Statistica Sinica Preprint No: SS-2023-0176	
Title	Estimation of A Distribution with A Bias and Its Applications to Competing Risks
Manuscript ID	SS-2023-0176
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202023.0176
Complete List of Authors	Hammou El Barmi and Zhiliang Ying
Corresponding Authors	Hammou El Barmi
E-mails	hammou.elbarmi@baruch.cuny.edu

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

Estimation of a distribution with a bias and its applications to competing risks

Hammou El Barmi[(1)] and Zhiliang Ying[(2)]

The City University of New York[(1)] *and Columbia University*[(2)]

Abstract: A random variable X is said to have a symmetric distribution function (DF) about zero if X and $-X$ have the same distribution. The estimation of such a distribution and tests for symmetry are widely studied in the literature. Some of the alternatives to symmetry describe some notion of skewness or one-sided bias in terms of an ordering of the distributions of X and $-X$. One such ordering is characterized by $r_{-X}(x) \leq r_X(x)$ for all $x > 0$ where $r_{-X}(x)$ and $r_X(x)$ are the hazard rates of $-X$ and X, respectively. This is equivalent to the ratio $P(X > x)/P(X < -x)$ being nondecreasing in $x > 0$. In this paper we derive the nonparametric maximum likelihood estimator (NPMLE) of F under this constraint and show that it is inconsistent. We then construct a new estimator and establish its consistency and weak convergence. We also develop a test for symmetry against this one-sided alternative and study the finite sample performance of this new estimator. We show through simulations that it outperforms the NPMLE in terms of mean squared error for all the distributions under consideration. We also show how to apply this approach to compare the conditional distributions (conditional on the risks) of two competing risks in a competing risks model.

Key words and phrases: Consistency, nonparametric likelihood estimator, restricted estimation, competing risks, weak convergence.

1. Introduction

A random variable X with DF F is said to have a symmetric distribution about zero if X and $-X$ have the same distribution. Symmetry of the underlying distribution is a commonly occurring assumption in many statistical analyses and the validity of some of frequently used procedures depends heavily on this assumption. This is particularly the case for several nonparametric procedures such as the Wilcoxon signed rank test. It is also the case that many statistical procedures that are based on normality are generally robust to this assumption when the underlying distribution is symmetric (Chaffin and Rhiel , 1993). For these reasons, a large number of nonparametric tests for symmetry have been developed. Many of these tests are variations of the sign test, Wilcoxon tests, Kolmogorov-Smirnov tests or Cramér-von Mises tests (Shorack and Wellner , 1986). The simplest, and commonest alternatives are one-sided or two-sided shifts.

More general alternatives have been considered by introducing the ordering

$$S_1(x) \equiv P(X > x) \geq S_2(x) \equiv P(X < -x), \quad x \geq 0, \tag{1.1}$$

or the ordering

$$P(0 < X \leq x) - P(-x \leq X < 0) \text{ is nondecreasing in } x \geq 0. \qquad (1.2)$$

The distribution of X is said to have a Type I positive bias under (1.1) and is said to have a Type II bias under (1.2). If X has a density f, then (1.2) corresponds to the density ordering, $f(-x) \leq f(x)$ for $x > 0$. For more on this, see Yanagimoto and Sibuya (1972).

Estimation under Type I positive bias was considered in Dykstra, Kochar and Robertson (1995), who studied the likelihood ratio test for symmetry against this type of bias in discrete or grouped data settings. They showed that the limiting null distribution of their test statistic is of a chi-bar square and they provided the expression of its weights. Their test is not asymptotically distribution-free but they were able to obtain the least favorable distribution. In addition, they also derived the NPMLE of F under this constraint and demonstrated that it is uniformly strongly consistent. Dykstra and Praestgaard (1996) studied its weak convergence and showed that it is asymptotically equivalent to the empirical distribution when Type I bias holds strictly. They also showed that, under symmetry, it is still $\sqrt{n}$ consistent but the limiting process is no longer Gaussian. Alfieri and El Barmi (2005) developed a new consistent and easy-to-compute estimator for the same situation. They also studied its weak convergence and devised a one-sided Kolmogorov-Smirnov type test for symmetry against this type of bias. This test is asymptotically distribution free when the underlying distribution is continuous. We note that the NPMLE under Type II bias is inconsistent. For this reason, El Barmi and Mukerjee (2004) developed a consistent estimator, established its weak convergence, and constructed a test for symmetry against this alternative.

In this paper, we consider instead the following alternative for symmetry:

$$\frac{S_1(x)}{S_2(x)} \quad \text{is nondecreasing on} \, \{x, S_2(x) > 0\}. \tag{1.3}$$

where S_1 and S_2 are defined in (1.1). Notice that

$$S(x) \equiv S_1(x) + S_2(x), \; x \geq 0,$$

defines the survival function of $|X|$. It turns out that, by viewing S_1 and S_2 as the subsurvival functions of competing risks in a two competing risks model, the theory we develop in this paper can be used to estimate these functions when their ratio is nondecreasing as well as compare the conditional distributions, conditioning on the risks. We will show how this is done and provide a real life example.

Our goal is to derive the NPMLE of F under (1.3) and show that it is inconsistent except in discrete or grouped data settings. We then provide a consistent estimator under this constraint when F is continuous. The estimator we propose is a projection-type estimator, similar to estimators employed by Rojo and Samaniego (1993) and Mukerjee (1996) for estimating two distribution functions under the uniform stochastic ordering constraint and by El Barmi and Mukerjee (2004) for estimating a distribution under the Type II bias. We will study the weak convergence of this estimator and show that when (1.3) is strict, this estimator is asymptotically equivalent to the empirical distribution. Otherwise we show that it is still $\sqrt{n}$ consistent but, the limiting process is not Gaussian.

This article is organized as follows. In Section 2, we derive the NPMLE under (1.3) and show that it is inconsistent. In Section 3, we introduce our estimator, prove its strong uniform consistency and study its weak convergence. We also provide a test for symmetry against this alternative. In Section 4, we consider an application of our approach to the estimation of two subsurvival functions under an ordering of the type in (1.3). We show that parallel results to those given in Section 3 continue to hold. In Section 5 we give two examples to illustrate the applicability of our theory. We also give simulation results to compare the finite sample performance of our estimator and the NPMLE. In Section 6 we offer some concluding remarks. Throughout we use $\xrightarrow{d}$, $\xrightarrow{P}$ and $\overset{w}{\Rightarrow}$ to denote convergence in distribution,

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

convergence in probability and weak convergence, respectively. We will also use $a \vee b$ and $a \wedge b$ to denote the maximum and the minimum of a and b, respectively. Because of the length of some proofs, the proof of Theorem 3.2 is relegated to the supplementary material.

2. Nonparametric maximum likelihood

Suppose that $\{X_1, X_2, \ldots, X_n\}$ is a random sample drawn from F which is unknown but satisfies the constraint constraint (1.3). First, we note that (1.3) is equivalent to

$$\frac{P(X > x|X > 0)}{P(-X > x| - X > 0)} \quad \text{is nondecreasing on } \{x, P(-X > x|-X > 0) > 0\}.$$

That is, $X|X > 0$ is uniformly stochastically larger than $-X| - X > 0$. Because of this and the inconsistency of the NPMLE of two uniformly stochastically ordered survival functions (Mukerjee , 1996), it has been conjunctured that the NMPLE of F in this case is inconsistent. We show this next.

Let $t_1 < t_2 < \ldots < t_m$ denote the corresponding distinct absolute values. Let also $t_0 = 0$ and $t_{m+1} = \infty$. The likelihood estimation is interpreted here in the generalized sense given in Kiefer and Wolfowitz (1956) in which

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

case the likelihood function is

$$L(F) = \left(\prod_{j=1}^{m}[F(t_j) - F(t_j^-)]^{n_{1j}}\right)\left(\prod_{j=1}^{m}[F((-t_j)) - F((-t_j)^-]^{n_{2j}}\right) \quad (2.1)$$

where $F(u-)$ is the left limit of F at u and n_{1j} and n_{2j} are the number of times t_j and $-t_j$ are observed, respectively. We take here $0^0 = 1$. In terms of S_1 and S_2, (2.1) is given by

$$L(S_1, S_2) = \left(\prod_{j=1}^{m}[S_1(t_j^-) - S_1(t_j)]^{n_{1j}}\right)\left(\prod_{j=1}^{m}[S_2(t_j^-) - S_2(t_j)]^{n_{2j}}\right) \quad (2.2)$$

Clearly if S_1 and S_2 satisfy the constraint (1.3), then the likelihood in (2.2) is not decreased if S_is are replaced by

$$S_i(t) = \sum_{j=0}^{m} S_i(t_j) I[t_j \le t < t_{j+1}], i = 1, 2.$$

Consequently it suffices to maximize

$$\begin{aligned} L(S_1, S_2) &= \left(\prod_{j=1}^{m}[S_1(t_{j-1}) - S_1(t_j)]^{n_{1j}}\right)\left(\prod_{j=1}^{m}[S_2(t_{j-1}) - S_2(t_j)]^{n_{2j}}\right) \\ &= \left(\prod_{j=1}^{m}\left[1 - \frac{S_1(t_j)}{S_1(t_{j-1})}\right]^{n_{1j}} S_1(t_{j-1})^{n_{1j}}\right)\left(\prod_{j=1}^{m}\left[1 - \frac{S_2(t_j)}{S_2(t_{j-1})}\right]^{n_{2j}} S_2(t_{j-1})^{n_{2j}}\right) \end{aligned}$$

subject to (1.3). Let $\theta_{ij} = S_i(t_j)/S_i(t_{j-1})$. Then

$$S_i(t_j) = S_i(t_0)\prod_{\ell=1}^{j}\theta_{i\ell}, i = 1, 2,$$

and (1.3) is equivalent to

$$\theta_{1j} \le \theta_{2j}, \quad j = 1, 2, \ldots, m, \quad (2.3)$$

with no constraints on $S_1(t_0)$ and $S_2(t_0)$. Using this new parametrization, we get, after rearranging the terms in $L(S_1, S_2)$,

$$L(S_1, S_2) = [S_1(t_0)]^{n_{1+}}[S_2(t_0)]^{n_{2+}} \prod_{j=1}^{m} \theta_{1j}^{n_{1j+}}(1-\theta_{1j})^{n_{1j}}\theta_{2j}^{n_{2j+}}(1-\theta_{2j})^{n_{2j}}$$

where $n_{i+} = \sum_{j=1}^{m} n_{ij}, i = 1, 2$, and $n_{ij+} = \sum_{\ell=j+1}^{m} n_{i\ell}$, for all (i, j). Here we assume that $\sum_{\emptyset} = 0$.

Since the constraints in (2.3) do not relate θ_{ij}s for different js and do not put any constraint on $(S_1(t_0), S_2(t_0))^T$, the likelihood factors into $m+1$ parts

$$L_0(S_1, S_2) \equiv [S_1(t_0)]^{n_{1+}}[S_2(t_0)]^{n_{2+}} \quad \text{and} \quad L_j(\theta_{1j}, \theta_{2j}) \equiv \theta_{1j}^{n_{1j+}}(1-\theta_{1j})^{n_{1j}}\theta_{2j}^{n_{2j+}}(1-\theta_{1j})^{n_{2j}},$$

$j = 1, 2, \ldots, m$, that can be maximized independently. Since $S_1(t_0)$ and $S_2(t_0)$ satisfy $S_1(t_k) + S_2(t_k) = 1, L_0(S_1, S_2)$ is maximized by $\tilde{S}_i(t_0) = \frac{n_{i+}}{n}, i = 1, 2$.

Notice that, under no constraints, $L_j(\theta_{1j}, \theta_{2j})$ is maximized by

$$(\hat{\theta}_{1j}, \hat{\theta}_{2j})^T \equiv \left(\frac{\hat{S}_1(t_j)}{\hat{S}_1(t_{j-1})}, \frac{\hat{S}_2(t_j)}{\hat{S}_2(t_{j-1})} \right)^T,$$

where $\hat{S}_1$ and $\hat{S}_2$ are the unrestricted empirical estimators of S_1 and S_2, respectively. Specifically, it is maximized by

$$\hat{S}_1(t) = \frac{1}{n}\sum_{j=1}^{m} n_{1j} I_{(t,\infty)}(t_j) = \frac{1}{n}\sum_{j=1}^{n} I_{(t,\infty)}(X_j) \quad \text{and}$$
$$\hat{S}_2(t) = \frac{1}{n}\sum_{j=1}^{m} n_{2j} I_{(-\infty,-t)}(-t_j) = \frac{1}{n}\sum_{i=j}^{n} I_{(-\infty,-t)}(X_j)$$

where $t \geq 0$.

Maximizing $L_j(\theta_{1j}, \theta_{2j})$ subject to $\theta_{1j} \leq \theta_{2j}$ is exactly the well known bioassay problem which is discussed in Robertson, Wright and Dykstra (1988). Its solution is

$$(\tilde{\theta}_{1j}, \tilde{\theta}_{2j})^T = E_{\boldsymbol{\omega}_j}[(\hat{\theta}_{1j}, \hat{\theta}_{2j})^T | \mathcal{K}]$$

where $E_{\boldsymbol{\omega}_j}[(\hat{\theta}_{1j}, \hat{\theta}_{2j})^T | \mathcal{K}]$ is the weighted least squares projection of $(\hat{\theta}_{1j}, \hat{\theta}_{2j})^T$ onto $\mathcal{K} = \{\mathbf{z} \in \mathbb{R}^2 : z_1 \leq z_2\}$, with weights $\boldsymbol{\omega}_j = (\hat{S}_1(t_{j-1}), \hat{S}_2(t_{j-1}))^T$. This solution can also be expressed as

$$\tilde{\theta}_{1j} = \frac{\hat{S}_1(t_j)}{\hat{S}_1(t_{j-1})} \wedge \frac{\hat{S}_1(t_{j-1})\frac{\hat{S}_1(t_1)}{\hat{S}_1(t_{j-1})} + \hat{S}_2(t_{j-1})\frac{\hat{S}_2(t_j)}{\hat{S}_2(t_{j-1})}}{\hat{S}_1(t_{j-1}) + \hat{S}_2(t_{j-1})} = \frac{\hat{S}_1(t_j)}{\hat{S}_1(t_{j-1})} \wedge \frac{\hat{S}(t_j)}{\hat{S}(t_{j-1})}$$

and

$$\tilde{\theta}_{2j} = \frac{\hat{S}_2(t_j)}{\hat{S}_2(t_{j-1})} \vee \frac{\hat{S}_1(t_{j-1})\frac{\hat{S}_1(t_1)}{\hat{S}_1(t_{j-1})} + \hat{S}_2(t_{j-1})\frac{\hat{S}_2(t_j)}{\hat{S}_2(t_{j-1})}}{\hat{S}_1(t_{j-1}) + \hat{S}_2(t_{j-1})} = \frac{\hat{S}_2(t_j)}{\hat{S}_2(t_{j-1})} \vee \frac{\hat{S}(t_j)}{\hat{S}(t_{j-1})}$$

where $\hat{S}(t) = \hat{S}_1(t) + \hat{S}_2(t)$ is the empirical estimator of $S(t)$. The NPMLE of F is then given by

$$\tilde{F}(t) = (1 - \tilde{S}_1(t))I_{[0,\infty)}(t) + \tilde{S}_2((-t)^-)I_{(-\infty,0)}(t)$$

where $\tilde{S}_i$ is obtained by setting

$$\tilde{S}_i(t_j) = \tilde{S}_i(t_0)\prod_{\ell=1}^{j}\tilde{\theta}_{i\ell}$$

and extending these estimates as right-continuous step functions.

The NPMLEs of S_1 and S_2 are therefore obtained through the isotonization of $\hat{\theta}_{1j}$ and $\hat{\theta}_{2j}$. When F is continuous, the probability of observations occurring at t_j and $-t_j$ is 0. As a result, if $\hat{\theta}_{1j} = 1$ then $\hat{\theta}_{2j} < 1$ and $\tilde{\theta}_{1j} = \tilde{\theta}_{2j}$ and if $\hat{\theta}_{1j} < 1$ then $\hat{\theta}_{2j} = 1$ and $(\tilde{\theta}_{1j}, \tilde{\theta}_{2j})^T = (\hat{\theta}_{1j}, \hat{\theta}_{2j})^T$. Therefore, for $t \geq 0$,

$$(\tilde{S}_1(t), \tilde{S}_2(t))^T = \left(\tilde{S}_1(t), \frac{\tilde{S}_1(t)\hat{S}_2(t)}{\hat{S}_1(t_0)}\right)^T.$$

Clearly these estimators are not consistent since if $\tilde{S}_1$ were consistent, then $\tilde{S}_2(t)$ will converge to $S_1(t)S_2(t)/S_1(t_0)$ and if $\tilde{S}_2$ were, then $\tilde{S}_1(t)$ will converge to $S_1(t_0)$. Hence the NPMLE $\tilde{F}$ of F is inconsistent. We note that this proof is closely related to a proof in Mukerjee (1996) of the inconsistency of NPMLEs of two DFs that are uniformly stochastically ordered. The inconsistency results from the fact that a positive observation makes equal contributions to both $\tilde{S}_1$ and $\tilde{S}_2$ whereas a negative observation affects only $\tilde{S}_2$.

Remark 1: In the discrete or grouped data cases, $\tilde{\theta}_{ij} \overset{as}{\to} \theta_{ij}$ for all (i, j) and since $\hat{S}_i(t_0) \overset{as}{\to} S_i(t_0), i = 1, 2,$ As a result $\tilde{F}$ is a consistent estimator of F.

doi:10.5705/ss.202023.0176

3. New estimator, consistency, algorithm and weak convergence

3.1 New estimators and consistency

In this section we consider the estimation of a continuous DF F under the constraint (1.3). We assume throughout that $P(X < 0) > 0$ and $P(X > 0) > 0$ to avoid trivialities.

Let $\hat{F}$ denote the empirical DF. It is clear that $\hat{F}$ is not guaranteed to satisfy (1.3). As far as we know there is no known consistent estimator of F under this constraint in the continuous case. Next we propose one and establish its weak convergence. Let $\hat{S}_1, \hat{S}_2$ and $\hat{S}$ be as defined before.

Clearly (1.3) holds if and only if

$$\psi_1(x) \equiv \frac{S_1(x)}{S(x)} \quad \text{is nondecreasing on } \{x > 0, S(x) > 0\}$$

or, equivalently,

$$\psi_2(x) \equiv \frac{S_2(x)}{S(x)} \quad \text{is nonincreasing on } \{x > 0, S(x) > 0\}.$$

This is can also be expressed as $\psi_1(x) = \sup_{0 \le y \le x} \psi_1(y)$ and $\psi_2(x) = \inf_{0 \le y \le x} \psi_2(y))$ for all $x > 0$. The estimators we propose for $\psi_1(x)$ and $\psi_2(x)$ employ the sample analog of this. Specifically, we estimate $\psi_1(x)$ and $\psi_2(x)$ by

$$\psi_1^*(x) = \sup_{0 \le y \le x} \hat{\psi}_1(y) \quad \text{and} \quad \psi_2^*(x) = \inf_{0 \le y \le x} \hat{\psi}_2(y),$$

respectively, where $\hat{\psi}_1(x) = \hat{S}_1(x)/\hat{S}(x)$ and $\hat{\psi}_2(x) = \hat{S}_2(x)/\hat{S}(x)$ are their corresponding unrestricted estimators.

Since $S_1(x) = \psi_1(x)S(x)$ and $S_2(x) = \psi_2(x)S(x)$, we estimate S_1 and S_2 by

$$S_1^*(x) = \psi_1^*(x)\hat{S}(x)I_{[0,\tau_{1-\hat{S}})}(x) \quad \text{and} \quad S_2^*(x) = \psi_2^*(x)\hat{S}(x)I_{[0,\tau_{1-\hat{S}})}(x) \quad (3.1)$$

where, for any DF G, $\tau_G = \inf\{x, G(x) = 1\}$. Our estimator of F is then given by

$$F^*(x) = S_2^*((-x)-)I_{(-\infty,0)}(x) + (1 - S_1^*(x))I_{[0,\infty)}(x). \quad (3.2)$$

Clearly F^* is right continuous. Next we show that it is nondecreasing. Suppose $x_1 \le x_2 \le 0$ in which case $\psi_2^*(-x_1) \le \psi_2^*(-x_2)$ and $\hat{S}(-x_1) \le \hat{S}(-x_2)$. As a result

$$F^*(x_1) = \lim_{z \nearrow -x_1} \psi_2^*(z)\hat{S}(z) \le \lim_{z \nearrow -x_2} \psi_2^*(z)\hat{S}(z) = F^*(x_2).$$

Now suppose $0 \le x_1 < x_2$. We have two cases to consider.

1. Case 1: $\psi_1^*(x_1) = \psi_1^*(x_2)$. Then

$$F^*(x_1) = 1 - \psi_1^*(x_1)\hat{S}(x_1) \le 1 - \psi_1^*(x_1)\hat{S}(x_2) = F^*(x_2).$$

2. Case 2: Suppose $\psi_1^*(x_1) < \psi_1^*(x_2)$. We can find $0 \le s_1 \le x_1 < s_2 \le x_2$ such that $\psi_1^*(x_1) = \hat{\psi}_1(s_1)$ and $\psi_1^*(x_2) = \hat{\psi}_1(s_2)$. Then

$$F^*(x_1) = 1 - \psi_1^*(s_1)\hat{S}(x_1) \le \hat{F}(x_1) \le \hat{F}(s_2) = F^*(s_2)$$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

but

$$F^*(s_2) = 1 - \psi_1^*(s_2)\hat{S}(s_2) \leq 1 - \psi_1^*(s_2)\hat{S}(x_2) = F^*(x_2)$$

and hence the desired conclusion.

Notice that since $\hat{S}_1(x) + \hat{S}_2(x) = \hat{S}(x)$, we have $\frac{\hat{S}_1(x)}{\hat{S}(x)} + \frac{\hat{S}_2(x)}{\hat{S}(x)} = 1$, whenever $\hat{S}(x) > 0$. As a consequence, we have $\psi_1^*(x) + \psi_2^*(x) = 1$ and $S_1^* + S_2^* = \hat{S}$. In addition, the following result holds.

Theorem 1. *For the estimator given in* (3.2), *we have*

$$P\left[\lim_{n\to\infty} \sup_x |F^*(x) - F(x)| = 0\right] = 1.$$

Proof. It suffices to show pointwise convergence since both F and F^* are nondecreasing and right continuous; see Chung (2005). For any $x > 0$ satisfying $S(x) > 0$, we have

$$\begin{aligned} |F^*(x) - F(x)| &= |\psi_1^*(x)\hat{S}(x) - \psi_1^*(x)S(x)| \\ &\leq \psi_1^*(x)|\hat{S}(x) - S(x)| + S(x)|\psi_1^*(x) - \psi_1(x)| \\ &\leq |\hat{S}(x) - S(x)| + |\psi_1^*(x) - \psi_1(x)|. \end{aligned}$$

Applying Lemma 1 in Rojo and Samaniego (1993), we get

$$\begin{aligned} |\psi_1^*(x) - \psi_1(x)| &\leq \sup_{0\leq y\leq x} \left|\frac{\hat{S}_1(y)}{\hat{S}(y)} - \frac{S_1(y)}{\hat{S}(y)}\right| \\ &\leq \frac{1}{\hat{S}(x)S(x)} \sup_y \left\{|\hat{S}_1(y) - S_1(y)| + |\hat{S}(y) - S(y)|\right\} \xrightarrow{as} 0, \end{aligned}$$

since $\hat{S}_1$ and $\hat{S}$ are uniformly strongly consistent estimators of S_1 and S, respectively. $\square$

3.2 Algorithm

In this section we extend an algorithm developed in El Barmi and Mukerjee (2016) to compute the estimators they proposed for uniformly stochastically ordered distributions to this case. First notice that the NPMLE of S_i is obtained sequentially by setting $\tilde{S}_i(t_0) = \hat{S}_i(t_0) = n_{i+}/n$ and $\tilde{S}_i(t_j) = \tilde{\theta}_{ij}\tilde{S}_i(t_{j-1}), j = 1, \ldots, m$, where $(\tilde{\theta}_{1j}, \tilde{\theta}_{2j})^T$ is the least squares projection of $(\hat{\theta}_{1j}, \hat{\theta}_{2j})^T \equiv (\hat{S}_1(t_j)/\hat{S}_1(t_{j-1}), \hat{S}_2(t_j)/\hat{S}_2(t_{j-1}))^T$ onto $\mathcal{K} = \{\mathbf{z} \in \mathbb{R}^2, z_1 \leq z_2\}$ with weights $(\omega_{1j}, \omega_{2j})^T \equiv (\hat{S}_1(t_{j-1}), \hat{S}_2(t_{j-1}))^T$ and extending these estimates as right-continuous step functions. The new estimator of S_i can also be obtained iteratively using similar steps except that, instead of $(\tilde{\theta}_{1j}, \tilde{\theta}_{2j})^T$, we use $(\theta^*_{1j}, \theta^*_{2j})^T$, the least squares projection of $(\tilde{\tilde{\theta}}_{1j}, \tilde{\tilde{\theta}}_{2j})^T \equiv (\hat{S}_1(t_j)/S^*_1(t_{j-1}), \hat{S}_2(t_j)/S^*_2(t_{j-1}))^T$ onto $\mathcal{K}$ with weights $(\tilde{\omega}_{1j}, \tilde{\omega}_{2j})^T \equiv (S^*_1(t_{j-1}), S^*_2(t_{j-1}))^T$. The new estimator updates at each step what we project and the weights that are used in the projection to take into account the adjustments made in the previous steps. The new estimation can be implemented using the following algorithm:

(1) Set $S^*_i(t_0) = n_{i+}/n, i = 1, 2$.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

(2) Define $(\tilde{\tilde{\theta}}_{11}, \tilde{\tilde{\theta}}_{21})^T = (\hat{S}_1(t_1)/S_1^*(t_0), \hat{S}_2(t_2)/S_2^*(t_0))^T$.

(3) Define $(\theta_{11}^*, \theta_{21}^*)^T$ as the least squares projection of $(\tilde{\tilde{\theta}}_{11}, \tilde{\tilde{\theta}}_{21})^T$ onto $\mathcal{K} = \{\mathbf{z} \in \mathbb{R}^2, z_1 \leq z_2\}$ with weights $(S_1^*(t_0), S_2^*(t_0))^T$ and set $S_i^*(t_1) = \theta_{i1}^* S_i^*(t_0)$.

(4) For $2 \leq \ell \leq m$, define sequentially $(\tilde{\tilde{\theta}}_{1\ell}, \tilde{\tilde{\theta}}_{2\ell})^T = (\hat{S}_1(t_\ell)/S_1^*(t_{\ell-1}), \hat{S}_2(t_\ell)/S_2^*(t_{\ell-1}))^T$, and $(\theta_{1\ell}^*, \theta_{2\ell}^*)^T$ as the least squares projection of $(\tilde{\tilde{\theta}}_{1\ell}, \tilde{\tilde{\theta}}_{2\ell})^T$ onto $\mathcal{I}$ with weights $(S_1^*(t_{\ell-1}), S_2^*(t_{\ell-1}))^T$ and set $S_i^*(t_j) = S_i^*(t_0)\prod_{\ell=1}^{j}\theta_{i\ell}^*$.

(5) Extend these estimators as right continuous step functions.

The proof of the equivalence of the estimators obtained using this algorithm and the estimators given in (3.1) is similar to a proof given in El Barmi and Mukerjee (2016) for the equivalence of their two estimator for uniformly stochastically ordered distributions and it is omitted.

Remark 2: Notice that

$$\tilde{\omega}_{1j}\tilde{\tilde{\theta}}_{1j} + \tilde{\omega}_{2j}\tilde{\tilde{\theta}}_{2j} = \tilde{\omega}_{1j}\theta_{1j}^* + \tilde{\omega}_{2j}\theta_{2j}^*$$

by the properties of isotonic regression (Robertson, Wright and Dykstra , 1988). Since $\tilde{\omega}_{1j}\tilde{\tilde{\theta}}_{1j} + \tilde{w}_{2j}\tilde{\tilde{\theta}}_{2j} = \hat{S}_1(t_j) + \hat{S}_2(t_j) = \hat{S}(t_j)$ and $\tilde{w}_{1j}\theta_{1j}^* + \tilde{w}_{2j}\theta_{2j}^* = S_1^*(t_j) + S_2^*(t_j)$ we have $S_1^*(t_j) + S_2^*(t_j) = \hat{S}(t_j)$.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

3.3 Weak convergence

Let $Z_{in} = \sqrt{n}[\hat{S}_i - S_i], i = 1, 2,$ and $Z_{3n} = \sqrt{n}[\hat{S} - S]$. It follows from standard weak convergence of empirical processes that

$$(Z_{1n}, Z_{2n}, Z_{3n})^T \stackrel{w}{\Longrightarrow} (Z_1, Z_2, Z_3)^T,$$

on $[0, \infty)^3$ where $(Z_1, Z_2, Z_3)^T$ is a zero mean tri-variate Gaussian process with the covariance function $x \leq y$ given by

$$Cov(Z_i(x), Z_j(y)) = S_i(y)[\delta_{ij} - S_j(x)], i, j = 1, 2,$$

$$Cov(Z_i(x), Z_3(y)) = S_i(y) - S_i(x)S(y), i = 1, 2,$$

$$Cov(Z_i(y), Z_3(x)) = S_i(y)S(x), i = 1, 2,$$

$$Cov(Z_3(x), Z_3(y)) = S(x)(1 - S(y)),$$

where $\delta_{ij} = I(i = j)$ is the Kronecker delta. In addition, $Z_i \stackrel{d}{=} B_i(S_i), i = 1, 2$, and $Z_3 \stackrel{d}{=} B_3(S)$, where B_1, B_2 and B_3 are dependent standard Brownian bridges.

Let

$$Z^*_{in} = \sqrt{n}[S^*_i - S_i],\ i = 1, 2. \tag{3.3}$$

The weak convergence of estimators similar to (3.1) for estimating two uniformly stochastically ordered DFs on the basis of independent samples was studied in Arcones and Samaniego (2001). Next we show that the weak convergence of the processes defined in (3.3) is a direct consequence of that of $(Z_{1n}, Z_{2n}, Z_{3n})^T$, the functional delta method and the continuous mapping theorem. Define the functional

$$\eta(x) = \inf\{y \leq x, \psi_1(y) = \psi_1(x)\}.$$

The following theorem whose proof is given the supplementary material holds.

Theorem 2. *For* (Z_{1n}^*, Z_{2n}^*) *defined by* (3.3), *we have*

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

on $[0, \tau_{(1-S)})^2$ *where*

$$\begin{aligned} Z_1^*(x) &= S(x) \sup_{\eta(x) \leq y \leq x} U_1(y) + \psi_1(x) Z_3(x), \\ Z_2^*(x) &= S(x) \inf_{\eta(x) \leq y \leq x} U_2(y) + \psi_2(x) Z_3(x) \end{aligned}$$

with

$$\begin{aligned} U_1(y) &= \frac{Z_1(y) - \psi_1(y) Z_3(y)}{S(y)} = \frac{(1-\psi_1(y))Z_1(y) - \psi_1(y)Z_2(y)}{S(y)}, \\ U_2(y) &= \frac{Z_2(y) - \psi_2(s) Z_3(y)}{S(y)} = \frac{(1-\psi_2(y))Z_2(y) - \psi_2(y)Z_1(y)}{S(y)}. \end{aligned}$$

Let $Z_n^* = \sqrt{n}[F^* - F]$. By virtue of Theorem 2 and (3.8), the following result holds.

Theorem 3. *The process*

$$Z_n^* \overset{w}{\Rightarrow} Z^*$$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

on $(-\infty, \infty)$ *where*

$$Z^*(x) = \begin{cases} -Z_1^*(x), & x \geq 0, \\ Z_2^*(-x), & x < 0. \end{cases}$$

Remark 3: Theorem 2 implies that when $\eta(x) = x$, $(Z_{1n}^*, Z_{2n}^*) \stackrel{w}{\Rightarrow} (Z_1, Z_2)$ where Z_1 and Z_2 are as defined before. The processes Z_1 and Z_2 have almost sure continuous paths while the paths of Z_1^*, Z_2^* are almost a.s. right continuous with a jump at the fixed points $\{\eta(x)\}$ where $\eta(x) = \inf\{y \leq x, \psi_1(y) = \psi_1(x)\}$.

The following theorem shows an interesting fact that, although $Z_i(t)$ and $Z_i^*(t)$ may not be equal in distribution for $i = 1, 2$, their absolute values are.

Theorem 4. *The processes* Z_1^* *and* Z_2^* *satisfy, for all* $x > 0$,

$$P[|Z_i^*(x)| > u] = P[|Z_i(x)| > u], \quad \text{for } i = 1, 2 \text{ and } u \geq 0.$$

That is, $|Z_i^*(x)| \stackrel{d}{=} |Z_i(x)|$.

Proof. We only show that the result holds for Z_1^*. Similar steps can be used to show that it holds also for Z_2^*. We have

$$\begin{aligned} Z_1^*(x) &= S(x) \sup_{\eta(x) \leq y \leq x} U_1(y) + \psi_1(x) Z_3(x) \\ &= S(x) \sup_{\eta(x) \leq y \leq x} [U_1(y) - U_1(\eta(x))] + S(x) U_1(\eta(x)) + \psi_1(x) Z_3(x). \end{aligned}$$

It is easy to check that the process

$$\{U_1(y)-U_1(\eta(x)), \eta(x) \le y \le x\} \stackrel{d}{=} \{\sqrt{\psi_1(x)(1-\psi_1(x))}B\left(\frac{1}{S(\eta(x))} - \frac{1}{S(y)}\right), \eta(x) \le y \le x\},$$

where B is a standard Brownian motion. In addition $\{U_1(y)-U_1(\eta(x)), \eta(x) \le y \le x\}, U_1(\eta(x))$ and $Z_3(x)$ are independent. We can easily show this by computing the relevant covariances. Now

$$\begin{aligned} Z_3(x) &\sim \sigma_1(x)Y_1 \\ U_1(\eta(x)) &\sim \sigma_2(x)Y_2 \\ \sup_{\eta(x)\le y\le x} [U_1(y) - U_1(\eta(x))] &\stackrel{d}{=} \sigma_3(x)|Y_3| \end{aligned}$$

where $\sigma_1(x) = \sqrt{S(x)(1-S(x))}$,

$\sigma_2(x) = \sqrt{\frac{\psi_1(x)(1-\psi_1(x))}{S(\eta(x))}}, \sigma_3(x) = \sqrt{\psi_1(x)(1-\psi_1(x))\left(\frac{1}{S(\eta(x))} - \frac{1}{S(x)}\right)}$

and Y_1, Y_2 and Y_3 are independent standard normal variables. The last equality follows from Billingsley (1968, p. 72). Using the fact that if X and Y are independent mean zero normal variates, then , $|X + |Y|| \stackrel{d}{=} |X + Y|$, we get

$$\begin{aligned} P(|Z_1^*(x)| > u) &= P(|\psi_1(x)\sigma_1(x)U_1 + S(x)\sigma_2(x)U_2 - S(x)\sigma_3(x)|U_3|| > u) \\ &= P(|\psi_1(x)\sigma_1(x)U_1 + S(x)\sigma_2(x)U_2 - S(x)\sigma_3(x)U_3| > u) \\ &= P(|N(0, \psi_1^2(x)\sigma_1^2(x) + S^2(x)\sigma_2^2(x) + S^2(x)\sigma_3^2(x))| > u). \end{aligned}$$

It is easy to verify that

$$\psi_1(x)^2\sigma_1^2(x) + S^2(x)\sigma_2^2(x) + S^2(x)\sigma_3^2(x) = S_1(x)(1 - S_1(x)).$$

Since $Z_1(x) \sim N(0, S_1(x)(1-S_1(x))$, we have $P(|Z_1^*(x)| > u) = P(|Z_1(x)| > u)$. □

As a corollary to this theorem we have the following result.

Corollary 1. *The process Z^* satisfies, for all $x > 0$,*

$$P[|Z^*(x)| > u] = P[|Z(x)| > u] \quad \textit{for all} \, u \geq 0.$$

That is, $|Z^*(x)| \stackrel{d}{=} |Z(x)|$.

Proof. The proof of this results follows readily from the proof of Theorem 4. □

3.4 A hypothesis test

Suppose that ψ_1 is nondecreasing. An interesting problem is then to test

$H_0 : S_1 = S_2$ versus $H_1 : S_1(x)/S_2(x)$, is nondecreasing in x with strict increasing at some x.

Note that H_1 holds if and only if

$$\sup_{y\geq 0} \sup_{0\leq x\leq y} [S_1(y)S(x) - S_1(x)S(y)] > 0.$$

A natural test criterion will then be to reject H_0 if the test statistic

$$T_n = \sqrt{n} \sup_{y\geq 0} \sup_{0\leq x\leq y} [\hat{S}_1(y)\hat{S}(x) - \hat{S}_1(x)\hat{S}(y)]$$

is large. Notice that

$$\begin{aligned} T_n &= \sqrt{n}\sup_{y\geq 0}\sup_{0\leq x\leq y}[(\hat{S}_1(y)Z_n(x)-S(y)Z_{1n}(x))-(\hat{S}_1(x)Z_n(y)-S(x)Z_{1n}(y))] \\ &+ \sqrt{n}[S_1(y)S(x)-S_1(x)S(y)]. \end{aligned}$$

Since under $H_0, S_1(y)S(x)-S_1(x)S(y)=0$,

$$\begin{aligned} T_n &\overset{d}{\to} \sup_{y\geq 0}\sup_{0\leq x\leq y}[(S_1(y)Z(x)-S(y)Z_1(x))-(S_1(x)Z(y)-S(x)Z_1(y))] \\ &\overset{d}{\to} \frac{1}{2}\sup_{y\geq 0}\sup_{0\leq x\leq y}[S(y)(Z(x)-2Z_1(x))-S(x)(Z(y)-2Z_1(y)))]. \end{aligned}$$

It is easy to check that, under H_0,

$$\{Z(x)-2Z_1(x), x\geq 0\} \overset{d}{=} \{B(S(x)), x\geq 0\}$$

where B is a standard Brownian motion. Therefore

$$\begin{aligned} T_n &\overset{d}{\to} \frac{1}{2}\sup_{y\geq 0}\sup_{0\leq x\leq y}[S(y)B(S(x))-S(x)B(S(y))] \\ &\overset{d}{\to} \frac{1}{2}\sup_{0\leq u\leq v\leq 1}[uB(v)-vB(u)]. \end{aligned}$$

To implement this test requires the critical values for T_n. The null distribution of T_n is clearly not tractable, even asymptotically, but it is asymptotically distribution free. These critical values are approximated by computing appropriate quantiles of $\{T(k), k=1,2,\ldots,10000\}$ where $T(k) = \max_{1\leq i\leq j\leq 1000}(u(i)B(v(j))-v(j)B(u(i)))$ with the values $\{B(u(i)), u(i)=i/1000, i=1,2,\ldots,1000\}$ simulated from a Brownian motion B.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

Table 1: Selected critical points of asymptotic distribution of T_n

Significance level α		
0.01	0.05	0.10
0.763	0.625	0.553

4. Applications to a competing risks model

Next we consider a competing risks model in which a unit or a subject is exposed to two risks at the same time but the actual failure (or death) is attributed to exactly one of them. Specifically, let X and Y denote the notional (or latent) lifetimes of a unit under these two risks. We do not assume that these variables are independent and we only observe (T, δ), where $T = \min(X, Y)$ is the time of failure and $\delta = 2 - I[X \leq Y]$ is the cause of failure. Throughout we assume that $P(X = Y) = 0$. The observed data in this case $(T_i, \delta_i), i = 1, 2, \ldots, n$.

For such data, we are interested in knowing whether or not the two risks are equal or one risk is greater than the other. Such comparisons are often made on the basis of $D = S_1 - S_2$ where S_i is the sub-survival function corresponding to the ith risk and is defined by

$$S_i(t) = P[T > t, \delta = i], i = 1, 2.$$

Note that $S(t) \equiv S_1(t) + S_2(t)$ is the survival function corresponding to T.

Clearly D is not a suitable measure if the goal is to compare the conditional distribution of T given $\delta = 1$ and its conditional distribution given $\delta = 2$. In this case we can consider the temporal function $r(t) \equiv S_1(t)/S_2(t)$. Notice that $r(t)$ is proportional to $P[T > t|\delta = 1]/P[T > t|\delta = 2]$ and it is nondecreasing in t if and only if $\psi_1(t) \equiv S_1(t)/S(t)(\psi_2(t) \equiv S_2(t)/S(t))$ is nondecreasing (nonincreasing) in t. Note also that when $P[T > t|\delta = 1]/P[T > t|\delta = 2]$ is nondecreasing in t, the conditional distribution of T given $\delta = 1$ is uniformly stochastically greater the conditional distribution of T given $\delta = 2$.

Suppose it is desired to estimate S_1 and S_2 under this constraint as well as develop a test for

$$H_0 : S_1 = S_2 \quad \text{against } H_1 - H_0, \text{ where } H_1 : S_1(t)/S_2(t) \quad \text{is nondecreasing on } \{t, S_2(t) > 0\}$$

Define the empirical estimators of S_1, S_2, S and $\psi_j, j = 1, 2$, by

$$\begin{aligned}
\hat{S}_j(t) &= \frac{1}{n}\sum_{i=1}^{n} I(T_i > t, \delta_i = j), \; j = 1, 2, \\
\hat{S}(t) &= \hat{S}_1(t) + \hat{S}_2(t), \quad \hat{\psi}_j(t) = \frac{\hat{S}_j(t)}{\hat{S}(t)}, j = 1, 2.
\end{aligned}$$

In addition, define their restricted estimators by

$$\psi_1^*(t) = \sup_{0\le s\le t} \hat{\psi}_1(s), \qquad \psi_2^*(t) = \inf_{0\le s\le t} \hat{\psi}_2(t),$$
$$S_1^*(t) = \psi_1^*(t)\hat{S}(t)I_{[0,\tau_{1-\hat{S}})}(t) \qquad \text{and} \qquad S_2^*(t) = \psi_2^*(t)\hat{S}(t)I_{[0,\tau_{1-\hat{S}})}(t). \tag{4.1}$$

Notice that the mathematical structure of the estimators in (3.1) and the estimators of S_1 and S_2 in (4.1) are identical even though their interpretations are different. Similar arguments to those used in the proof of Theorem 1 can be used to show that the uniform strong consistency of the S_i^*s follows from that of the ψ_i^*s and of $\hat{S}$. Therefore we have the following theorem.

Theorem 5. *For the estimators given in* (4.1) *we have*

$$P\left[\lim_{n\to\infty} \sup_t |S_j^*(t) - S_j(t)| = 0, \quad j = 1, 2\right] = 1.$$

If we define the processes $(Z_{1n}, Z_{2n}, Z_{3n})^T$ and $(Z_{1n}^*, Z_{2n}^*, Z_{3n}^*)^T$ using the same symbols as in section 3.2, then the conclusions of Theorem 2 and Theorem 4 continue to hold here also. Finally, we note that the test in Section 3.4 can be used to test H_0 against $H_1 - H_0$.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

5. Examples and simulations

5.1 Example 1:

In our first example we use data in Table 3 below from Moore and McCabe (1993). This data gives the pretest and posttest scores on the MLA listening test in Spanish for 20 high school teachers who attended an intensive course in Spanish.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

Table 2: 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

Let X_i = (posttest score - pretest score) corresponding to the ith subject and assume that the resulting twenty random differences form a sample from a distribution F that satisfies (1.3). The restricted estimator F^* of this distribution is given in Table 3. Although the sample size is only 20, if we use our asymptotic test in Section 3.2 to test for symmetry about zero, which can be interpreted as a test for the efficacy of the training method, we get a test statistic value of 0.045 with a p-value > 0.10 indicating that the training method is not effective. On the other hand, a one sided paired

Table 3: Estimate of F under (1.3)

-7.00	0.0313
-5.00	0.0781
-4.00	0.1250
-3.00	0.1875
-2.00	0.2500
-1.00	0.3500
1.00	0.45000
2.00	0.5875
3.00	0.7250
4.00	0.8281
5.00	0.9313
7.00	1.0000

t-test for comparing the mean of pretestand the mean of posttest yields a p-value of 0.029 and leads to different conclusion. Thus it appears that the assumption of normality used in the t-test is crucial in concluding that the training method is effective.

5.2 Example 2

In this example 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. To illustrate the applicability of our results is Section 4, we consider "other causes" as cause 1 and cancer as cause 2. Intuitively, one would expect that S_1/S_2 is nondecreasing. The unrestricted and restricted estimators of S_1 and S_2 are displayed in Figure 1. We also considered the large sample test of $H_0 : S_1 = S_2$ against $H_1 - H_0$, where $H_1 : S_1/S_2$ nondecreasing in t, using the test described in Section 2.4. The value of the test statistic is 0.615 corresponding to a p-value less

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

that 0.10 based on Table 1.

5.3 Simulations

Here we present the results of a simulation designed to compares the finite sample performance of the NMPLE and the new estimators in terms of their mean square errors at different quantiles of the distribution F. We take F to be

$$F(x) = \begin{cases} \frac{1}{2}e^{\beta x}, & x < 0 \\ 1 - \frac{1}{2}e^{-x}, & x \geq 0 \end{cases}$$

where $\beta \geq 1$. In this case, for $x \geq 0$,

$$S_1(x) = \frac{1}{2}e^{-x} \quad \text{and} \quad S_2(x) = \frac{1}{2}e^{-\beta x}.$$

Clearly (1.3) is satisfied in this case.

Figure 2 and 3 show the ratio of the mean square error (MSE) of the NPMLE to that of the new estimator of F for $b = 1, 1.1, 1.2$ and 1.3 when $n = 30$ and $n = 50$, respectively, and the number of replications is 3000. This graph shows that the new estimator outperforms the NPMLE in terms of MSE.

Next, we view S_1 and S_2 as the sub-survival functions corresponding to two competing. Tables 4 and 5 give the results for $\beta = 1, 1.1, 1.2$ and 1.3 and different sample sizes. We use again 3,000 replications to compute the bias and the MSE of the new estimator, S_i^*, of $S_i, i = 1, 2$, and its corresponding NPMLE $\tilde{S}_i$. As expected, S_1^* and $\tilde{S}_1$ exhibit a negative bias while S_2^* and $\tilde{S}_2$ show positive bias, although not very much in the case of the new estimators. The MSEs corresponding to the new estimators are uniformly smaller than the corresponding MSEs of the NPMLEs in all the cases that we consider. Finally we note that the gain in terms of MSE goes up as S_2 gets closer to S_1 which makes sense since this corresponds to the scenario where reversals are more likely to occur.

Table 4: comparison of bias (B) and MSE of $S_1^*, \tilde{S}_1, S_2^*$ and $\tilde{S}_2$ of the q-quantiles of S with $n = 30$ and 3,000 replications.

q	$B(\tilde{S}_1)$	$B(S_1^*)$	$\frac{MSE(\tilde{S}_1)}{MSE(S_1^*)}$	$B(\tilde{S}_2)$	$B(S_2^*)$	$\frac{MSE(\tilde{S}_2)}{MSE(S_2^*)}$
		$\beta = 1.0$			$\beta = 1.1$	
0.1	-0.0618	-0.0372	1.5467	-0.0606	-0.0357	1.5369
0.2	-0.0846	-0.0445	1.7872	-0.0808	-0.0413	1.7489
0.3	-0.0824	-0.0441	1.6452	-0.0779	-0.0410	1.5853
0.4	-0.0625	-0.0387	1.3182	-0.0591	-0.0366	1.2848
0.5	-0.0219	-0.0140	1.0104	-0.0256	-0.0178	1.0153
0.6	-0.0424	-0.0167	1.1072	-0.0465	-0.0199	1.1146
0.7	-0.0829	-0.0273	1.7806	-0.0836	-0.0271	1.7305
0.8	-0.1111	-0.0293	3.0845	-0.1088	-0.0267	2.9452
0.9	-0.1169	-0.0244	5.5311	-0.1098	-0.0204	5.0060
		$\beta = 1.2$			$\beta = 1.3$	
0.1	-0.0594	-0.0341	1.5148	-0.0568	-0.0315	1.4856
0.2	-0.0794	-0.0406	1.7088	-0.0757	-0.0382	1.6502
0.3	-0.0772	-0.0413	1.5550	-0.0737	-0.0394	1.4953
0.4	-0.0582	-0.0369	1.2677	-0.0559	-0.0358	1.2420
0.5	-0.0260	-0.0187	1.0180	-0.0269	-0.0204	1.0204
0.6	-0.0449	-0.0179	1.1007	-0.0452	-0.0177	1.1021
0.7	-0.0805	-0.0219	1.7070	-0.0805	-0.0217	1.6975
0.8	-0.1037	-0.0209	2.8692	-0.1017	-0.0204	2.7320
0.9	-0.1034	-0.0163	4.9215	-0.0971	-0.0154	4.4617

Table 5: Comparison of bias (B) and MSE of $S_1^*, \tilde{S}_1, S_2^*$ and $\tilde{S}_2$ at the q-quantiles of S with $n = 50$ and 3,000 replications.

q	$B(\tilde{S}_1)$	$B(S_1^*)$	$\frac{MSE(\tilde{S}_1)}{MSE(S_1^*)}$	$B(\tilde{S}_2)$	$B(S_2^*)$	$\frac{MSE(\tilde{S}_2)}{MSE(S_2^*)}$
	$\beta = 1.0$			$\beta = 1.1$		
0.1	-0.0607	-0.0297	2.1668	0-.0587	-0.0271	2.1765
0.2	-0.0825	-0.0357	2.4373	-0.0778	-0.0313	2.4058
0.3	-0.0792	-0.0353	2.1035	-0.0741	-0.0313	2.0263
0.4	-0.0562	-0.0297	1.4852	-0.0517	-0.0268	1.4334
0.5	-0.0161	-0.0112	1.0107	-0.0158	-0.0111	1.0161
0.6	-0.0469	-0.0179	1.3078	-0.0465	-0.0166	1.2820
0.7	-0.0854	-0.0231	2.5721	-0.0836	-0.0197	2.4538
0.8	-0.1127	-0.0232	5.0379	-0.1096	-0.0196	4.7347
0.9	-0.1188	-0.0185	9.4214	-0.1131	-0.0163	8.3859
	$\beta = 1.2$			$\beta = 1.3$		
0.1	-0.0561	-0.0241	2.1103	-0.0536	-0.0214	2.0886
0.2	-0.0731	-0.0276	2.3305	-0.0688	-0.0241	2.1835
0.3	-0.0687	-0.0283	1.9085	-0.0642	-0.0249	1.8251
0.4	-0.0497	-0.0259	1.3983	-0.0449	-0.0225	1.3499
0.5	-0.0149	-0.0104	1.0147	-0.0156	-0.0114	1.0084
0.6	-0.0457	-0.0149	1.2743	-0.0457	-0.0145	1.2808
0.7	-0.0836	-0.0194	2.4241	-0.0824	-0.0169	2.4042
0.8	-0.1078	-0.0188	4.4397	-0.1054	-0.0165	4.4895
0.9	-0.1064	-0.0144	7.4076	-0.1004	-0.0118	7.2727

6. Conclusion

In this paper we showed that the NPMLE of a DF of a random variable X that satisfies $P(X > x)/P(X < -x)$ is nondecreasing in $x > 0$ is inconsistent. We then proposed a projection type estimator for it that is strongly uniformly consistent and derived its weak convergence. We also showed that this new estimator improves in finite samples on the NPMLE in terms of the MSE using simulations. In addition, we developed a procedure for testing symmetry against this restriction. It turns out that the theoretical results we developed can be used to estimate the sub-survival functions of competing risks in a two competing risks model under the restriction that their ratio is nondecreasing. We showed how this can be done and provided a real life example to illustrate our procedures.

7. Supplementary material

A detailed proof of Theorem 3.2 is given in the Supplementary Material.

Acknowledgments

The authors are grateful to an associate editor and a referee for their comments that have resulted in a much improved paper. The work of Hammou El Barmi acknowledges the support of the City University of New York through a PSC-CUNY grant and the work of Zhiliang Ying is supported in part by US NSF Grant DMS-2015417.

8. References

Alfieri, A. and El Barmi, H. (2005). Nonparametric estimation of a distribution with Type I bias with application to competing risks. *Journal of Nonparameteric Statistics*, 17, 319-33.

Arcones, M. A. and Samaniego, F. J. (2001). On the asymptotic distribution theory of a class of consistent estimators of a distribution satisfying a uniform stochastic ordering constraint. *Annals of Statistics* **28** 116-150.

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

Chaffin, W. W, and Rhiel S. G. (1993). The effect of skewness and kurtosis on the one-sample T test and the impact of knowledge of the population standard deviation. *Journal of Statistical Computation in Simulation*, 45, 79-90.

Chung, K. L. (2001). *A course in probability theory*, Third edition, Academic Press, New York.

Dykstra, D., Kochar, S. and Robertson, T. (1995). Likelihood ratio tests for symmetry against some one-sided alternatives. *Annals of the Institute of Statistical Mathematics* 54 163-174.

Dykstra, R. and Praestgaard, J. T. (1996) The asymptotics for the NPMLE of a distribution with positive bias. *personal communication.*

El Barmi, H. and Mukerjee, H. (2004). Consistent estimation of distributions with type II bias with applications to competing risks. *Annals of Statistics*, 32, 245-267.

El Barmi, H. and Mukerjee, H. (2016). Consistent estimation of survival functions under uniform stochastic ordering: the K-sample case. *Journal of Multivariate Analysis*, 144, 99-109.

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

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Annals of Mathematical Statistics*, 27, 887-906.

Moore, S. D. and McCabe, P. G. (1993). *Introduction to the Practice of Statistics.* W. H. Freeman, New York.

Mukerjee, H. (1996). Estimation of survival functions under uniform stochastic ordering. *Journal of the American Statistical Association.* 91 1684-1689.

Robertson, T., Wright, F. T. and Dykstra R. L. (1988). Order Restricted Inference. Wiley, New York.

Rojo, J. and Samaniego, F. J. (1993). On estimating a survival curve subject to a uniform stochastic ordering constraint. *Journal of the American Statistical Association* 88 566-572.

Shorack, G. R. and Wellner, J. A. (1986).*Empirical Processes with Applications to Statistics.* Wiley, New York. Corrections posted at http://www.stat.washington.edu/jaw/RESEARCH/BOOKS/book1.html

Yanagimoto, T. and Sibuya, M. (1972). Test of symmetry of a one-dimensional distribution against positive biasedness. *Annals of the Institute of Statistical Mathematics*, 24, 423-434.

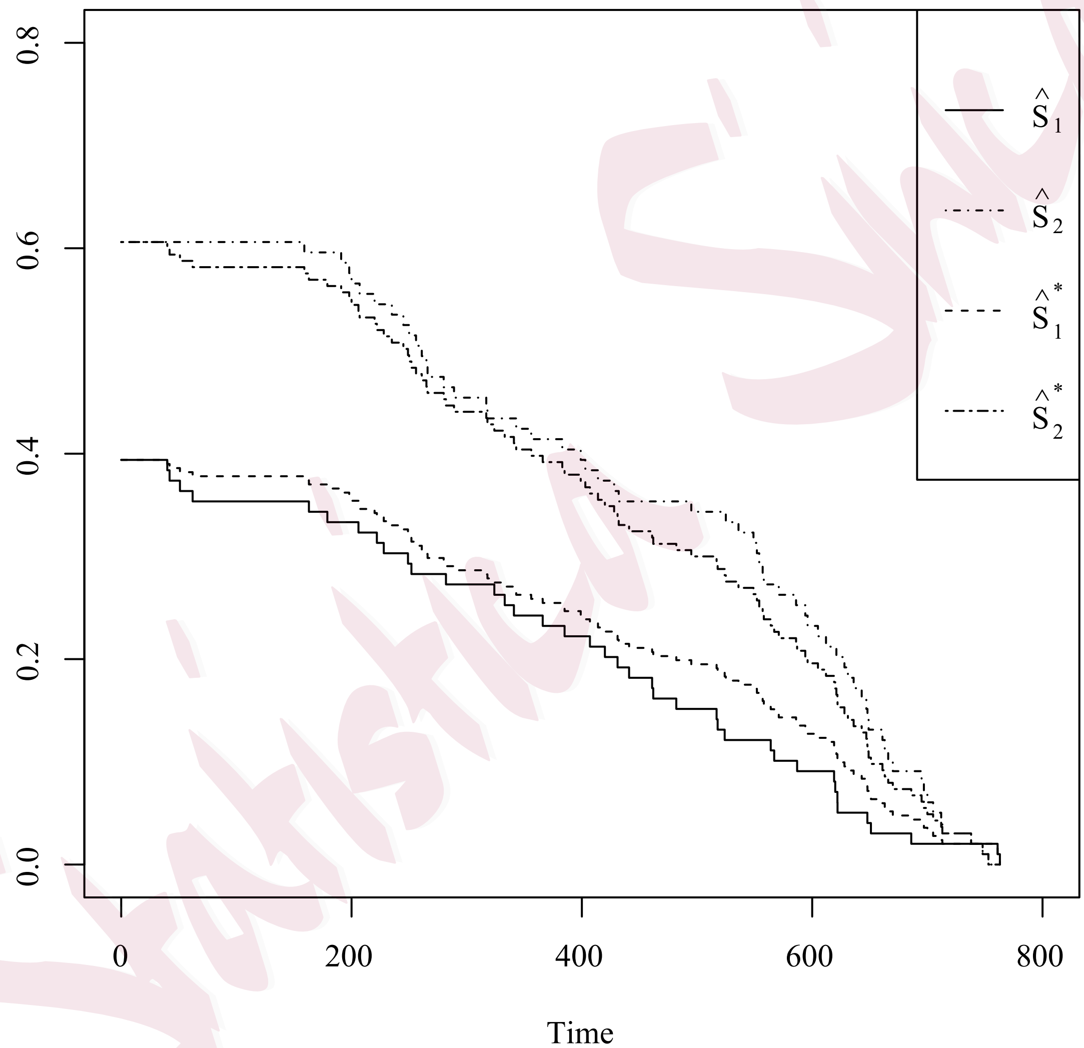

Figure 1: Unrestricted and restricted estimators of S_1 and S_2.

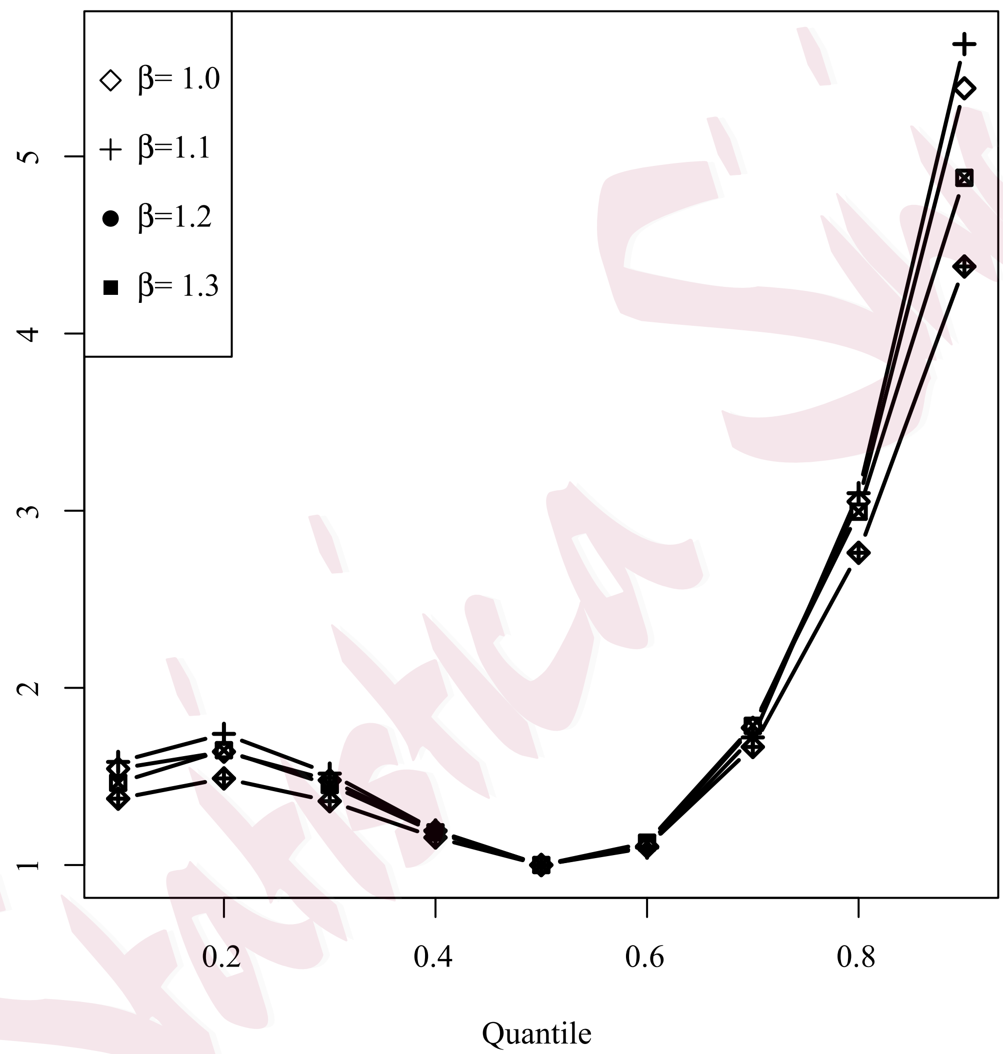

Figure 2: Ratio of the MSE of the NPMLE of F to that of its new estimator ($n = 30$).

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0176

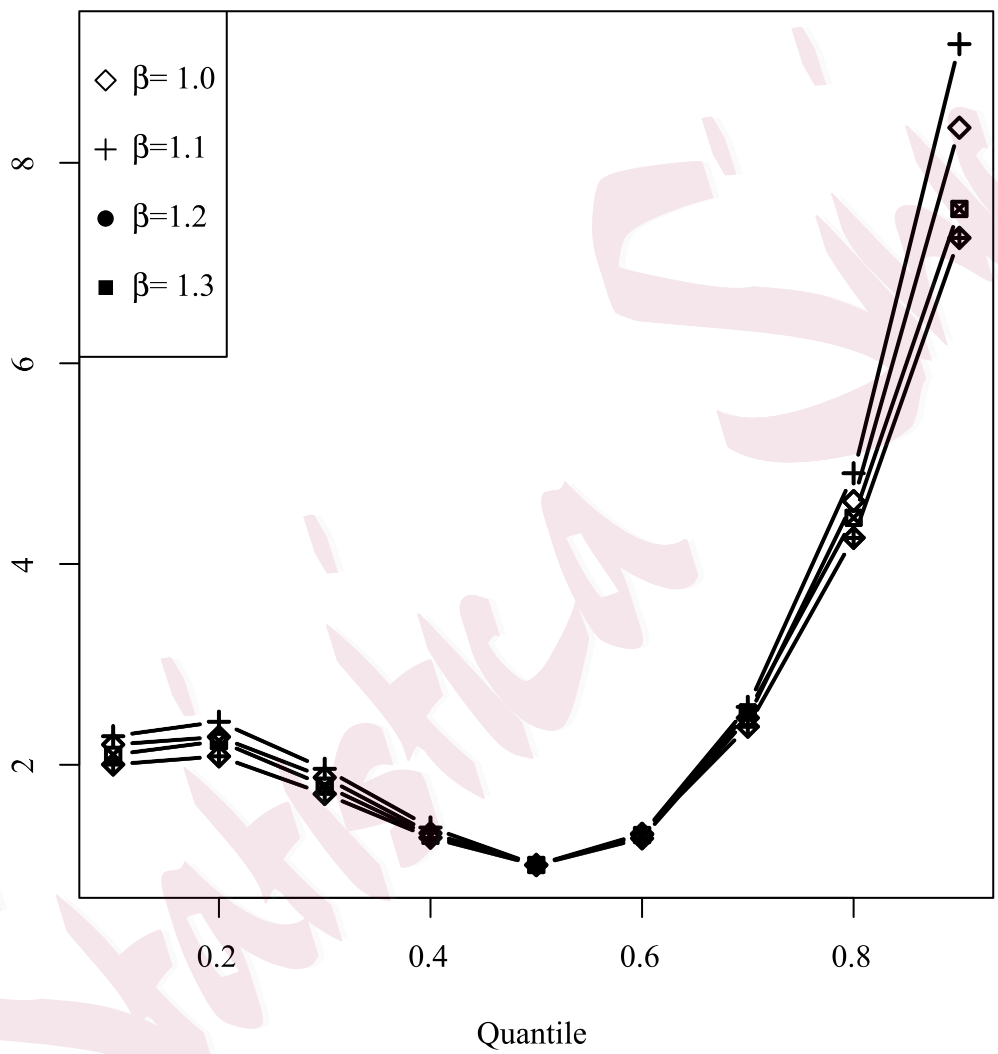

Figure 3: Ratio of the MSE of the NPMLE of F to that of its new estimator ($n = 50$).