isid/ms/2002/19
July 23, 2002
http://www.isid.ac.in/~statmath/eprints

Estimation of cumulative incidence functions in competing risks studies under an order restriction

Hammou El Barmi
Subhash C. Kochar
Hari Mukerjee
and
Francisco J. Samaniego

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Estimation of Cumulative Incidence Functions in Competing Risks Studies Under an Order Restriction

(Abbreviated Title: Restricted Estimation)

Hammou El Barmi
Kansas State University

Subhash C. Kochar
Indian Statistical Institute

Hari Mukerjee
Wichita State University

Francisco J. Samaniego
University of California, Davis

Abstract

In the competing risks problem an important role is played by the cumulative incidence function (CIF), whose value at time t is the probability of failure by time t for a particular type of failure in the presence of other risks. Its estimation and asymptotic distribution theory have been studied by many. In some cases there are reasons to believe that the CIF's due to two types of failure are order restricted. Several procedures have appeared in the literature for testing for such orders. In this paper we initiate the study of estimation of two CIF's subject to a type of stochastic ordering, both when there are just two causes of failure and when there are more than two causes of failure, treating those other than the two of interest as a censoring mechanism. We do not assume independence of the two types of failure of interest, however, these are assumed to be independent of the other causes in the censored case. Weak convergence results for the estimators have been derived. It is shown that when the order restriction is strict, the asymptotic distributions are the same as those for the empirical estimators without the order restriction. Thus we get the restricted estimators "free of charge", at least in the asymptotic sense. When the two CIF's are equal, the asymptotic MSE is reduced by using the order restriction. For finite sample sizes simulations seem to indicate that the restricted estimators have uniformly smaller MSE's than the unrestricted ones in all cases.

Key Words: Cumulative incidence functions; Competing risks; Order restriction; Estimation; Weak convergence.
AMS 2000 Subject Classifications. Primary, 62G05; secondary, 60F17, 62G30.

1 Introduction

In the competing risks model a unit or subject is exposed to several risks at the same time, but the actual failure or death is caused by exactly one of these. Let T denote the lifetime, that we assume to be continuous, with distribution function (DF) F and survival function (SF) S, and let δ denote the cause of death, i.e., $\{\delta = j\}$ is the event that the cause of death is j. An important biometric function of interest is the cumulative incidence function (CIF) due to cause j, a sub-distribution function (SDF), defined by

$$F_j(t) = P[T \leq t,\ \delta = j],\ j = 1, 2, \dots, \tag{1.1}$$

with $F(t) = \sum_j F_j(t)$. The cause specific hazard rate due to cause j is defined by

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathrm{P}[t \leq T < t + \Delta t,\ \delta = j \mid T \geq t],\ j = 1, 2, \dots;$$

the overall hazard rate is $\lambda(t) = \sum_j \lambda_j(t)$. The CIF, $F_j(t)$, may be written as

$$F_j(t) = \int_0^t \lambda_j(u) S(u)\, du. \tag{1.2}$$

In some cases experience and empirical evidence indicate an ordering of two CIF's. Hu and Tsai (1999) consider the two causes of death of HIV positive non-Hodgkin's lymphoma (NHL) patients,- opportunistic infection (OI) and NHL. They show strong evidence for $\lambda_{NHL}(t) \geq \lambda_{OI}(t)$ at all t, which implies that $F_{OI}(t) \leq F_{NHL}(t)$, but not vice versa. For inference purposes, use of the latter assumption generalizes the former the same way that the increasing failure rate average (IFRA) assumption generalizes the increasing failure rate (IFR) assumption. For an HIV vaccine efficacy trial in a region, McKeague, Gilbert and Kanki (2001) compare the risks due to two types of HIV, HIV-1 and HIV-2, so that the genotypes of the HIV antigens in a vaccine matches those of the type with the higher risk. This comparison of risks could be done in several ways, comparison of the CIF's perhaps makes the weakest assumptions. In an example to be analyzed later, there are two causes of death of mice irradiated in infancy,- cancer and other causes. Mice are highly susceptible to thymic lymphoma when irradiated (Kamisaku, *et al.* (1997)). Thus, it is reasonable to assume that the two corresponding CIF's are ordered.

A test for this ordering of SDF's, that may be called stochastic ordering of SDF's, has been developed by Aly, Kochar and McKeague (1994). Dykstra, Kochar and Robertson (1995) tested for ordering of the cause specific hazard rates. Our aim in this paper is the estimation of two CIF's, F_1 and F_2, $F_1 + F_2 = F$, under the restriction $F_1 \leq F_2$. We do not assume that the two risks are independent. In Section 2 we describe our estimators, prove their strong consistency, and derive the weak convergence of the estimators. It is shown that when the order restriction is strict, the asymptotic distributions are Gaussian, the same as those for the unrestricted case. Thus we get the restricted estimators "free of charge", at least in the asymptotic sense. For finite samples our estimators do obey the order restriction, and simulations seem to indicate that they have uniformly smaller mean square errors (MSE's) than the unrestricted estimators. When $F_1 = F_2$ the asymptotic distribution is non-Gaussian. Weak convergence over the entire range may fail in other cases. For $F_1 = F_2$, and even for small violations of the ordering along some contiguous alternatives, we show that the asymptotic MSE is reduced by using the restricted estimators. We also derive a test for $H_0 : F_1 = F_2$ vs $H_1 - H_0$, where $H_1 : F_1 \leq F_2$, and give formulas for asymptotic confidence bands of our estimators. In Section 3 we consider the case when there are risks other than the two of interest, serving as a censoring mechanism. We obtain the estimators in this censored case under the order restriction, prove their strong consistency, derive their weak convergence results, and perform the same test as in the uncesored case. Gray (1988), Pepe and Mori (1993) and Lin (1997) have compared the CIF of one competing risk in different groups. Of these, Lin's work on estimation of the CIF of interest is similar to ours, although his interest was in comparing the single CIF in two groups in the presence of another competing risk, and under censoring. Our problem is comparing two CIF's in the same group in the presence of censoring. As in the uncensored case when the order restriction is strict, our asymptotic results coincide with those of Lin (1997), when generalized to the estimation problem of both CIF's. In Section 4 we provide some simulation results. In Section 5 we consider some examples. In Section 6 we provide a discussion of our results.

2 Uncensored case

Suppose that we have n items exposed to two risks. We observe (T_i, δ_i), the time and cause of failure of the ith item, $1 \leq i \leq n$. We make no assumptions about the independence of the two risk factors. We wish to estimate the CIF's, F_1 and F_2, defined by (1.1) or (1.2), subject to $F_1 \leq F_2$.

2.1 The estimators

The nonparametric maximum likelihood estimator (NPMLE) in the unrestricted case is given by (see Peterson, 1977)

$$\widehat{F}_j(t) = \frac{1}{n} \sum_{i=1}^{n} I(T_i \leq t,\, \delta_i = j),\ \ j = 1, 2. \tag{2.1}$$

Our estimators simply average $\widehat{F}_1(t)$ and $\widehat{F}_2(t)$ when the restriction is violated. Since $\widehat{F}_1 + \widehat{F}_2 = \widehat{F}$, we can define our estimators by

$$F_1^* = \widehat{F}_1 \wedge \widehat{F}/2 \quad \text{and} \quad F_2^* = \widehat{F}_2 \vee \widehat{F}/2, \text{ where } \widehat{F}(t) = \frac{1}{n} \sum_{i=1}^{n} I(T_i \leq t). \tag{2.2}$$

This type of simple estimator was studied by Rojo and Ma (1996) and Rojo (1995) for estimating two stochastically ordered distributions based on independent samples. They have simple asymptotic distributions, and, simulations showed that in terms of MSE they are typically superior to the NPMLE's that were first derived by Brunk *et al.* (1966), and whose (complicated) asymptotic distributions were found only recently by Praestgaard and Huang (1996).

2.2 Consistency

Let $||\cdot||$ denote the sup-norm. For any real functions, f_1, f_2, g_1, g_2, it is easy to check that

$$||\mathrm{m}\{f_1, f_2\} - \mathrm{m}\{g_1, g_2\}|| \leq ||f_1 - g_1|| \vee ||f_2 - g_2||,$$

where "m" can stand for either minimum or maximum. Since $F_1 = F_1 \wedge F/2$ and $F_2 = F_2 \vee F/2$, strong consistency of our estimators follows from those of $\widehat{F}_1, \widehat{F}_2$, and $\widehat{F}$, as stated in the following theorem.

Theorem 2.1 *For the estimators given in (2.2),*

$$P[||F_1^* - F_1|| \to 0,\ ||F_2^* - F_2|| \to 0] = 1.$$

2.3 Weak convergence

The weak convergence of the unrestricted estimators, $\widehat{F}_1$ and $\widehat{F}_2$, follows from Breslow and Crowley (1974) (see also Gill (1983) for a minor correction and some extensions); those for our restricted estimators are derivable from these. Let

$$Z_n = \sqrt{n}[\widehat{F} - F],\ Z_{jn} = \sqrt{n}[\widehat{F}_j - F_j] \text{ and } Z_{jn}^* = \sqrt{n}[F_j^* - F_j],\ j = 1, 2. \tag{2.3}$$

We consider the weak convergence of the bivariate process (Z_{1n}^*, Z_{2n}^*).

Let $b < \infty$ be any constant such that $F(b) < 1$. Breslow and Crowley (1974) have shown that, on $D[0, b] \times D[0, b]$, the bivariate process (Z_n, Z_{1n}) converges weakly to (Z, Z_1), a bivariate Gaussian process with mean 0 and a covariance structure given by

$$\begin{aligned} Cov(Z(s), Z(t)) &= F(s)(1 - F(t)), \\ Cov(Z_1(s), Z_1(t)) &= F_1(s)(1 - F_1(t)), \\ Cov(Z(s), Z_1(t)) &= F_1(s) - F_1(t)F(s), \\ Cov(Z_1(s), Z(t)) &= F_1(s)(1 - F(t)), \end{aligned} \tag{2.4}$$

where $s \leq t$. Since $Z_{1n} + Z_{2n} = Z_n$, by the continuous mapping theorem (Billingsley, 1968), the trivariate process (Z_n, Z_{1n}, Z_{2n}) converges weakly to the mean-zero Gaussian process, (Z, Z_1, Z_2) on $D[0, b] \times D[0, b] \times D[0, b]$ with

$$Cov(Z_i(s), Z_j(t)) = F_i(s)(\delta_{ij} - F_j(t)),\ i, j \in \{1, 2\},\ s \leq t, \tag{2.5}$$

where $\delta_{ij} = I(i = j)$, the Kronecker delta function; the other covariances can be obtained from (2.4). Marginally, $(Z_{1n}, Z_{2n}) \stackrel{w}{\Longrightarrow} (Z_1, Z_2)$ with the covariance structure given by (2.5). Let

$T_F = \inf\{x : F(x) = 1\}$ denote the right end point of the support of F, T_F may be ∞. By Stone (1963) and Lindvall (1973), weak convergence on $[0, b]$ for all $b < T_F$ is equivalent to weak convergence on $[0, T_F)$. Thus $[0, b]$ may be substituted by $[0, T_F)$ above.

Now

$$
\begin{aligned}
Z_{1n}^* &= Z_{1n} + 0 \wedge [Z_n/2 - Z_{1n} + \sqrt{n}(F/2 - F_1)] \\
&= Z_{1n} + 0 \wedge [(Z_{2n} - Z_{1n})/2 + \sqrt{n}(F/2 - F_1)], \qquad (2.6) \\
Z_{2n}^* &= Z_{2n} + 0 \vee [(Z_{1n} - Z_{2n})/2 + \sqrt{n}(F/2 - F_2)]. \qquad (2.7)
\end{aligned}
$$

Note that, if $F_1 < F/2$ (implying $\sqrt{n}(F/2 - F_1) \to \infty$) on $(0, b]$, then $(Z_{1n}^*, Z_{2n}^*) \overset{w}{\Longrightarrow} (Z_1, Z_2)$, and, if $F_1 = F/2$ on $[0, b]$, then

$$(Z_{1n}^*, Z_{2n}^*) = (Z_{1n} + 0 \wedge [Z_n/2 - Z_{1n}], Z_{2n} + 0 \vee [Z_n/2 - Z_{2n}])$$

.

Theorem 2.2 *For (Z_{1n}^*, Z_{2n}^*) defined by (2.3), we have the following results.*
(1) If $F_1 < F/2$ on $(0, b]$, then

$$(Z_{1n}^*, Z_{2n}^*) \overset{w}{\Longrightarrow} (Z_1, Z_2).$$

(2) If $F_1 = F/2$ on $[0, b]$, then

$$(Z_{1n}^*, Z_{2n}^*) \overset{w}{\Longrightarrow} (W_1, W_2),$$

where

$$W_1 = Z_1 + 0 \wedge (Z_2 - Z_1)/2 \text{ and } W_2 = Z_2 - 0 \wedge (Z_2 - Z_1)/2.$$

(3) If $F_1(t_0) = F(t_0)/2$ for some $t_0 \in (0, b)$ and, for some $\gamma > 0$, $F_1 < F/2$ on $(t_0, t_0 + \gamma]$ with $t_0 + \gamma < b$, then (Z_{1n}^*, Z_{2n}^*) does not converge weakly.
(4) If $F_1 < F/2$ on $(0, t_0)$ for some $t_0 < b$ and $F_1 = F_2$ on $[t_0, b]$, then (Z_{1n}^*, Z_{2n}^*) does converge weakly, but has paths that are discontinuous at t_0 with positive probability.

PROOF. The proofs of parts (1) and (2) follow from the arguments given before the theorem and the continuous mapping theorem.

The proof of part (3) follows from an argument establishing lack of tightness. WLOG assume that $F_1 < F_2$ on $(t_0, t_0 + \gamma]$, $\gamma > 0$. Note that

$$Z_{1n}^*(t_0) = Z_{1n}(t_0) + 0 \wedge (Z_{2n}(t_0) - Z_{1n}(t_0))/2 \equiv Z_{1n}(t_0) + 0 \wedge U_n(t_0),$$

where

$$U_n(t_0) = (Z_{2n}(t_0) - Z_{1n}(t_0))/2 \xrightarrow{d} N(0, F((t_0)/4).$$

Since F_1 is continuous, Z_1 has continuous paths $a.s.$ Since $Z_{1n} \overset{w}{\Longrightarrow} Z_1$ on $[0,b]$, $\{Z_{1n}\}$ is tight on $[0,b]$. Let $\epsilon > 0$, $\eta > 0$ be arbitrary. Then, for every $0 < \rho < \gamma$, there exists $0 < \delta < \rho$ such that

$$P[\sup_{t_0 \le s \le t_0+\delta} |Z_{1n}(t_0) - Z_{1n}(s)| \le \epsilon] \ge 1 - \eta$$

for all n large enough. If Z_{1n}^* is tight on $[0,b]$ then, for all n large enough we must have

$$P[\sup_{t_1 \le s \le t < t_2} |Z_{1n}^*(t) - Z_{1n}^*(s)| \le \epsilon] \ge 1 - \eta$$

for some $t_1 \le t_0 < t_2$ (Billingsley (1968)). Now, from (2.6), if $F_1(t) < F_2(t)$, then, for all n large enough,

$$P[Z_{1n}^*(t) = Z_{1n}(t)] > 1 - \eta.$$

Assume that $\{Z_{1n}^*\}$ is tight. Then there exists $0 < \delta' < \delta$ such that, for all n large enough, we have

$$P[\sup_{t_0 \le s \le t_0+\delta'} |Z_{1n}^*(t_0) - Z_{1n}^*(s)| \le \epsilon] \ge 1 - \eta.$$

Now fix $t_0 < t < t_0 + \delta'$. Then, for all n large enough

$$\begin{aligned} P[\sup_{t_0 \le s \le t_0+\delta'} |Z_{1n}^*(t_0) - Z_{1n}^*(s)| \le \epsilon] &\le P[|Z_{1n}^*(t_0) - Z_{1n}^*(t)| \le \epsilon] \\ &\le P[|Z_{1n}^*(t_0) - Z_{1n}(t)| \le \epsilon] + \eta \\ &\le P[|0 \wedge U_n(t_0)| \le 2\epsilon] + 2\eta \\ &\to 1 - \Phi(-4\epsilon/\sqrt{F(t_0)}) + 2\eta, \end{aligned}$$

where Φ is the standard normal DF. This is a contradiction to our assumption that $\{Z_{1n}^*\}$ is tight. Note that the paths fail to be continuous at t_0 when $Z_1(t_0) > Z_2(t_0)$, which occurs with probability 1/2, which is also the limiting value of the last expression in the display above as $\epsilon, \eta \to 0$.

A crucial feature in the proof of part (3) is that every interval of the form $[t_1, t_2)$ containing t_0 must also contain (t_0, t_2) where $F_1 < F_2$. Under the condition in part (4), this problem is avoided by choosing intervals of the form

$$[0 = t_0, t_1), \cdots, [t_{i-1}, t_i = t_0), [t_i = t_0, t_{i+1}), \cdots,$$

and using the tightness of (Z_{1n}, Z_{2n}). However, modifying the proof of part (3) to the interval $[t_0 - \gamma, t_0]$ shows that the paths of the limiting distribution of Z_{1n}^* are discontnuous at t_0 when $Z_1(t_0) > Z_2(t_0)$, which occurs with probability 1/2 since $F_1(t_0) = F_2(t_0)$. ■

REMARK. The lack of tightness of Z_{1n}^* in case (3) of Theorem 2.2 stems from the fact that $Z_{1n}^*(t_0) = Z_{1n}(t_0) + U_n(t_0)$, where $U_n(t_0)$ is asymptotically $N(0, F(t_0)/4)$ in distribution, whereas $Z_{1n}^*(t) = Z_{1n}(t)$ with arbitrarily high probability for all t in some right neigborhood of t_0 if n is large enough.

2.4 A stochastic ordering result

From Theorem 2.2, W_j and Z_j have the same distribution for $j = 1, 2$, if $F_1 < F_2$ on $(0, b)$. The following theorem shows that $|W_j(t)|$ is stochastically smaller than $|Z_j(t)|$ when $F_j(t) = F(t)/2$, $j = 1, 2$, and $0 < F(t) < 1$. However, it contains an even more general result. We show that this stochastic ordering still holds when the order restriction is violated by a small amount along some contiguous alternatives. To this end we assume that the CIF's are changing with the sample size, given by

$$F_{1n} = F/2 + \Delta_n(F)/\sqrt{n} \text{ and } F_{2n} = F/2 - \Delta_n(F)/\sqrt{n}, \tag{2.8}$$

where $||\Delta_n - \Delta|| \to 0$ as $n \to \infty$ for some nonnegative measurable functions, Δ_n and Δ. Let $\widehat{F}_1, \widehat{F}_2, F_1^*, F_2^*$ and Z_n be defined as before, but define

$$\widetilde{Z}_{jn} = \sqrt{n}[\widehat{F}_j - F_{jn}] \text{ and } \widetilde{Z}_{jn}^* = \sqrt{n}[F_j^* - F_{jn}],\ j = 1, 2.$$

By Theorem 3.2.1 in Shorack and Wellner (1986),

$$(Z_n, \widetilde{Z}_{1n}, \widetilde{Z}_{2n}) \overset{w}{\Longrightarrow} (Z, Z_1, Z_2),$$

where (Z, Z_1, Z_2) is the same Gaussian process as when $\Delta \equiv 0$.

From the definition of F_j^* in (2.2), we have

$$\begin{aligned}
\widetilde{Z}_{1n}^* &= \sqrt{n}[F_1^* - F_{1n}] \\
&= \sqrt{n}[\widehat{F}_1 - F_{1n}] + 0 \wedge \sqrt{n}[\widehat{F}_2 - \widehat{F}_1]/2 \\
&= \widetilde{Z}_{1n} + 0 \wedge \sqrt{n}[(\widehat{F}_2 - F_{2n}) - (\widehat{F}_1 - F_{1n}) + (F_{2n} - F_{1n})]/2 \\
&= \widetilde{Z}_{1n} + 0 \wedge [\widetilde{Z}_{2n} - \widetilde{Z}_{1n} - 2\Delta(F)]/2.
\end{aligned}$$

Similarly,

$$\widetilde{Z}_{2n}^* = \widetilde{Z}_{2n} - 0 \wedge [\widetilde{Z}_{2n} - \widetilde{Z}_{1n} - 2\Delta(F)]/2.$$

By the continuous mapping theorem,

$$(\widetilde{Z}_{1n}^*, \widetilde{Z}_{2n}^*) \overset{w}{\Longrightarrow} (Z_1^*, Z_2^*),$$

where

$$Z_1^* = Z_1 + 0 \wedge [Y - \Delta(F)] \text{ and } Z_2^* = Z_2 - 0 \wedge [Y - \Delta(F)],\ \text{with } Y = (Z_2 - Z_1)/2.$$

Theorem 2.3 *Suppose that the contiguous alternatives, (2.8), hold. Let $\sigma^2(t) = F(t)[1 - F(t)]/4$ and let $t > 0$ be arbitrary with $0 < F(t) < 1$. Then,*

$$P(|Z_j^*(t)| \le u) > P(|Z_j(t)| \le u),\ j = 1, 2,\ \textit{for all } u > 0,$$

if $\Delta(F(t)) \le r_0\, \sigma(t)$ for some $r_0 > 0$ that does not depend on t.

PROOF. We prove only the case when $j = 1$. Fix t such that $0 < F(t) < 1$. We have

$$\sigma_1^2(t) \equiv Var(Z_1(t)) = \frac{F(t)}{2}\left[1 - \frac{F(t)}{2}\right], \ \sigma_2^2(t) \equiv Var(Y(t)) = \frac{F(t)}{4},$$

$$\sigma_{12}(t) \equiv Cov(Z_1(t), Y(t)) = -\frac{F(t)}{4}, \text{ and}$$

$$\rho(t) \equiv Corr(Z_1(t), Y(t)) = \frac{\sigma_{12}(t)}{\sigma_1(t)\sigma_2(t)} = -\frac{\sigma_2(t)}{\sigma_1(t)} = -\frac{1}{\sqrt{2 - F(t)}}.$$

Since $(Z_1(t), Y(t))$ has a bivariate normal distribution, the conditional distribution of $Z_1(t)$ given $Y(t) = y$ is

$$N(-y, \sigma^2(t)), \text{ where } \sigma^2(t) = \sigma_1^2(t)(1 - \rho^2(t)) = [F(t)][1 - F(t)]/4.$$

Since $0 < \sigma(t) < \infty$, we can write $\Delta(F(t)) = r(t)\sigma(t)$ for some $r(t) \geq 0$, and replace the line with the probability inequality in the theorem to read

$$P(|Z_j^*(t)| \leq u\sigma(t)) > P(|Z_j(t)| \leq u\sigma(t)), \ j = 1, 2, \ for \ all \ u > 0,$$

which simplifies the expressions below. For further simplification, we suppress the t-dependence in the remainder of the proof. For $u > 0$,

$$P(|Z_1^*| \leq u\sigma) = P(|Z_1| \leq u\sigma, Y \geq r\sigma) + P(|Z_1 + Y - r\sigma| \leq u\sigma, Y < r\sigma).$$

Thus, it is sufficient to show that

$$P(|Z_1 + Y - r\sigma| \leq u\sigma, Y < r\sigma) > P(|Z_1| \leq u\sigma, Y < r\sigma) \text{ for all } u > 0.$$

Let $n_{(\mu,V)}$ dentote the density of a $N(\mu, V)$ random variable. Noting that the conditional distribution of $Z_1 + y$ given $Y = y$ is $N(0, \sigma^2)$, we have

$$\begin{aligned}
&P(|Z_1 + Y - r\sigma| \leq u\sigma, Y < r\sigma) \\
&= \int_{-\infty}^{r\sigma} P(|Z_1 + y - r\sigma| \leq u\sigma | Y = y)\, n_{(0,\sigma_2^2)}(y)dy \\
&= \int_{-\infty}^{r\sigma}\int_{\sigma(r-u)}^{\sigma(r+u)} n_{(0,\sigma^2)}(x)\, n_{(0,\sigma_2^2)}(y)\, dx\, dy \\
&= \int_{-\infty}^{r\sigma/\sigma_2} [\Phi(r+u) - \Phi(r-u)]\phi(z)\, dz \\
&= [\Phi(r+u) - \Phi(r-u)]\Phi(r\alpha),
\end{aligned} \tag{2.9}$$

where $\alpha = \sigma/\sigma_2$, Φ is the standard normal DF, and ϕ is its density. Now,

$$\begin{aligned}
&P(|Z_1| \leq u\sigma, Y < r\sigma) \\
&= \int_{-\infty}^{r\sigma} P(|Z_1| \leq u\sigma | Y = y) n_{(0,\sigma_2^2)}(y)\, dy
\end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{r\sigma} \left[\Phi\left(\frac{y+u\sigma}{\sigma}\right) - \Phi\left(\frac{y-u\sigma}{\sigma}\right)\right] n_{(0,\sigma_2^2)}(y)\, dy \\
&= \int_{-\infty}^{r\sigma/\sigma_2} \left[\Phi\left(\frac{\sigma_2 x+u\sigma}{\sigma}\right) - \Phi\left(\frac{\sigma_2 x-u\sigma}{\sigma}\right)\right] \phi(x)\, dx \\
&= \frac{\sigma}{\sigma_2}\int_{-\infty}^{r} [\Phi(z+u) - \Phi(z-u)]\phi(z\sigma/\sigma_2)\, dz \\
&= [\Phi(r+u) - \Phi(r-u)]\Phi(r\alpha) \\
&\qquad - \int_{-\infty}^{r} [\phi(z+u) - \phi(z-u)]\Phi(z\alpha)\, dz,
\end{aligned}
$$

where we have used integration by parts in the last step. Note that

$$\alpha = \sigma/\sigma_2 = \sqrt{1-F(t)}, \text{ and hence, } 0 \le \alpha \le 1.$$

Denote the integral above by $q(\alpha, r, u)$. Comparing the last term above with (2.9), it is sufficient to show that $q(\alpha, r, u) > 0$ for all $u > 0$ and $0 < \alpha < 1$, if $r > 0$ is sufficiently small. Since $[\phi(z+u) - \phi(z-u)] > (<)\ 0$ for $z < (>)\ 0$ for all $u > 0$, and $\Phi(z\alpha) > (<)\ \Phi(z)$ for all $z < (>)\ 0$, it is sufficient to show that

$$q^*(r,u) \equiv \int_{-\infty}^{r} [\phi(z+u) - \phi(z-u)]\Phi(z)\, dz > 0 \text{ for all } u > 0,$$

if $r > 0$ is sufficiently small. Now, using integration by parts,

$$
\begin{aligned}
q^*(0,u) &= \int_{-\infty}^{0} [\phi(z+u) - \phi(z-u)]\Phi(z)\, dz \\
&= [\Phi(u) - \Phi(-u)]/2 - \int_{-\infty}^{0} [\Phi(z+u) - \Phi(z-u)]\phi(z)\, dz.
\end{aligned}
$$

If (X, Z) is a standard bivariate normal then $(X, Z) \stackrel{d}{=} (-X, -Z)$. Thus, the integral in the expression above may be written as

$$
\begin{aligned}
P(|X-Z| \le u, Z \le 0) &= P(|-X+Z| \le u, Z \ge 0) = P(|X-Z| \le u)/2 \\
&= [\Phi(u/\sqrt{2}) - \Phi(-u/\sqrt{2}]/2.
\end{aligned}
$$

Hence,

$$q^*(0,u) = \Phi(u) - \Phi\left(\frac{u}{\sqrt{2}}\right) = \int_{u/\sqrt{2}}^{u} \phi(z)\, dz > u\phi(u)\left(1 - \frac{1}{\sqrt{2}}\right) \ \forall u > 0.$$

Now, $\partial_r q^*(r,u) = [\phi(r+u) - \phi(r-u)]\Phi(r)$. Thus,

$$\partial_r q^*(0,u) \equiv 0 \text{ and } \partial_r^2 q^*(0,u) = [\phi'(u) - \phi'(-u)]/2 = \phi'(u) = -u\phi(u).$$

Using Taylor's theorem, there exists $r_0 > 0$, such that

$$q^*(r,u) > u\phi(u)\left(1 - \frac{1}{\sqrt{2}} - \frac{r^2}{2}\right) > 0 \text{ for all } 0 \le r \le r_0 \text{ and } u > 0.$$

An estimate of 0.77 is obtained for r_0 by ignoring any remainder term in Taylor's theorem; extensive numerical computations of $q^*(r,u)$ seem to indicate that r_0 is slightly over 0.70. ■

Theorem 2.3 tells us that, if $F_{1n} = F/2 + \Delta_n(F)$, $\sup_t |\Delta_n(F(t)) - \Delta(F(t))| \to 0$, and $\Delta(F(t)) \leq r_0\sigma(t)$, then $P(|Z_j^*(t)| \leq u) > P(|Z_j(t)| \leq u)$ for all $u > 0$, and for all t such that $0 < F(t) < 1$. Note that $\sigma(t) = \sqrt{F(t)[1-F(t)]}/2$ turns out to be the natural unit for measuring the limiting allowable perturbations of $F_1(t)$ above $F(t)/2$ (in units of $1/\sqrt{n}$) for this result to hold.

2.5 Asymptotic bias and MSE

We continue using the notation and results of Section 2.4, and write $\delta(t)$ for $\Delta(F(t))$. From Theorem 2.2, the restricted and the empirical estimators have the same asymptotic distribution when $F_1(t) < F_2(t)$. We can compute the asymptotic bias of the restricted estimators and compare their MSE's when $F_1(t) = F_2(t)$, or under the contiguous alternatives, (2.8). It is well known that for any DF or SDF, H, and its empirical estimator, H_n, $E[\sqrt{n}\sup_t |H_n(t) - H(t)|]^r$ is a bounded sequence for all $r > 0$. By Theorem 2.3, this is also true for $|F_j^*(t) - F_{jn}(t)|$, $j = 1,2$, if $\delta(t) \geq 0$ is small enough. Hence, $E[\sqrt{n}[F_j^*(t) - F_{jn}(t)]^r]$ converges to $E[Z_j^*(t)]^r$, $j = 1,2$, for all $r > 0$.

2.5.1 Asymptotic bias

Utilizing the fact that $Y(t) - \delta(t) \sim N(-\delta(t), \sigma_2^2(t))$, where $\sigma_2^2(t) = Var(Y(t)) = F(t)/4$, the asymptotic bias of $F_1^*(t)$ is given by

$$\frac{E[Z_j^*(t)]}{\sqrt{n}} = \frac{E[Z_1 + 0 \wedge [Y(t) - \delta(t)]]}{\sqrt{n}} = -\frac{\sigma_2(t)}{\sqrt{2\pi n}}\, e^{-\delta(t)^2/2\sigma_2^2(t)} - \frac{\delta(t)}{\sqrt{n}}\,\Phi\left(\frac{\delta(t)}{\sigma_2(t)}\right);$$

the asymptotic bias of $F_2^*(t)$ is the negative of this. When $\delta(t) = 0$, the asymptotic bias of $F_1^*(t)$ is $-\sigma_2(t)/\sqrt{2\pi n} = -\sqrt{F(t)/8\pi n}$.

2.5.2 Asymptotic MSE

From the stochastic ordering result given by Theorem 2.3, the asymptotic MSE (AMSE) of $F_j^*(t)$ is less than that of $\widehat{F}_j(t)$, $j = 1,2$, if $0 \leq \Delta(F(t)) \leq r_0\sigma(t)$ and $0 < F(t) < 1$. The following theorem establishes more precise results.

Theorem 2.4 *Let $t > 0$ be fixed such that $0 < F(t) < 1$. Let s_0 be the unique solution of $\int_0^s u\Phi(u)\,du = 1/4$. Under the contiguous alternatives, (2.8),*

$$AMSE(F_j^*(t)) < AMSE(\widehat{F}_j(t)) = \sigma_j^2(t),\ j = 1,2,\ if\ \Delta(F(t)) < s_0\sigma_2(t).$$

PROOF. We prove only the case when $j = 1$. Let $\Delta(F(t)) = s(t)\sigma_2(t)$ for some $s(t) \geq 0$. Let U denote a $N(0, \sigma^2(t))$ random variable, which is the distribution of $Z_1 + y$ given $Y = y$, and recall that

$$\sigma^2(t) = \sigma_2^2(t)[1 - F(t)],\ \sigma_2^2(t) = \frac{\sigma_1^2(t)}{2 - F(t)},\ \text{and } \sigma_1^2(t) = \frac{F(t)[2 - F(t)]}{4}.$$

We suppress the t-dependence in the remainder of the proof except for $F(t)$. Now,

$$\begin{aligned}
E[Z_1^*]^2 &= E[Z_1 + 0 \wedge (Y - s\sigma_2))]^2 = E[Z_1 + Y - Y \vee s\sigma_2]^2 \\
&= \int_{-\infty}^{\infty} E[U - y \vee s\sigma_2]^2 n_{(0,\sigma_2^2(t))}(y)\, dy = \int_{-\infty}^{\infty} [\sigma^2 + (\sigma_2 u \vee s\sigma_2)^2]\phi(u)\, du \\
&= \sigma^2 + \int_{-\infty}^{\infty} [s^2\sigma_2^2 I(u \le s) + \sigma_2^2 u^2 I(u > s))]\phi(u)\, du \\
&= \sigma^2 + s^2\sigma_2^2\Phi(s) + \frac{\sigma_2^2}{2} - \sigma_2^2 \int_0^s u^2\phi(u)\, du \\
&= \sigma_2^2[1 - F(t)] + s^2\sigma_2^2\Phi(s) + \frac{\sigma_2^2}{2} - \sigma_2^2 \int_0^s u^2\phi(u)\, du \\
&= \sigma_2^2 \left[1 - F(t) + s^2\Phi(s) + \frac{1}{2} - \int_0^s u^2\phi(u)\, du\right] \\
&= \frac{\sigma_1^2}{2 - F(t)} \left\{\frac{3}{2} - F(t) + s^2\Phi(s) - \left[s^2\Phi(s) - \int_0^s 2u\Phi(u)\, du\right]\right\} \\
&= \sigma_1^2 \left\{1 - \frac{1}{2 - F(t)}\left[\frac{1}{2} - \int_0^s 2u\Phi(u)\, du\right]\right\},
\end{aligned}$$

where we have used integration by parts to obtain the second from the last equality. Note that the last expression is a strictly increasing function of s; it is less than σ_1^2 when $s = 0$ and equals σ_1^2 when $s = s_0$, where s_0 is the unique solution of $\int_0^s u\Phi(u)\, du = 1/4$. Evaluating numerically, we get $s_0 = 0.84$. This corresponds to

$$\Delta(F(t)) = s_0\sigma_2 = \sigma_1 \frac{s_0}{\sqrt{2 - F(t)}} = s_0\sqrt{F(t)}/2 = 0.42\sqrt{F(t)}.$$

Thus, if $0 \le F_{1n}(t) - F_{2n}(t) < 0.84\sqrt{F(t)}/\sqrt{n}$, and F_{1n} and F_{2n} are CIF's, then the restricted estimators reduce the AMSE if $0 < F(t) < 1$. ∎

We note that $\sigma_2(t) = \sqrt{F(t)}/2$ turns out to be the natural unit for measuring $\Delta(F(t))$ for reduction of AMSE's for all t such that $0 < F(t) < 1$; the natural unit was $\sigma(t) = \sqrt{F(t)[1 - F(t)]}/2$, which is less than $\sigma_2(t)\ \forall t > 0$, for the stochastic ordering result in Theorem 2.3. A possible explanation for this is that the probability inequality in Theorem 2.3 seems to hold for larger pertubations, but only for larger u's, as evidenced by numerical computations of $q^*(r, u)$.

2.6 A hypothesis test

Let

$$H_0 : F_1 = F_2 \text{ and } H_1 : F_1 \le F_2, \tag{2.10}$$

and consider testing H_0 against $H_1 - H_0$. It is natural to reject H_0 for large values of $\sqrt{n}\,\sup_t [F_2^*(t) - F_1^*(t)]$, which is equal to $\sup_t [Z_{2n}^*(t) - Z_{1n}^*(t)]$ when $F_1 = F_2$. Assume that the null hypothesis is true. For an asymptotic test we use the limiting distribution of $Z_{2n}^* - Z_{1n}^*$ under H_0. Using part (2) of Theorem 2.2, and the covariance formula (2.5) with $F_1 = F_2 = F/2$, we see that $Z_{2n}^* - Z_{1n}^* \overset{w}{\Longrightarrow} W_2 - W_1 = (Z_2 - Z_1) \vee 0$, where $Z_2 - Z_1$ is a mean-zero Gaussian process with

$$Cov(Z_2(s) - Z_1(s), Z_2(t) - Z_1(t)) = F(s) \text{ for } s \leq t.$$

Thus $Z_{2n}^*(t) - Z_{1n}^*(t)$ has asymptotically the same distribution as $B(F(t)) \vee 0$, where B is a standard Brownian motion. Now, for $y \geq 0$,

$$P\left[\sup_{0\leq t\leq 1} B(t) \vee 0 > y\right] = 2(1 - \Phi(y)), \tag{2.11}$$

where Φ is the DF of a standard normal variable, and we get a distribution-free test. Aly, Kochar and McKeague (1994) derived the same asymptotic test using the fact that T and δ are independent when $F_1 = F_2$ (Kochar and Proschan, 1991), and a martingale formulation.

2.7 Confidence bands

It may be of interest to find $(1-\alpha)$-coefficient *simultaneous* confidence bands for F_1 and F_2 on $[0, b]$ for some $b < T_F$. We consider the Kolmogorov-Smirnov bands (cf. equation (2.11)) and use the Bonferroni procedure. Using the stochastic ordering result in Theorem 2.3, we could use the bands $[L_1, U_1]$ and $[L_2, U_2]$, respectively, where

$$L_j(\cdot)(U_j(\cdot)) = F_j^*(\cdot) - (+) z_{\alpha/4}\sqrt{\widehat{F}_j(\cdot)[1 - \widehat{F}_j(\cdot)]/n}, \; j = 1, 2, \tag{2.12}$$

with $z_\beta = \Phi^{-1}(1-\beta)$, which provide more conservative bands than in the unrestricted case obtained by replacing F_j^* by $\widehat{F}_j$ in (2.12). Note that a simultaneous confidence region of the form $[L_1, U_1] \times [L_2, U_2]$ could be possibly reduced by intersecting it with the set $A = \{(x_1, x_2) \in \mathcal{R}^2 : x_1 \leq x_2\}$. For a rectangular region we use

$$[L_1, U_1 \wedge U_2] \times [L_1 \vee L_2, U_2]. \tag{2.13}$$

3 Censored case

Here we consider the case when there is censoring in addition to the two competing risks. It is important that the censoring mechanism, that may be a combination of other competing risks, be independent of the two of interest. We now identify three causes of failure, $\delta = 0, 1$, or 2, where $\{\delta = 0\}$ is the event that the subject was censored.

Let C_i denote the censoring time, assumed continuous, for the ith subject, and let $L_i = T_i \wedge C_i$. We assume that $\{C_i\}$ are identically and independently distributed (IID) with SF, S_C, and are independent of the life distributions, $\{T_i\}$. For the ith subject we observe (L_i, δ_i), the time and cause of the failure of the ith subject. Here the $\{L_i\}$ are IID by assumption.

3.1 The estimators and consistency

For $j = 1, 2$, let Λ_j be the cumulative hazard function for risk j, and let $\Lambda = \Lambda_1 + \Lambda_2$ be the cumulative hazard function of the life distribution T. For the censored case, the natural (unrestricted) estimators of the CIF's are the sample equivalents of (1.2) using the Kaplan-Meier (1958) estimator,$\widehat{S}$, of S:

$$\widehat{F}_j(t) = \int_0^t \widehat{S}(u)\, d\widehat{\Lambda}_j(u), \; j = 1, 2, \text{ with } \widehat{F} = \widehat{F}_1 + \widehat{F}_2, \tag{3.1}$$

where $\widehat{S}$ is chosen to be the left-continuous version for technical reasons, and $\widehat{\Lambda}_j$ is the Nelson-Aalen estimator (see, e.g., Fleming and Harrington, 1991) of Λ_j, given by

$$\widehat{\Lambda}_j(t) = \sum_{i=1}^{n} \frac{I(L_i \le t, \delta_i = j)}{\sum_{s=1}^{n} I(L_s \ge L_i)}, \; j = 1, 2. \tag{3.2}$$

Although our estimators use the Kaplan-Meier estimator of S rather than the empirical, we continue to use the same notation for the various estimators and related entities as in the uncensored case for notational simplicity.

As in the unrestricted case, we define our restricted estimators by

$$F_1^* = \widehat{F}_1 \wedge \widehat{F}/2, \; F_2^* = \widehat{F}_2 \vee \widehat{F}/2, \text{ and } \widehat{F}^* = F_1^* + F_2^* = \widehat{F}. \tag{3.3}$$

Consistency. Let T_C denote the right endpoint of the support of C. We assume that $T_C \le T_F$. Then it is well known that $\widehat{F}_1$ and $\widehat{F}_2$ given by (3.1) are strongly uniformly consistent. Thus the consistency of the restricted estimators follow using the same argument as in the uncensored case in Section (2.2).

3.2 Weak convergence

Let

$$Z_n = \sqrt{n}[\widehat{F} - F], \; Z_{jn} = \sqrt{n}[\widehat{F}_j - F_j] \text{ and } Z_{jn}^* = \sqrt{n}[F_j^* - F_j], \; j = 1, 2,$$

that have the same form as (2.3) in the uncensored case except that the unresticted estimators have been obtained via (3.1). Let

$$\pi(t) = P[L_i \ge t] = P[T_i \ge t, C_i \ge t] = S(t) S_C(t).$$

Using a counting process-martingale formulation, Lin (1997) has shown that $Z_{1n} \overset{w}{\Longrightarrow} Z_1$ on $[0, b]$ for all $b < T_F$, where Z_1 is a mean-zero Gaussian process with

$$\begin{aligned} Cov(Z_1(s), Z_1(t)) &= \int_o^s [1 - F_1(s) - F_2(u))][1 - F_1(t) - F_2(u))] \frac{d\Lambda_1(u)}{\pi(u)} \\ &\quad + \int_0^s [F_1(u) - F_1(s)][F_1(u) - F_1(t)] \frac{d\Lambda_2(u)}{\pi(u)}. \end{aligned} \tag{3.4}$$

From this one can show that $Z_{2n} \overset{w}{\Longrightarrow} Z_2$ and $Z_n \overset{w}{\Longrightarrow} Z$, both mean-zero Gaussian processes, with covariances found by switching subscripts 1 and 2 in (3.4) in the first case, and by setting $F_2 = 0$, $\Lambda_2 = 0$, $F_1 = F$ and $\Lambda_1 = \Lambda$, in the second case. Now, using the argument of Breslow and Crowley (1974), these marginal convergences imply the following joint convergence, as in Section 2.3.

Theorem 3.1 *The trivariate process* $(Z_{1n}, Z_{2n}, Z_n) \overset{w}{\Longrightarrow} (Z_1, Z_2, Z = Z_1 + Z_2)$ *on* $[0,b]\times[0,b]\times[0,b]$ *for all* $b < T_F$*, where* Z_1, Z_2*, and* Z *are mean-zero Gaussian precesses, and, for* $s \le t$*, their covariances are given by*

$$\begin{aligned} Cov(Z(s), Z(t)) &= S(s)S(t)\int_0^s \frac{d\Lambda(u)}{\pi(u)}, & (3.5)\\ Cov(Z_1(s), Z_2(t)) &= \int_0^s [1 - F_1(s) - F_2(u)][F_2(u) - F_2(t)]\frac{d\Lambda_1(u)}{\pi(u)} \\ &\quad + \int_0^s [1 - F_2(t) - F_1(u)][F_1(u) - F_1(s)]\frac{d\Lambda_2(u)}{\pi(u)}; & (3.6) \end{aligned}$$

the other covariances being computable from (3.4)-(3.6), and the fact that $Z = Z_1 + Z_2$*.*

Again, using the same notation as in the unrestricted case, we have

$$\begin{aligned} Z^*_{1n} &= Z_{1n} + 0 \wedge [(Z_{2n} - Z_{1n})/2 + \sqrt{n}(F/2 - F_1)], \\ Z^*_{2n} &= Z_{2n} + 0 \vee [(Z_{1n} - Z_{2n})/2 + \sqrt{n}(F/2 - F_2)], \end{aligned}$$

and, using the same proof as that of Theorem 2.2, we get

Theorem 3.2 *The conclusions of Theorem 2.2 hold using the Kaplan-Meier estimator of* S *instead of the empirical, with the covariance structure obtainable from Theorem 3.1 the same way as the covariance structure of Theorem 2.2 was derived from (2.4).*

We obtain the same stochastic ordering result in Section 2.4 in the censored case under the contiguous alternatives, (2.8), by observing the following. From (3.4)-(3.6), with $Y = (Z_2 - Z_1)/2$, it can be shown after some algebra that

$$\sigma_1^2(t) \equiv Var(Z_1(t)) = \frac{1}{2}\int_0^t [S^2(t) + S^2(u)]\frac{d\Lambda(u)}{\pi(u)},$$

$$\sigma_2^2(t) \equiv Var(Y(t)) = \frac{1}{2}\int_0^t S^2(u)\frac{d\Lambda(u)}{\pi(u)},$$

$$\sigma_{12}(t) \equiv Cov(Z_1(t), Y(t)) = -\frac{1}{2}\int_0^t S^2(u)\frac{d\Lambda(u)}{\pi(u)}, \text{ and}$$

$$\rho(t) \equiv Corr(Z_1(t), Y(t)) = -\frac{\sigma_2(t)}{\sigma_1(t)}.$$

This shows that the conditional distribution of $Z_1(t)$ given $Y(t) = y$ is $N(-y, \sigma^2(t))$, where $\sigma^2(t) = \sigma_1^2(t)(1 - \rho^2(t))$, which is exactly the same result as in the uncensored case except for the fact that the formulas for $\sigma_1^2(t), \sigma_2^2(t), \sigma_{12}(t), \rho(t)$ and $\sigma(t)$ are different. However, the proof of Theorem 2.3 is independent of the values of these parameters. By a similar argument, the results of Section 2.5 continue to hold for the censoring case using the formulas for the parameters in this case.

As Lin (1997) has noted, the form of the covariance in (3.4) is too complicated to compute $P[\sup_t Z_1(t) > x]$ to construct exact asymptotic confidence bands as was done in the unrestricted case using (2.11). He has given several procedures for constructing approximate confidence bands for F_1. Naturally, these could also be used for confidence bands for F_2. As in the uncensored case, we can get more conservative bands by centering at F_j^* instead of $\widehat{F}_j$ in these bands, and we could possibly reduce the widths of the *simultaneous* confidence bands using (2.13).

3.3 Hypothesis test

Consider the same test as in (2.10) using censored observations. Again it is natural to reject H_0 for large values of $\sup_t[Z_{2n}^*(t) - Z_{1n}^*(t)]$. Now $Z_{2n}^*(t) - Z_{1n}^*(t)$ has the limiting mean-zero Gaussian distribution of $W_2 - W_1 = (Z_2 - Z_1) \vee 0$ under $H_0 : F_1 = F_2 = F/2$ in the notation of Theorem 2.2, but using the Kaplan-Meier estimators. ¿From the covariance formulas in (3.4)-(3.6), we get a somewhat simplified expression for

$$Cov(Z_1(s) - Z_2(s), Z_1(t) - Z_2(t)) = \int_0^s S(u)\, \frac{d\Lambda(u)}{S_C(u)},$$

using the fact that $\pi(u) = S(u)S_C(u)$ by our independence assumption. Since the Kaplan-Meier estimator of S_C, $\widehat{S}_C$, converges uniformly to S_C wp1, we have

$$\sqrt{n}\int_0^t \sqrt{\widehat{S}_C(u)}\, d[\widehat{F}_2(u) - \widehat{F}_1(u)] \to_d B(F(t)),\ 0 \le t < T_F. \tag{3.7}$$

Note that the test statistic on the *lhs* of (3.7) reduces to that of the uncensored case when $S_C \equiv 1$. Thus we can use the same asymptotic test as in the unrestricted case with the addition of the term $\sqrt{\widehat{S}_C(u)}$ in (3.7). This is the same test derived by Aly, Kochar and McKeague (1994) using a different method.

4 Simulations

Since there are no mathematically tractable parametric families of continuous non-independent bivariate distributions where the CIF's are ordered but not the hazard rates, we have chosen Block and Basu's (1974) absolutely continuous bivariate exponential distribution and one of Gumbel's (1960) bivariate exponential distributions for our simulation study.

If (X, Y) has Block and Basu's distribution, then its joint density is given by

$$f(x,y) = \begin{cases} \frac{\alpha_1\alpha(\alpha_2+\alpha_0)}{\alpha_1+\alpha_2}\exp\{-\alpha_1 x - (\alpha_2+\alpha_0)y\}, & x < y, \\ \frac{\alpha_2\alpha(\alpha_1+\alpha_0)}{\alpha_1+\alpha_2}\exp\{-\alpha_2 y - (\alpha_1+\alpha_0)x\}, & x > y, \end{cases}$$

where $\alpha_0 \geq 0$, $\alpha_1 > 0$ and $\alpha_2 > 0$ are parameters, and $\alpha = \alpha_0 + \alpha_1 + \alpha_2$. The cause specific hazard rates are

$$\lambda_j(t) \equiv \frac{\alpha_j \alpha}{\alpha_1 + \alpha_2},\ j = 1, 2,$$

are proportional, and α_0 is the dependence parameter, with X and Y independent if and only if $\alpha_0 = 0$. Since the problem is scale independent, we have kept α_1 fixed at 1, and carried out the simulations for various values of α_0 and α_2, computed the biases of all the estimators, and the ratios of the MSE's of the unrestricted to those of the corresponding restricted estimators at all deciles of the distribution of $(X \wedge Y)$. Many different sample sizes were used and all simulations were done with 10,000 iterations. For brevity, we report only a few of the cases. In both the uncensored case and the censored case we present the simulation results for $\alpha_0 =$ 1 and $\alpha_2 = 1$, 1.5 and 2, with a sample size of 50. The censoring variable was the standard exponential, corresponding to a censoring of 25%, 22.2% and 20%, respectively, in the 3 cases.

For the Gumbel distribution, (X, Y) has the survival function,

$$S(x, y) = \exp\{-ax - by - cxy\},$$

with nonnegative values of the parameters. In both the uncensored and the censored cases we consider the parameter values of $a = c = 1$ with $b = 1$ $(F_1 = F_2)$, 1.5 and 2. A censoring distribution of $\exp\{0.85\}$ was used; the amounts of censoring are 25.18%, 22.23% and 19.86%, respectively. In all cases a sample size of 50 was used with 10,000 iterations.

As expected, F_1^* shows a negative bias while F_2^* shows a positive bias (except for one case that is probably due to sampling error), although not very much. The MSE for the restricted estimators are uniformly smaller than the unrestricted ones in all cases. The gain in terms of MSE goes up as F_2 gets closer to F_1 when order reversals are more likely.

We also carried out the same simulations for Block and Basu's (1974) bivariate exponential distribution with $\alpha_0 = \alpha_2 = 1$, and $\alpha_1 = 1 + .01j$, $j = 1, \ldots, 10$, without censoring, to assess the effect of the violation of our order restriction on the MSE. For $\alpha_1 = 1.01, 1.02$ and 1.03, the ratio $\text{MSE}(\widehat{F}_j)/\text{MSE}(F_j^*)$ seems to be uniformly larger than 1; then it starts falling off starting at $\alpha_1 = 1.04$. We present the results for the cases of $\alpha_1 = 1.03, 1.04$ and 1.05 in Table 3.

Table 1.
Comparison of bias (B) and MSE of $\widehat{F}_1$, F_1^*, $\widehat{F}_2$ and F_2^* at q-quantiles of $(X \wedge Y)$ for Block and Basu distribution, with $n = 50$ and 10,000 iterations.

q	$B(\widehat{F}_1)$	$B(F_1^*)$	$\frac{MSE(\widehat{F}_1)}{MSE(F_1^*)}$	$B(\widehat{F}_2)$	$B(F_2^*)$	$\frac{MSE(\widehat{F}_2)}{MSE(F_2^*)}$
		uncensored	$\alpha_0 = 1,\ \alpha_1 = 1,\ \alpha_2 = 1$			
.1	0.0000	-.0086	1.5322	-.0001	0.0085	1.2117
.2	-.0001	-.0124	1.5192	0.0002	0.0122	1.2737
.5	-.0003	-.0202	1.5655	0.0001	0.0199	1.4461
.8	-.0003	-.0254	1.7499	0.0001	0.0253	1.6874
.9	-.0002	-.0268	1.8631	0.0006	0.0272	1.8010
		uncensored	$\alpha_0 = 1,\ \alpha_1 = 1,\ \alpha_2 = 1.5$			
.1	0.0002	-.0044	1.3940	0.0006	0.0052	1.1942
.2	0.0005	-.0044	1.3207	0.0001	0.0049	1.1972
.5	0.0007	-.0031	1.2089	0.0004	0.0042	1.1567
.8	0.0003	-.0025	1.1551	-.0001	0.0027	1.1384
.9	0.0006	-.0019	1.1451	-.0004	0.0021	1.1359
		uncensored	$\alpha_0 = 1,\ \alpha_1 = 1,\ \alpha_2 = 2$			
.1	-.0001	-.0028	1.2895	-.0002	0.0026	1.1492
.2	-.0002	-.0023	1.1748	-.0002	0.0019	1.1079
.5	-.0008	-.0017	1.0615	0.0000	0.0009	1.0450
.8	-.0014	-.0017	1.0221	0.0011	0.0014	1.0195
.9	-.0013	-.0015	1.0179	0.0013	0.0015	1.0164
		25% censoring	$\alpha_0 = 1,\ \alpha_1 = 1,\ \alpha_2 = 1$			
.1	0.0002	-.0085	1.5328	0.0004	0.0090	1.2062
.2	-.0002	-.0128	1.4989	0.0004	0.0129	1.2710
.5	0.0008	-.0204	1.5613	0.0002	0.0213	1.4288
.8	0.0006	-.0274	1.7404	0.0011	0.0291	1.6090
.9	-.0003	-.0308	1.8190	0.0008	0.0313	1.7159
		22.2% censoring	$\alpha_0 = 1,\ \alpha_1 = 1,\ \alpha_2 = 1.5$			
.1	-.0001	-.0047	1.3971	0.0000	0.0047	1.2028
.2	0.0000	-.0051	1.3304	0.0003	0.0053	1.2009
.5	-.0003	-.0050	1.2433	-.0003	0.0044	1.1733
.8	-.0001	-.0042	1.2016	-.0012	0.0029	1.1741
.9	0.0000	-.0042	1.2040	-.0010	0.0032	1.1855
		20% censoring	$\alpha_0 = 1,\ \alpha_1 = 1,\ \alpha_2 = 2$			
.1	0.0001	-.0026	1.2945	0.0000	0.0027	1.1413
.2	0.0000	-.0023	1.1853	0.0002	0.0025	1.1097
.5	0.0001	-.0011	1.0759	-.0001	0.0011	1.0541
.8	0.0002	-.0004	1.0409	0.0008	0.0014	1.0356
.9	0.0002	-.0004	1.0350	-.0001	0.0004	1.0324

TABLE 2.
Comparison of bias (B) and MSE of $\widehat{F}_1$, F_1^*, $\widehat{F}_2$ and F_2^* at q-quantiles of $(X \wedge Y)$ for Gumbel distribution, with $n = 50$ and 10,000 iterations.

q	$B(\widehat{F}_1)$	$B(F_1^*)$	$\frac{MSE(\widehat{F}_1)}{MSE(F_1^*)}$	$B(\widehat{F}_2)$	$B(F_2^*)$	$\frac{MSE(\widehat{F}_2)}{MSE(F_2^*)}$
		uncensored	$a = 1, b = 1, c = 1$			
.1	-.0004	-.0089	1.5308	0.0004	0.0089	1.2020
.2	-.0003	-.0128	1.5169	0.0004	0.0129	1.2974
.5	-.0005	-.0201	1.5486	0.0001	0.0198	1.4542
.8	0.0002	-.0246	1.7263	0.0002	0.0250	1.6821
.9	0.0003	-.0264	1.8459	-.0004	0.0263	1.8282
		uncensored	$a = 1, b = 1.5, c = 1$			
.1	-.0025	-.0072	1.3635	-.0004	0.0043	1.2067
.2	0.0002	-.0048	1.3521	-.0001	0.0049	1.2102
.5	-.0001	-.0048	1.2367	-.0003	0.0045	1.1922
.8	0.0002	-.0042	1.2294	-.0003	0.0042	1.2130
.9	0.0002	-.0041	1.2322	-.0005	0.0038	1.2222
		uncensored	$a = 1, b = 2, c = 1$			
.1	0.0000	-.0027	1.2830	0.0006	0.0033	1.1390
.2	0.0002	-.0020	1.1846	0.0007	0.0030	1.1086
.5	-.0004	-.0015	1.0725	0.0008	0.0018	1.0526
.8	-.0015	-.0021	1.0463	-.0003	0.0003	1.0399
.9	-.0008	-.0014	1.0427	0.0009	0.0015	1.0398
		25.18% censoring	$a = 1, b = 1, c = 1$			
.1	-.0002	-.0091	1.5524	-.0001	-.0088	1.2163
.2	0.0001	-.0128	1.4901	-.0003	0.0127	1.2815
.5	-.0005	-.0216	1.5431	-.0003	0.0209	1.4266
.8	-.0005	-.0285	1.7140	0.0002	0.0282	1.6457
.9	-.0002	-.0309	1.8081	0.0000	0.0308	1.7570
		22.23% censoring	$a = 1, b = 1.5, c = 1$			
.1	-.0018	-.0066	1.3681	0.0008	0.0055	1.1992
.2	0.0009	-.0043	1.4048	0.0008	0.0060	1.2049
.5	0.0009	-.0047	1.2831	-.0006	0.0050	1.2085
.8	0.0006	-.0051	1.2674	-.0005	0.0052	1.2298
.9	0.0001	-.0059	1.2860	0.0002	0.0063	1.2517
		19.86% censoring	$a = 1, b = 2, c = 1$			
.1	-.0003	-.0030	1.2924	-.0001	0.0027	1.1398
.2	-.0001	-.0024	1.1860	0.0004	0.0019	1.0684
.5	0.0004	-.0011	1.0931	0.0004	0.0019	1.0684
.8	-.0009	-.0019	1.0685	-.0015	-.0004	1.0569
.9	-.0001	-.0012	1.0696	0.0000	0.0011	1.0611

TABLE 3.
Comparison of bias (B) and MSE of $\widehat{F}_1$, F_1^*, $\widehat{F}_2$ and F_2^* at q-quantiles of $(X \wedge Y)$ for Block and Basu distribution with order reversals, with $n = 50$ and 10,000 iterations.

q	$B(\widehat{F}_1)$	$B(F_1^*)$	$\frac{MSE(\widehat{F}_1)}{MSE(F_1^*)}$	$B(\widehat{F}_2)$	$B(F_2^*)$	$\frac{MSE(\widehat{F}_2)}{MSE(F_2^*)}$
		uncensored	$\alpha_0 = 1,\ \alpha_1 = 1.03,\ \alpha_2 = 1$			
.1	0.0003	-.0710	1.2012	0.0000	0.0713	1.1910
.2	0.0000	-.0648	1.2428	-.0002	0.0646	1.1947
.5	-.0010	-.0455	1.3850	0.0003	0.0448	1.1797
.8	-.0010	-.0239	1.5016	0.0004	0.0233	1.1377
.9	-.0015	-.0154	1.5605	0.0004	0.0143	1.1004
		uncensored	$\alpha_0 = 1,\ \alpha_1 = 1.04,\ \alpha_2 = 1$			
.1	0.0006	-.0840	0.9728	-.0011	0.0836	0.9390
.2	0.0001	-.0764	1.0403	-.0015	0.0745	0.9876
.5	0.0005	-.0517	1.2877	-.0011	0.0511	1.0574
.8	-.0005	-.0260	1.5016	-.0002	0.0253	1.0797
.9	0.0002	-.0157	1.6625	-.0003	0.0156	1.0461
		uncensored	$\alpha_0 = 1,\ \alpha_1 = 1.05,\ \alpha_2 = 1$			
.1	0.0003	-.0964	0.7813	0.0003	0.0970	0.7351
.2	-.0007	-.0871	0.8914	0.0006	0.0871	0.7756
.5	0.0006	-.0576	1.1936	-.0008	0.0573	0.9054
.8	0.0001	-.0275	1.4988	0.0007	0.0282	1.0101
.9	0.0002	-.0163	1.6310	0.0004	0.0169	1.0035

5 Examples

We analyze a set of mortality data provided by Dr. H. E. Walburg, Jr. of the Oak Ridge National Laboratory and reported by Hoel (1972). The data were obtained from a laboratory experiment on 99 RFM strain male mice who had received a radiation dose of 300 rads at 5-6 weeks of age, and were kept in a conventional laboratory environment. After autopsy, the causes of death were classified as cancer, of which there were two types,- thymic lymphoma, reticulum cell sarcoma, and other causes, 39 of the 99 being classified in the last category. Since mice are known to be highly susceptible to cancer when irradiated (Kamisaku, *et al.* (1997)), we illustrate our procedure for the uncensored case considering "other causes" as cause 1 and cancer as cause 2 in our terminology of Section 2, making the assumption that $F_1 \leq F_2$. The unrestricted estimators are displayed in Figure 1, the restricted estimators and their simultaneous 80% confidence bands are displayed in Figure2.

We also considered the large sample test of $H_0 : F_1 = F_2$ against $H_1 - H_0$, where $H_1 : F_1 \leq F_2$, using the test described in Section 2.4. The value of the test statistic is 2.316 corresponding to a p-value of .0206. For the same data, Dykstra, Kochar and Robertson (1995) tested $H_0 : \lambda_1 = \lambda_2$ against $H_1 - H_0$, where $H_1 : \lambda_1 \leq \lambda_2$. They

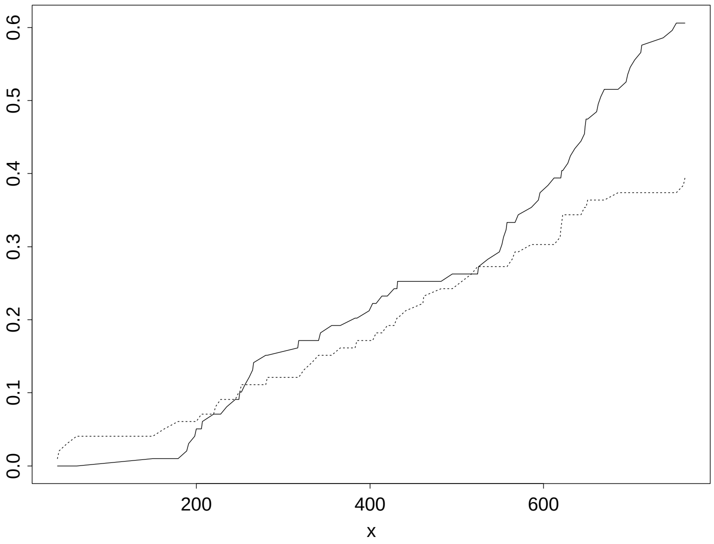

Figure 1: Unresticted estimators of cumulative incidence functions: $\widehat{F}_1$ (dotted line) and $\widehat{F}_2$ (solid line).

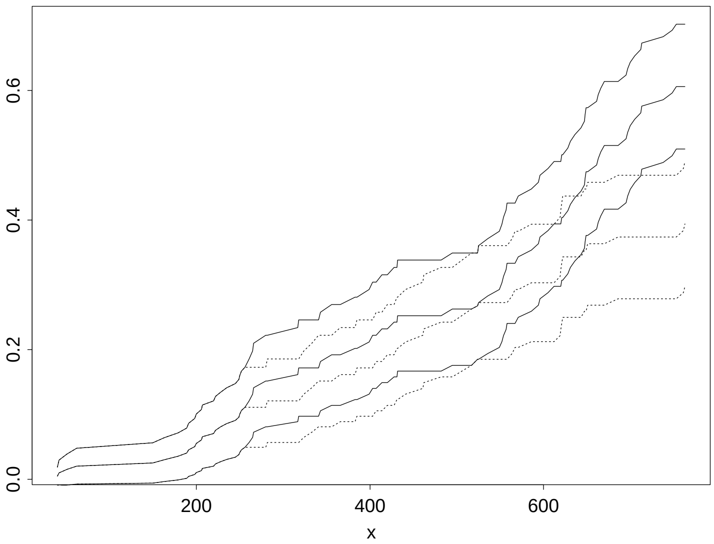

Figure 2: Restricted estimators of cumulative incidence functions and their 80% simultaneous confidence bands: F_1^* (dotted lines) and F_2^* (solid lines).

grouped the data into 6 groups and performed an asymptotic likelihood ratio test (LRT). Recently, El Barmi and Kochar (2000) used a LRT to test for ordering of cause specific survival functions (integral over (t, ∞) instead of $[0, t]$ in (1.2)) using the same grouped data. Using their technique we have performed the LRT on this grouped data for testing ordering of the CIF's. This yielded a p-value of .0447 based on the least favorable distribution. Although the results are comparable, we do not recommend grouping of continuous data for inference when methods for analyzing the complete data are available, since grouping introduces uncontrollable bias.

The same data were analyzed by Aly, Kochar and McKeague (1994) in testing whether cumulative incidence for lymphoma was larger than that for sarcoma, treating the "other causes" as a censoring mechanism, assumed to be independent of the incidence of cancer. From (3.7), their procedure and ours are identical in this case.

6 Concluding remarks

In this paper we have provided estimators of the CIF's of two competing risks under a stochasting ordering constraint, with and without censoring. The two risks of interest are not assumed to be independent, but the censoring mechanism, when existent, is assumed to be independent of the two causes of failure. We have shown that the estimators are uniformly strongly consistent. The weak convergence of the estimators have been derived. These have been used to develop asymptotic tests for equality of the CIF's against the ordered alternative. We have also provided formulas for constructing asymptotic simultaneous confidence bands that are more conservative than in the unrestricted case, and whose widths could possibly be smaller than those without the order restriction. Exact asymptotic bands are difficult to construct in the censored case because of the complicated covariance structure. However, Lin (1997) has provided several methods for approximating the bands, and we have shown how to take advantage of the order restriction to produce more conservative bands, and possibly reduce the bandwidths. Simulations seem to indicate that the MSE's of the restricted estimators are uniformly smaller than those of the unrestricted ones in all cases for finite samples when the order restriction holds. We have also analyzed a real life data set and compared our results with those in the literature.

ACKNOWLEDGMENT

The authors are grateful to a referee whose suggestions and comments substantially improved the content and the presentation of the paper.

References

Aly, E.A.A., Kochar, S.C. and McKeague, I.W. (1994). Some tests for comparing cumulative incidence functions and cause-specific hazard rates. *Journal of the American Statistal Association* **89**, 994-999.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Block, H.W. and Basu, A.P. (1974). A continuous bivariate exponential extension. *Journal of the American Statistical Association* **69**, 1031-1037.

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimate under random censorship. *Annals of Statistics* **2**, 437-453.

Brunk, H.D., Franck, W.E., Hanson, D.L. and Hogg, R.V. (1966). Maximum likelihood estimation of the distributions of two stochastically ordered random variables. *Journal of the American Statistal Association* **61**, 1067-1080.

Dykstra, D., Kochar, S. and Robertson, T. (1995). Likelihood based inference for cause specific hazard rates under order restrictions. *Journal of Multivariate Analysis* **54**, 163-174.

El Barmi, H. and Kochar, S. (2000). Inference for subsurvival functions under order restrictions. *Unpublished manuscript.*

Fleming, T.R. and Harrington, D.P. (1991). *Counting Processes and Survival Analysis.* Wiley, New York.

Gill, R. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Annals of Statistics* **11**, 49-58.

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

Gumbel, E. J. (1960). Bivariate exponential distribution. *Journal of the American Statistical Association* **55**, 698-707.

Hoel, D. G. (1972). A representation of mortality data by competing risks. *Biometrics* **28**, 475-478.

Hu, X. S. and Tsai, W. Y. (1999). Linear rank tests for competing risks model. *Statistica Sinica* **9**, 971-983.

Kamisaku, M, Aizawa, S., Kitagawa, M., Ikarashi, Y. and Sado, T. (1997). Limiting dilution analysis of T-cell progenitors in the bone marrow of thymic lymphoma susceptible B10 and resistant C3H mice after fractionated whole-body radiation. *International Journal of Radiation Biology* **72**, 191-199.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimator from incomplete observations. *Journal of the American Statistal Association* **53**, 457-481.

Kochar, S.C and Proschan, F. (1991). Independence of time and cause of failure in dependent competing risk. *Statistica Sinica* **1**, 295-299.

Lin, D.Y. (1997). Non-parametric inference for cumulative incidence functions in competing risks studies. *Statistics in Medicine* **16**, 901-910.

Lindvall, T. (1973). Weak convergence of probability measures and random functions in the function space $D[0,\infty)$. *Journal of Applied Probability* **10**, 109-121.

McKeague, I. W., Gilbert, P. B. and Kanki, P. J. (2001). Omnibus tests for comparison of competing risks. Unpublished manuscript.

Pepe, M.S. and Mori, M. (1993). Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Statistics in Medicine* **12**, 737-751.

Peterson, A.V. (1977). Expressing the Kaplan-Meier estimator as a function of empirical survival functions. *Journal of the American Statistal Association* **72**, 854-858.

Praestgaard, J. T. and Huang, J. (1996). Asymptotic theory of nonparametric estimation of survival curves under order restrictions. *Annals of Statistics* **24**, 1679-1716.

Rojo, J. (1995). On the weak convergence of certain estimators of stochastically ordered survival functions. *Nonparametric Statistics* **4**, 349-363.

Rojo, J. and Ma, Z. (1996). On the estimation of stochastically ordered survival functions. *Journal of Statistical Computation and Simulation* **55**, 1-21.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

Stone, C. (1963). Weak convergence of stochastic processes defined on a semifinite time interval. *Proceedings American Mathematical Society* **14**, 694-696.