应用概率统计　第二十九卷
第三期　2013年6月

Chinese Journal of Applied Probability
and Statistics Vol.29 No.3 Jun. 2013

A Method for Estimating the Association Parameter in the Clayton Model *

Zhang Qiaozhen

(*Department of Statistics, School of Mathematical Sciences and LPMC, Nankai University, Tianjin, 300071*)

He Shuyuan

(*School of Mathematical Sciences, Capital Normal University, Beijing, 100048*)

Abstract

Based on the estimates of bivariate hazard functions, for right censored data, we give an estimator of association parameter in Clayton model in the paper. The consistency and asymptotic distribution are derived for the estimator. Simulation studies show that this procedure is effective.

Keywords: Right censorship, Clayton model, association parameter, asymptotic theory.

AMS Subject Classification: 62F10, 62F12.

§1. Introduction

Bivariate survival time data arises when a sample consists of two variables. The analysis of bivariate survival time must reflect the non-independence of failures between the two variables. Let T_1 and T_2 be survival times with continuous probability density $f(t_1, t_2)$. For (t_1, t_2) such that $f(t_1, t_2) > 0$, Clayton model defined by

$$f(t_1,t_2)\int_{t_1}^{\infty}\int_{t_2}^{\infty} f(u,v)\mathrm{d}u\mathrm{d}v = \theta\int_{t_1}^{\infty} f(u,t_2)\mathrm{d}u\int_{t_2}^{\infty} f(t_1,v)\mathrm{d}v$$

is an appealing representation for such data. The parameter θ, called association parameter, measures the degree of association between T_1 and T_2. Independence of T_1 and T_2 is implied by $\theta = 1$ and positive association is implied by $\theta > 1$. Inverse association is implied by $\theta < 1$, but this case seems to have little practical importance (see Clayton (1978)).

*The work was supported by National Natural Science Foundation of China (11171230, 11231010) and the Fundamental Research Funds for the Central Universities (65011481).

Received September 12, 2012. Revised January 3, 2013.

Define the following bivariate hazard functions:

$$\lambda_{10}(t_1,t_2)=\frac{\displaystyle\int_{t_2}^{\infty}f(t_1,v)\mathrm{d}v}{F(t_1-,t_2)},\quad \lambda_{01}(t_1,t_2)=\frac{\displaystyle\int_{t_1}^{\infty}f(u,t_2)\mathrm{d}u}{F(t_1,t_2-)},\quad \lambda_{11}(t_1,t_2)=\frac{f(t_1,t_2)}{F(t_1-,t_2-)},$$

where $F(t_1,t_2)=\mathsf{P}(T_1>t_1,T_2>t_2)$ is the joint survival function of (T_1,T_2). Using the functions, we have an equivalent form of the Clayton model,

$$\frac{\lambda_{11}(t_1,t_2)}{\lambda_{10}(t_1,t_2)\lambda_{01}(t_1,t_2)}=\theta. \tag{1.1}$$

Many authors considered the inference of θ. Clayton (1978) gave an estimator of θ for uncensored data. Oakes (1986) derived the asymptotic variance of Clayton's estimator and obtained a simple explicit formula for uncensored data. He indicated that a modification is necessary for random censored data. Genest et al. (1995) estimated θ from a pseudo likelihood with nonparametric estimation of the marginal survival functions for complete data. Shih and Louis (1995) investigated two-stage parametric and two-stage semi-parametric estimation procedure in copula model where censoring was allowed. Glidden (2000) extended the above approach and proposed a two-stage estimator of θ, the estimator is consistent and asymptotically normal under mild regularity conditions. Nan et al. (2006) used Clayton model to describe association of age at a marker event and age at menopause. Ghosh (2008) used the model to solve problem of surrogate endpoints in clinical research. Emura et al. (2010) extended an existing method suitable for the Clayton model to general Archimedean copula models and derived the asymptotic properties of the proposed test statistics.

This paper is organized as follows. In Section 2, we give a new estimator of θ by the ratio of hazard functions. In Section 3, we present the asymptotic properties of the estimator. In Section 4, we present simulation studies.

§2. The Estimation of Association

Let $T=(T_1,T_2)$ be a 2-vector of failure times with continuous survival function $F(t_1,t_2)$. Let $C=(C_1,C_2)$ be the censoring time independent of T with survival function

$$G(t_1,t_2)=\mathsf{P}(C_1>t_1,C_2>t_2).$$

Under right censorship, the data consist of n realizations of $(X_1,X_2,\delta_1,\delta_2)$, where $X_i=\min(T_i,C_i)$, $\delta_i=I\{T_i\le C_i\}$. Note that the survival function of (X_1,X_2) is $H=F\cdot G$.

The observation

$$(X_{1j}, X_{2j}, \delta_{1j}, \delta_{2j}), \qquad j = 1, \dots, n$$

are i.i.d. copies of $(X_1, X_2, \delta_1, \delta_2)$.

Define bivariate cumulative functions:

$$\begin{aligned}
\Lambda_{10}(t_1,t_2) &= \int_0^{t_1} \lambda_{10}(u,t_2)\mathrm{d}u = \int_0^{t_1} \frac{-F(\mathrm{d}u,t_2)}{F(u-,t_2)},\\
\Lambda_{01}(t_1,t_2) &= \int_0^{t_2} \lambda_{01}(t_1,v)\mathrm{d}v = \int_0^{t_2} \frac{-F(t_1,\mathrm{d}v)}{F(t_1,v-)},\\
\Lambda_{11}(t_1,t_2) &= \int_0^{t_2}\int_0^{t_1} \lambda_{11}(u,v)\mathrm{d}u\mathrm{d}v = \int_0^{t_2}\int_0^{t_1} \frac{F(\mathrm{d}u,\mathrm{d}v)}{F(u-,v-)}.
\end{aligned}$$

Let

$$\begin{aligned}
\widehat{H}(t_1,t_2) &= \frac{1}{n}\sum_{j=1}^n I\{X_{1j}>t_1, X_{2j}>t_2\},\\
\widehat{K}_1(t_1,t_2) &= \frac{1}{n}\sum_{j=1}^n I\{X_{1j}>t_1, X_{2j}>t_2, \delta_{1j}=1, \delta_{2j}=1\},\\
\widehat{K}_2(t_1,t_2) &= \frac{1}{n}\sum_{j=1}^n I\{X_{1j}>t_1, X_{2j}>t_2, \delta_{1j}=1\},\\
\widehat{K}_3(t_1,t_2) &= \frac{1}{n}\sum_{j=1}^n I\{X_{1j}>t_1, X_{2j}>t_2, \delta_{2j}=1\},
\end{aligned}$$

and

$$H = \mathsf{E}(\widehat{H}), \qquad K_1 = \mathsf{E}(\widehat{K}_1), \qquad K_2 = \mathsf{E}(\widehat{K}_2), \qquad K_3 = \mathsf{E}(\widehat{K}_3).$$

It can be checked that

$$\begin{aligned}
\Lambda_{10}(t_1,t_2) &= -\int_0^{t_1} \frac{K_2(\mathrm{d}u,t_2)}{H(u-,t_2)},\\
\Lambda_{01}(t_1,t_2) &= -\int_0^{t_2} \frac{K_3(t_1,\mathrm{d}v)}{H(t_1,v-)},\\
\Lambda_{11}(t_1,t_2) &= \int_0^{t_1}\int_0^{t_2} \frac{K_1(\mathrm{d}u,\mathrm{d}v)}{H(u-,v-)}.
\end{aligned}$$

Thus, the estimates of the bivariate cumulative hazard functions can be defined as follows,

$$\begin{aligned}
\widehat{\Lambda}_{10}(t_1,t_2) &= -\int_0^{t_1} \frac{\widehat{K}_2(\mathrm{d}u,t_2)}{\widehat{H}(u-,t_2)} = \sum_{j=1}^n \frac{I\{X_{1j}\le t_1, X_{2j}>t_2, \delta_{1j}=1\}}{n\widehat{H}(X_{1j}-,t_2)},\\
\widehat{\Lambda}_{01}(t_1,t_2) &= -\int_0^{t_2} \frac{\widehat{K}_3(t_1,\mathrm{d}v)}{\widehat{H}(t_1,v-)} = \sum_{j=1}^n \frac{I\{X_{1j}>t_1, X_{2j}\le t_2, \delta_{2j}=1\}}{n\widehat{H}(t_1,X_{2j}-)},\\
\widehat{\Lambda}_{11}(t_1,t_2) &= \int_0^{t_1}\int_0^{t_2} \frac{\widehat{K}_1(\mathrm{d}u,\mathrm{d}v)}{\widehat{H}(u-,v-)} = \sum_{j=1}^n \frac{I\{X_{1j}\le t_1, X_{2j}\le t_2, \delta_{1j}=1, \delta_{2j}=1\}}{n\widehat{H}(X_{1j}-,X_{2j}-)}.
\end{aligned}$$

Fermanian (1997) proposed the following kernel estimators of $\lambda_{10}, \lambda_{01}, \lambda_{11}$,

$$
\begin{aligned}
\widehat{\lambda}_{10}(t_1,t_2) &= \int K_h^{(1)}(t_1-u)\widehat{\Lambda}_{10}(\mathrm{d}u,t_2) \\
&= \sum_{i=1}^{n} K_h^{(1)}(t_1-X_{1i})\frac{I\{\delta_{1i}=1, X_{2i}>t_2\}}{n\widehat{H}(X_{1i}-,t_2)}, \\
\widehat{\lambda}_{01}(t_1,t_2) &= \int K_h^{(1)}(t_2-v)\widehat{\Lambda}_{01}(t_1,\mathrm{d}v) \\
&= \sum_{i=1}^{n} K_h^{(1)}(t_2-X_{2i})\frac{I\{\delta_{2i}=1, X_{1i}>t_1\}}{n\widehat{H}(t_1,X_{2i}-)}, \\
\widehat{\lambda}_{11}(t_1,t_2) &= \int\int K_h^{(2)}(t_1-u,t_2-v)\widehat{\Lambda}_{11}(\mathrm{d}u,\mathrm{d}v) \\
&= \sum_{i=1}^{n} K_h^{(2)}(t_1-X_{1i},t_2-X_{2i})\frac{I\{\delta_{1i}=1, \delta_{2i}=1\}}{n\widehat{H}(X_{1i}-,X_{2i}-)},
\end{aligned}
$$

where we use $\int$ for $\int_{-\infty}^{\infty}$ and

$$K_h^{(1)}(u)=\frac{1}{h}K^{(1)}(u/h), \qquad K_h^{(2)}(u,v)=\frac{1}{h^2}K^{(2)}(u/h,v/h).$$

Let $K^{(1)}$ be a bounded Lebesgue-integrable kernel function with integral 1 defined on the real line R and$K^{(2)}$ be a bounded Lebesgue-integrable kernel function with integral 1 defined on the plane R^2.

In view of (1.1), for any (t_1,t_2) such that $\widehat{\lambda}_{10}(t_1,t_2)\widehat{\lambda}_{01}(t_1,t_2)>0$, we can define an estimator of θ by

$$\widehat{\theta}(t_1,t_2)=\frac{\widehat{\lambda}_{11}(t_1,t_2)}{\widehat{\lambda}_{10}(t_1,t_2)\widehat{\lambda}_{01}(t_1,t_2)}. \tag{2.1}$$

§3. Asymptotic Theory

Consider the subset $\tau=[0,\tau_1]\times[0,\tau_2]$ of R^2 and a positive ϵ such that $H(\tau_1+\epsilon,\tau_2+\epsilon)>0$. Select the bandwidth $(h_n)_{n\to\infty}$ such that $h_n\to 0$.

General assumptions on the kernel functions $K^{(1)}, K^{(2)}$ are:

(K1) $K^{(1)}$ is compactly supported with support $[-A,A]$, $K^{(2)}$ is compactly supported with support $[-A_1,A_1]\times[-A_2,A_2]$;

(K2) $K^{(i)}$is symmetric.

In the following, by saying a function is C^2 in a set, we mean it is twice continuous differentiable in the set.

Theorem 3.1 Suppose $K^{(1)}, K^{(2)}$ satisfy conditions (K1) and (K2), f is continuous on τ, for all $(t_1,t_2)\in\tau$, $\lambda_{10}(\cdot,t_2)$ is C^2 in a neighborhood of t_1, $\lambda_{01}(t_1,\cdot)$ is C^2 in a neighborhood of t_2 and $\lambda_{11}(\cdot,\cdot)$ is C^2 in a neighborhood of (t_1,t_2). If $\lambda_{01}(t_1,t_2)\lambda_{01}(t_1,t_2) > 0$ for $(t_1,t_2)\in\tau$ and $nh_n^2/\ln n\to\infty$, then $\widehat{\theta}(t_1,t_2)\to\theta$ in probability for all $(t_1,t_2)\in\tau$, where $\widehat{\theta}(t_1,t_2)$ is defined in (2.1).

Proof From Proposition 2.5 of Fermanian (1997), we can derive for $(t_1,t_2)\in\tau$,

$$\begin{aligned}\widehat{\lambda}_{10}(t_1,t_2)&\to\lambda_{10}(t_1,t_2), &&\text{in prob},\\ \widehat{\lambda}_{01}(t_1,t_2)&\to\lambda_{01}(t_1,t_2), &&\text{in prob},\\ \widehat{\lambda}_{11}(t_1,t_2)&\to\lambda_{11}(t_1,t_2), &&\text{in prob}.\end{aligned}$$

It follows that

$$\frac{\widehat{\lambda}_{11}}{\widehat{\lambda}_{01}\widehat{\lambda}_{10}}(t_1,t_2)\to\frac{\lambda_{11}}{\lambda_{01}\lambda_{10}}(t_1,t_2)=\theta,\qquad\text{in prob.}\qquad\Box$$

Theorem 3.2 Under the conditions of Theorem 3.1, if $H(t_1,t_2)$ is continuous on τ, and $nh_n^5=o(1)$, $nh_n^2/\ln n\to\infty$, then for any $(t_1,t_2)\in\tau$,

$$(nh_n^2)^{1/2}(\widehat{\theta}(t_1,t_2)-\theta)\Rightarrow N(0,\sigma^2(t_1,t_2)),$$

where $\widehat{\theta}(t_1,t_2)$ is defined in (2.1), "$\Rightarrow$" denotes convergence in distribution and

$$\sigma^2(t_1,t_2)=\frac{\theta^2}{H(t_1,t_2)\lambda_{11}(t_1,t_2)}\int\int[K^{(2)}(u,v)]^2\mathrm{d}u\mathrm{d}v. \tag{3.1}$$

Proof Firstly, we have

$$\begin{aligned}&\sqrt{nh_n^2}(\widehat{\theta}(t_1,t_2)-\theta)\\ =&\sqrt{nh_n^2}\Big(\frac{\widehat{\lambda}_{11}(t_1,t_2)}{\widehat{\lambda}_{10}(t_1,t_2)\widehat{\lambda}_{01}(t_1,t_2)}-\frac{\lambda_{11}(t_1,t_2)}{\lambda_{10}(t_1,t_2)\lambda_{01}(t_1,t_2)}\Big)\\ =&\sqrt{nh_n^2}\Big(\frac{\widehat{\lambda}_{11}}{\widehat{\lambda}_{10}\widehat{\lambda}_{01}}-\frac{\lambda_{11}}{\widehat{\lambda}_{10}\widehat{\lambda}_{01}}+\frac{\lambda_{11}}{\widehat{\lambda}_{10}\widehat{\lambda}_{01}}-\frac{\lambda_{11}}{\lambda_{10}\widehat{\lambda}_{01}}+\frac{\lambda_{11}}{\lambda_{10}\widehat{\lambda}_{01}}-\frac{\lambda_{11}}{\lambda_{10}\lambda_{01}}\Big)(t_1,t_2)\\ =&\frac{1}{\widehat{\lambda}_{10}(t_1,t_2)\widehat{\lambda}_{01}(t_1,t_2)}\big[\sqrt{nh_n^2}(\widehat{\lambda}_{11}(t_1,t_2)-\lambda_{11}(t_1,t_2))\big]\\ &+\sqrt{h_n}\frac{\lambda_{11}(t_1,t_2)}{\widehat{\lambda}_{01}(t_1,t_2)}\Big[\sqrt{nh_n}\Big(\frac{1}{\widehat{\lambda}_{10}(t_1,t_2)}-\frac{1}{\lambda_{10}(t_1,t_2)}\Big)\Big]\\ &+\sqrt{h_n}\frac{\lambda_{11}(t_1,t_2)}{\lambda_{10}(t_1,t_2)}\Big[\sqrt{nh_n}\Big(\frac{1}{\widehat{\lambda}_{01}(t_1,t_2)}-\frac{1}{\lambda_{01}(t_1,t_2)}\Big)\Big].\end{aligned} \tag{3.2}$$

Using Proposition 2.6 of Fermanian (1997), for any $(t_1,t_2)\in\tau$ we have

$$\sqrt{nh_n^2}(\widehat{\lambda}_{11}(t_1,t_2)-\lambda_{11}(t_1,t_2))\Rightarrow N(0,\Phi(t_1,t_2)),$$

where

$$\Phi(t_1,t_2)=\frac{\lambda_{11}(t_1,t_2)}{H(t_1,t_2)}\int\int[K^{(2)}(u,v)]^2\mathrm{d}u\mathrm{d}v.$$

Proposition 2.6 of Fermanian (1997) also gives the asymptotic property of $\widehat{\lambda}_{10}(t_1,t_2)$ and $\widehat{\lambda}_{01}(t_1,t_2)$, that is

$$\sqrt{nh_n}(\widehat{\lambda}_{10}(t_1,t_2)-\lambda_{10}(t_1,t_2)) \qquad \text{and} \qquad \sqrt{nh_n}(\widehat{\lambda}_{01}(t_1,t_2)-\lambda_{10}(t_1,t_2))$$

are asymptotically normal and hence they are bounded in probability.

Using Theorem 3.1, we can derive both

$$\sqrt{nh_n}\Big(\frac{1}{\widehat{\lambda}_{10}(t_1,t_2)}-\frac{1}{\lambda_{10}(t_1,t_2)}\Big) \qquad \text{and} \qquad \sqrt{nh_n}\Big(\frac{1}{\widehat{\lambda}_{01}(t_1,t_2)}-\frac{1}{\lambda_{10}(t_1,t_2)}\Big)$$

are bounded in probability. Using the fact that $h_n=o(1)$, we get that the last two terms of (3.2) are both $o_p(1)$.

At last, using the consistency of $\widehat{\lambda}_{10}(t_1,t_2)$ and $\widehat{\lambda}_{01}(t_1,t_2)$, we know that the first term of (3.2) converges in distribution to a normal random variable with mean zero and variance

$$\begin{aligned}\sigma^2(t_1,t_2) &= \frac{\Phi_{11}(t_1,t_2)}{\lambda_{10}^2(t_1,t_2)\lambda_{01}^2(t_1,t_2)}\\ &= \frac{1}{\lambda_{10}^2(t_1,t_2)\lambda_{01}^2(t_1,t_2)}\frac{\lambda_{11}(t_1,t_2)}{H(t_1,t_2)}\int\int[K^{(2)}(u,v)]^2\mathrm{d}u\mathrm{d}v\\ &= \frac{\theta^2}{H(t_1,t_2)\lambda_{11}(t_1,t_2)}\int\int[K^{(2)}(u,v)]^2\mathrm{d}u\mathrm{d}v. \qquad \Box\end{aligned}$$

Note that $\sigma^2(t_1,t_2)$ can be estimated by

$$\widehat{\sigma}^2(t_1,t_2)=\frac{\widehat{\theta}(t_1,t_2)^2}{\widehat{H}(t_1,t_2)\widehat{\lambda}_{11}(t_1,t_2)}\int\int[K^{(2)}(u,v)]^2\mathrm{d}u\mathrm{d}v.$$

The consistency of $\widehat{\sigma}^2(t_1,t_2)$ is proved by Theorem 3.1 and the consistency of $\widehat{H}(t_1,t_2)$. So to find a more efficient estimator, we only need to use$(\dot{t}_1,\dot{t}_2)$ such that $\widehat{H}(t_1,t_2)\widehat{\lambda}_{11}(t_1,t_2)$ hits its maximum value. In this case, the asymptotic variance of $\widehat{\theta}(\dot{t}_1,\dot{t}_2)$ may be smaller.

§4. Simulation Study

Simulation studies are conducted to examine the properties of the estimator $\widehat{\theta}(t_1,t_2)$. We choose $n=100$ and $n=200$ to conduct 2000 simulations at each of $\theta=1.2$, 1.4, 1.6, 1.8 and 2.0 with data generated from

$$F(t_1,t_2)=\big(e^{(\theta-1)t_1}+e^{(\theta-1)t_2}-1\big)^{-1/(\theta-1)}.$$

In addition, three types of censoring are explored:

1) No censoring.

2) Both survival variables are independently censored at the fixed time $C = 2$, that is we set $C_1 = C_2 = 2$ and this yields approximately 13.53% marginal censoring;

3) Two censoring variables are independently and identically distributed uniformly over $[0, 2.3]$, and this giving approximately 39.12% censorship on each marginal random variable.

The probability of both survival variables being censored increases with θ varying from 1.2 to 2.0. For example, for the first type of censoring, when $\theta = 1.2$, $\mathsf{P}(T_1 > C_1, T_2 > C_2) = 16.84\%$, for $\theta = 2.0$, the probability is 20.68%.

In the simulation study, the bandwidth is $h_n = n^{-1/6}$, the kernel is selected to be the Epanechnikov's kernel which was used in Fermanian (1997),

$$K^{(1)}(u) = \frac{3}{4\sqrt{5}}\Big(1 - \frac{u^2}{5}\Big) I\{u \in [-\sqrt{5}, \sqrt{5}]\},$$
$$K^{(2)}(u, v) = \Big(\frac{3}{4\sqrt{5}}\Big)^2 \Big(1 - \frac{u^2}{5}\Big)\Big(1 - \frac{v^2}{5}\Big) \cdot I\{u \in [-\sqrt{5}, \sqrt{5}]\} I\{v \in [-\sqrt{5}, \sqrt{5}]\}.$$

From the formula of $\sigma^2(t_1, t_2)$, we know that to find an estimator with smaller variance, one only need to find (t_1, t_2) which maximize $H(t_1, t_2)\lambda(t_1, t_2)$. However, $H(t_1, t_2) \cdot \lambda(t_1, t_2)$ is unknown, we use $(\dot{t}_1, \dot{t}_2)$ such that $\widehat{H}(t_1, t_2)\widehat{\lambda}_{11}(t_1, t_2)$ hits its maximum value. Since

$$H(t_1, t_2)\lambda(t_1, t_2) = G(t_1, t_2) f(t_1, t_2),$$

in our example, $G(t_1, t_2) f(t_1, t_2)$ is a decreasing function, so $\sigma^2(0, 0)$ is the maximum value, and $(\dot{t}_1, \dot{t}_2)$ is closing to $(0, 0)$. The kernel function is symmetrical, so for $\widehat{\theta}_1 \triangleq \widehat{\theta}(\dot{t}_1, \dot{t}_2)$, there is little observation since the left side of the point, and now only 60% of the sample can be used to compute the kernel estimates. Here we also estimate the association parameter at $(t_1^*, t_2^*) = (\text{median}\{X_{1j}\}, \text{median}\{X_{2j}\})$ and the estimator is denoted by $\widehat{\theta}_2$. The simulation results are displayed in Table 1.

In Table 1, "$(\widehat{\theta}_1 + \widehat{\theta}_2)/2$" refers the average of $\widehat{\theta}_1$ and $\widehat{\theta}_2$, "$\widetilde{\theta}$" stands for the two-stage estimator of Glidden (2000). "Mean" denotes the mean of relating estimator from the 2000 simulations. "SD" is the standard error, for $\widehat{\theta}_1$, "SD" equals $(nh_n^2)^{-1/2}\widehat{\sigma}(\dot{t}_1, \dot{t}_2)$, for $\widehat{\theta}_2$, "SD" is equal to $(nh_n^2)^{-1/2}\widehat{\sigma}(t_1^*, t_2^*)$. However, there isn't the theoretical formula for the variance of $(\widehat{\theta}_1 + \widehat{\theta}_2)/2$, we use sample standard deviation to compute its "SD".

It can be seen from Table 1 that the standard error is decreasing with n increasing, all estimators perform worse with the level of association increasing or the larger censoring proportion. The bias of $\widehat{\theta}_1$ is a little bit larger than $\widehat{\theta}_2$ and $\widetilde{\theta}$, the reason is $\widehat{\theta}_1$ using less sample as $(\dot{t}_1, \dot{t}_2)$ approaches $(0, 0)$. However, the sample standard deviation of $\widehat{\theta}_1$ is noticeably smaller than the others, especially when θ is relatively large. The estimators $\widehat{\theta}_2$

and $\widetilde{\theta}$ performs quite similar, when θ is small, their sample stand deviations differ little, it seems $\widetilde{\theta}$ performs a little better. However with θ increasing, especially when $\theta = 2$, the "SD" of $\widehat{\theta}_2$ is smaller than $\widetilde{\theta}$. So the table indicates that our estimator performs robustly.

Table 1 Summary of simulation results

θ	n	Cens.	$\widehat{\theta}_1$		$\widehat{\theta}_2$		$(\widehat{\theta}_1+\widehat{\theta}_2)/2$		$\widetilde{\theta}$	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
1.20	100	0.0%	1.25	0.10	1.19	0.18	1.22	0.12	1.21	0.14
		13.53%	1.26	0.12	1.22	0.19	1.22	0.13	1.21	0.14
		39.12%	1.15	0.15	1.22	0.20	1.23	0.15	1.22	0.18
	200	0.0%	1.23	0.08	1.21	0.11	1.21	0.08	1.20	0.10
		13.53%	1.22	0.10	1.21	0.13	1.21	0.10	1.21	0.10
		39.12%	1.18	0.15	1.19	0.15	1.22	0.13	1.21	0.13
1.40	100	0.0%	1.37	0.07	1.42	0.16	1.41	0.10	1.42	0.16
		13.53%	1.45	0.13	1.43	0.18	1.43	0.14	1.44	0.16
		39.12%	1.33	0.17	1.45	0.23	1.43	0.18	1.46	0.20
	200	0.0%	1.42	0.06	1.39	0.14	1.41	0.09	1.40	0.14
		13.53%	1.42	0.11	1.41	0.17	1.41	0.12	1.41	0.15
		39.12%	1.43	0.15	1.41	0.18	1.41	0.12	1.41	0.16
1.60	100	0.0%	1.57	0.08	1.58	0.18	1.58	0.12	1.62	0.20
		13.53%	1.63	0.15	1.62	0.23	1.62	0.17	1.63	0.22
		39.12%	1.66	0.17	1.64	0.26	1.64	0.20	1.65	0.27
	200	0.0%	1.62	0.06	1.60	0.14	1.61	0.09	1.60	0.15
		13.53%	1.58	0.13	1.61	0.18	1.61	0.14	1.61	0.17
		39.12%	1.64	0.16	1.59	0.21	1.61	0.17	1.61	0.19
1.80	100	0.0%	1.77	0.08	1.81	0.20	1.78	0.13	1.82	0.22
		13.53%	1.85	0.16	1.78	0.23	1.82	0.18	1.84	0.26
		39.12%	1.87	0.18	1.85	0.25	1.83	0.20	1.83	0.29
	200	0.0%	1.78	0.07	1.80	0.16	1.79	0.12	1.81	0.18
		13.53%	1.82	0.15	1.79	0.19	1.81	0.15	1.81	0.20
		39.12%	1.83	0.16	1.78	0.23	1.81	0.18	1.82	0.24
2.00	100	0.0%	2.05	0.12	1.97	0.22	2.02	0.16	1.97	0.24
		13.53%	2.05	0.17	2.03	0.25	2.03	0.19	2.04	0.27
		39.12%	2.07	0.20	2.04	0.31	2.04	0.24	2.04	0.32
	200	0.0%	2.05	0.14	2.02	0.18	2.01	0.14	1.98	0.19
		13.53%	2.04	0.16	1.97	0.23	2.03	0.18	2.03	0.24
		39.12%	2.04	0.18	2.03	0.25	2.03	0.20	2.03	0.26

The estimator $\widehat{\theta}(t_1,t_2)$ is a function of (t_1,t_2), so if we want, we can estimate the association parameter in many points. Considering the more information we use, perhaps the estimator will be more accurate, thus a weighted estimator $\sum\limits_{(t_1,t_2)} w(t_1,t_2)\widehat{\theta}(t_1,t_2)$ is a better choice. In Table 1, "$(\widehat{\theta}_1+\widehat{\theta}_2)/2$" is a simple weighted estimator, comparing the bias and standard error with $\widetilde{\theta}$, we can see that the simple weighted estimator performs better than $\widetilde{\theta}$, especially the difference between their "SD" becomes larger with θ increasing. However, the choice of weight $w(t_1,t_2)$ needs further investigation.

§5. Conclusion

This paper applies the results of Fermanian (1997) and derives a ratio estimator of the association parameter θ in Clayton model, then the new estimator is shown to be consistent and asymptotically normal. Simulation studies indicate that our estimator is effective.

Since $\widehat{\theta}(t_1,t_2)$ is a function of (t_1,t_2), we need to choose an appropriate point to give a good estimator. Simulation results tell us the appropriate point can be $(\dot{t}_1,\dot{t}_2)$, however if $(\dot{t}_1,\dot{t}_2)$ is closing to $(0,0)$, the bias of estimator is large, in the situation, one need to choose other points to estimate parameter in terms of the balance between bias and stand deviation. Simulation results tell us that a weighted estimator is a good choice, thought the weight needs further study.

Oakes (1989) introduced a local association parameter

$$\theta^*(t_1,t_2)=\frac{f(t_1,t_2)F(t_1,t_2)}{\displaystyle\int_{t_1}^{\infty} f(u,t_2)\mathrm{d}u\int_{t_2}^{\infty} f(t_1,v)\mathrm{d}v}.$$

Since $\theta^*(t_1,t_2)$ is a function of (t_1,t_2) rather than a constant, so our $\widehat{\theta}(t_1,t_2)$ can also be used to estimate $\theta^*(t_1,t_2)$.

In addition, as we know, the problem of estimating of bivariate survivor function $F(t_1,t_2)=\mathsf{P}(T_1>t_1,T_2>t_2)$ when the data are subject to censoring in either or both components is surprisingly difficult. Various proposals for the estimation have been made. However, the estimators suffer the drawback that the estimates are not survival functions because they are not monotone. But if the Clayton model holds, then we can use

$$\widehat{F}(t_1,t_2)=\Big[\Big(\frac{1}{\widehat{F}_1(t_1)}\Big)^{\widehat{\theta}(t_1,t_2)-1}+\Big(\frac{1}{\widehat{F}_2(t_2)}\Big)^{\widehat{\theta}(t_1,t_2)-1}-1\Big]^{-1/[\widehat{\theta}(t_1,t_2)-1]}$$

as an estimator of $F(t_1,t_2)$, where $\widehat{F}_1(t_1),\widehat{F}_2(t_2)$ are the Kaplan-Meier estimators of the marginal survival functions $F_1(t_1)$ and $F_2(t_2)$.

References

[1] Clayton, D.G., A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, **65(1)**(1978), 141–151.

[2] Emura, T., Lin, C.W. and Wang, W.J., A goodness-of-fit test for Archimedean copula models in the presence of right censoring, *Computational Statistics & Data Analysis*, **54(12)**(2010), 3033–3043.

[3] Fermanian, J.D., Multivariate hazard rates under random censorship, *Journal of Multivariate Analysis*, **62(2)**(1997), 273–309.

[4] Genest, C., Ghoudi, K. and Rivest, L.-P., A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, **82(3)**(1995), 543–552.

[5] Glidden, D.V., A two-stage estimator of the dependence parameter for the Clayton-Oakes model, *Lifetime Data Analysis*, **6(2)**(2000), 141–156.

[6] Ghosh, D., Semiparametric inference for surrogate endpoints with bivariate censored data, *Biometrics*, **64(1)**(2008), 149–156.

[7] Nan, B., Lin, X.H., Lisabeth, L.D. and Harlow, S.D., Piecewise constant cross-ratio estimation for association of age at a marker event and age at menopause, *Journal of the American Statistical Association*, **101(473)**(2006), 65–77.

[8] Oakes, D., Semiparametric inference in a model for association in bivanate survival data, *Biometrika*, **73(2)**(1986), 353–361.

[9] Oakes, D., Bivariate survival models induced by frailties, *Journal of the American Statistical Association*, **84(406)**(1989), 487–493.

[10] Shih, J.H. and Louis, T.A., Inferences on the association parameter in copula models for bivariate survival data, *Biometrics*, **51(4)**(1995), 1384–1399.

一种估计Clayton模型中关联参数的方法

张巧真

(南开大学数学科学学院统计系及核心数学与组合数学实验室, 天津, 300071)

何书元

(首都师范大学数学科学学院, 北京, 100048)

在右删失情形下, 基于二元风险函数的核估计, 我们对Clayton模型中的关联参数给出了一种新的估计方法. 新的估计量具有相合性和渐近分布, 随机模拟也显示这种估计方法是非常有效的.

关键词: 右删失, Clayton模型, 关联参数, 渐近理论.

学科分类号: O212.7.