

## CONSISTENT ESTIMATION OF DISTRIBUTIONS WITH TYPE II BIAS WITH APPLICATIONS IN COMPETING RISKS PROBLEMS

BY HAMMOU EL BARMİ AND HARI MUKERJEE

*Baruch College, City University of New York and Wichita State University*

A random variable  $X$  is symmetric about 0 if  $X$  and  $-X$  have the same distribution. There is a large literature on the estimation of a distribution function (DF) under the symmetry restriction and tests for checking this symmetry assumption. Often the alternative describes some notion of skewness or one-sided bias. Various notions can be described by an ordering of the distributions of  $X$  and  $-X$ . One such important ordering is that  $P(0 < X \leq x) - P(-x \leq X < 0)$  is increasing in  $x > 0$ . The distribution of  $X$  is said to have a Type II positive bias in this case. If  $X$  has a density  $f$ , then this corresponds to the density ordering  $f(-x) \leq f(x)$  for  $x > 0$ . It is known that the nonparametric maximum likelihood estimator (NPMLE) of the DF under this restriction is inconsistent. We provide a projection-type estimator that is similar to a consistent estimator of two DFs under uniform stochastic ordering, where the NPMLE also fails to be consistent. The weak convergence of the estimator has been derived which can be used for testing the null hypothesis of symmetry against this one-sided alternative. It also turns out that the same procedure can be used to estimate two cumulative incidence functions in a competing risks problem under the restriction that the cause specific hazard rates are ordered. We also provide some real life examples.

**1. Introduction.** Let  $X$  be a random variable with the distribution function (DF)  $F$ . A common assumption in many statistical analyses is that the distribution, possibly after a transposition of the problem, is symmetric (about 0). Many procedures are heavily dependent on this assumption. The literature contains a large number of nonparametric tests for testing this assumption. Many of these tests are variations of the sign test, Wilcoxon tests, Kolmogorov–Smirnov tests or Cramér–von Mises tests [see Chapter 22 in Shorack and Wellner (1986)]. The simplest and commonest alternatives are one-sided or two-sided shifts. More general alternatives can be obtained by introducing some ordering between the subdistribution functions (SDFs),

$$(1.1) \quad F_1(x) = F(0^-) - F(-x^-) \quad \text{and} \quad F_2(x) = F(x) - F(0), \quad x \geq 0.$$

---

Received December 2000; revised October 2002.

*AMS 2000 subject classifications.* Primary 62G05, 62G30; secondary 62G10, 62P99.

*Key words and phrases.* Type II bias, estimation, weak convergence, cumulative incidence functions, hypothesis testing, confidence bands.

Note that

$$(1.2) \quad H(x) = F_1(x) + F_2(x), \quad x \geq 0,$$

defines the DF of the life distribution,  $|X|$ , on  $[0, \infty)$  if  $P(X = 0) = 0$ .

Dykstra, Kochar and Robertson (1995) considered the testing problems

$$(1.3) \quad H_0 : F_1(x) = F_2(x) \quad \forall x$$

against the two alternatives  $H_1 - H_0$  and  $H_2 - H_0$ , where

$$(1.4) \quad H_1 : F_2(x) \geq F_1(x) \quad \forall x$$

and

$$(1.5) \quad H_2 : F_2(x) - F_1(x) \text{ is increasing in } x \quad \forall x.$$

Throughout we use increasing (decreasing) for nondecreasing (nonincreasing). Yanagimoto and Sibuya (1972) discuss the two notions of positive biasedness represented by  $H_1$ , referred to as Type I bias, and  $H_2$ , referred to as Type II bias. Dykstra, Kochar and Robertson (1995) consider only the discrete multinomial problem where the observations are assumed to be restricted to a finite grid symmetric about 0. Although their description is not couched in the language of (1.4) and (1.5), it can be seen that they are the same as (1.4) and (1.5) specialized to their discrete case. They obtain the nonparametric maximum likelihood estimators (NPMLEs) under the three hypotheses and develop the likelihood ratio tests, the test statistics having chi-bar squared (mixture of chi-squares) distributions asymptotically. They also prove strong consistencies for their estimators and point out that strong consistency still holds for the NPMLE in the continuous case under Type I bias. However, it fails to hold under Type II bias because the NPMLE assigns the mass  $1/n$  at each observation,  $X_i$ , if  $X_i > 0$ , but assigns the mass  $1/2n$  at each of  $X_i$  and  $-X_i$  if  $X_i < 0$ .

A common way for Type II bias to occur is when a distribution with a symmetric unimodal density shifts. In applications, these situations are frequently treated as slippage problems under the assumption of normality. For example, the difference in a paired comparison is often assumed to be a shifted normal. The assumption of Type II bias greatly generalizes this distributional assumption, and at the same time it improves on the purely empirical estimate as shown later. Our aim in this paper is to provide a consistent estimator under Type II bias when  $F$  is continuous. The estimator is a projection-type estimator, similar to those employed by Rojo and Samaniego (1993), Mukerjee (1996) and Arcones and Samaniego (2000) for estimating two distributions under uniform stochastic ordering, where the NPMLE fails to be consistent, as shown by Rojo and Samaniego (1991) and Mukerjee (1996). Kochar, Mukerjee and Samaniego (2000) have also used these ideas for estimating a monotone mean residual life function. We show that the estimator is strongly uniformly consistent. We also derive its weak convergence properties, find confidence bands for  $F$  and develop tests for testing  $H_0$  against  $H_2 - H_0$ .

In Section 2 we define our estimator under the order restriction in (1.5) and prove its strong uniform consistency. In Section 3 we derive the asymptotic properties of our estimator. In Section 4 we consider an application of these methods to the estimation of two cumulative incidence functions (defined in that section) under an ordering of the cause-specific hazard rates. It turns out that the mathematical analysis of this problem can be borrowed entirely from Sections 2 and 3. We also consider the same problem in the presence of censoring. In Section 5 we give some results of simulations. In Section 6 we give some real life examples. In Section 7 we provide a discussion of our results.

**2. Estimator and consistency.** We consider the estimation of  $F$  under the order restriction (1.5). To avoid trivialities we assume that  $P(X < 0) > 0$ . Throughout we assume that  $X$  is continuous. If it has a density  $f$ , then our order restriction is equivalent to the statement that  $f(x) \geq f(-x)$  for all  $x \geq 0$ . By our continuity assumption,

$$(2.1) \quad H(x) = F_1(x) + F_2(x) = F(x) - F(-x) = P(-x \leq X \leq x), \quad x \geq 0.$$

Define  $\Delta$  by

$$(2.2) \quad \Delta(x) = F_2(x) - F_1(x) = F(x) + F(-x) - 2F(0), \quad x \geq 0.$$

We can write  $F$  in terms of  $H$ ,  $\Delta$  and  $F(0)$  by

$$F(x) = \begin{cases} (\Delta(x) + H(x))/2 + F(0), & x \geq 0, \\ (\Delta(-x) - H(-x))/2 + F(0), & x < 0. \end{cases}$$

Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$  from  $F$ . Let  $F_n$  denote the usual empirical DF of  $F$ . For  $x \geq 0$ , define the following ‘‘empirical’’ estimates:

$$\begin{aligned} F_{1n}(x) &= F_n(0) - F_n(-x^-) = \frac{1}{n} \sum_i I(-x \leq X_i \leq 0), \\ F_{2n}(x) &= F_n(x) - F_n(0) = \frac{1}{n} \sum_i I(0 < X_i \leq x), \\ (2.3) \quad H_n(x) &= F_{1n}(x) + F_{2n}(x) = F_n(x) - F_n(-x^-) = \frac{1}{n} \sum_i I(-x \leq X_i \leq x), \\ \Delta_n(x) &= F_{2n}(x) - F_{1n}(x) = F_n(x) + F_n(-x^-) - 2F_n(0) \\ &= \frac{1}{n} \sum_i [I(0 < X_i \leq x) - I(-x \leq X_i \leq 0)]. \end{aligned}$$

For convenience, we have defined the estimators above to be right continuous. Under our continuity assumption, with probability 1,  $\Delta_n(0) = H_n(0) = 0$ , and

$\Delta_n$  and  $H_n$  are step functions with jumps of  $1/n$  at each  $|X_i|$ . The estimate  $F_n$  may be recovered by

$$(2.4) \quad F_n(x) = \begin{cases} (\Delta_n(x) + H_n(x))/2 + F_n(0), & x \geq 0, \\ (\Delta_n(-x^-) - H_n(-x^-))/2 + F_n(0), & x < 0. \end{cases}$$

Note that  $(F_n, F_{1n}, F_{2n}, \Delta_n, H_n)$  are unbiased estimators of  $(F, F_1, F_2, \Delta, H)$ . Also note that

$$F(0) = F_1(\infty) = [H(\infty) - \Delta(\infty)]/2 = [1 - \Delta(\infty)]/2.$$

Our order restriction corresponds to  $\Delta(x)$  is increasing in  $x$ . Thus  $\Delta(x) = \sup_{y \leq x} \Delta(y)$  for all  $x$ . Our restricted estimator of  $\Delta$  employs the sample analog of this. We define the restricted estimators by

$$(2.5) \quad \begin{aligned} H_n^*(x) &= H_n(x), & x \geq 0, \\ \Delta_n^*(x) &= \sup_{y \leq x} \Delta_n(y), & x \geq 0, \\ F_{1n}^*(x) &= (H_n(x) - \Delta_n^*(x))/2, & x \geq 0, \\ F_{2n}^*(x) &= (H_n(x) + \Delta_n^*(x))/2, & x \geq 0, \\ F_n^*(x) &= \begin{cases} (\Delta_n^*(x) + H_n(x) + 1 - \Delta_n^*(\infty))/2, & x \geq 0, \\ (\Delta_n^*(-x^-) - H_n(-x^-) + 1 - \Delta_n^*(\infty))/2, & x < 0. \end{cases} \end{aligned}$$

Note that  $F_{1n}^*(\infty) = [1 - \Delta_n^*(\infty)]/2$  is our restricted estimator of  $F(0)$ .

2.1. *Consistency.* The strong uniform consistency of  $F_n^*$  follows essentially from the Glivenko–Cantelli lemma and the triangle inequality of the sup-norm.

**THEOREM 2.1.** *Let  $F_n^*$  be defined as in (2.5). Then*

$$\sup_x |F_n^*(x) - F(x)| \xrightarrow{a.s.} 0.$$

**PROOF.** Since  $H_n$  is the empirical DF of  $|X|$ ,  $H_n^* \xrightarrow{a.s.} H$  uniformly. We also have

$$\Delta_n(x) = \frac{1}{n} \sum_i [I(0 < X_i \leq x) - I(-x \leq X_i < 0)] \xrightarrow{a.s.} \Delta(x),$$

uniformly for  $x \in [0, \infty)$ , from an application of the Glivenko–Cantelli lemma. Since  $\Delta$  is increasing on  $[0, \infty)$ , from the triangle inequality of the sup-norm [see, e.g., Lemma 2 in Rojo and Samaniego (1993)] we have

$$\sup_{x \geq 0} \left| \sup_{y \leq x} \Delta_n(y) - \sup_{y \leq x} \Delta(y) \right| \leq \sup_{y \geq 0} |\Delta_n(y) - \Delta(y)|.$$

Thus  $\sup_{x \geq 0} |\Delta_n^*(x) - \Delta(x)| \xrightarrow{a.s.} 0$ . Putting these results in (2.5) and using the continuity of  $\Delta_n$  and  $H_n$  completes the proof of the theorem.  $\square$

**3. Weak convergence.** Let  $Z_n = \sqrt{n}[F_n - F]$  on  $(-\infty, \infty)$ , and let  $Z_{1n} = \sqrt{n}[F_{1n} - F_1]$ ,  $Z_{2n} = \sqrt{n}[F_{2n} - F_2]$ ,  $Z_{3n} = \sqrt{n}[\Delta_n - \Delta]$  and  $Z_{4n} = \sqrt{n}[H_n - H]$  on  $[0, \infty)$ . Using the special construction in Theorem 3.1.1 in Shorack and Wellner (1986) for convergence of the uniform empirical and its distributional equivalence to an arbitrary empirical process in their Theorem 1.1.2, it can be seen that

$$(3.1) \quad (Z_n, Z_{1n}, Z_{2n}, Z_{3n}, Z_{4n}) \xrightarrow{w} (Z, Z_1, Z_2, Z_3, Z_4)$$

in  $(D(R), \mathcal{D}(R), \|\cdot\|) \times (D(R^+), \mathcal{D}(R^+), \|\cdot\|)^4$ , where  $\mathcal{D}$  is the  $\sigma$ -algebra generated by the finite-dimensional distributions. The limiting five-variate process is a mean-zero Gaussian process with a.s. continuous paths. Here

$$Z_1 = (Z_4 - Z_3)/2, \quad Z_2 = (Z_4 + Z_3)/2,$$

and

$$Z(x) = \begin{cases} (Z_4(x) + Z_3(x) - Z_3(\infty))/2, & x \geq 0, \\ (Z_3(-x) - Z_4(-x) - Z_3(\infty))/2, & x < 0. \end{cases}$$

Their covariances can be computed using (2.3) and (2.4). The following is a partial list with  $s \leq t$ , the latter ones being derivable from the first:

$$(3.2) \quad \begin{aligned} \text{Cov}(Z_i(s), Z_j(t)) &= F_i(s)[\delta_{ij} - F_j(t)], & i, j \in \{1, 2\}, \\ \text{Cov}(Z_3(s), Z_3(t)) &= H(s) - \Delta(s)\Delta(t), \\ \text{Cov}(Z_4(s), Z_4(t)) &= H(s)[1 - H(t)], \\ \text{Cov}(Z_3(s), Z_4(t)) &= \Delta(s)[1 - H(t)], \\ \text{Cov}(Z_4(s), Z_3(t)) &= \Delta(s) - H(s)\Delta(t), \\ \text{Cov}(Z_1(s), Z_3(t)) &= -F_1(s)F_2(t) - F_1(s)[1 - F_1(t)], \\ \text{Cov}(Z_3(s), Z_1(t)) &= -F_2(s)F_1(t) - F_1(s)[1 - F_1(t)], \\ \text{Cov}(Z_2(s), Z_3(t)) &= F_2(s)[1 - F_2(t)] + F_2(s)F_1(t), \\ \text{Cov}(Z_3(s), Z_2(t)) &= F_2(s)[1 - F_2(t)] + F_1(\infty)F_2(t), \end{aligned}$$

where  $\delta_{ij} = 1$  if  $i = j$  and 0 otherwise. We know that  $Z(x) \stackrel{d}{=} B^0(F(x))$  and  $Z_4(x) \stackrel{d}{=} B^0(H(x))$ , where  $B^0$  is a (standard) Brownian bridge. The process  $Z_3$  does not have a simple description in terms of a Brownian bridge or a Brownian motion. However, if  $\Delta$  is constant on  $[c, d]$ , then

$$(3.3) \quad \{Z_3(y) - Z_3(c) : c \leq y \leq d\} \stackrel{d}{=} \{B(H(y) - H(c)) : c \leq y \leq d\},$$

where  $B$  is a (standard) Brownian motion, and

$$(3.4) \quad \{Z_3(y) - Z_3(c) : c \leq y \leq d\} \perp \{\{Z_3(y) : y \leq c\}, \{Z_4(y) : y \geq 0\}\},$$

where  $X \perp Y$  means  $X$  and  $Y$  are independent. This can be verified by computing the covariances. In particular, when  $\Delta \equiv 0$ , that is, when  $F$  is symmetric,  $Z_3$  is a Brownian motion rescaled by  $H$ , and  $Z_3$  and  $Z_4$  are independent processes.

It is worth noting that, if  $F$  has the compact support  $[\tau_L, \tau_R]$ , with  $\tau_R \geq -\tau_L$  by our order restriction, then the process  $Z \equiv 0$  on  $(-\infty, \tau_L]$  and on  $[\tau_R, \infty)$ ,  $Z_i \equiv 0$  on  $(-\infty, 0]$  for  $1 \leq i \leq 4$ ,  $Z_4 \equiv 0$  on  $[\tau_R, \infty)$ ,  $Z_3 = Z_3(\tau_R) \equiv V \sim N(0, 1 - \Delta^2(\infty))$  on  $[\tau_R, \infty)$ , and  $Z_1 = -V/2$  and  $Z_2 = V/2$  on  $[\tau_R, \infty)$ .

Let

$$Z_n^* = \sqrt{n}[F_n^* - F], \quad Z_{in}^* = \sqrt{n}[F_{in}^* - F_i], \quad i = 1, 2,$$

and

$$Z_{3n}^* = \sqrt{n}[\Delta_n^* - \Delta].$$

Note that

$$Z_{1n}^* = (Z_{4n} - Z_{3n}^*)/2 \quad \text{and} \quad Z_{2n}^* = (Z_{4n} + Z_{3n}^*)/2.$$

We consider the weak convergence of  $(Z_n^*, Z_{1n}^*, Z_{2n}^*, Z_{3n}^*, Z_{4n})$  on  $(-\infty, \infty) \times [0, \infty)^4$ . We first derive the weak convergence of  $Z_{3n}^* = \sqrt{n}(\Delta_n^* - \Delta)$ . Define the functions  $l$  and  $u$  from  $[0, \infty)$  to  $[0, \infty)$  by

$$l(x) = \inf\{y \leq x : \Delta(y) = \Delta(x)\} \quad \text{and} \quad u(x) = \sup\{y \geq x : \Delta(y) = \Delta(x)\}.$$

By continuity of  $\Delta$ , the infimum (supremum) is actually a minimum [maximum, if  $u(x) < \infty$ ].

**THEOREM 3.1.** *Let  $Z_3^*(x) = \sup_{l(x) \leq y \leq x} Z_3(y)$ ,  $x \geq 0$ . Then*

$$Z_{3n}^* \xrightarrow{w} Z_3^* \quad \text{on } [0, \infty).$$

**PROOF.** By Lindvall (1973) and Stone (1963), it is sufficient to prove the weak convergence on  $[0, b]$  for all  $b > 0$ . Fix  $b > 0$ . For  $x \leq b$  we have

$$\begin{aligned} Z_{3n}^*(x) &= \sqrt{n} \max \left[ \sup_{l(x) \leq y \leq x} \Delta_n(y) - \Delta(x), \sup_{y \leq l(x)} \Delta_n(y) - \Delta(x) \right] \\ &= \sqrt{n} \max \left[ \sup_{l(x) \leq y \leq x} (\Delta_n(y) - \Delta(y)), \sup_{y \leq l(x)} \Delta_n(y) - \Delta(l(x)) \right]. \end{aligned}$$

By the continuous mapping theorem, the process

$$\sqrt{n} \left\{ \sup_{l(x) \leq y \leq x} (\Delta_n(y) - \Delta(y)) : 0 \leq x \leq b \right\} \xrightarrow{w} Z_3^* \quad \text{on } [0, b].$$

Since  $Z_3$  has a.s. continuous paths, a path of  $Z_3^*$  is a.s. left continuous with a jump down at the fixed points  $\{u(x)\}$  when  $l(x) < u(x) < \infty$ , and is right continuous

at 0 with a value of 0 at 0. We will show that the second term has the same limiting distribution as  $\sqrt{n}(\Delta_n \circ l - \Delta \circ l)$ , and, since

$$\sqrt{n}(\Delta_n \circ l - \Delta \circ l) \xrightarrow{w} Z_3 \circ l$$

by the continuous mapping theorem, it is dominated by  $Z_3^*$ . The paths of  $Z_3 \circ l$  are similar to those of  $Z_3^*$  except that they are constant on  $[l(x), u(x)]$  for all  $x$ , and may jump up or down at  $u(x)$ . Now, for  $0 \leq x \leq b$  and arbitrary  $\delta > 0$ ,

$$\begin{aligned} & \sqrt{n}(\Delta_n(l(x)) - \Delta(l(x))) \\ & \leq \sqrt{n} \sup_{y \leq l(x)} (\Delta_n(y) - \Delta(l(x))) \\ & \leq \sqrt{n} \sup_{(l(x)-\delta) \vee 0 \leq y \leq l(x)} (\Delta_n(y) - \Delta(y)) \\ & \quad \vee \sqrt{n} \sup_{y \leq (l(x)-\delta) \vee 0} (\Delta_n(y) - \Delta(l(x))) \\ & \leq [\sqrt{n}(\Delta_n(l(x)) - \Delta(l(x))) + w(\sqrt{n}(\Delta_n - \Delta); \delta)] \\ & \quad \vee \sqrt{n} \sup_{y \leq (l(x)-\delta) \vee 0} (\Delta_n(y) - \Delta(l(x))), \end{aligned}$$

where the modulus of oscillation,

$$w(\sqrt{n}(\Delta_n - \Delta); \delta) = \sup_{\{x,y \in [0,b]: |y-x| \leq \delta\}} |\sqrt{n}[\Delta_n(y) - \Delta(y) - \Delta_n(x) + \Delta(x)]|,$$

satisfies

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P[w(\sqrt{n}(\Delta_n - \Delta); \delta) \geq \varepsilon] = 0 \quad \forall \varepsilon > 0.$$

Let  $\tau_L$  be the left endpoint of the support of  $\Delta$ . Then  $\Delta_n(l(x)) = \Delta(l(x)) = 0$  for  $x \leq \tau_L$ . For  $x > \tau_L$ ,

$$\begin{aligned} & \sqrt{n} \sup_{y \leq (l(x)-\delta) \vee 0} (\Delta_n(y) - \Delta(x)) \\ & \leq \sqrt{n} \sup_{y \leq (l(x)-\delta) \vee 0} (\Delta_n(y) - \Delta(y)) + \sqrt{n}[\Delta((l(x) - \delta) \vee 0) - \Delta(l(x))] \\ & = O_p(1) + D_n(x, \delta), \end{aligned}$$

where the  $O_p(1)$  term is bounded above by  $\|Z_{3n}\| \xrightarrow{d} \|Z_3\|$  uniformly in  $x$ , and  $D_N(x, \delta) \rightarrow -\infty$  as  $n \rightarrow \infty$  uniformly on  $[a, b]$  for all  $\tau_L < a \leq b$ . Thus, the process  $\{O_p(1) + D_n(x, \delta) : 0 \leq x \leq b\} \xrightarrow{P} -\infty I(\tau_L, b]$ , which proves the claimed convergence of the second term.  $\square$

**THEOREM 3.2.** *The process*

$$(Z_n^*, Z_{1n}^*, Z_{2n}^*, Z_{3n}^*, Z_{4n}) \xrightarrow{w} (Z^*, Z_1^*, Z_2^*, Z_3^*, Z_4)$$

on  $(-\infty, \infty) \times [0, \infty)^4$ , where

$$(3.5) \quad Z_1^* = (Z_4 - Z_3^*)/2 \quad \text{and} \quad Z_2^* = (Z_3^* + Z_4)/2 \quad \text{on } [0, \infty),$$

and

$$(3.6) \quad Z^*(x) = \begin{cases} (Z_3^*(x) + Z_4(x) - Z_3^*(\infty))/2, & x \geq 0, \\ (Z_3^*(-x) - Z_4(-x) - Z_3^*(\infty))/2, & x < 0. \end{cases}$$

Also, for  $x \geq 0$  and for all  $t > 0$ ,

$$P(|Z_i^*(x)| > t) = P(|Z_i(x)| > t), \quad i = 1, 2, 3.$$

PROOF. The proof of the first part follows from the weak convergence of the unstarred process in (3.1) with a.s. continuous paths in the limit, Theorem 3.1 and the continuous mapping theorem.

We first prove the second part for  $i = 3$ . Note that

$$Z_3^*(x) = Z_3(l(x)) + \sup_{l(x) \leq y \leq x} [Z_3(y) - Z_3(l(x))].$$

Now,

$$Z_3(l(x)) \sim \sqrt{H(l(x)) - \Delta^2(x)} U_1,$$

where  $U_1 \sim N(0, 1)$ , and using (3.3) and (3.4), the second term is distributed as

$$(3.7) \quad \sup_{l(x) \leq y \leq x} B(H(y) - H(l(x))) \stackrel{d}{=} \sqrt{H(x) - H(l(x))} |U_2|,$$

where  $U_1$  and  $U_2$  are independent standard normals [Billingsley (1968), page 72]. We note that, if  $X$  and  $Y$  are independent mean-zero normals, then for  $t > 0$ ,

$$\{|X + |Y|| \leq t\} = \{|X + Y| \leq t, Y \geq 0\} \cup \{|X - Y| \leq t, Y < 0\},$$

and  $Y \stackrel{d}{=} -Y$ , implying  $P(|X + |Y|| \leq t) = P(|X + Y| \leq t)$ . Thus, if  $U$  is a standard normal, then

$$\begin{aligned} |Z_3^*(x)| &\stackrel{d}{=} \left| \sqrt{H(l(x)) - \Delta^2(x)} U_1 + \sqrt{H(x) - H(l(x))} |U_2| \right| \\ &\stackrel{d}{=} \sqrt{H(x) - \Delta^2(x)} |U| \\ &\stackrel{d}{=} |Z_3(x)|. \end{aligned}$$

For  $i = 1$ , from (3.5) we have

$$2Z_1^*(x) = Z_4(x) - Z_3(l(x)) - \sup_{l(x) \leq y \leq x} [Z_3(y) - Z_3(l(x))].$$



From (3.4),  $Z_4(x) - Z_3(l(x))$  is a mean-zero normal, independent of the last term on the right-hand side above. Writing  $2F_1(x) = H(x) - \Delta(x)$  in (3.2) we get

$$\text{Var}(2Z_1(x)) - \text{Var}(Z_4(x) - Z_3(l(x))) = H(x) - H(l(x)).$$

Using this, an analysis similar to the case of  $i = 3$  shows that  $|Z_1^*(x)| \stackrel{d}{=} |Z_1(x)|$ . The proof the case  $i = 2$  parallels that of the case  $i = 1$ .  $\square$

The second part of this theorem points out the interesting fact that, although  $Z_i^*(x)$  may not be equal to  $Z_i(x)$  in distribution for  $i = 1, 2, 3$ , their absolute values are. In Section 3.1 we show that  $P(|Z^*(x)| > t) \leq P(|Z(x)| > t)$  for all  $x$  and for all  $t > 0$ , the proof of which is a little more complicated.

3.1. *Asymptotic bias and MSE.* It is well known that for any DF (or SDF)  $G$ , and its empirical estimator  $G_n$ ,  $E[\sqrt{n} \sup_x |G_n(x) - G(x)|^r]$  is a bounded sequence for all  $r > 0$ . Thus Theorem 3.2 allows us to compute the asymptotic bias and the asymptotic MSE (AMSE) of the various estimators.

From (3.7), we have

$$E[Z_3^*(x)] = \sqrt{\frac{2}{\pi}} \sqrt{H(x) - H(l(x))},$$

which is zero when  $l(x) = x$  and approaches the maximum possible value of  $\sqrt{(2/\pi)}$  when  $\Delta \equiv 0$ , corresponding to  $l(x) = 0$  for all  $x$  and  $x \rightarrow \infty$ . From this we see that

$$E[Z_1^*(x)] = -\sqrt{\frac{H(x) - H(l(x))}{2\pi}} \quad \text{and} \quad E[Z_2^*(x)] = \sqrt{\frac{H(x) - H(l(x))}{2\pi}}.$$

Let  $x \geq 0$  and

$$L = \inf\{y : \Delta(y) = \Delta(\infty)\}.$$

Since  $\Delta$  is increasing, either  $L = l(x)$  or  $L > x$ . Note that  $L$  may be  $\infty$ . From (3.6), we have

$$E[Z^*(x)] = \frac{1}{\sqrt{2\pi}} [\sqrt{H(x) - H(l(x))} - \sqrt{1 - H(L)}].$$

If  $\Delta \equiv 0$ , implying  $l(x) = L = 0$ , then  $E[Z^*(x)] = -\sqrt{(1/2\pi)}$  at  $x = 0$ , and then monotonically increases to 0 as  $x \rightarrow \infty$ . If  $H(L) = 1$  then  $E[Z^*(x)] > 0$  if  $l(x) < x$ , and is equal to zero otherwise. If  $\Delta(x)$  is strictly increasing in  $x$  then  $Z^*(x) \stackrel{d}{=} Z(x)$  for all  $x$ . A similar analysis applies in the case  $x < 0$ .

From the second part of Theorem 3.2, the AMSEs of  $F_{1n}^*(x)$ ,  $F_{2n}^*(x)$  and  $\Delta_n^*(x)$  are the same as those of  $F_{1n}(x)$ ,  $F_{2n}(x)$  and  $\Delta_n(x)$ , respectively. We compare  $E[Z^*(x)]^2$  and  $E[Z(x)]^2$  via the following theorem.

THEOREM 3.3. *For all  $x \in \mathbb{R}$ , we have*

$$P(|Z^*(x)| > t) \leq P(|Z(x)| > t) \quad \forall t > 0,$$

where the inequality is strict if  $l(|x|) < |x|$  and  $0 < F(x) < 1$ .

PROOF. First, assume that  $x \geq 0$ . Now  $Z(x)$  is a mean-zero Gaussian process whose variance can be computed from the following general result that holds for  $x \leq u \leq \infty$ :

$$\begin{aligned} \sigma^2(u) &\equiv E[Z_4(x) + Z_3(x) - Z_3(u)]^2 \\ &= H(x)[1 - H(x)] + [H(x) - \Delta^2(x)] + [1 - \Delta^2(u)] \\ (3.8) \quad &+ 2\{[\Delta(x) - H(x)\Delta(x)] \\ &\quad - [\Delta(x) - H(x)\Delta(u)] - [H(x) - \Delta(x)\Delta(u)]\} \\ &= H(u) - [H(x) - \Delta(u) + \Delta(x)]^2. \end{aligned}$$

In comparing the tail probabilities of  $|Z(x)|$  and  $|Z^*(x)|$  we use the fact that, if  $U$  is a mean-zero normal, then  $P(|U - b| > t)$  is increasing in  $b > 0$  for all  $t > 0$ , as can be verified by direct computation.

If  $L = l(x)$  then  $\Delta(x) = \Delta(\infty)$  and  $\text{Var}(2Z(x)) = \sigma^2(\infty) = 1 - H^2(x)$ . Note that

$$\begin{aligned} 0 &\leq Z_3^*(\infty) - Z_3^*(x) \\ &= \left[ Z_3^*(x) \vee \sup_{y \geq x} Z_3(y) \right] - Z_3^*(x) \\ (3.9) \quad &= 0 \vee \sup_{y \geq x} [Z_3(y) - Z_3^*(x)] \\ &\leq \sup_{y \geq x} [Z_3(y) - Z_3(x)], \end{aligned}$$

where the inequality is strict with positive probability if  $l(x) < x$  and  $F(x) < 1$ . Using the independence of  $Z_4(x)$  and  $\sup_{y \geq x} [Z_3(y) - Z_3(x)]$  from (3.4), we have

$$\begin{aligned} &P(|2Z^*(x)| > t) \\ &= P(|Z_4(x) - [Z_3^*(\infty) - Z_3^*(x)]| > t) \\ &\leq P\left(\left| Z_4(x) - \sup_{y \geq x} [Z_3(y) - Z_3(x)] \right| > t\right) \\ &= P(|\sqrt{H(x)[1 - H(x)]}U_1 - \sqrt{1 - H(x)}|U_2|| > t) \\ &= P(\sqrt{1 - H^2(x)}|U| > t) \\ &= P(|2Z(x)| > t) \end{aligned}$$

for all  $t > 0$ , where  $U, U_1$  and  $U_2$  are standard normals with  $U_1 \perp U_2$ .

If  $L > x$  then, using  $Z_3^*(L) \stackrel{d}{=} Z_3(L)$ ,  $\Delta(L) = \Delta(\infty)$  and (3.9) with  $L$  in place of  $x$ , we have

$$\begin{aligned} P(2|Z^*(x)| > t) &= P(|Z_4(x) - [Z_3^*(\infty) - Z_3^*(x)]| > t) \\ &\leq P(|Z_4(x) + Z_3(x) - Z_3^*(\infty)| > t) \\ &= P(|Z_4(x) + Z_3(x) - Z_3(L) - [Z_3^*(\infty) - Z_3^*(L)]| > t) \\ &\leq P\left(\left|Z_4(x) + Z_3(x) - Z_3(L) - \sup_{y \geq L} [Z_3(y) - Z_3(L)]\right| > t\right) \\ &= P(|\sigma(L)U_1 - \sqrt{1 - H(L)}|U_2| > t) \\ &= P(\sqrt{\sigma^2(L) + 1 - H(L)}|U| > t) \\ &= P(\sigma(\infty)|U| > t) = P(2|Z(x)| > t), \end{aligned}$$

where  $U, U_1$  and  $U_2$  are standard normals with  $U_1 \perp U_2$ , and the first inequality is strict with positive probability if  $l(x) < x$ .

The proof for the case  $x < 0$  with  $F(x) > 0$  is analogous, using

$$\begin{aligned} \tau^2(u) &\equiv E[Z_3(-x) - Z_4(-x) - Z_3(u)]^2 \\ &= H(u) - [H(-x) + \Delta(u) - \Delta(-x)]^2 \end{aligned}$$

for  $-x \leq u \leq \infty$  in place of  $\sigma^2(u)$  in (3.9). However, if  $x < 0$  and  $F(x) = 0$ , then  $\text{Var}(2Z(x)) = \tau^2(\infty) = 0$ ; that is,  $Z^*(x) \stackrel{d}{=} Z(x) \stackrel{d}{=} 0$ , using the covariance formulas in (3.2). Thus the strict inequality in the theorem holds only when  $F(x) > 0$  and  $l(-x) < -x$ .  $\square$

**COROLLARY 3.4.** *For all  $x \in \mathbb{R}$ , we have  $E[Z^*(x)]^2 \leq E[Z(x)]^2$ , with strict inequality if  $l(|x|) < |x|$  and  $F(x) > 0$ .*

The proof follows directly from Theorem 3.3.

**3.2. Asymptotic tests and confidence bands.** We now consider some asymptotic inference procedures. Consider the test of  $H_0$  against  $H_2 - H_0$  as given by (1.3) and (1.5). Equivalent expressions for  $H_0$  and  $H_2$  are

$$H_0: \Delta(x) = 0 \ \forall x \geq 0 \quad \text{and} \quad H_2: \Delta(x) \text{ is increasing in } x \ \forall x \geq 0.$$

Since  $\Delta$  is nondecreasing, it has at least one point of increase if and only if

$$\sup_{x \geq 0} \sup_{0 \leq y \leq x} [\Delta(x) - \Delta(y)] > 0,$$

and a natural test criterion based on a sample of size  $n$  will be to reject  $H_0$  if the test statistic

$$T_n \equiv \sup_{x \geq 0} \sup_{0 \leq y \leq x} \sqrt{n}[\Delta_n(x) - \Delta_n(y)] \text{ is large.}$$

Since  $\Delta \equiv 0$  under  $H_0$ ,

$$\begin{aligned} T_n &\xrightarrow{d} \sup_{x \geq 0} \sup_{0 \leq y \leq x} [B(H(x)) - B(H(y))] \\ &= \sup_{0 \leq v \leq 1} \left[ B(v) - \inf_{0 \leq u \leq v} B(u) \right] \\ &\stackrel{d}{=} \sup_{0 \leq v \leq 1} |B(v)|, \end{aligned}$$

the last equality following from Lévy (1948). The limiting distribution is well known [Billingsley (1968), page 79]. Hall and Wellner (1979) have shown that the approximation of  $P[\sup_{0 \leq v \leq 1} |B(v)| \leq a]$  by  $1 - 4(1 - \Phi(a))$  gives a three-decimal place accuracy for  $a > 1.4$ ; here  $\Phi$  is the DF of a standard normal. Using this approximation, it is easy to carry out a test for  $H_0$  against  $H_2 - H_0$  at the usual levels of significance.

One might also be interested in finding an asymptotic confidence interval for  $F$  at a point or a confidence band over some interval  $[c, d]$ . Note that  $P(|Z^*(x)| \leq t) \geq P(|Z(x)| \leq t)$  for all  $t > 0$  by Theorem 3.3. Thus, if the usual confidence region of the form  $F_n(\cdot) \pm t_n(\cdot)$  in the unrestricted case is replaced by  $F_n^*(\cdot) \pm t_n(\cdot)$ , we get a conservative confidence region with the same confidence coefficient while ensuring that the order restriction (1.5) holds for the point estimator of  $F$ .

**4. An application in competing risks problems.** In this section we apply the theory developed to a problem in competing risks theory. Suppose that  $T$  is the continuous lifetime of a unit with the DF  $H$  and the survival function (SF)  $S = 1 - H$ , subject to several competing causes of failure. Let  $\delta$  denote the cause of failure; that is,  $\{\delta = j\}$  is the event that the cause of failure is  $j$ . The cumulative incidence function (CIF) due to cause  $j$  is a SDF, defined by

$$(4.1) \quad F_j(t) = P[T \leq t, \delta = j], \quad j = 1, 2, \dots,$$

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

$$(4.2) \quad \lambda_j(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} P[t \leq T < t + \Delta t, \delta = j | T \geq t], \quad j = 1, 2, \dots,$$

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

$$(4.3) \quad F_j(t) = \int_0^t \lambda_j(u) S(u) du.$$

Often it is reasonable to assume that the hazard rates are ordered,  $\lambda_1 \leq \lambda_2$ , implying risk two progresses faster than risk one. In the case of just two causes of failure, it can be seen from (4.3) that this corresponds to  $\Delta(t) \equiv F_2(t) - F_1(t)$  is increasing in  $t$ . Aly, Kochar and McKeague (1994) have developed a test for

$$(4.4) \quad H_0: \lambda_1 = \lambda_2 \quad \text{against} \quad H_2 - H_0, \quad \text{where} \quad H_2: \lambda_1 \leq \lambda_2.$$

We carry out the corresponding estimation problem under the order restriction  $H_2$ . Let  $T_1, T_2, \dots, T_n$  be a random sample from  $H$ . For  $x \geq 0$ , define the empirical estimators of  $F_1, F_2, \Delta$  and  $H$  by

$$(4.5) \quad F_{jn}(t) = \frac{1}{n} \sum_i I(T_i \leq t, \delta_i = j), \quad j = 1, 2,$$

$$H_n = F_{1n} + F_{2n} \quad \text{and} \quad \Delta_n = F_{2n} - F_{1n}.$$

Define the restricted estimators by

$$(4.6) \quad \begin{aligned} H_n^*(t) &= H_n(t), & \Delta_n^*(t) &= \sup_{s \leq t} \Delta_n(s), \\ F_{1n}^*(t) &= \frac{H_n(t) - \Delta_n^*(t)}{2} & \text{and} & \quad F_{2n}^*(t) = \frac{H_n(t) + \Delta_n^*(t)}{2}. \end{aligned}$$

By comparing the estimators in (4.5) and (4.6) with those in (2.3) and (2.5), we note that the mathematical structures of the estimators of  $F_1, F_2, \Delta$  and  $H$  are identical, although their interpretations are different. Thus, the restricted estimators are strongly uniformly consistent; the normalized processes  $(Z_{1n}, Z_{2n}, Z_{3n}, Z_{4n})$  and  $(Z_{1n}^*, Z_{2n}^*, Z_{3n}^*, Z_{4n}^*)$ , defined with the same symbols as in Section 3, have the same weak convergence properties as given by Theorems 3.1 and 3.2 (with  $Z_n^*$  and  $Z^*$  omitted); and the asymptotic bias and MSE of  $F_{1n}^*, F_{2n}^*, \Delta_n^*$  and  $H_n^*$  are the same as given in Section 3.1.

The testing problem in (4.4) is the same as the one in Section 3.2, which coincides with that obtained by Aly, Kochar and McKeague (1994). Since  $|Z_i^*|$  and  $|Z_i|$  have the same distributions by the second part of Theorem 3.2, by replacing the confidence intervals or confidence bands of the form  $F_{jn}(\cdot) \pm t_n(\cdot)$  for the unrestricted case by  $F_{jn}^*(\cdot) \pm t_n(\cdot)$ ,  $j = 1, 2$ , we get confidence intervals and confidence bands with the same confidence coefficient, while ensuring that the order restriction  $H_2$  in (4.4) holds.

4.1. *Censoring case.* We now consider the case when there is censoring in addition to the two competing risks. We identify three causes of failure,  $\delta = 0, 1$  and  $2$ , where  $\{\delta = 0\}$  is the event that the subject was censored.

Let  $C_i$  denote the censoring time for the  $i$ th subject, and let  $L_i = T_i \wedge C_i$ . We assume that the  $C_i$ s are continuous, and identically and independently distributed (i.i.d.) with SF  $S_C$ , and are independent of the life distributions  $\{T_i\}$ . We also

assume that the right endpoint of the support of  $S$ , if finite, is less than or equal to that of  $S_C$ . For the  $i$ th subject we observe  $(L_i, \delta_i)$ , the time and cause of failure of the  $i$ th subject. Here the  $L_i$ s are i.i.d. by our assumptions. Let  $\pi$  denote the SF of the  $L_i$ s; that is,  $\pi(t) = S(t)S_C(t)$  for  $t \geq 0$ , by our independence assumption.

For  $j = 1, 2$ , let  $\Lambda_j$  be the cumulative cause-specific hazard function for risk  $j$  and let  $\Lambda = \Lambda_1 + \Lambda_2$  be the cumulative hazard function for the life distribution  $T$ . Let  $\widehat{S}_n$  and  $\widehat{S}_C$  denote the Kaplan and Meier (1958) estimators of  $S$  and  $S_C$ , respectively, chosen to be left continuous for technical reasons. Define the (unrestricted) estimators of  $F_1, F_2, \Delta = F_2 - F_1$  and  $H = F_1 + F_2$  by

$$(4.7) \quad F_{jn}(t) = \int_0^t \widehat{S}_n(u) d\Lambda_{jn}(u), \quad j = 1, 2,$$

$$\Delta_n = F_{2n} - F_{1n} \quad \text{and} \quad H_n = F_{1n} + F_{2n},$$

where  $\Lambda_{jn}$  is the Nelson–Aalen estimator of  $\Lambda_j$  [see, e.g., Fleming and Harrington (1991)], given by

$$\Lambda_{jn} = \sum_{i=1}^n \frac{I(L_i \leq t, \delta_i = j)}{\sum_{s=1}^n I(L_s \geq L_i)}, \quad j = 1, 2.$$

Although we are using the Kaplan–Meier estimator  $\widehat{S}_n$  for  $S = 1 - H$  in the censored case, we continue using the same notation for the various other estimators and related entities as in the uncensored case for notational simplicity. Note that  $F_{1n}, F_{2n}, \Delta_n$  and  $H_n$  are right continuous. We define our restricted estimators under the assumption that  $\Delta$  is increasing by

$$(4.8) \quad \begin{aligned} H_n^* &= H_n, & \Delta_n^*(t) &= \sup_{s \leq t} \Delta_n(s), \\ F_{1n}^*(t) &= \frac{H_n(t) - \Delta_n^*(t)}{2} & \text{and} & \quad F_{2n}^*(t) = \frac{H_n(t) + \Delta_n^*(t)}{2}. \end{aligned}$$

Strong uniform consistency of the restricted estimators on  $[0, b]$  for all  $b$  with  $\pi(b) > 0$  follows from that of the unrestricted ones [see, e.g., Shorack and Wellner (1986), page 306, and the corrections posted on the website given in the reference] under our assumptions, as in the proof of Theorem 2.1.

We define the normalized processes  $(Z_{1n}, Z_{2n}, Z_{3n}, Z_{4n})$  and  $(Z_{1n}^*, Z_{2n}^*, Z_{3n}^*, Z_{4n}^*)$  for the unrestricted and the restricted estimators, respectively, the same way as in Section 2, and using the same symbols. In the censoring case we consider the weak convergence of these processes on  $[0, b]^4$ , where  $b$  is any constant with  $\pi(b) > 0$ . As in the uncensored case,

$$(Z_{1n}, Z_{2n}, Z_{3n}, Z_{4n}) \xrightarrow{w} (Z_1, Z_2, Z_3, Z_4) \quad \text{on } [0, b]^4,$$

where  $(Z_1, Z_2, Z_3, Z_4)$  is a four-variate Gaussian process with a.s. continuous paths.

Lin (1997) considered the estimation of  $F_1$  only in the same context, but without the order restriction. We follow the martingale formulation of the problem similar to that given in Lin (1997). For  $s \leq t \leq b$ , letting  $d\Lambda'_j(u) = d\Lambda_j(u)/\pi(u)$ ,  $j = 1, 2$ , and  $d\Lambda'(u) = d\Lambda(u)/\pi(u)$  for  $u \geq 0$ , a partial list of the covariances is given by

$$\begin{aligned}
 & \text{Cov}(Z_1(s), Z_1(t)) \\
 &= \int_0^s [1 - F_1(s) - F_2(u)][1 - F_1(t) - F_2(u)] d\Lambda'_1(u) \\
 & \quad + \int_0^s [F_1(u) - F_1(s)][F_1(u) - F_1(t)] d\Lambda'_2(u), \\
 & \text{Cov}(Z_1(s), Z_2(t)) \\
 &= \int_0^s [1 - F_1(s) - F_2(u)][F_2(u) - F_2(t)] d\Lambda'_1(u) \\
 & \quad + \int_0^s [1 - F_2(t) - F_1(u)][F_1(u) - F_1(s)] d\Lambda'_2(u), \\
 & \text{Cov}(Z_3(s), Z_3(t)) \\
 &= \int_0^s [S(u) - \Delta(u) + \Delta(t)][S(u) - \Delta(u) + \Delta(s)] d\Lambda'_1(u) \\
 (4.9) \quad & \quad + \int_0^s [S(u) + \Delta(u) - \Delta(t)][S(u) + \Delta(u) - \Delta(s)] d\Lambda'_2(u), \\
 & \text{Cov}(Z_4(s), Z_4(t)) \\
 &= S(s)S(t) \int_0^s d\Lambda'(u), \\
 & \text{Cov}(Z_3(s), Z_4(t)) \\
 &= S(t) \int_0^s [\Delta(u) - S(u) - \Delta(s)] d\Lambda'_1(u) \\
 & \quad + S(t) \int_0^s [\Delta(u) + S(u) - \Delta(s)] d\Lambda'_2(u), \\
 & \text{Cov}(Z_4(s), Z_3(t)) \\
 &= S(s) \int_0^s [\Delta(u) - S(u) - \Delta(t)] d\Lambda'_1(u) \\
 & \quad + S(s) \int_0^s [\Delta(u) + S(u) - \Delta(t)] d\Lambda'_2(u).
 \end{aligned}$$

Lin’s (1997) derivation of the weak convergence of  $Z_{1n}$  to  $Z_1$  can be adapted to find the first two covariances in (4.9).  $\text{Cov}(Z_2(s), Z_2(t))$  and  $\text{Cov}(Z_2(s), Z_1(t))$  can be obtained by interchanging the subscripts in the first two formulas. The other covariances follow from the facts that  $\Delta_n = F_{2n} - F_{1n}$  and  $H_n = F_{1n} + F_{2n}$ , using  $2F_1 = 1 - S - \Delta$ ,  $2F_2 = 1 - S + \Delta$ ,  $1 - F_1 = S + F_2$  and  $1 - F_2 = S + F_1$ .

The weak convergence results for  $Z_{3n}^*$  and  $(Z_{1n}^*, Z_{2n}^*, Z_{3n}^*, Z_{4n})$  are the same as in Theorem 3.1 and the first part of Theorem 3.2 (omitting  $Z_n^*$  and  $Z^*$ ), using the same arguments. To prove the second part of Theorem 3.2, suppose that  $\Delta$  is constant on  $[c, d]$  for some  $0 \leq c \leq d \leq b$ . Then, for  $c \leq y \leq x \leq d$ , by splitting the integrals  $\int_0^y(\cdot)$  as  $\int_0^c(\cdot) + \int_c^y(\cdot)$  in  $\text{Cov}(Z_3(y), Z_3(x))$  and using the fact that  $\Delta(u)$  is the constant  $\Delta(c)$ , for  $c \leq u \leq x$ , we have

$$\text{Cov}(Z_3(y) - Z_3(c), Z_3(x) - Z_3(c)) = \int_c^y S^2(u) d\Lambda'(u)$$

so that

$$(4.10) \quad \{Z_3(y) - Z_3(c) : c \leq y \leq d\} \stackrel{d}{=} \{B(V(y) - V(c)) : c \leq y \leq d\},$$

where  $B$  is a Brownian motion and

$$V(z) = \int_0^z S^2(u) d\Lambda'(u), \quad z \leq b.$$

This result reduces to (3.3) when there is no censoring, that is, when  $\pi = S$  and  $V = H$ . By computing covariances, we also get

$$(4.11) \quad \{Z_3(y) - Z_3(c) : c \leq y \leq d\} \perp \{\{Z_3(y) : y \leq c\}, \{Z_4(y) : y \geq 0\}\},$$

which is the same as (3.4), repeated for convenience.

Let  $\sigma^2(s)$  denote the variance of  $Z_3(s)$ ,  $s \leq b$ , and note that  $V(y) - V(s)$  in (4.10) could also be written as  $\sigma^2(y) - \sigma^2(s)$ . The proof of  $|Z_3^*(x)| \stackrel{d}{=} |Z_3(x)|$  follows by the same arguments as in Theorem 3.2 by changing  $\text{Var}(Z_3(x))$  from  $H(l(x)) - \Delta^2(x)$  to  $\sigma^2(l(x))$  and by changing (3.8) to read

$$\sup_{l(x) \leq y \leq x} B(V(y) - V(l(x))) \stackrel{d}{=} \sqrt{\sigma^2(x) - \sigma^2(l(x))} |U_2|.$$

The proof of  $|Z_j^*(x)| \stackrel{d}{=} |Z_j(x)|$ ,  $j = 1, 2$ , follows in a similar manner. Summarizing, we have the following theorem.

**THEOREM 4.1.** *For any  $b > 0$  with  $\pi(b) > 0$ ,*

$$(Z_{1n}^*, Z_{2n}^*, Z_{3n}^*, Z_{4n}) \xrightarrow{w} (Z_1^*, Z_2^*, Z_3^*, Z_4) \quad \text{on } [0, b]^4,$$



where  $Z_1^* = (Z_4 - Z_3^*)/2$  and  $Z_2^* = (Z_3^* + Z_4)/2$ . Also, for  $x \leq b$  and for all  $t > 0$ ,

$$P(|Z_j^*(x)| > t) = P(|Z_j(x)| > t), \quad j = 1, 2, 3.$$

4.1.1. *Asymptotic tests and confidence bands under censoring.* The test in Section 3.2 needs to be modified when there is censoring. In the martingale formulation, when  $\Delta = 0$ ,

$$Z_{3n}(x) = \sqrt{n} \int_0^x \widehat{S}_n(u) d(\Lambda_{2n} - \Lambda_{1n})(u)$$

which converges weakly to  $Z_3(x) \stackrel{d}{=} B(V(x))$  on  $[0, b]$  as opposed to  $B(H(x))$  in the uncensored case. This greatly inflates the argument of  $B$ . However, a rescaled version,

$$(4.12) \quad \widehat{Z}_{3n}(x) = \sqrt{n} \int_0^x \sqrt{\widehat{S}_C(u)} \widehat{S}_n(u) d(\Lambda_{2n} - \Lambda_{1n})(u)$$

converges weakly to  $B(H(x))$ , as may be verified by computing the covariances. Thus, we reject  $H_0$  in favor of  $H_2 - H_0$  if

$$\sup_{x \geq 0} \sup_{0 \leq y \leq x} [\widehat{Z}_{3n}(x) - \widehat{Z}_{3n}(y)],$$

which converges in distribution to  $\sup_{0 \leq y \leq 1} |B(y)|$ , is large.

As Lin (1997) has noted, the confidence bands for  $F_1$  and  $F_2$  using the covariance formulas in (4.9) appear intractable. He has proposed several methods of constructing confidence bands, centered at the estimators,  $F_{1n}$  and  $F_{2n}$ . From Theorem 4.1, we can obtain confidence bands with the same confidence coefficient by changing the centering to  $F_{1n}^*$  and  $F_{2n}^*$ , while retaining the order restriction.

**5. Simulations.** In this section we present a small set of a large number of simulations we have carried out comparing the bias and MSE of the restricted and the unrestricted estimators. Table 1 compares the simulated bias and MSE of  $F_n$  and  $F_n^*$  for the DF  $F$  of  $X \sim N(0.1, 1)$  at  $p$ -percentiles of the distribution using a sample size of  $n = 30$  with 10,000 replications.

In this simulation  $F(0) \doteq 0.46$ , and the bias of  $F_n^*(x)$  is largest around  $x = 0$ , decreasing as  $|x|$  becomes larger, which is commensurate with our analysis in Section 3.1. The gain in MSE by using the restricted estimator appears to be quite substantial.

For the application in Section 4 we present some results in the uncensored case using Block and Basu's (1974) absolutely continuous bivariate exponential distribution for our simulation study. If  $(X, Y)$  has this distribution, then the joint

TABLE 1  
 Comparison of bias ( $B$ ) and MSE of  $F_n$  and  $F_n^*$  at  
 $p$ -percentiles of  $X$  with  $n = 30$  and 10,000 replications

$p$	$B(F_n)$	$B(F_n^*)$	$\frac{MSE(F_n)}{MSE(F_n^*)}$
05	0.0002	0.0023	1.6994
10	0.0001	0.0020	1.7138
15	0.0008	0.0009	1.7618
20	0.0009	-0.0018	1.8262
25	0.0006	-0.0059	1.8410
30	0.0009	-0.0107	1.8780
35	0.0005	-0.0173	1.8585
40	0.0009	-0.0263	1.7668
45	0.0008	-0.0391	1.4769
50	0.0010	-0.0314	1.4689
55	0.0008	-0.0207	1.5102
60	0.0006	-0.0138	1.5123
65	0.0007	-0.0085	1.4718
70	0.0006	-0.0046	1.4412
75	0.0008	-0.0014	1.4415
80	-0.0003	0.0000	1.4286
85	-0.0003	0.0016	1.4395
90	-0.0000	0.0023	1.4700
95	0.0005	0.0024	1.5070

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

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

are proportional, and  $\alpha_0$  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  $\alpha_1$  fixed at 1, and carried out the simulations for various values of  $\alpha_0$  and  $\alpha_2$ , computed the biases of all the estimators, and the ratios of the MSEs of the unrestricted to those of the corresponding restricted estimators at all deciles of the distribution of  $(X \wedge Y)$ . We present only the cases with  $\alpha_0 = 1$  and  $\alpha_2 = 1.0, 1.5, 2.0$  and  $2.5$ . The simulations are based on 10,000 iterations in each case.

With the exception of the case  $\alpha_0 = 1$ ,  $\alpha_1 = 1$ ,  $\alpha_2 = 1$ ,  $q = 0.1$  for estimating  $F_2$ , the restricted estimators give a small improvement in the MSE over the unrestricted ones.

TABLE 2  
 Comparison of bias ( $B$ ) and MSE of  $F_{1n}$ ,  $F_{1n}^*$ ,  $F_{2n}$  and  $F_{2n}^*$  at  $q$ -quantiles of  $(X \wedge Y)$ ,  
 uncensored case with  $n = 30$  and 10,000 replications

$q$	$B(F_{1n})$	$B(F_{1n}^*)$	$\frac{MSE(F_{1n})}{MSE(F_{1n}^*)}$	$B(F_{2n})$	$B(F_{2n}^*)$	$\frac{MSE(F_{2n})}{MSE(F_{2n}^*)}$
$\alpha_0 = 1, \alpha_1 = 1, \alpha_2 = 1.0$						
0.1	0.0001	0.0003	1.4263	0.0003	0.0117	0.9490
0.2	0.0001	0.0000	1.2876	0.0000	0.0193	1.0058
0.5	-0.0007	0.0002	1.1671	0.0002	0.0344	1.0687
0.8	0.0005	-0.0012	1.1529	-0.0012	0.0448	1.1129
0.9	0.0009	-0.0007	1.1572	-0.0006	0.0483	1.1147
$\alpha_0 = 1, \alpha_1 = 1, \alpha_2 = 1.5$						
0.1	-0.0001	0.0001	1.4123	0.0001	0.0080	1.0266
0.2	0.0003	-0.0002	1.3002	-0.0002	0.0118	1.0928
0.5	-0.0002	-0.0012	1.2209	-0.0016	0.0147	1.1496
0.8	-0.0004	0.0001	1.2431	0.0001	0.0176	1.1933
0.9	-0.0002	-0.0004	1.2445	-0.0004	0.0179	1.2297
$\alpha_0 = 1, \alpha_1 = 1, \alpha_2 = 2.0$						
0.1	-0.0003	0.0003	1.3738	0.0003	0.0061	1.0309
0.2	-0.0004	0.0012	1.2721	0.0010	0.0089	1.0708
0.5	0.0002	0.0013	1.1916	0.0013	0.0107	1.0899
0.8	0.0003	0.0002	1.1415	0.0003	0.0103	1.1035
0.9	0.0001	-0.0003	1.1347	-0.0003	0.0097	1.1182
$\alpha_0 = 1, \alpha_1 = 1, \alpha_2 = 2.5$						
0.1	0.0000	-0.0001	1.3482	-0.0001	0.0044	1.0411
0.2	-0.0002	0.0000	1.2342	0.0000	0.0060	1.0722
0.5	-0.0009	0.0005	1.1194	0.0006	0.0072	1.0599
0.8	-0.0008	0.0001	1.0956	0.0002	0.0070	1.0704
0.9	-0.0006	-0.0002	1.0771	-0.0002	0.0066	1.0648

**6. Examples.** For our first example we consider the data in Table 3 from Moore and McCabe (1993) giving the pretest and posttest scores on the MLA listening test in Spanish for 20 high school teachers who attended an intensive course in Spanish.

We assume that the 20 random variables  $X_1, \dots, X_{20}$  are i.i.d. with the DF  $F$  that has a Type II positive bias, where  $X_i = (\text{posttest score} - \text{pretest score})$  for the  $i$ th subject. Estimating  $F$  by  $F_n^*$  yields Table 4. Although the sample size is only 20, if we use our asymptotic test in Section 3.2 for testing the efficacy of the training method, we get the test statistic = 1.7888 with a  $p$ -value = 0.148. A one-sided paired  $t$ -test for comparing the mean of pretest and the mean of posttest yields a  $p$ -value of 0.029. Thus it appears that the assumption of normality used in the  $t$ -test is crucial in concluding that the training method is effective.

For our second example we consider a set of mortality data [Hoel (1972)] of 99 RMF strain mice subjected to radiation in a laboratory experiment. Causes of death

TABLE 3  
Pretest and posttest data

Subject	Pretest	Posttest	Subject	Pretest	Posttest
1	30	29	11	30	32
2	28	30	12	29	28
3	31	32	13	31	34
4	26	30	14	29	32
5	20	16	15	34	32
6	30	25	16	20	27
7	34	31	17	26	28
8	15	18	18	25	29
9	28	33	19	31	32
10	20	25	20	29	32

were classified into thymic lymphoma, reticulum cell sarcoma, and other causes. For purposes of illustration we consider only the deaths from sarcoma ( $\delta = 1$ ) and lymphoma ( $\delta = 2$ ), and assume that the deaths, 39 of them, due to other causes are independent of these. The restricted and unrestricted estimators of the CIFs,  $F_1$  and  $F_2$ , are given in Figure 1.

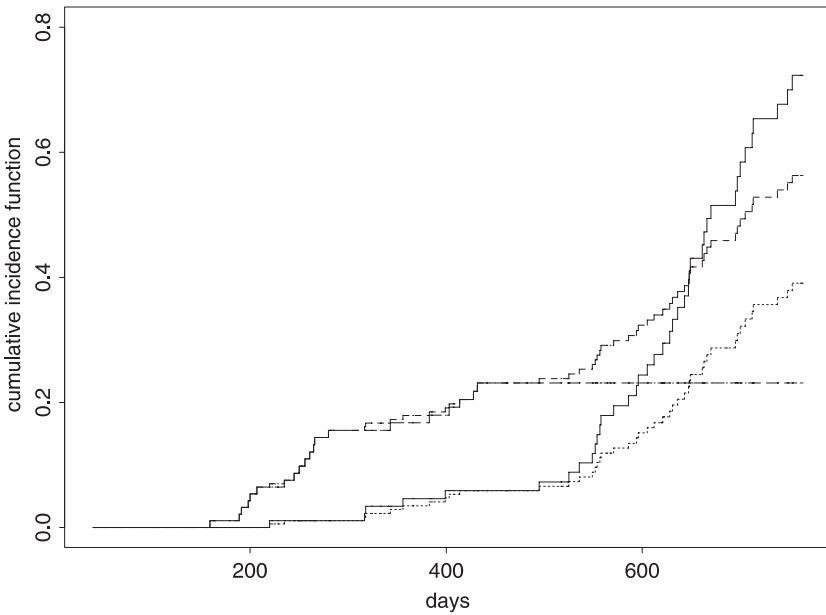


FIG. 1. Estimators of  $F_1$  and  $F_2$ :  $F_{1n}$  (solid),  $F_{2n}$  (long dash-dots),  $F_{1n}^*$  (dots),  $F_{2n}^*$  (short dash-dots).

TABLE 4  
*Estimation of  $F$  under  
 Type II positive bias*

$x$	$F_n^*(x)$
-7.00	0.000
-5.00	0.050
-4.00	0.100
-3.00	0.150
-2.00	0.200
-1.00	0.300
0.00	0.300
1.00	0.400
2.00	0.550
3.00	0.750
4.00	0.850
5.00	0.950
7.00	1.000

The unrestricted estimators appear to show that  $F_2 - F_1$  does increase until the last death due to lymphoma ( $\delta = 2$ ). For larger sample sizes the same trend might have continued as indicated by the restricted estimators. The test for this was carried out by Aly, Kochar and McKeague (1994) whose test procedure was the same as ours.

**7. Discussion.** Type II bias is a generalization of the parametric problem of a shift of the median of a symmetric distribution, and as such could be utilized in many applications. For continuous distributions, it is known that the NPMLE of a distribution function with Type II bias is inconsistent. In this paper we have proposed a projection type estimator that is strongly uniformly consistent. We have derived its weak convergence properties, and have shown that the restricted estimator improves on the empirical in terms of MSE. Asymptotic confidence bands for the DF using the unrestricted bands, but by changing the centering from the unrestricted to the restricted estimator, have been shown to be conservative for the same confidence coefficient. We have also developed procedures for testing symmetry against Type II positive bias.

As an application of the techniques, we have shown that the mathematical structure of the estimation of two cumulative incidence functions in competing risks problems, ordered by their cause specific hazard rates, is similar. Consistency and asymptotic distributions of the estimators have been derived, and asymptotic tests and confidence bands have been discussed. In this case, the asymptotic MSE is the same for the restricted and the (unrestricted) empirical estimators. We have also considered the same problem under censoring using the Kaplan–Meier estimator, and the results are similar.

We have provided some simulation results to compare the MSEs of the restricted and unrestricted estimators for finite samples for several cases. It appears that the restricted estimators are uniformly better in terms of MSE. We have also given some real life examples to illustrate our procedures.

**Acknowledgments.** The comments of two referees substantially improved the paper. The authors are particularly grateful to one referee for pointing out several deficiencies and suggestions for improvements.

## 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. *J. Amer. Statist. Assoc.* **89** 994–999.
- ARCONES, M. A. and SAMANIEGO, F. J. (2000). On the asymptotic distribution theory of a class of consistent estimators of a distribution satisfying a uniform stochastic ordering constraint. *Ann. Statist.* **28** 116–150.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- BLOCK, H. W. and BASU, A. P. (1974). A continuous bivariate exponential extension. *J. Amer. Statist. Assoc.* **69** 1031–1037.
- DYKSTRA, R., KOCHAR, S. and ROBERTSON, T. (1995). Likelihood ratio tests for symmetry against some one-sided alternatives. *Ann. Inst. Statist. Math.* **47** 719–730.
- FLEMING, T. R. and HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York.
- HALL, W. J. and WELLNER, J. A. (1979). Estimation of a mean residual life. Unpublished manuscript.
- HOEL, D. G. (1972). A representation of mortality data by competing risks. *Biometrics* **28** 475–478.
- KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.
- KOCHAR, S. C., MUKERJEE, H. and SAMANIEGO, F. J. (2000). Estimation of a monotone mean residual life. *Ann. Statist.* **28** 905–921.
- LÉVY, P. (1948). *Processus stochastiques et mouvement Brownien*. Gauthier-Villars, Paris.
- LIN, D. Y. (1997). Nonparametric 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  $D[0, \infty)$ . *J. Appl. Probab.* **10** 109–121.
- MOORE, D. S. and MCCABE, G. P. (1993). *Introduction to the Practice of Statistics*, 2nd ed. Freeman, New York.
- MUKERJEE, H. (1996). Estimation of survival functions under uniform stochastic ordering. *J. Amer. Statist. Assoc.* **91** 1684–1689.
- ROJO, J. and SAMANIEGO, F. J. (1991). On nonparametric maximum likelihood estimation of a distribution uniformly stochastically smaller than a standard. *Statist. Probab. Lett.* **11** 267–271.
- ROJO, J. and SAMANIEGO, F. J. (1993). On estimating a survival curve subject to a uniform stochastic ordering constraint. *J. Amer. Statist. Assoc.* **88** 566–572.
- SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York. (Corrections posted at [www.stat.washington.edu/jaw/RESEARCH/BOOKS/book1.html](http://www.stat.washington.edu/jaw/RESEARCH/BOOKS/book1.html).)
- STONE, C. (1963). Weak convergence of stochastic processes defined on semifinite time intervals. *Proc. Amer. Math. Soc.* **14** 694–696.

YANAGIMOTO, T. and SIBUYA, M. (1972). Test of symmetry of a one-dimensional distribution against positive biasedness. *Ann. Inst. Statist. Math.* **24** 423–434.

BARUCH COLLEGE  
CITY UNIVERSITY OF NEW YORK  
NEW YORK, NEW YORK  
USA

DEPARTMENT OF MATHEMATICS  
AND STATISTICS  
WICHITA STATE UNIVERSITY  
WICHITA, KANSAS 67260-0033  
USA  
E-MAIL: mukerjee@math.twsu.edu