Journal of the Korean Data & Information Science Society
2017, **28**(4), 947–955

http://dx.doi.org/10.7465/jkdi.2017.28.4.947
한국데이터정보과학회지

Nonpararmetric estimation for interval censored competing risk data[†]

Yang-Jin Kim[1] · Do young Kwon[2]

[12]Department of Statistics, Sookmyung Women's University

Received 3 July 2017, revised 24 July 2017, accepted 26 July 2017

Abstract

A competing risk analysis has been applied when subjects experience more than one type of end points. Geskus (2011) showed three types of estimators of CIF are equivalent under left truncated and right censored data. We extend his approach to an interval censored competing risk data by using a modified risk set and evaluate their performance under several sample sizes. These estimators show very similar results. We also suggest a test statistic combining Sun's test for interval censored data and Gray's test for right censored data. The test sizes and powers are compared under several cases. As a real data application, the suggested method is applied a data where the feasibility of the vaccine to HIV was assessed in the injecting drug uses.

Keywords: Competing risks, interval censored data, inverse probability weighting, log rank test, product limit estimator.

1. Introduction

Interval censored failure time data are composed of two inspection times including a failure time. For example, in an AIDS cohort study, a HIV infection is detected through patient's blood tests. The last test time with a negative response and the first test time showing a positive response became an interval censored data in order to estimate the distribution of the infection time. Sun (2006) provided a comprehensive overview for interval censored failure time data including current status data and panel count data. In many clinical and epidemiological studies, subjects can experience more than one causes and the related data is formatted as competing risk data. Competing risks can be represented either with a latent failure approach or with a multistate model (Kalbfleisch and Prentice, 2002; Andersen *et al.*, 1993). Two commonly referred statistics are the cause specific hazard (CSH) and the cumulative incidence function (CIF). The CSH is the occurrence rate of a particular cause in the presence of all causes of failure and given $Z = z$, the CSH is defined as $\lambda_k(t|z) = lim_{h\to 0} h^{-1}\Pr(t < T \leq t + h, \epsilon = k|T \geq t, z)$. The CIF is the cumulative probability of a

[†] This work was supported by Korea research grant NRF-2014R1A2A2A01003567.

[1] Corresponding author: Associate Professor, Department of Statistics, Sookmyung Women's University, Seoul 04310, Korea, E-mail: yjin@sookmyung.ac.kr

[2] Master Student, Department of Statistics, Sookmyung Women's University, Seoul 04310, Korea.

particular cause event occurring by a certain time t and is defined as $F_k(t) = \Pr(T \leq t, \epsilon = k)$. Since $lim_{t\to\infty} F_k(t) = \Pr(\epsilon = k)$, the CIF is often called as a subdistribution function. For an inference procedure, while the CSH regards other causes-related failures as a censoring, the CIF depends on the estimation of other CSHs as follows, $\hat{F}_k(t) = \int_0^t \hat{\lambda}_k(s)\hat{S}(s)ds$, where $\hat{S}(t) = \exp(-\int_0^t \sum_{k=1}^K \hat{\lambda}_k(s)ds)$. Given $Z = z$, a subdistribution hazard (SH) has a following relation with the CIF,

$$\gamma_k(t|z) = -\frac{d}{dt}\log\{1 - F_k(t|z)\} = \frac{\lambda_k(t|z)S(t|z)}{1 - F_k(t|z)}.$$

Gray (1988) proposed a weighted risk set In order to estimate the SH and the log rank test. Fine and Gray (1999) considered a proportional hazard regression model under several censoring schemes and under a randomly right censored data, they applied an inverse probability censoring weight (IPCW) technique assigning weights to the subjects experiencing competing risks where the weight is determined by both competing event time and the censoring time distribution. As an alternative method, Ruan and Gray (2008) considered the imputation of potential censoring times for the subjects experiencing competing events.

A few studies have been done with an interval censored competing risk data. Jewell *et al.* (2003) developed a nonparametric maximum likelihood estimator (NPMLE) and a pseudo MLE of CIF and Jewell and Kalbfleisch (2004) suggested a modified pool adjacent violator algorithm (PAVA) for type I interval censoring (Nam and Kang, 2014). For general interval censored data, Hudgens *et al.* (2001) suggested the NPMLEs. Sun and Shen (2009) proposed a two-stage estimation procedure for both a cause specific event incidence probability and a hazard function. Hudgens *et al.* (2014) suggested a parametric approach by extending Jeong and Fine's method (2006).

Recently, Geskus (2011) showed three types of CIF estimators are equivalent under left truncated and right censored data and derived a regression model. In this paper, we apply his approach to interval censored data and construct a test statistic to compare CIFs. In Section 2, we introduce notations and present two types of NPMLEs of CIF, the product limit estimator and inverse probability weighted estimator. Section 3 proposes a test statistic combining Sun's test statistic (Zhou and Sun, 2004) and Gray's one. Section 4 shows simulation results and real data analysis. In Section 5, we conclude with some remarks.

2. Estimation of NPMLE

Observed data comprise $\{O_i = (L_i, R_i, \delta_i, \epsilon_i, Z_i), i = 1, \cdots, n\}$, where $\delta_i = I(R_i < \infty)$ is a censoring indicator, ϵ_i a cause indicator and Z_i is a group indicator. Hudgens *et al.* (2001) extended a Turnbull's self consistent algorithm to competing risk setting. Let $s_1 < \cdots < s_m < s_{m+1} = \infty$ be the ordered distinct time points derived from the interval censored data. We assume that inspection times are independent of failure times and causes. Define d_{kj} as the number of failures related with cause k at s_j and n_j as at risk number at s_j, respectively. With $\alpha_{ij} = I(s_j \in (L_i, R_i])$,

$$d_{kj} = \sum_{i=1}^{n} d_{ikj}, \qquad d_{ikj} = \delta_i I(\epsilon_i = k)\frac{\alpha_{ij}\hat{f}_{kj}}{\sum_{u=1}^{m+1}\alpha_{iu}\hat{f}_{ku}}, \tag{2.1}$$

$$n_j = \sum_{i=1}^{n} \left[\delta_i \sum_{r=j}^{m+1} d_{ijr} + (1-\delta_i) I(L_i > s_j) \right],$$

where $\hat{f}_{kj} = [\hat{F}_k(s_j) - \hat{F}_k(s_j-)]$.

Geskus (2011) showed two alternative representations of F_k under left truncated and right censored data. One is an inverse probability weighted estimator (IPW), $\hat{F}_k^{IPW}$ and the other is a product limit estimator (PL), $\hat{F}_k^{PL}$.

In this section, our interest is to derive these two representations in a context of interval censored data. First, we assume there exist a right censoring process working independently with an inspection process. Let $C_i \sim G$ be a right censoring time and C_i^* be defined as an observable censoring time. In a context of interval censored data, set $C_i^* = L_i$ for $\delta_i = 0$ and $C_i^* = R_i$ for $\delta_i = 1$. Then using $(C^*, 1-\delta)$, the product limit estimator of $\bar{G} = 1 - G$ has a following form,

$$\hat{\bar{G}}(t) = \prod_{j:c_j \le t} \left(1 - \frac{m_j}{r_c(c_j)} \right), \tag{2.2}$$

where $c_1 < \cdots < c_r$ are the ordered distinct observed censoring times, m_l denotes the number of censored subjects at c_l and $r_c(c) = \sum_{i=1}^{n} I(C_i^* > c)$. Then a weighted empirical cumulative distribution function is calculated as follows

$$\hat{F}_k^{IPW}(t) = \frac{1}{n} \sum_{j:s_j<t} \frac{d_{kj}}{\hat{\bar{G}}(s_j)}.$$

Next, for defining the product-limit estimator, if a risk set should be redefined as

$$n_{kj}^* = \sum_{i=1}^{n} n_{ikj}^*,$$

$$\begin{aligned} n_{ikj}^* \quad = \quad & I(\epsilon_i = k) \sum_{l=j}^{m+1} d_{ikl} + I(\epsilon_i \neq k) \Big[I(s_j > c_i^*) \frac{\hat{\bar{G}}(s_j)}{\hat{\bar{G}}(c_i^*)} \\ & + I(s_j \le c_i^*) \Big] + (1-\delta_i) I(s_j < L_i), \end{aligned} \tag{2.3}$$

then $\hat{F}_k^{PL}$ has the following form

$$\hat{F}_k^{PL}(t) = \prod_{j:s_j<t} \left(1 - \frac{d_{kj}}{n_{kj}^*} \right), \tag{2.4}$$

where $\frac{d_{kj}}{n_{kj}^*} = \hat{\gamma}_{kj}$ is interpreted as the estimated subdistribution hazard at s_j. Since d_{kj} includes the calculation of $\hat{F}_k^{IPW}$ and $\hat{F}_k^{PL}$, it is also an extension of a self-consistency algorithm where iterate the calculations of d_{kj} in the E-step and that of $\hat{F}_k^{IPW}(\hat{F}_k^{PL})$ in the M-step until a convergence criterion satisfies.

3. Two sample Testing

We consider the comparison of the CIFs of two groups in this section. For brevity, we assume there exist two causes. A Gray test (1988) is commonly applied because it reduces to the ordinary log rank test in a case of no competing risk. For interval censored data, Sun (1996) suggested a log rank test reducing to an ordinary log rank test for right censored data. Our goal is to construct a test statistic which becomes either a Gray test statistic under no interval censoring or a Sun test statistic with no competing risk.

Firstly, we review a Sun's test statistic for interval censored data with no competing risk. Denote F_l as a distribution function of a group l. Then $H_0 : F_1 = F_2 = F_0$ or $H_0 : S_1 = S_2 = S_0$. Let $w_1 < \cdots < w_m < w_{m+1} = \infty$ be ordered distinct time points derived from the interval censored data. With an indicator $\eta_{ij} = I(w_j \in (L_i, R_i])$, define $\tilde{d}_{jl}$ as the overall observed failure number and $\tilde{n}_{jl}$ as at risk number for group $l(= 1, 2)$ at w_j and are also expressed as

$$\tilde{d}_{jl} = \sum_{i=1}^{n} \tilde{d}_{ijl}, \quad \tilde{d}_{ijl} = I(z_i = l)\delta_i \frac{\eta_{ij}\hat{h}_j}{\sum_{u=1}^{m+1} \eta_{iu}\hat{h}_u},$$

$$\tilde{n}_{jl} = \sum_{i=1}^{n} I(z_i = l) \left[\delta_i \sum_{r=j}^{m+1} \tilde{d}_{ijr} + (1 - \delta_i) I(L_i > w_j) \right],$$

where $\hat{h}_j = [\hat{S}_0(w_j-) - \hat{S}_0(w_j)]$ and $\hat{S}_0$ is the estimated survival function using a pooled data under H_0. Using these quantities, Sun (1996) derived the following log rank test,

$$U = \sum_{j=1}^{m} \frac{\tilde{n}_{j1}\tilde{n}_{j2}}{\tilde{n}_{j1} + \tilde{n}_{j2}} \left(\frac{\tilde{d}_{j1}}{\tilde{n}_{j1}} - \frac{\tilde{d}_{j2}}{\tilde{n}_{j2}} \right), \tag{3.1}$$

where $\tilde{n}_j = \tilde{n}_{j1} + \tilde{n}_{j2}$ and $\tilde{d}_j = \tilde{d}_{j1} + \tilde{d}_{j2}$. With interval censored competing risk data, the null hypothesis to be tested is $H_0 : F_{11}(t) = F_{12}(t) = F_{10}(t)$, for all t, where F_{1l} is a CIF of cause 1 in a group $l(= 1, 2)$ and F_{10} denotes a common CIF. We assume that each group has a distinct censoring distribution $G_l(l = 1, 2)$ and a corresponding estimator is obtained from (2) with $(C^*_{il}, 1 - \delta_i I(Z_i = l))$ where $C^*_{il} = L_i \times I(Z_i = l)$ for $\delta_i = 0$ and $C^*_{il} = R_i \times I(Z_i = l)$ for $\delta_i = 1$. In a context of right censored competing risk data, Gray (1988) expressed a test statistic as a sum of the differences of subdistribution hazards. Extending his idea to interval censored data, a test statistic using the subdistribution derived at (4) is defined as follows,

$$\tilde{U}^* = \sum_{j=1}^{m} \frac{\tilde{n}^*_{j1}\tilde{n}^*_{j2}}{\tilde{n}^*_{j1} + \tilde{n}^*_{j2}} \left(\frac{\tilde{d}^*_{j1}}{\tilde{n}^*_{j1}} - \frac{\tilde{d}^*_{j2}}{\tilde{n}^*_{j2}} \right), \tag{3.2}$$

where $\tilde{d}^*_{jl}$ is modified as

$$\tilde{d}^*_{jl} = \sum_{i=1}^{n} \tilde{d}^*_{ijl}, \qquad \tilde{d}^*_{ijl} = I(Z_i = l)\delta_i I(\epsilon_i = 1) \frac{\alpha_{ij}\hat{f}^*_j}{\sum_{u=1}^{r+1} \tilde{\alpha}_{iu}\hat{f}^*_j},$$

where $\hat{f}_j^* = [\hat{F}_{10}(w_j) - \hat{F}_{10}(w_j-)]$ and $\hat{F}_{10}$ is the common CIF of cause 1 estimated with a pooled data of group 1 and group 2 under H_0. Once calculating $\tilde{d}_{ijl}^*$, $\tilde{n}_{jl}^*$ is redefined as

$$\begin{aligned}\tilde{n}_{jl}^* &= \sum_{i=1}^{n}\Bigg\{I(\epsilon_i=1)\sum_{q=j}^{r+1}\tilde{d}_{iql}^* + I(\epsilon_i\neq 1)\Bigg[I(s_j > C_{il}^*)\frac{\hat{\tilde{G}}_l(s_j)}{\hat{\tilde{G}}_l(C_{il}^*)} + I(s_j \le C_{il}^*)\Bigg] \\ &\quad +(1-\delta_i)I(s_j < C_{il}^*)\Bigg\}.\end{aligned}$$

If there is no competing risk ($K=1$), $\tilde{n}_{jl}^*$ and $\tilde{d}_{jlk}^*$ become $\tilde{n}_{jl}$ and $\tilde{d}_{jl}$, respectively and (3.2) reduces to (3.1). For testing the hypothesis H_0, the variance matrix $\tilde{V}^*$ of $\tilde{U}^*$ is estimated using a multiple imputation. The detail procedure is as follows,

Step 1: For $q = 1, \cdots, M$, repeat

(a) For $\delta_i = 0$, let $T_i^{(q)} = L_i$

(b) For $\delta_i = 1$, generate $T_i^{(q)}$ satisfying the condition $L_i < T_i^{(q)} \le R_i$ where according to a cause indicator ϵ_i, $T_i^{(q)}$ is a random sample drawn using
$u = \frac{\hat{F}_{10}(T_i^{(q)}) - \hat{F}_{10}(L_i)}{\hat{F}_{10}(R_i) - \hat{F}_{10}(L_i)}, u \sim U(0,1)$ for $\epsilon_i = 1$,
$v = \frac{\hat{F}_{20}(T_i^{(q)}) - \hat{F}_{20}(L_i)}{\hat{F}_{20}(R_i) - \hat{F}_{20}(L_i)}, v \sim U(0,1)$ for $\epsilon_i = 2$,
respectively.

Step 2: With a data obtained from Step 1 $\{T_i^{(q)}, \epsilon_i^{(q)}, \delta_i^{(q)}, Z_i^{(q)}\}$, a Gray test is applied for the test statistic $U_1^{(q)}$ and the variance matrix $V_1^{(q)}$.

Step 3: Iterate M' times Step 1 and Step 2. Define

$$\tilde{V}^* = \frac{1}{M}\sum_{q=1}^{M} V_1^{(q)} + (1+\frac{1}{M})\sum_{q=1}^{M}(U_1^{(q)} - \bar{U})^2,\ \bar{U} = \sum_{q=1}^{M} U_1^{(q)}.$$

Then the hypothesis H_0 can be tested by using $\chi^2 = \tilde{U}^{*2}/\tilde{V}^* \sim \chi^2(df=1)$ or $Z^* = \tilde{U}^*/\sqrt{\tilde{V}^*} \sim N(0,1)$.

4. Simulation

Simulation studies were performed in order to assess the unbiasedness of the proposed estimators and to evaluate the size and the power of the test statistics. To generate failure times, the cause specific hazard rates are fixed as (λ_1, λ_2) following Beyersmann *et al.* (2009),

(i) Simulate a failure time T using all-cause hazards with $\lambda_1 + \lambda_2$.

(ii) For determining the cause of T generated in (i), run a binomial experiment with a probability $\lambda_1/(\lambda_1+\lambda_2)$ on cause 1.

(iii) For generating a right censoring times, a uniform distribution is applied.

(iv) For $T < C$, determine the number of inspection times from a discrete uniform distribution, $b \sim U(10, 15)$. Then generate R's discrete inspection times from $w_l \sim U(0, 5)$ and sort $w_1 < \cdots < w_b$. Set (L, R) satisfying $w_{l-1} = L < T < R = w_l$. For $T > C$, set $L = C$ and $R = \infty$.

500's random samples were generated at $n = 50, 100$ and 200. The means (standard deviations) of two estimators are shown in Table 1. Two estimators give almost identical values and are unbiased for every t. As the sample size increases, the MSEs are getting smaller. Figure 1 shows the two estimates of F_1^{PL} and F_1^{IPW} under $n = 50, 100$ and 200.

For testing $H_0 : F_{11} = F_{12} = F_{10}$, 1000 simulated datasets are generated under two different censoring rates, 10% and 30%. There are five scenarios according to H_0 and H_1's where the cause specific hazard rates of group 1 are fixed as $(\lambda_{11}, \lambda_{21}) = (0.3, 0.2)$ and those of group 2 are varying according to alternative hypotheses. Interval censored time and censoring time were generated in a similar manner to the one-sample case. Figure 2 shows the CIFs of F_{11} and F_{12} (the cause 1 CIFs of group 1 and group 2) under four alternative hypotheses. Table 2 shows the estimated sizes at the nominal level of 0.05 and powers. The estimated sizes show somewhat smaller type 1 error at $n = 50$ but maintain as the sample size increases. When $(\lambda_{12}, \lambda_{22}) = (0.8, 0.2)$, the power has the largest value which coincide with the difference between F_{11} and F_{12} at Figure 2. Meanwhile, $(\lambda_{12}, \lambda_{22}) = (0.3, 0.1)$ has the smallest power among four alternative hypotheses. As censoring rate increases, the power becomes smaller.

In order to illustrate the suggested method, we analyzed a HIV vaccine study (Hudgens *et al.*, 2011). This project was established to assess the feasibility of the vaccine to HIV in the injecting drug users (IDU) in Bangkok, Thailand. 1209 HIV seronegative IDU were enrolled and they were supposed to visit about every four months for counseling and assessment of HIV seroconversion. 1124 people had at least one visit and 133 ones had been diagnosed with HIV seroconversion among them. In this study, two subtypes strains such as subtype B strain and subtype E strain can be occurred and the occurrence of one subtype censors one of the other subtype. In detail, of the 133 converts, 27 and 99 subjects have subtype B and subtype E, respectively, and the remaining seven patients' subtypes were unknown. By treating two subtypes as competing risks, the suggested method is applied. We investigated two genders' CIFs for subtype E and applied a suggested test statistic. $\tilde{U}^* = 5.58$ with $\tilde{V}^* = 6.19$ gives $p - value = 0.025$ which means two groups have significantly different CIFs.

5. Concluding remarks

In this paper, we proposed the IPW estimator and the PL estimator of CIF for interval censored competing risk data by extending Geskus's estimators. Simulation results show the estimates are unbiased over all time period and are almost equivalent. The convergence rate was so fast and it took less than 10 second for F^{IPW} and 20 one for F^{PL} at $n = 200$, respectively. We also suggested a test statistic for comparing two CIFs and several simulation schemes were considered to evaluate the performance of a test statistic. According to simulation result, test sizes maintain nominal levels and powers are getting smaller as censoring rate increases. The suggested method assumes there exists a right censoring variable generating independently with an inspection process composing interval censoring data.

Table 4.1 NPMLEs and standard deviations of F_1 and F_2

	$n=50$		$n=100$		$n=200$	
F_1	$\hat{F}_1^{IPW}$	$\hat{F}_1^{PL}$	$\hat{F}_1^{IPW}$	$\hat{F}_1^{PL}$	$\hat{F}_1^{IPW}$	$\hat{F}_1^{PL}$
0.04	0.036(0.029)	0.034(0.029)	0.043(0.025)	0.040(0.024)	0.040(0.017)	0.039(0.018)
0.08	0.079(0.040)	0.083(0.040)	0.080(0.031)	0.080(0.032)	0.080(0.022)	0.080(0.023)
0.10	0.097(0.041)	0.096(0.041)	0.102(0.035)	0.101(0.034)	0.098(0.023)	0.098(0.023)
0.12	0.116(0.044)	0.115(0.044)	0.122(0.036)	0.122(0.037)	0.118(0.024)	0.119(0.024)
0.16	0.152(0.053)	0.151(0.053)	0.159(0.038)	0.159(0.038)	0.157(0.025)	0.157(0.025)
F_2	$\hat{F}_2^{IPW}$	$\hat{F}_2^{PL}$	$\hat{F}_2^{IPW}$	$\hat{F}_2^{PL}$	$\hat{F}_2^{IPW}$	$\hat{F}_2^{PL}$
0.16	0.171(0.072)	0.157(0.070)	0.161(0.053)	0.154(0.053)	0.163(0.036)	0.157(0.037)
0.32	0.332(0.072)	0.319(0.077)	0.322(0.054)	0.316(0.053)	0.320(0.042)	0.316(0.043)
0.40	0.410(0.079)	0.399(0.077)	0.399(0.054)	0.394(0.053)	0.400(0.042)	0.397(0.042)
0.48	0.492(0.076)	0.495(0.079)	0.477(0.055)	0.475(0.057)	0.481(0.040)	0.478(0.041)
0.64	0.641(0.078)	0.638(0.081)	0.633(0.052)	0.635(0.053)	0.634(0.038)	0.636(0.038)

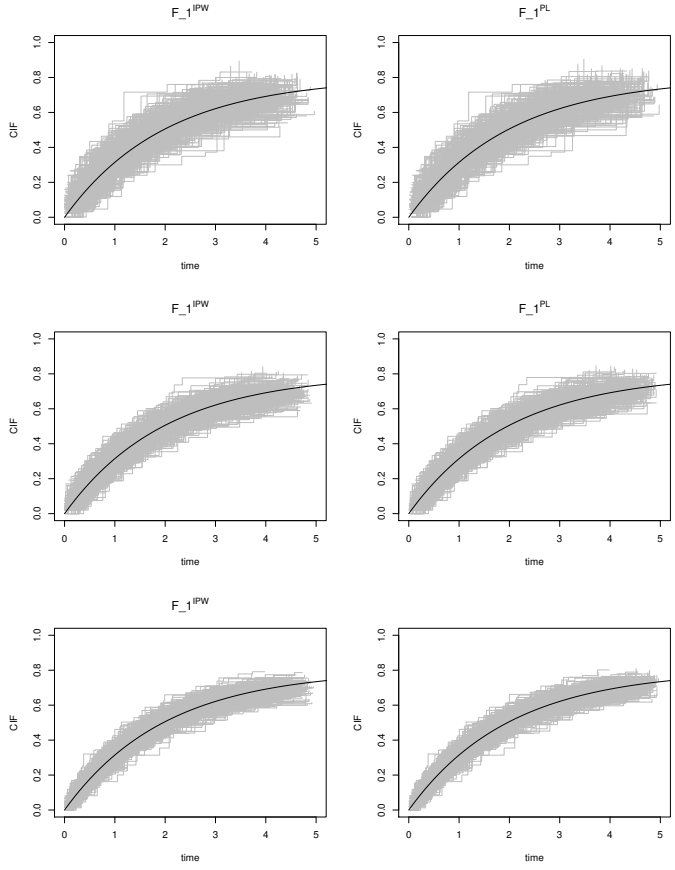

Figure 4.1 F^{IPW} and F^{PL} at $n = 50, 100$ and 200

Table 4.2 Empirical sizes and powers at $\lambda_{11} = 0.3, \lambda_{21} = 0.2$

	censoring=10			censoring=30		
$(\lambda_{12}, \lambda_{22})$	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
(0.3,0.2)	0.047	0.047	0.050	0.047	0.046	0.051
(0.5,0.2)	0.2670	0.520	0.822	0.253	0.458	0.800
(0.8,0.2)	0.760	0.960	1.000	0.746	0.968	1.000
(0.3,0.1)	0.140	0.251	0.461	0.080	0.110	0.180
(0.3,0.5)	0.330	0.682	0.910	0.181	0.350	0.590

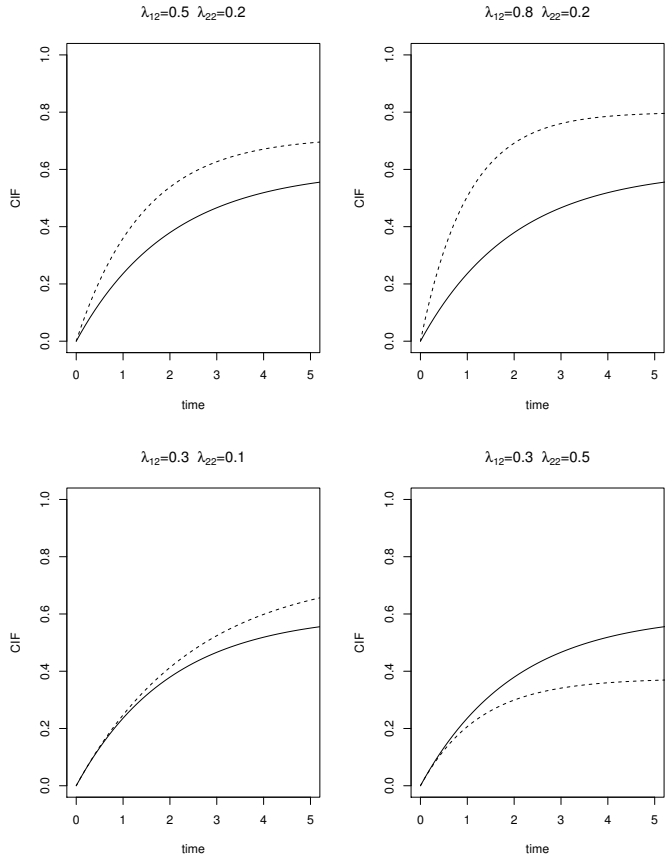

Figure 4.2 F_{11} and F_{12} when $(\lambda_{11}, \lambda_{21}) = (0.3, 0.2)$

In this paper, however, a regression model for estimating the effects of covariates has not been considered. Very similar concepts can be extended to a regression models. Another direction for future research would be a missing cause problem. The HIV dataset includes seven subjects with unidentified subtypes. Even though the number of the subjects with missing cause is small in this dataset, this problem has commonly occurred in competing risk models thus suitable methods should be developed for interval censored data (Goetghebeur

and Ryan, 1995; Lu and Tsiatis, 2001; Moreno-Betancur and Latouche, 2013). Do and Kim (2017) applied a pseudo value approach to interval censored data with missing cause.

References

Andersen, P. K., Klein, J. P. and Rosthøj, S. (2003). Generalized linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika,* **90,**15-27.

Beyersmann, J., Latouche, A., Buchholz, A. and Schumacher M. (2009). Simulating competing risks data in survival analysis. *Statistics in Medicine,* **28,** 956-971.

De Gruttola, V. and Lagakos, S. W. (1989). Analysis of doubly-censored survival data, with application to AIDS. *Biometrics,* **45,** 1-11.

Do, G., Kim, S. and Kim, Y-J. (2015). Statistical analysis of economic activity state of workers with industrial injuries using a competing risk model. *Journal of the Korean Data & Information Science Society,* **26,** 1271-1281.

Do, G. and Kim, Y-J. (2017). Analysis of interval censored competing risk data with missing causes of failure using pseudo values approach. *Journal of Statistical Computation and Simulation,* **87,** 631-639.

Fine, J. P. and Gray, R. J. (1999). Proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association,* **94,** 496-509.

Geskus, R. B. (2011). Cause-specific cumulative incidence estimation and the fine and gray model under both left truncation and right censoring. *Biometrics,* **67,** 39-49.

Goetghebeur, E and Ryan, L. P. (1995). Analysis of competing risks survival data when some failure types are missing. *Biometrika,* **82,** 821-834.

Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics,* **16,** 1141-1154.

Hudgens, M. G., Li, C., and Fine, J. P. (2014). Parametric likelihood inference for interval censored competing risks data. *Biometrics*, **70,** 1-9.

Hudgens, M. G., Satten, G. A. and Longini, I. M. (2001). Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation. *Biometrics*, **57,** 74-80.

Jeong, J. and Fine, J. P. (2006). Direct parametric inference for the cumulative incidence function. *Journal of Royal Statistical Society*, **55,** 187-200.

Jewell, N. P. and Van der Laan, M., and Henneman, T. (2003). Nonparametric estimation from current status data with competing risks. *Biometrika*, **90,** 183-197.

Jewell, N. P. and Kalbfleisch, J. D. (2004). Maximum likelihood estimation of ordered multinominal parameters. *Biostatistics*, **5,** 291-306.

Kalbfleisch, J. D. and Prentice, R. (2002). *The statistical analysis of failure time data*, Wiley, New York.

Lu, K, and Tsiatis, A. A. (2001). Multiple imputation methods for estimating regression coefficients in the competing risks model with missing cause of failure. *Biometrics*, **57,** 1191-1197.

Moreno-Betancur, M. and Latouche, A. (2013). Regression modeling of the cumulative incidence function with missing cause of failure using pseudo-values. *Statistics in Medicine*, **32,** 3206-3223.

Nam, S. J. and Kang, S. B. (2014). Estimation for the extreme value distribution under progressive type-I interval censoring. *Journal of the Korean Data & Information Science Society*, **25,** 643-653.

Ruan, P. K and Gray, R. J. (2008). Analysis of cumulative incidence functions via non-parametric multiple imputation. *Statistics in Medicine*, **27,** 5709-5724.

Sun, J. (1996). A non-parametric test for interval-censored failure time data with application to AIDS studies. *Statistics in Medicine*, **15,** 1387-1395.

Sun, J. (2006). *The statistical analysis of interval-censored failure time data*, Springer-Verlag, New-York.

Sun, J. and Shen, J. (2009). Efficient estimation for the proportional hazard model with competing risks and current status data. *Canadian Journal of Statistics*, **37,** 592-606.

Zhao, Q and Sun, J. (2004). Generalized log-rank test for mixed interval-censored failure time data. *Statistics in Medicine*, **23,** 1621-1629.