) developed the artificial technique for inference on the covariate effects on X under these assumptions. Similarly, Lin et al. (1996) imposed location-shift covariate effects on X and Y separately for the case of no covariate effect on the dependence structure between X and Y. The artificial censoring technique is then used to derive an asymptotic normal estimator of treatment effect of Z on X. Further generalization of the artificial censoring techniques in Chang (2000), Ghosh and Lin (2003), Lin and Ying (2003) and Peng and Fine (2006) all assumed that covariate Z does not affect the dependence structure.

Other papers not using artificial censoring technique also are based on the assumption that covariate Z does not affect the dependence structure. For example, Peng and Fine (2007) proposed fitting marginal regression with time-varying coefficients and the dependent structure from a copula family with time-varying copula parameter. However, the time-varying copula parameter is not affected by covariate Z.

Hsieh et al. (2008) proposed an alternative method that allows covariate Z to affect the dependence structure by assuming that the dependence structure between X and Y belongs to a copula family whose copula parameter changes with Z. The joint distribution of X and Y can be separated into three components: marginal distributions of X and Y, and a copula representing the dependence structure. This decomposition allows easy interpretation of covariate effects on X directly. This paper spells out the identifiability condition from the view of this copula approach.

* Tel.: +1 617 3735231; fax: +1 617 3735658.
E-mail address: ding@neu.edu.

0167-7152/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2010.01.027

Section 2 discusses the identifiability condition. We specify the covariate model through its effects on the three model components: two marginals and the copula. The identifiable conditions are derived from this point of view. We also show that many copula families satisfy this identifiable condition. Section 3 gives a corresponding version of Robins' assumption for causal interpretation. Summary discussions are contained in Section 4.

2. Identifiable conditions for covariate model

2.1. Model notations

The joint survival function of X and Y is denoted by $S(x, y) = pr(X \geq x, Y \geq y)$. Generally the joint survival function of two variables can be decomposed into three components: the marginal distributions of each variable and the correlation structure represented by a copula — a joint survival function on the unit square $[0, 1] \times [0, 1]$. That is, $S(x, y) = C\{S_1(x), S_2(y)\}$, where $S_1(x) = pr(X \geq x)$ and $S_2(y) = pr(Y \geq y)$ are the marginal survival functions of X and Y respectively, and $C(s_1, s_2)$ is the copula.

Using this decomposition, we can consider the covariate effects separately on three components. The covariate effect on the survive time X can be represented by monotone non-decreasing functions $m_{1,z}(t)$ so that $m_{1,z}(X)$ given $Z = z$ has the same marginal distribution as X given $Z = 0$. Therefore, $pr(X \geq t|Z = z) = pr(X \geq m_{1,z}(t)|Z = 0)$. That is, $S_1(t; z) = S_1\{m_{1,z}(t); 0\}$. The covariate effects model is represented by a family $\mathcal{M}_1$, the collection of the possible $m_{1,z}$. Similarly, the covariate effects model on Y is modeled as $S_2(t; z) = S_2\{m_{2,z}(t); 0\}$ with $m_{2,z} \in \mathcal{M}_2$. The copula given $Z = z$ is indexed by $\alpha_z \in \mathcal{A}$. That is, $C(s_1, s_2; z) = C(s_1, s_2|Z = z) = C_{\alpha_z}(s_1, s_2)$. Thus the overall covariate effects on the bivariate survival times are denoted by $(m_{1,z}, m_{2,z}, \alpha_z)$, with the possible values of the three components belong to families $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{A}$ respectively for all $z \in \mathcal{Z}$.

Remark. Notice the above covariate effects model on the marginal include all possible univariate covariate effects models and $m_{1,0}(t) = t$ always. For example, a nonparametric covariate model is represented by $\mathcal{M}_1 = \{m(t) : m(0) = 0, m(\cdot)$ non-decreasing left-continuous mapping from $[0, \infty)$ to $[0, \infty)\}$. The accelerated failure times model corresponds to $\mathcal{M}_1 = \{m(t) = \exp(\theta_1 z)t : -\infty < \theta_1 < \infty\}$. The proportional hazards model corresponds to $\mathcal{M}_1 = \{m(t) = S_0^{-1}[S_0(t)^{\exp(\theta_1 z)}] : -\infty < \theta_1 < \infty,\ S_0(t)$ the baseline survive function, a non-decreasing left-continuous mapping from $[0, \infty)$ to $[0, 1]\}$.

2.2. Identifiable conditions

Due to the copula decomposition

$$S(x, y; z) = C_{\alpha_z}\{S_1(x; z), S_2(y; z)\}, \tag{1}$$

it is obvious that the identifiability of $S(x, y; z)$ requires all three components are identifiable: (I1) If $S_1\{m^*_{1,z}(t); 0\} = S_1\{m_{1,z}(t); 0\}$ for all $t \geq 0$, then $m^*_{1,z} = m_{1,z}$; (I2) If $S_2\{m^*_{2,z}(t); 0\} = S_2\{m_{2,z}(t); 0\}$ for all $t \geq 0$, then $m^*_{2,z} = m_{2,z}$; (I3) If $C_{\alpha^*}(s, t) = C_\alpha(s, t)$ for all $(s, t) \in [0, 1] \times [0, 1]$, then $\alpha^* = \alpha$.

When the distribution of X and Y are continuous, the identifiability of the three components will imply the identifiability of the joint distribution $S(x, y; z)$ for $x \geq 0, y \geq 0$ and $z \in \mathcal{Z}$. However, when X is informatively censored by Y, then we only observe $T = min(X, Y)$, $\delta = I\{T = X\}$ and $D = Y$. Therefore, the observable region for the joint survival distribution is only $y \geq x \geq 0$. So we need the identifiability of $S(x, y; z)$ for $y \geq x \geq 0$ and $z \in \mathcal{Z}$. Let $y_z = \inf[t : S_2\{m_{2,z}(t); 0\} = 0]$, then y_z denotes the upper bound of the support of Y given $Z = z$ under $m_{2,z}$. Note that y_z may be ∞. Since no data can be observed for $Y > y_z$, the identifiability will require (a) $S_1\{m_{1,z}(y_z); 0\} < 1$; (Otherwise, all α_z values result in the same joint distribution on the observable region, thus unidentifiable.) and (b) if $S_1\{m^*_{1,z}(t); 0\} = S_1\{m_{1,z}(t); 0\}$ for all $0 \leq t \leq y_z$, then $m^*_{1,z} = m_{1,z}$. So (I1) needs to be replaced by the stronger necessary condition (I1*): For any $m_{2,z}$ and $y_z = \inf[t : S_2\{m_{2,z}(t); 0\} = 0]$, we have $S_1\{m_{1,z}(y_z); 0\} < 1$, and that $S_1\{m^*_{1,z}(t); 0\} = S_1\{m_{1,z}(t); 0\}$ for all $0 \leq x \leq y_z$ implies $m^*_{1,z} = m_{1,z}$.

The conditions (I1*), (I2) and (I3), while necessary, still do not ensure the identifiability under the informative censoring. We also need to enhance (I3) to a stronger condition of a lower-right tail identifiable copula family.

Definition 1. A copula family $\{C_\alpha(s, t), \alpha \in \mathcal{A}\}$ is called lower-right tail identifiable if

$$C_\alpha(s, t) = C_{\alpha^*}(s^*(s), t)$$

on an open set $s_0 < s < 1$ and $0 < t < t_0$ implies that $\alpha = \alpha^*$ and $s^*(s) = s$ on $s_0 < s < 1$.

We will assume (I3*): $C_{\alpha_z}(s, t)$ belongs to a lower-right tail identifiable copula family $\{C_\alpha(s, t), \alpha \in \mathcal{A}\}$, for all $z \in \mathcal{Z}$.

Theorem 1. *Assume that the distributions of X and Y are absolutely continuous, and the identifiable condition* (I1*) *and* (I2) *are satisfied. Then* (I3*) *is a sufficient condition for the identifiability of joint covariate effects model under informative censoring. When the univariate covariate models $\mathcal{M}_1$ and $\mathcal{M}_2$ are nonparametric, the condition* (I3*) *is also necessary.*

Download English Version:

https://daneshyari.com/en/article/1153227

Download Persian Version:

https://daneshyari.com/article/1153227

Daneshyari.com