Journal of Multivariate Analysis 132 (2014) 229–241

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

An improved nonparametric estimator of sub-distribution function for bivariate competing risk models

Takeshi Emura [a,*], Fan-Hsuan Kao [a], Hirofumi Michimae [b]

[a] Graduate Institute of Statistics, National Central University, Taiwan
[b] School of Pharmacy, Department of Clinical Medicine (Biostatistics), Kitasato University, Japan

ARTICLE INFO

Article history:
Received 18 July 2013
Available online 6 September 2014

AMS subject classifications:
62N01
62N02
62G05
62G20
62C05

Keywords:
Bivariate survival function
Right censoring
Survival analysis

ABSTRACT

For competing risks data, it is of interest to estimate the sub-distribution function of a particular failure event, which is the failure probability in the presence of competing risks. However, if multiple failure events per subject are available, estimation procedures become challenging even for the bivariate case. In this paper, we consider nonparametric estimation of a bivariate sub-distribution function, which has been discussed in the related literature. Adopting a decision-theoretic approach, we propose a new nonparametric estimator which improves upon an existing estimator. We show theoretically and numerically that the proposed estimator has smaller mean square error than the existing one. The consistency of the proposed estimator is also established. The usefulness of the estimator is illustrated by the salamander data and mouse data.

1. Introduction

Statistical analysis of competing risks data is common in biology, where individuals experience multiple failure causes. For instance, larvae grown in a cage may experience either metamorphosis or death, whichever comes first [15]. These two failure causes are mutually exclusive in that each larva exhibits only one of the two causes at the time of failure. This type of data is popular, especially in biomedical research involving human and animal subjects [2]. Competing risks models are used to analyze such data. An overview of competing risks data analysis is referred to Crowder [8] and Bakoyannis and Touloumi [5].

In competing risks data analysis, the sub-distribution function plays a fundamental role. Let T be a failure time and $C \in \{1, 2, \ldots, \gamma\}$ be the failure cause for γ distinct causes. The *sub-distribution function* (also known as cumulative incidence function) is defined as

$$F_j(t) = \Pr(T \le t, C = j), \quad j = 1, 2, \ldots, \gamma.$$

This is the proportion of failure events occurring due to cause j before time t. The sub-distribution function is easy to interpret and is often the target for estimation [8,5,13].

In applications, bivariate competing risks arise naturally. For instance, a pair of larvae in an experimental cage shares unobserved environmental or genetic factors [15]. In analysis of such data, the univariate competing risks models need to be generalized to bivariate models.

* Corresponding author.
E-mail addresses: emura@stat.ncu.edu.tw, takeshiemura@gmail.com (T. Emura).

http://dx.doi.org/10.1016/j.jmva.2014.08.009
0047-259X/

The bivariate competing risk models have been recently considered by many authors [4,19,17,18,20]. Under the bivariate competing risks models, the target of estimation is the *bivariate sub-distribution function* (formally defined in Section 2). Several nonparametric estimators have been proposed under various censoring and truncation schemes. Antony and Sankaran [4] and Sankaran et al. [19] developed nonparametric estimators under right-censoring. Sankaran and Antony [18] considered a similar problem where the censoring times are missing. Sankaran and Antony [17] proposed a nonparametric estimator under left-truncation and right-censoring. Shen [20] considered nonparametric estimation under double censoring.

This paper considers nonparametric estimation of the bivariate sub-distribution functions under right censoring as in [4,19]. Note that nonparametric estimation under bivariate competing risks is much more challenging than its univariate counterpart. Especially in small sample sizes, the estimator of Sankaran et al. [19] will generally be a crude step function and will have a large mean squared error (MSE). In light of this problem, the main objective of this paper is to propose a new nonparametric estimator that aims to improve upon the existing estimator. The proposed estimator not only reduces the MSE but also smoothes out the crude step function estimator in some degree.

The paper is organized as follow. Section 2 introduces basic notations and the estimator of Sankaran et al. [19]. Section 3 proposes a new estimator for the bivariate sub-distribution function. Section 4 verifies the consistency of the proposed estimator. Section 5 presents simulations comparing the proposed method with the existing one. Section 6 analyzes the mouse data and the salamander data. Section 7 concludes the paper.

2. Preliminary

This section defines basic notations for bivariate competing risks models and then introduces the nonparametric estimator of Sankaran et al. [19] for estimating a bivariate sub-distribution function.

Let $S(t_1, t_2) = \Pr(T_1 > t_1, T_2 > t_2)$ be the survivor function of bivariate failure times (T_1, T_2). Also, let $(C_1, C_2) \in \{1, 2, \ldots, \gamma_1\} \times \{1, 2, \ldots, \gamma_2\}$ be the corresponding bivariate failure causes. For $(i, j) \in \{1, 2, \ldots, \gamma_1\} \times \{1, 2, \ldots, \gamma_2\}$, the cause-specific hazard is

$$\Lambda_{ij}(dt_1, dt_2) = \frac{\Pr(T_1 \in dt_1, T_2 \in dt_2, C_1 = i, C_2 = j)}{\Pr(T_1 \geq t_1, T_2 \geq t_2)}.$$

Also, the sub-distribution function is

$$F_{ij}(t_1, t_2) = \Pr(T_1 \leq t_1, T_2 \leq t_2, C_1 = i, C_2 = j).$$

The cause-specific hazard and the sub-distribution functions are related through

$$F_{ij}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} S(u^-, v^-) \Lambda_{ij}(du, dv). \tag{1}$$

The above identity is useful for estimating the sub-distribution F_{ij} under right-censoring. If (T_1, T_2) are censored by a pair of independent censoring times (Z_1, Z_2), one observes (Y_1, Y_2) and (δ_1, δ_2), where $Y_k = \min(T_k, Z_k)$ and $\delta_k = \mathbf{I}(T_k = Y_k)$ for $k = 1, 2$, where $\mathbf{I}(\cdot)$ is the indicator function. If $\delta_k = 0$, then we set $C_k = 0$ since the value of C_k is not available. If $H(t_1, t_2) \equiv \Pr(Y_1 > t_1, Y_2 > t_2) > 0$, Eq. (1) becomes

$$F_{ij}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \frac{S(u^-, v^-) F^*_{ij}(du, dv)}{H(u^-, v^-)}, \quad i = 1, 2, \ldots, \gamma_1, j = 1, 2, \ldots, \gamma_2, \tag{2}$$

where

$$F^*_{ij}(t_1, t_2) = \Pr(T_1 \leq t_1, T_2 \leq t_2, \delta_1 = 1, \delta_2 = 1, C_1 = i, C_2 = j).$$

Sankaran et al. [19] used Eq. (2) to estimate F_{ij} based on observations (Y_{1u}, Y_{2u}), (C_{1u}, C_{2u}), and $(\delta_{1u}, \delta_{2u})$, $u = 1, 2, \ldots, n$, which are i.i.d. replications of (Y_1, Y_2), (C_1, C_2), and (δ_1, δ_2). They consider an estimator of $H(t_1, t_2)$ as

$$\hat{H}(t_1, t_2) = \frac{1}{n} \sum_{u=1}^{n} \mathbf{I}(Y_{1u} > t_1, Y_{2u} > t_2),$$

and an estimate of $F^*_{ij}(t_1, t_2)$ as

$$\hat{F}^*_{ij}(t_1, t_2) = \frac{1}{n} \sum_{u=1}^{n} \mathbf{I}(Y_{1u} \leq t_1, Y_{2u} \leq t_2, \delta_{1u} = 1, \delta_{2u} = 1, C_{1u} = i, C_{2u} = j).$$

Under $\hat{H}(t_1, t_2) > 0$, they obtain the nonparametric estimator for $F_{ij}(t_1, t_2)$ as

$$\hat{F}_{ij}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \frac{\hat{S}(u^-, v^-) \hat{F}^*_{ij}(du, dv)}{\hat{H}(u^-, v^-)}, \quad i = 1, 2, \ldots, \gamma_1, j = 1, 2, \ldots, \gamma_2. \tag{3}$$

Sankaran et al. [19] proposed to apply the Dabrowska estimator [9] for $\hat{S}$. Other estimators are also available, such as the estimators of Prentice and Cai [16] and Wang and Wells [22]. The strong consistency and weak convergence for $\hat{F}_{ij}$ are studied by [19].

3. Proposed estimator

3.1. The independence estimator

Before deriving the proposed estimator, we first consider a simplified estimator under the independence assumption. The idea is motivated by Wang and Wells [22].

If (T_1, C_1, Z_1) and (T_2, C_2, Z_2) are independent, it is easy to show the identity

$$F_{ij}(t_1, t_2) = \left\{ \int_0^{t_1} S_1(u^-)\Lambda_{1i}(du) \right\} \times \left\{ \int_0^{t_2} S_2(v^-)\Lambda_{2j}(dv) \right\}, \quad i = 1, 2, \ldots, \gamma_1, j = 1, 2, \ldots, \gamma_2$$

where $S_k(u) = \Pr(T_k > u)$ and $\Lambda_{ki}(du) = \Pr(T_k \in du, C_k = i | T_k \geq u)$ for $k = 1, 2$. We define the "independence estimator" of $F_{ij}(t_1, t_2)$ as

$$\hat{F}^0_{ij}(t_1, t_2) = \hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2), \tag{4}$$

where $\hat{F}_{1i}(t_1)$ and $\hat{F}_{2j}(t_2)$ are the two univariate estimators of the form

$$\hat{F}_{ki}(t_k) = \int_0^{t_k} \frac{\hat{S}_k(u^-)}{\hat{H}_k(u^-)} \hat{F}^*_{ki}(du), \quad i = 1, 2, \ldots, \gamma_k, k = 1, 2,$$

where $\hat{F}^*_{ki}(u) = \sum_{u=1}^n \mathbf{I}(Y_{ku} \leq u, \delta_{ku} = 1, C_{ku} = i)/n$, $\hat{H}_k(u^-) = \sum_{u=1}^n \mathbf{I}(Y_{ku} \geq u)/n$, and $\hat{S}_k(\cdot)$ is the Kaplan–Meier (KM) estimator of T_k.

Obviously, the estimator in Eq. (4) is inconsistent except for the independence case. However, estimation of two univariate function in Eq. (4) is much easier than estimation of a bivariate function in the existing estimator of Sankaran et al. [19]. Also, due to the trade-off between bias and variance, the independence estimator often has smaller MSE than the existing estimator. In the following, we take advantage of the independence estimator to refine the estimator of Sankaran et al. [19].

3.2. The proposed estimator

We propose a new nonparametric estimator which refines the estimator of Sankaran et al. [19] for estimating the bivariate sub-distribution function.

We define a class of estimators that combine the existing estimator and independence estimator as follow:

$$\hat{F}^a_{ij}(t_1, t_2) = a\,\hat{F}_{ij}(t_1, t_2) + (1 - a)\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2), \quad a \in [0, 1].$$

If $a = 1$, this is the estimator of Sankaran et al. [19] in Eq. (3); if $a = 0$ this is the independence estimator in Eq. (4). Hence, the class of estimators includes the two estimators as special cases.

Now we consider how to choose the optimal value of a. In the statistical decision theory, one often searches the estimator that minimizes the MSE within a class of shrinkage estimators (e.g., Khan [12]; Wencheko and Wijekoon [23]). In estimating a bivariate survival function, this approach was taken by Akritas and Keilegom [1]. By adopting this approach, we find a that archives the smallest MSE. The MSE is calculated as

$$\begin{aligned} \mathrm{MSE}[\hat{F}^a_{ij}(t_1, t_2)] &= E[\hat{F}^a_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)]^2 \\ &= a^2\, E[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)]^2 + (1 - a)^2 E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]^2 \\ &\quad + 2a\,(1 - a)E\{[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)][\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]\}. \end{aligned}$$

The minimizer is obtained by solving

$$0 = \frac{d}{da}\mathrm{MSE}[\hat{F}^a_{ij}(t_1, t_2)] = 2\{a(x + y - 2z) + z - y\}.$$

This results in

$$a^*\,(t_1, t_2) = \underset{a}{\mathrm{argmin}}\,\mathrm{MSE}[\hat{F}^a_{ij}(t_1, t_2)] = \frac{y\,(t_1, t_2) - z(t_1, t_2)}{x(t_1, t_2) + y\,(t_1, t_2) - 2z(t_1, t_2)} \tag{5}$$

where

$$\begin{aligned} x &= x(t_1, t_2) \equiv E[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)]^2, \\ y &= y(t_1, t_2) \equiv E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]^2, \\ z &= z(t_1, t_2) \equiv E\{\,[\,\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)][\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]\}. \end{aligned}$$

The next theorem shows that, at the optimal value of a^*, $\mathrm{MSE}[\hat{F}^{a^*}_{ij}(t_1, t_2)]$ is strictly smaller than both $\mathrm{MSE}[\hat{F}_{ij}(t_1, t_2)]$ and $\mathrm{MSE}[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2)]$.

Theorem 1. *If $x(t_1, t_2) \neq z(t_1, t_2)$ and $y(t_1, t_2) \neq z(t_1, t_2)$ then $0 < a^*\,(t_1, t_2) < 1$ and*

$$\mathrm{MSE}[\hat{F}^{a^*}_{ij}(t_1, t_2)] < \min\{\mathrm{MSE}[\hat{F}_{ij}(t_1, t_2)], \mathrm{MSE}[\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2)]\}.$$

Proof. Note that the denominator of $a^*(t_1, t_2)$ in Eq. (5) is well-defined since

$$x(t_1, t_2) + y(t_1, t_2) - 2z(t_1, t_2) = E[\hat{F}_{ij}(t_1, t_2) - \hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2)]^2 > 0.$$

Also, the MSE function is strictly convex since

$$\frac{d^2}{da^2}\mathrm{MSE}[\hat{F}_{ij}^{a}(t_1, t_2)] = 2(x + y - 2z) > 0.$$

Hence, $a^*(t_1, t_2)$ is the unique minimizer. One can verify $0 < a^*(t_1, t_2) < 1$ under the assumptions $x(t_1, t_2) \neq z(t_1, t_2)$ and $y(t_1, t_2) \neq z(t_1, t_2)$ since

$$0 \neq a^*(t_1, t_2) \Leftrightarrow 0 \neq \frac{y(t_1, t_2) - z(t_1, t_2)}{x(t_1, t_2) + y(t_1, t_2) - 2z(t_1, t_2)} \Leftrightarrow 0 \neq y(t_1, t_2) - z(t_1, t_2),$$

$$a^*(t_1, t_2) \neq 1 \Leftrightarrow 0 \neq \frac{x(t_1, t_2) - z(t_1, t_2)}{x(t_1, t_2) + y(t_1, t_2) - 2z(t_1, t_2)} \Leftrightarrow 0 \neq x(t_1, t_2) - z(t_1, t_2). \quad \square$$

The conditions of Theorem 1 exclude some extreme cases. When (t_1, t_2) are very close to (0, 0), one may have $\hat{F}_{ij}(t_1, t_2) = \hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) = 0$ with probability one. In such cases, the conditions do not hold. The same is true if (t_1, t_2) are too large. In usual cases, $x(t_1, t_2)$, $y(t_1, t_2)$ and $z(t_1, t_2)$ cannot be equal since $\hat{F}_{ij}$ and $\hat{F}_{1i}\hat{F}_{2j}$ have different distributions.

When applying the proposed estimator to real data, one needs to estimate $a^*(t_1, t_2)$. We suggest nonparametric bootstrap to estimate $x(t_1, t_2)$, $y(t_1, t_2)$ and $z(t_1, t_2)$ as in the manner of Akritas and Keilegom [1], Sankaran and Antony [17] and Shen [20].

The bootstrap method for estimating $a^*(t_1, t_2)$:

Step 1. Calculate $\hat{F}_{ij}(t_1, t_2)$ from observed data.

Step 2. Let $\{(Y_{1u}^{*(b)}, Y_{2u}^{*(b)}, \delta_{1u}^{*(b)}, \delta_{2u}^{*(b)}, C_{1u}^{*(b)}, C_{2u}^{*(b)}) : u = 1, 2, \ldots, n)\}$ be a random sample with replacement from the observed data $\{(Y_{1u}, Y_{2u}, \delta_{1u}, \delta_{2u}, C_{1u}, C_{2u}) : u = 1, 2, \ldots, n\}$ for $b = 1, 2, \ldots, B$, where B is the bootstrap number.

Step 3. Calculate $\hat{F}_{ij}^{*(b)}(t_1, t_2)$ and $\hat{F}_{1i}^{*(b)}(t_1)\hat{F}_{2j}^{*(b)}(t_2)$ based on the resampled data $\{(Y_{1,u}^{*(b)}, Y_{2,u}^{*(b)}, \delta_{1,u}^{*(b)}, \delta_{2,u}^{*(b)}, C_{1,u}^{*(b)}, C_{2,u}^{*(b)}) : u = 1, 2, \ldots, n)\}$, $b = 1, 2, \ldots, B$, and then compute the bootstrap approximations to $x(t_1, t_2)$, $y(t_1, t_2)$ and $z(t_1, t_2)$, defined as

$$\hat{x}(t_1, t_2) = \frac{1}{B}\sum_{b=1}^{B}[\hat{F}_{ij}^{*(b)}(t_1, t_2) - \hat{F}_{ij}(t_1, t_2)]^2,$$

$$\hat{y}(t_1, t_2) = \frac{1}{B}\sum_{b=1}^{B}[\hat{F}_{1i}^{*(b)}(t_1)\hat{F}_{2j}^{*(b)}(t_2) - \hat{F}_{ij}(t_1, t_2)]^2,$$

$$\hat{z}(t_1, t_2) = \frac{1}{B}\sum_{b=1}^{B}[\hat{F}_{1i}^{*(b)}(t_1)\hat{F}_{2j}^{*(b)}(t_2) - \hat{F}_{ij}(t_1, t_2)][\hat{F}_{ij}^{*(b)}(t_1, t_2) - \hat{F}_{ij}(t_1, t_2)].$$

Then, we obtain the estimator of $a^*(t_1, t_2)$ as

$$\hat{a}(t_1, t_2) = \frac{\hat{y}(t_1, t_2) - \hat{z}(t_1, t_2)}{\hat{x}(t_1, t_2) + \hat{y}(t_1, t_2) - 2\hat{z}(t_1, t_2)}.$$

Remark 1. By simulations not shown here, we have checked that $\hat{x}(t_1, t_2)$, $\hat{y}(t_1, t_2)$ and $\hat{z}(t_1, t_2)$ are all good estimators of $x(t_1, t_2)$, $y(t_1, t_2)$ and $z(t_1, t_2)$, respectively.

Remark 2. The case of $\hat{a}(t_1, t_2) < 0$ or $\hat{a}(t_1, t_2) > 1$ could occur in small samples. In such a case, one can set $\hat{a}(t_1, t_2) = 0$ or $\hat{a}(t_1, t_2) = 1$, respectively.

4. Asymptotic theory

For fixed (t_1, t_2), we shall prove the consistency of the proposed estimator

$$\hat{F}_{ij}^{\hat{a}}(t_1, t_2) = \hat{a}(t_1, t_2)\hat{F}_{ij}(t_1, t_2) + \{1 - \hat{a}(t_1, t_2)\}\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2).$$

The proof differs between two cases:

(i) (T_1, C_1, Z_1) and (T_2, C_2, Z_2) are independent;
(ii) (T_1, C_1, Z_1) and (T_2, C_2, Z_2) are not independent.

Under the case (i), both $\hat{F}_{ij}$ and $\hat{F}_{1i}\hat{F}_{2j}$ are consistent for F_{ij} under the conditions of Theorem 2 below. Then, the consistency of their middle point $\hat{F}_{ij}^{\hat{a}}$ follows immediately. In this case, we even do not need to establish the consistency of the bootstrap estimator $\hat{a}$.

The case (ii) needs more careful study. As a special case of a more general theory of product-limit estimator (Theorem IV.4.1, Andersen et al. [3]), we have

Lemma 1. *For fixed t_k with $G_k(t_k) > 0$, where $G_k(t) = \Pr(Z_k > t)$, as $n \to \infty$,*

$$\hat{F}_{kj}(t_k) \xrightarrow{P} F_{kj}(t_k) \quad \text{for } k = 1, 2.$$

As a direct consequence from Lemma 1 and the fact

$$G(t_1, t_2) \equiv \Pr(Z_1 > t_1, Z_2 > t_2) \le \min\{G_1(t_1), G_2(t_2)\},$$

we also have the following Lemma:

Lemma 2. *For fixed (t_1, t_2) with $G(t_1, t_2) > 0$, as $n \to \infty$,*

$$\hat{F}_{1i}(t_1)\hat{F}_{2i}(t_2) \xrightarrow{P} F_{1i}(t_1)F_{2i}(t_2).$$

The next theorem establishes the convergence of $a^*(t_1, t_2)$ defined in Eq. (5).

Theorem 2. *Suppose that (t_1, t_2) satisfies $G(t_1, t_2) > 0$, and that (T_1, C_1, Z_1) and (T_2, C_2, Z_2) are not independent, Then, as $n \to \infty$,*

$$a^*(t_1, t_2) \to 1.$$

Proof. Sankaran et al. [19] gave the consistency $\hat{F}_{ij}(t_1, t_2) \xrightarrow{P} F_{ij}(t_1, t_2)$ under the condition of Theorem 2. One also verifies $\hat{F}_{1i}(t_1)\hat{F}_{2i}(t_2) \xrightarrow{P} F_{1i}(t_1)F_{2j}(t_2)$ in Lemma 2. Note that these two estimators are uniformly bounded sequences on [0, 1]. Hence, these two estimators converge in L_p for $0 < p < \infty$ (p. 71 of Chung [6]). To complete the proof the theorem, we need the following three claims:

Claim 1. $\lim_{n\to\infty} E[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)]^2 = 0$.

Proof. The proof follows by the definition of convergence in L_P at $p = 2$. □

Claim 2. $\lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]^2 = [F_{1i}(t_1)F_{2j}(t_2) - F_{ij}(t_1, t_2)]^2$.

Proof. Since $\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) \xrightarrow{L_P} F_i(t_1)F_j(t_2)$, $0 < p < \infty$, it follows that

$$\lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{1i}(t_1)F_{2j}(t_2)] = 0,$$

$$\lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{1i}(t_1)F_{2j}(t_2)]^2 = 0.$$

Then we obtain

$$\begin{aligned}\lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]^2 &= \lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{1i}(t_1)F_{2j}(t_2)]^2 + [F_{1i}(t_1)\,F_{2j}(t_2) - F_{ij}(t_1, t_2)]^2 \\ &\quad + 2[F_{1i}(t_1)\,F_{2j}(t_2) - F_{ij}(t_1, t_2)] \lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) - F_i(t_1)F_j(t_2)] \\ &= [F_{1i}(t_1)F_{2j}(t_2) - F_{ij}(t_1, t_2)]^2. \quad \square\end{aligned}$$

Claim 3. $\lim_{n\to\infty} E\{[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)]\,[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]\} = 0$.

Proof. Since $\hat{F}_{ij}(t_1, t_2) \xrightarrow{P} F_{ij}(t_1, t_2)$ and $\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) \xrightarrow{P} F_i(t_1)F_j(t_2)$ holds under the conditions of Theorem 2, Slutsky's theorem shows

$$\hat{F}_{ij}(t_1, t_2)\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2) \xrightarrow{P} F_{ij}(t_1, t_2)F_{1i}(t_1)F_{2j}(t_2).$$

Note that the sequence $\{\hat{F}_{ij}(t_1, t_2)\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2)\}$ is also uniformly bounded, the above convergence implies the convergence in L_P at $p = 2$. Then one can verify that

$$\begin{aligned}&\lim_{n\to\infty} E\{[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)][\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]\} \\ &= \lim_{n\to\infty} E[\hat{F}_{ij}(t_1, t_2)\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2)] - F_{ij}(t_1, t_2) \lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\,\hat{F}_{2j}(t_2)] - F_{ij}(t_1, t_2) \lim_{n\to\infty} E[\hat{F}_{ij}(t_1, t_2)] + F_{ij}(t_1, t_2)^2\end{aligned}$$

$$= \lim_{n\to\infty} E[\hat{F}_{ij}(t_1, t_2)\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2)] - F_{ij}(t_1, t_2)^2 - F_{ij}(t_1, t_2)F_{1i}(t_1)F_{2j}(t_2) + F_{ij}(t_1, t_2)^2$$
$$= \lim_{n\to\infty} E[\hat{F}_{ij}(t_1, t_2)\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)F_{1i}(t_1)F_{2j}(t_2)]$$
$$= 0. \quad \square$$

We use Claims 1–3 to conclude

$$\lim_{n\to\infty} x(t_1, t_2) = \lim_{n\to\infty} E[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)]^2 = 0,$$
$$\lim_{n\to\infty} y(t_1, t_2) = \lim_{n\to\infty} E[\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)]^2 = [F_{1i}(t_1)F_{2j}(t_2) - F_{ij}(t_1, t_2)]^2,$$
$$\lim_{n\to\infty} z(t_1, t_2) = \lim_{n\to\infty} E\,[\hat{F}_{ij}(t_1, t_2) - F_{ij}(t_1, t_2)][\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) - F_{ij}(t_1, t_2)] = 0.$$

By the definition of $a^*(t_1, t_2)$,

$$\lim_{n\to\infty} a^*(t_1, t_2) = \lim_{n\to\infty} \frac{y(t_1, t_2) - z(t_1, t_2)}{x(t_1, t_2) + y(t_1, t_2) - 2z(t_1, t_2)} = \frac{\lim_{n\to\infty} y(t_1, t_2) - \lim_{n\to\infty} z(t_1, t_2)}{\lim_{n\to\infty} x(t_1, t_2) + \lim_{n\to\infty} y(t_1, t_2) - \lim_{n\to\infty} 2z(t_1, t_2)} = 1.$$

The proof of Theorem 2 is complete. $\square$

Note that $a^*(t_1, t_2)$ needs to be estimated by the bootstrap estimator $\hat{a}(t_1, t_2)$. The formal asymptotic results for the bootstrap estimator $\hat{a}$ are fairly difficult to derive. In general, empirical process techniques are used to study the bootstrap consistency [21]. Since our bootstrap uses the same algorithm as Dabrowska [10], the consistency may follow arguments similar to pp. 313–314 of Dabrowska [10].

Conjecture 1. *Suppose that* (t_1, t_2) *satisfies* $G(t_1, t_2) > 0$. *Then, as* $n \to \infty$,

$$\hat{a}(t_1, t_2) \xrightarrow{P} a^*(t_1, t_2).$$

We do not have a proof. Instead, we use simulations to verify that $E[\hat{a}(t_1, t_2)]$ goes to $a^*(t_1, t_2)$ and $\mathrm{MSE}[\hat{a}(t_1, t_2)]$ goes to zero as $n \to \infty$ (see Section 5).

Theorem 3. *Suppose that* (t_1, t_2) *satisfies* $G(t_1, t_2) > 0$. *Then, as* $n \to \infty$,

$$\hat{F}^{\hat{a}}_{ij}(t_1, t_2) = \hat{a}\,(t_1, t_2)\hat{F}_{ij}(t_1, t_2) + \{1 - \hat{a}(t_1, t_2)\}\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2) \xrightarrow{P} F_{ij}(t_1, t_2).$$

Proof. We only need to consider the case (ii) (T_1, C_1, Z_1) and (T_2, C_2, Z_2) are not independent. With this case, by Theorem 2 and Conjecture 1, one has $\hat{a}(t_1, t_2) \xrightarrow{P} 1$. The proof completes by Slutsky's theorem with $\hat{F}_{ij}(t_1, t_2) \xrightarrow{P} F_{ij}(t_1, t_2)$. $\square$

5. Simulations

Simulations are conducted to study the performance of the proposed estimator and to compare it with the estimator of Sankaran et al. [19] and the independence estimator.

5.1. Simulation design

We carry out a series of 500 simulations with $n = 100$ based on data generated from the Clayton model [7]:

$$\Pr(T_1 \le t_1, T_2 \le t_2) = \max\left[\{F_1(t_1)^{-(\alpha-1)} + F_2(t_2)^{-(\alpha-1)} - 1\}^{\frac{-1}{\alpha-1}}, 0\right], \quad \alpha \in [0, \infty) \setminus \{1\}.$$

When $\alpha \in [0, 1)$, T_1 and T_2 have negative correlation; when $\alpha \in (1, \infty]$, T_1 and T_2 have positive correlation. The Kendall's tau between T_1 and T_2 is $\tau = (\alpha-1)/(\alpha+1)$. The marginal distributions are the unit exponential distribution $F_k(t) = 1-e^{-t}$ for $k = 1, 2$. The censoring times Z_1 and Z_2 independently follow the unit exponential distributions. In this way, censoring percentages for T_1 and T_2 are both 50%. The causes C_1 and C_2 are independent and take values 1 or 2 with equal probability. The number of bootstrap replicates to obtain the estimator $\hat{a}$ is taken to be $B = 500$.

We have done the same set of simulations for the lognormal distribution $F_k(t) = \Phi(\ln t)$ for $k = 1, 2$ where Φ is the distribution function of the standard normal distribution. The results parallel the case of exponential distributions and therefore not shown.

Table 1
Simulation results for estimating $F_{11}(t_1, t_2)$ based on four different estimators with $n = 100$.

(t_1, t_2)	$F_{11}(t_1, t_2)$	Estimator	$E(\hat{F}_{11})$	MSE$(\hat{F}_{11})$	a^*	$E(\hat{a})$
(i) Independent case (i.e., $\alpha = 1$)						
(1, 2)	0.1366	$\hat{F}_{11}(t_1, t_2)$	0.1290	0.00438		
		$\hat{F}_1(t_1)\hat{F}_1(t_2)$	0.1219	0.00113		
		$\hat{F}_{11}^{a^*}(t_1, t_2)$	0.1221	0.00113	0.023	
		$\hat{F}_{11}^{\hat{a}}(t_1, t_2)$	0.1213	0.00277		0.360
(0.5, 0.5)	0.0387	$\hat{F}_{11}(t_1, t_2)$	0.0386	0.00060		
		$\hat{F}_1(t_1)\hat{F}_1(t_2)$	0.0376	0.00016		
		$\hat{F}_{11}^{a^*}(t_1, t_2)$	0.0376	0.00016	0.000	
		$\hat{F}_{11}^{\hat{a}}(t_1, t_2)$	0.0360	0.00038		0.388
(ii) Dependent case with $\alpha = 5 (\tau = 0.667)$						
(1, 2)	0.1534	$\hat{F}_{11}(t_1, t_2)$	0.1514	0.00457		
		$\hat{F}_1(t_1)\hat{F}_1(t_2)$	0.1221	0.00200		
		$\hat{F}_{11}^{a^*}(t_1, t_2)$	0.1274	0.00184	0.181	
		$\hat{F}_{11}^{\hat{a}}(t_1, t_2)$	0.1348	0.00290		0.372
(0.5, 0.5)	0.0830	$\hat{F}_{11}(t_1, t_2)$	0.0816	0.00117		
		$\hat{F}_1(t_1)\hat{F}_1(t_2)$	0.0372	0.00228		
		$\hat{F}_{11}^{a^*}(t_1, t_2)$	0.0675	0.00093	0.683	
		$\hat{F}_{11}^{\hat{a}}(t_1, t_2)$	0.0688	0.00115		0.604

5.2. Simulation results

Table 1 summarizes the results. In all cases, the proposed estimator using the true a^* shows the best performance in terms of the MSE, as supported by Theorem 1. When T_1 and T_2 are independent (i.e., $\alpha = 1$), the proposed estimator using the true a^* performs the best, but the performance is nearly identical to the independence estimator. When a^* is estimated by $\hat{a}$, the performance is no longer the best, but still better than the existing estimator. When T_1 and T_2 are dependent with $\alpha = 5$ (Kendall's tau $= 0.667$), the proposed estimator using the true a^* and estimator $\hat{a}$ outperforms the existing estimator in terms of the MSE. In spite of the superior performance in terms of the MSE, the proposed estimator is biased compared to the existing estimator. This is a typical phenomenon of the trade-off between bias and variance, seen in shrinkage estimators. The results for negative correlation with $\alpha = 1/2$ (Kendall's tau $= -0.333$) are similar and not shown here.

We examine the properties of the bootstrap estimator $\hat{a}$ of a^*. Table 2 shows the results when the association parameter takes $\alpha = 5$ (Kendall's tau $= 0.667$). When n increase from 100 to 300, $E[\hat{a}(t_1, t_2)]$ approaches to $a^*(t_1, t_2)$ and MSE$[\hat{a}(t_1, t_2)]$ approaches to zero. Hence, the estimator $\hat{a}$ appears to be consistent. Although there need more extensive studies, the results give some numerical support for Conjecture 1.

Fig. 1 compares the bias of the three methods under $\alpha = 1.1 \sim 7$ (Kendall's tau $= 0.05 \sim 0.75$). We see that the existing estimator of Sankaran et al. [19] has the smallest bias while the proposed estimator and the independence estimator have downward bias.

Fig. 2 compares the MSE of the three methods when the association parameter takes $\alpha = 1.1 \sim 7$ (Kendall's tau $= 0.05 \sim 0.75$). In spite of the downward bias, the proposed estimator with the true $a = a^*$ has the smallest MSE. This is the consequence of the trade-off between bias and variance. Also, the proposed estimator with estimate $\hat{a}$ of a^* does not change the performance much, as the bootstrap estimator $\hat{a}$ is a good approximation to the true a^*. As a result, the proposed estimator using $\hat{a}$ performs better than the existing estimator.

We examine how the MSE curve of $\hat{F}_{ij}^{a}(t_1, t_2)$ changes as the value of a varies from 0 to 1. Fig. 3 shows the results under $\alpha = 5$ (Kendall's tau $= 0.667$). It is seen that the MSE curve attains the minimal value at a^*, at which it is strictly smaller than those of the MSEs of the existing estimator and independence estimator. This observation provides a numerical support for Theorem 1.

6. Data analysis

We illustrate our proposal using the mouse data and salamander data.

6.1. Mouse data analysis

We consider the mouse data concerning time to tumor appearance or death for 50 pairs of mice from the same litter in a tumor genesis experiment [14,24]. In this data, T_1 and T_2 are failure times (in weeks) for a pair of mice from the same litter,

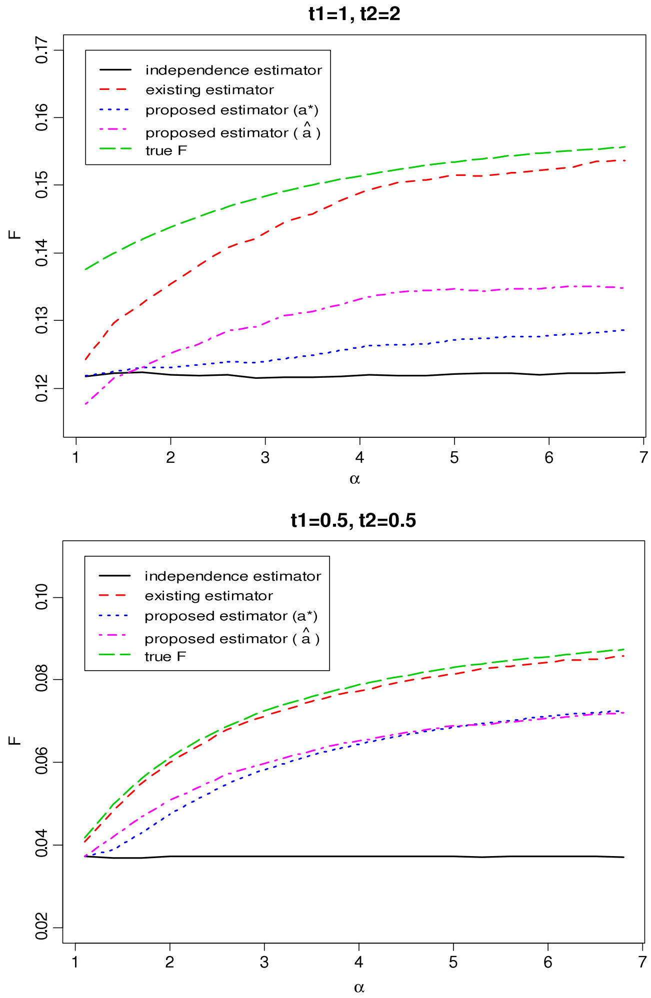

Fig. 1. Simulation results for $E[\hat{F}_{11}(t_1, t_2)]$ based on four different estimators under Clayton model with $\alpha = 1.1 \sim 7(\tau = 0.05 \sim 0.75)$. Their values are compared with the true $F_{11}(t_1, t_2)$.

Table 2
Simulation results for estimating the true a^* using the bootstrap estimator $\hat{a}$. The corresponding pair of causes is $(i, j) = (1, 1)$.

Dependent case with $\alpha = 5(\tau = 0.667)$				
n	(t_1, t_2)	$a^*(t_1, t_2)$	$E[\hat{a}(t_1, t_2)]$	$\text{MSE}[\hat{a}(t_1, t_2)]$
100	(1, 2)	0.1853	0.3737	0.09094
200		0.2540	0.3812	0.07169
300		0.2832	0.3905	0.06837
100	(0.5, 0.5)	0.6754	0.5939	0.05670
200		0.8107	0.7670	0.02110
300		0.8833	0.8376	0.00975

and the corresponding causes C_1 and C_2 express the appearance of a tumor ($C_k = 1$), death happened before the tumor appearance ($C_k = 2$), or censoring ($C_k = 0$). A common censoring (Type I censoring) occurs at 104 weeks for all subjects.

Figs. 4 and 5 compare estimates of sub-distributions calculated by the existing estimator $\hat{F}_{ij}(t_1, t_2)$ and the proposed estimator $\hat{F}_{ij}^{\hat{a}}(t_1, t_2)$ for $i, j = 1, 2$. The surface of the existing estimator $\hat{F}_{ij}(t_1, t_2)$ is a fairly crude step function (Fig. 4). This is because the jumps of $\hat{F}_{ij}(t_1, t_2)$ occur only when the joint failure events correspond to the pair (i, j) occurs. On the other hand, the surface of $\hat{F}_{ij}^{\hat{a}}(t_1, t_2)$ is smoother than the surface of $\hat{F}_{ij}(t_1, t_2)$ (Fig. 5). This smoothness of $\hat{F}_{ij}^{\hat{a}}(t_1, t_2)$ is achieved by borrowing information from the independence estimator $\hat{F}_{1i}(t_1)\hat{F}_{2j}(t_2)$, where the jumps occur when the marginal events correspond to either cause i or j occur.

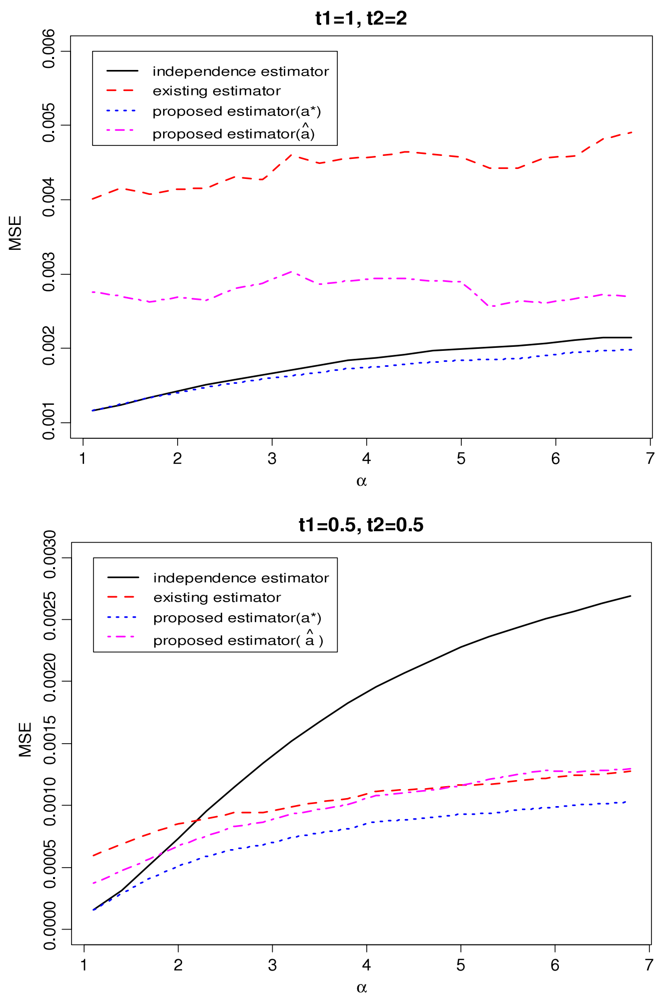

Fig. 2. Simulation results for $\mathrm{MSE}[\hat{F}_{11}(t_1, t_2)] = E[\hat{F}_{11}(t_1, t_2) - F_{11}(t_1, t_2)]^2$ based on four different estimators under Clayton model with $\alpha = 1.1 \sim 7$ ($\tau = 0.05 \sim 0.75$).

6.2. Salamander data analysis

We analyze time to completion of metamorphosis on the salamander larvae living in Hokkaido, Japan [15]. We consider a subset of the larvae that are grown under the high-water level. The resultant data consists of $n = 90$ egg clutches, and each clutch contains 2 larvae (Fig. 6). The times to events for a pair of larvae are denoted as T_1 and T_2. Since a pair of larvae belongs to the same clutch (same parent), they share unobserved characteristics, which induces correlation. The real data on the 90 pair of larvae is shown in Table 3. A pair of causes C_1 and C_2 indicates whether the failure event is metamorphosis ($C_k = 1$), or the death prior to metamorphosis ($C_k = 2$). There is no censoring in this data. We focus on the estimation of $F_{11}(t_1, t_2)$ since the major biological interest is on the time-to-metamorphosis.

Table 4 shows the estimates $\hat{F}_{11}(t_1, t_2)$, $\hat{F}_1(t_1)\hat{F}_1(t_2)$, $\hat{a}$ and $\hat{F}_{11}^{\hat{a}}(t_1, t_2)$. Since T_1 and T_2 are from the same clutch, the correlation between them is fairly strong. Due to this reason, the estimates $\hat{a}$ are close to 1 in many cases. Therefore, the proposed estimator $\hat{F}_{ij}^{\hat{a}}(t_1, t_2)$ is very close to $\hat{F}_{ij}(t_1, t_2)$. Fig. 7 draw the plot for $\hat{F}_{11}^{\hat{a}}(t_1, t_2)$. Note that Michimae and Emura [15] treated all larvae in the clutch as independent observations and plotted the univariate cause-specific distribution function. Although such simplified analysis is useful for biological studies, we propose to redo their analysis taking into account the dependency between T_1 and T_2.

7. Conclusion and discussion

In this paper, we have developed a new nonparametric estimator of sub-distribution function for bivariate competing risks models. The new method with the optimal choice of the tuning parameter improves upon the existing estimator of Sankaran et al. [19]. To be practical, we suggest the bootstrap to estimate the optimal tuning parameter. For large sample analysis, we prove the pointwise consistency of the proposed estimator. Simulation results show that proposed estimator

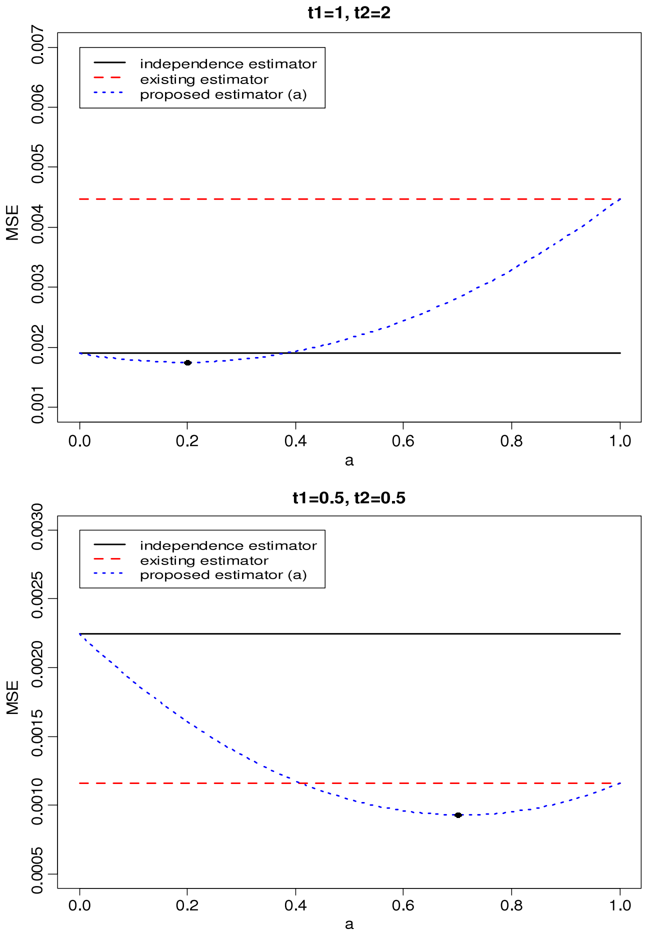

Fig. 3. Simulation results for $\text{MSE}[\hat{F}_{11}(t_1, t_2)] = E[\hat{F}_{11}(t_1, t_2) - F_{11}(t_1, t_2)]^2$ based on three different estimators. The mark "·" indicates the optimal value $a^*(t_1, t_2) = \underset{a}{\text{argmin}}\, \text{MSE}[\hat{F}^a_{11}(t_1, t_2)]$ for the proposed estimator.

Table 3
The real data on 90 pairs of salamander larvae grown under the high-water level from Michimae and Emura (2012).

(T_1, T_2)	(C_1, C_2)	(T_1, T_2)	(C_1, C_2)	(T_1, T_2)	(C_1, C_2)	(T_1, T_2)	(C_1, C_2)	(T_1, T_2)	(C_1, C_2)
(89, 80)	(1, 1)	(69, 40)	(1, 2)	(68, 81)	(1, 1)	(72, 79)	(1, 1)	(83, 85)	(1, 1)
(79, 28)	(1, 2)	(75, 69)	(1, 1)	(72, 71)	(1, 1)	(76, 62)	(1, 1)	(84, 82)	(1, 1)
(82, 80)	(1, 1)	(80, 89)	(1, 1)	(70, 71)	(1, 1)	(74, 79)	(1, 1)	(81, 80)	(1, 1)
(78, 88)	(1, 1)	(78, 76)	(1, 1)	(75, 77)	(1, 1)	(65, 69)	(1, 1)	(77, 74)	(1, 1)
(74, 78)	(1, 1)	(75, 79)	(1, 1)	(74, 81)	(1, 1)	(78, 35)	(1,2)	(73, 74)	(1, 1)
(72, 80)	(1, 1)	(84, 81)	(1, 1)	(81, 74)	(1, 1)	(77, 70)	(1, 1)	(79, 77)	(1, 1)
(71, 73)	(1, 1)	(80, 88)	(1, 1)	(80, 83)	(1, 1)	(86, 83)	(1, 1)	(74, 77)	(1, 1)
(73, 71)	(1, 1)	(77, 64)	(1, 1)	(81, 85)	(1, 1)	(83, 87)	(1, 1)	(72, 73)	(1, 1)
(71, 74)	(1, 1)	(81, 66)	(1, 1)	(83, 82)	(1, 1)	(83, 81)	(1, 1)	(82, 88)	(1, 1)
(81, 69)	(1, 1)	(72, 70)	(1, 1)	(67, 65)	(1, 1)	(84, 82)	(1, 1)	(83, 86)	(1, 1)
(85, 85)	(1, 1)	(74, 73)	(1, 1)	(71, 66)	(1, 1)	(80, 83)	(1, 1)	(85, 85)	(1, 1)
(84, 80)	(1, 1)	(74, 69)	(1, 1)	(78, 74)	(1, 1)	(76, 79)	(1, 1)	(81, 74)	(1, 1)
(80, 85)	(1, 1)	(74, 88)	(1, 1)	(76, 74)	(1, 1)	(77, 78)	(1, 1)	(75, 85)	(1, 1)
(78, 79)	(1, 1)	(80, 83)	(1, 1)	(76, 74)	(1, 1)	(78, 75)	(1, 1)	(80, 73)	(1, 1)
(72, 77)	(1, 1)	(77, 74)	(1, 1)	(87, 80)	(1, 1)	(73, 74)	(1, 1)	(71, 73)	(1, 1)
(70, 62)	(1, 1)	(75, 76)	(1, 1)	(78, 86)	(1, 1)	(79, 73)	(1, 1)	(76, 74)	(1, 1)
(75, 74)	(1, 1)	(83, 89)	(1, 1)	(72, 48)	(1, 2)	(79, 81)	(1, 1)	(65, 44)	(1,2)
(73, 71)	(1, 1)	(66, 72)	(1, 1)	(78, 84)	(1, 1)	(82, 79)	(1, 1)	(79, 79)	(1, 1)

Note: The failure times of the pair of larvae are denoted as T_1 and T_2; Causes C_1 and C_2 indicate whether the failure event is metamorphosis ($C_k = 1$), or the death prior to metamorphosis ($C_k = 2$). The data do not have censored observations.

has smaller MSE than the existing estimator in finite sample. Real data analyses demonstrate that the proposed estimator of the cause-specific distribution is smoother than the existing estimator, where the tuning parameter is regarded as

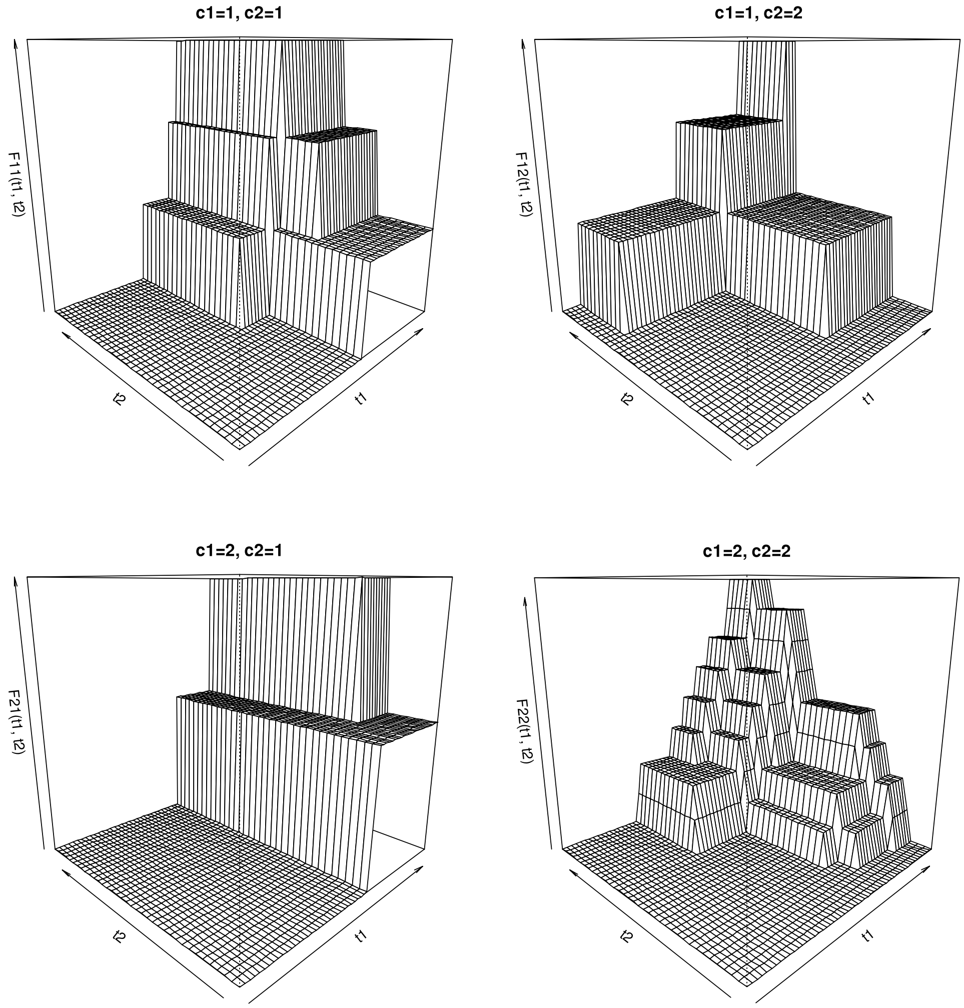

Fig. 4. The existing estimator $\hat{F}_{ij}(t_1, t_2)$ for a bivariate cause-specific distribution function with the mouse data. Four plots correspond to pairs of 4 different causes.

Table 4
Estimates of the cause-specific distribution function $F_{11}(t_1, t_2)$ using the salamander data.

(t_1, t_2)	$\hat{F}_{11}(t_1, t_2)$	$\hat{F}_1(t_1)\hat{F}_1(t_2)$	$\hat{a}$	$\hat{F}^{\hat{a}}_{11}(t_1, t_2)$
(73.25, 73.00)	0.144	0.065	1.000	0.144
(73.25, 77.00)	0.189	0.119	0.922	0.183
(73.25, 81.75)	0.222	0.176	0.818	0.214
(77.00, 73.00)	0.211	0.131	1.000	0.211
(77.00, 77.00)	0.356	0.239	1.000	0.356
(77.00, 81.75)	0.456	0.352	1.000	0.456
(81.00, 73.00)	0.256	0.204	1.000	0.256
(81.00, 77.00)	0.467	0.373	1.000	0.467
(81.00, 81.75)	0.611	0.551	1.000	0.611

Note: The selected values for (t_1, t_2) are the 25% point, median, and 75% point of the observed values of T_1 and T_2, respectively.

a smoothing parameter. Therefore, our proposed estimator achieves two goals in a single framework: the improvement in the MSE and the smoothing of the estimator. These advantages are important, especially in the present bivariate function estimations, where traditional nonparametric estimator can be a crude step function.

Although the idea of combining two estimators of bivariate survival functions to improve the MSE has been considered by Akritas and Keilegom [1], our proposal has the fundamental difference from their approach. Their method combines two consistent estimators that have similar performance. However, our proposal combines a consistent estimator with an inconsistent estimator which has smaller variance. The small variance exploits the reduction of a bivariate functional esti-

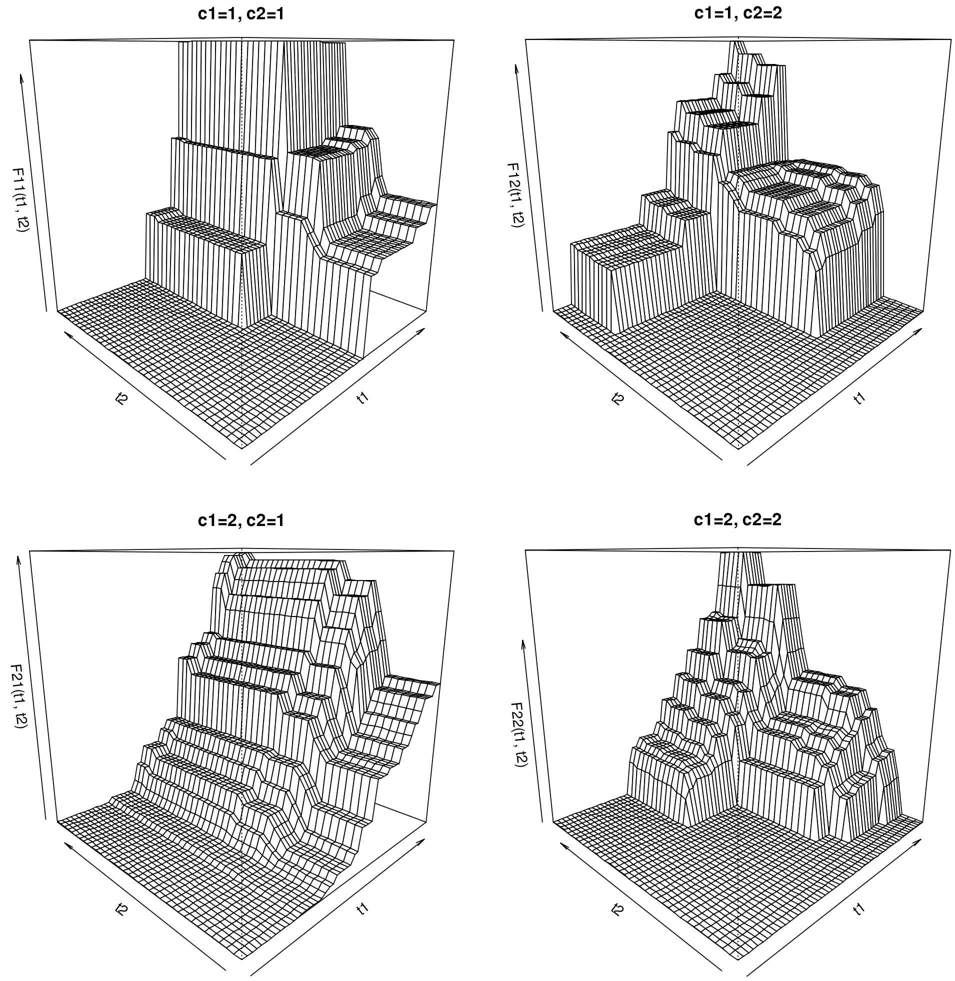

Fig. 5. The proposed estimator $\hat{F}_{ij}^{\hat{a}}(t_1, t_2)$ for a bivariate cause-specific distribution function with the mouse data. Four plots correspond to pairs of 4 different causes.

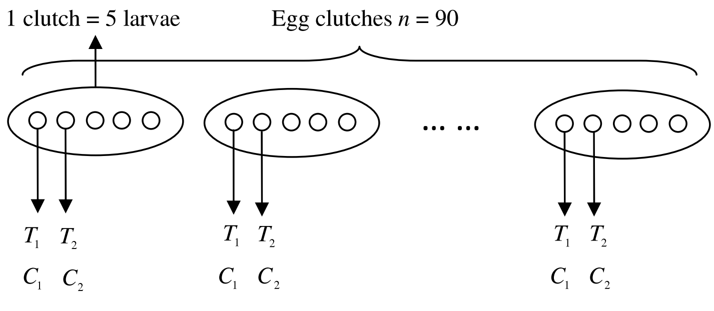

Fig. 6. Bivariate competing risks data from the salamander data of Michimae and Emura (2012). To adapt to the bivariate setting, we chose the first two larvae from the 5 larvae in egg clutches.

mator to a pair of two univariate function estimators under the independence. This trade-off between bias and variance is the key in our approach, which makes it different from the existing approach. In the dimension reduction point of view, our construction of the improved estimator is similar to Emura et al. [11] who propose to reduce the large variability of the multivariate Cox's partial likelihood estimators under high-dimensional covariates. They utilize the univariate partial likelihood estimator which is the biased (inconsistent) estimator but substantially reduces the variance. Combining the multivariate likelihood with the univariate likelihood, the resultant estimators acquire both consistency and reduced variability due to the high-dimensionality. In light of this result, our proposal for the bivariate competing risks data can be extended to higher dimensional competing risks data.

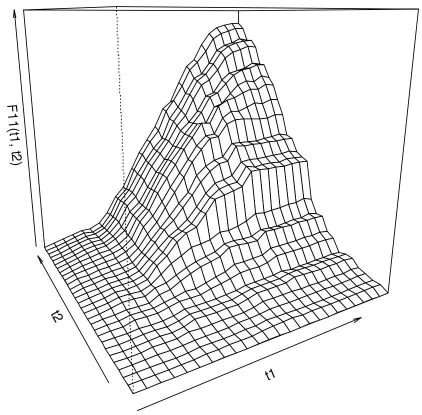

Fig. 7. The plot for the proposed estimator $\hat{F}_{11}^{\hat{a}}(t_1, t_2)$ for a bivariate cause-specific distribution function with the salamander data of Michimae and Emura [15].

Acknowledgments

We would like to thank the anonymous reviewer for his/her helpful comments that greatly improved the manuscript. The work is supported by the grant NSC101-2118-M008-002-MY2.

References

[1] M.G. Akritas, I. van Keilegom, Estimation of bivariate and marginal distributions with censored data, J. R. Stat. Soc. Ser. B 65 (2003) 457–471.
[2] P.K. Andersen, Competing risks as a multi-state model, Stat. Methods Med. Res. 11 (2002) 203–215.
[3] P.K. Andersen, O. Borgan, R.D. Gill, N. Keiding, Statistical Models Based on Counting Processes, Springer, New York, 1993.
[4] A.A. Antony, P.G. Sankaran, Estimation of bivariate survivor function of competing risk models under censoring, J. Stat. Theory Appl. 4 (2005) 401–423.
[5] G. Bakoyannis, G. Touloumi, Practical methods for competing risks data: a review, Statistical Method in Medical Research 21 (2012) 257–272.
[6] K.L. Chung, A Course in Probability Theory, Academic Press, 2001.
[7] D.G. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika 65 (1978) 141–151.
[8] M.J. Crowder, Classical Competing Risks, Chapman and Hall/CRC, 2001.
[9] D.M. Dabrowska, Kaplan–Meier estimate on the plane, Ann. Statist. 18 (1988) 1475–1489.
[10] D.M. Dabrowska, Kaplan–Meier estimate on the plane: weak convergence, LIL, and the bootstrap, J. Multivariate Anal. 29 (1989) 308–325.
[11] T. Emura, Y.-H. Chen, H.-Y. Chen, Survival prediction based on compound covariate under Cox proportional hazard models, PLoS One 7 (10) (2012). http://dx.doi.org/10.1371/joumal.pone.0047627.
[12] R.A. Khan, A note on estimating the mean of a normal distribution with known coefficient of variation, J. Amer. Statist. Assoc. (1968) 1039–1041.
[13] J.P. Klein, M.L. Moeschberger, Survival Analysis: Techniques for Censored and Truncated Data, second ed., Springer, New York, 2003.
[14] N. Mantel, J.L. Ciminera, Use of log rank series in the analysis of litter-matched data on time to tumour appearance, Cancer Res. 39 (1979) 4308–4315.
[15] H. Michimae, T. Emura, Correlated evolution of phenotypic plasticity in metamorphic timing, J. Evol. Biol. 25 (2012) 1331–1339.
[16] R.L. Prentice, J. Cai, Covariate and survivor function estimation using censored multivariate failure time data, Biometrika 79 (1992) 495–512.
[17] P.G. Sankaran, A.A. Antony, Bivariate competing risks models under random left truncation and right censoring, Sankhya 69 (2007) 425–447.
[18] P.G. Sankaran, A.A. Antony, Non-parametric estimation of lifetime distribution of competing risk models when censoring times are missing, Statist. Papers 50 (2009) 339–361.
[19] P.G. Sankaran, J.F. Lawless, B. Abraham, A.A. Antony, Estimation of distribution function in bivariate competing risk models, Biom. J. 48 (2006) 399–410.
[20] P. Shen, Estimation of the bivariate cause-specific distribution function with doubly censored competing risks data, J. Statist. Plann. Inference 141 (2011) 2614–2621.
[21] A.W. Van Der Vaart, J.A. Wellner, Weak Convergence and Empirical Process, Springer, New York, 1996.
[22] W. Wang, M.T. Wells, Nonparametric estimations of the bivariate survival function under simplified censoring conditions, Biometrika 84 (1997) 863–883.
[23] E. Wencheko, P. Wijekoon, Improved estimation of the mean in one-parameter exponential families with known coefficient of variation, Statist. Papers 46 (2005) 101–115.
[24] Z. Ying, L.J. Wei, The Kaplan–Meier estimate for dependent failure time observations, J. Multivariate Anal. 50 (1994) 17–29.