Statistics and Probability Letters 208 (2024) 110045

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Estimating several survival functions under uniform stochastic ordering

Sebastian Arnold [a,*], Hammou El Barmi [b], Hari Mukerjee [c], Johanna Ziegel [a]

[a] *University of Bern, Bern, Switzerland*
[b] *City University of New York, New York, USA*
[c] *Wichita State University, Wichita, USA*

ARTICLE INFO

Keywords:
Consistency
Estimation
Uniform stochastic order
Hazard rate order
Order constraint

ABSTRACT

El Barmi and Mukerjee (2016, Journal of Multivariate Analysis 144, 99–109) studied the estimation of survival functions of k samples under uniform stochastic ordering constraints. There were two crucial errors in the consistency proof. Here, we provide alternative estimators and show consistency.

1. Introduction

Two distributions on the positive half-line with survival functions (SFs) S_1 and S_2 are uniformly stochastically ordered (USO), $S_1 \leq_{uso} S_2$, if S_1/S_2 is nonincreasing. The uniform stochastic order is equivalently known as the hazard rate order (Shaked and Shanthikumar, 2007, Section 1.B.1). We use the term uniform stochastic order, since some parts of the literature restrict the definition of the hazard rate order to absolutely continuous distributions, and we do not make such a restriction. If $S_1 \leq S_2$ holds pointwise, then S_1 is smaller than S_2 in the (usual) stochastic order. The terminology USO is motivated by the fact that $S_1 \leq_{uso} S_2$ is equivalent to the stochastic ordering of certain conditional distributions, that is, for $X_1 \sim S_1$ and $X_2 \sim S_2$, it holds that

$$P(X_1 > t+s | X_1 > t) \leq P(X_2 > t+s | X_2 > t), \quad s,t \geq 0. \tag{1.1}$$

In particular, USO always implies the stochastic ordering of the distributions. Eq. (1.1) shows that USO is of interest in reliability and life testing if e.g. X_1 and X_2 describe the life times of two different items. We refer to Rojo and Samaniego (1993) or Mukerjee (1996) for a more extensive illustration of situations where USO seems to be a natural order constraint.

With this paper, we are contributing the first consistent distribution estimators for k samples under an USO constraint. Statistical inference under USO is a classical topic in non-parametric modeling and inference. For k life distributions, it was started by Dykstra et al. (1991). Based on independent random samples, they derived the nonparametric maximum likelihood estimators (NPMLEs) for general SFs, with and without censoring. Rojo and Samaniego (1991) and Mukerjee (1996) gave counterexamples to show that the NPMLEs are inconsistent for $k = 2$ in the 1-sample (one S_i known) and the 2-sample cases, respectively, when the SFs are continuous. When $S_1 \leq_{uso} S_2$, Rojo and Samaniego (1993) provided a consistent estimator of one SF when the other is known by using the sample analog of the fact that S_1/S_2 is nonincreasing if and only if $S_1(x)/S_2(x) = \inf_{y \leq x}[S_1(y)/S_2(y)]$ for $S_2(x) > 0$, or equivalently, $S_2(x)/S_1(x) = \sup_{y \leq x}[S_2(y)/S_1(y)]$ for $S_1(x) > 0$. For any SF S of a life distribution, let b_S denote the right endpoint of

* Corresponding author.
E-mail addresses: sebastian.arnold@unibe.ch (S. Arnold), hammou.elbarmi@baruch.cuny.edu (H. El Barmi), mukerjee@math.wichita.edu (H. Mukerjee), johanna.ziegel@unibe.ch (J. Ziegel).

https://doi.org/10.1016/j.spl.2024.110045
Received 28 September 2022; Received in revised form 16 March 2023; Accepted 11 January 2024
Available online 19 January 2024

the support of S. Denoting the empiricals by $\hat{S}_1$ and $\hat{S}_2$ based on independent random samples of sizes n_1 and n_2, respectively, the restricted estimators are given by

$$\begin{aligned} S_1^*(x) &= \inf_{y\le x} \frac{\hat{S}_1(y)}{S_2(y)} S_2(x) I(x < b_{S_2}) \quad \text{when } S_2 \text{ is known,} \\ S_2^*(x) &= \sup_{y\le x} \frac{\hat{S}_2(y)}{S_1(y)} S_1(x) I(x < b_{S_1}) + \hat{S}_2(x) I(x \ge b_{S_1}) \quad \text{when } S_1 \text{ is known.} \end{aligned} \tag{1.2}$$

In the 2-sample case, when both SFs are unknown, they suggested setting $\hat{S}_1$ or $\hat{S}_2$ fixed and estimating the other SF under USO as in the 1-sample case. Mukerjee (1996) showed that the 2-sample estimators can be improved by holding the combined empirical, $\hat{S}_{1:2} \equiv (n_1\hat{S}_1 + n_2\hat{S}_2)/(n_1+n_2)$ fixed and estimating S_1 and S_2 under the constraint $S_1 \le_{uso} \hat{S}_{1:2} \le_{uso} S_2$ as two 1-sample estimators:

$$\begin{aligned} S_1^*(x) &= \inf_{y\le x} \frac{\hat{S}_1(y)}{\hat{S}_{1:2}(y)} \hat{S}_{1:2}(x) I(\hat{S}_{1:2}(x) > 0), \\ S_2^*(x) &= \sup_{y\le x} \frac{\hat{S}_2(y)}{\hat{S}_{1:2}(y)} \hat{S}_{1:2}(x) I(\hat{S}_{1:2}(x) > 0). \end{aligned} \tag{1.3}$$

Mukerjee (1996) showed that these estimators are strongly uniformly consistent if $n_i/(n_1+n_2) \to \alpha_i > 0$ for $i = 1,2$ as $n_1, n_2 \to \infty$.

In the k-sample case, El Barmi and Mukerjee (2016) introduced what they thought to be consistent estimators. We found some mistakes in that paper, which are discussed in Section 3. The purpose of this paper is to provide new consistent estimators.

We assume that for k populations under the order constraint

$$S_1 \le_{uso} S_2 \le_{uso} \cdots \le_{uso} S_k, \tag{1.4}$$

we have independent random samples of sizes $n_1, \ldots, n_k$ that satisfy

$$\hat{\alpha}_i = n_i \Big/ \sum_{j=1}^k n_j \to \alpha_i > 0 \text{ for all } 1 \le i \le k \text{ as } \min_{i:1\le i\le k} n_i \to \infty. \tag{1.5}$$

For $1 \le i \le k$, assume that we have consistent but unrestricted estimators of S_i available which we denote by $\hat{S}_i$. Here $\hat{S}_i$ will typically be the empirical survival function or the Kaplan–Meier estimator in case of randomly right censoring. For $1 \le r \le s \le k$, let

$$\alpha_{r:s} = \sum_{j=r}^s \alpha_j, \quad \hat{\alpha}_{r:s} = \sum_{j=r}^s \hat{\alpha}_j, \quad S_{r:s} = \frac{\sum_{j=r}^s \alpha_j S_j}{\alpha_{r:s}}, \quad \hat{S}_{r:s} = \frac{\sum_{j=r}^s \hat{\alpha}_j \hat{S}_j}{\hat{\alpha}_{r:s}}.$$

Note that $(S_{i:i}\ \hat{S}_{i:i})$ is simply $(S_i\ \hat{S}_i)$ for all $1 \le i \le k$. For $1 \le r < u < s \le k$, we have $S_{r:s} = (\alpha_{r:u}S_{r:u} + \alpha_{u+1:s}S_{u+1:s})/(\alpha_{r:u} + \alpha_{u+1:s})$.

2. Suggested estimators and their consistency

We suggest a set of estimators where all the order restrictions are active. In Section B of the Supplementary Material, we discuss other consistent estimators that are simpler to implement, but that do not exploit all the order constraints; heuristically and in simulations, these are suboptimal. We start by assuming complete observations with $\hat{S}_i$ as the estimator of S_i. The censored case using the Kaplan–Meier estimators is actually simpler even though consistency will be limited to a narrower range. We discuss it at the end of this section.

On numerous occasions, we examine properties of ratios of nonnegative functions of the form $f(x)/g(x)$ with the multiplier $I(g(x) > 0)$. Instead of tedious repetition of this qualifier, we will assume this without explicit mention.

Shaked and Shanthikumar (2007, Theorem 1.B.22) show that if $U_1 \le_{uso} U_2$ are SFs and $0 < \alpha < 1$. Then $U_1 \le_{uso} U_{12} \equiv \alpha U_1 + (1-\alpha)U_2 \le_{uso} U_2$. We will frequently use this result without explicit mention. An example of an application is showing $S_{1:4} \le_{uso} S_{3:5}$ by arguing sequentially $S_{1:2} \le_{uso} S_2 \le_{uso} S_3 \le_{uso} S_{3:4}$, implying $S_{1:4} \le_{uso} S_{3:4} \le_{uso} S_4 \le_{uso} S_5$, which in turn implies $S_{1:4} \le_{uso} S_{3:4} \le_{uso} S_{3:5}$.

We build up a pyramidal structure of $\{S_{r:s}\}$ with $S_{1:k}$ at the apex, and with fewer and fewer components going down until we hit $\{S_{i:i} = S_i,\ 1 \le i \le k\}$ in the bottom row. By our claim above,

$$\begin{aligned} &S_{1:k-1} \le_{uso} S_{1:k} \le_{uso} S_{2:k} \\ &S_{1:k-2} \le_{uso} S_{1:k-1} \le_{uso} S_{2:k-1} \le_{uso} S_{2:k} \le_{uso} S_{3:k} \\ &\vdots \\ &S_{1:k-m+1} \le_{uso} S_{1:k-m+2} \le_{uso} S_{2:k-m+2} \le_{uso} S_{2:k-m+3} \le_{uso} \cdots \le_{uso} S_{m-1:k} \le_{uso} S_{m:k} \\ &\vdots \\ &S_{1:1} \le_{uso} S_{1:2} \le_{uso} S_{2:2} \le_{uso} \cdots \le_{uso} S_{k-1:k} \le_{uso} S_{k:k} \end{aligned}$$

Fig. 1 gives a graphical presentation for $k = 4$ with an $\rightarrow$ representing $\le_{uso}$ and where the gray path illustrates the mth row for $m = 3$.

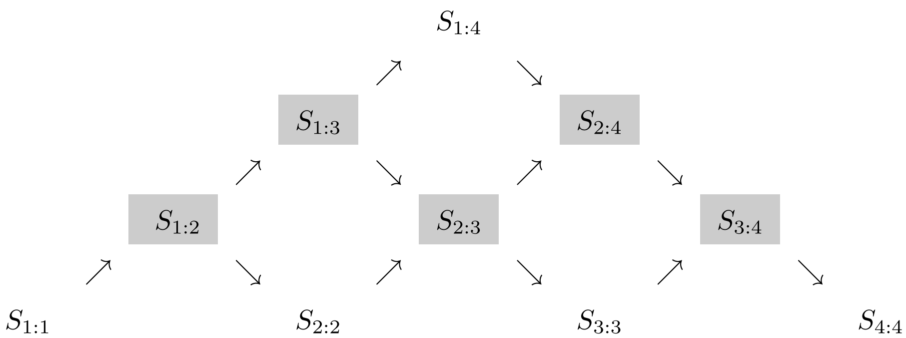

Fig. 1. Graphical presentation of $\{S_{r:s}\}$ for $k = 4$..

We estimate from top to bottom of the pyramidal structure. There are no restrictions on $S_{1:k}$, and we set $\bar{S}_{1:k} = \hat{S}_{1:k}$. In the outward boundaries of mth row of the pyramid, $S_{1:k-m+1} \leq_{uso} S_{1:k-m+2}$ and $S_{m:k} \geq_{uso} S_{m-1:k}$ have only single ordering restrictions from the row above for $2 \leq m \leq k$, and we define their 1-sample type restricted estimators sequentially:

$$\begin{aligned}
\bar{S}_{1:k-m+1}(x) &= \inf_{y \leq x} \frac{\hat{S}_{1:k-m+1}(y)}{\bar{S}_{1:k-m+2}(y)} \bar{S}_{1:k-m+2}(x),\ 2 \leq m \leq k, \\
\bar{S}_{m:k}(x) &= \sup_{y \leq x} \frac{\hat{S}_{m:k}(y)}{\bar{S}_{m-1:k}(y)} \bar{S}_{m-1:k}(x),\ 2 \leq m \leq k.
\end{aligned} \tag{2.1}$$

For $2 \leq r \leq s \leq k-1$, we have a double ordering restriction from above: $S_{r-1:s} \leq_{uso} S_{r:s} \leq_{uso} S_{r:s+1}$. We define our restricted estimators inductively. First note that $\bar{S}_{1:k}, \bar{S}_{1:k-1}$ and $\bar{S}_{2:k}$ are defined by (2.1). Assume that for some $3 \leq m \leq k$, the restricted estimators, $\bar{S}_{1:k-j+1}, \bar{S}_{2:k-j}, \ldots, \bar{S}_{j:k}$ have been defined for all $j = 3, \ldots, m-1$. For $2 \leq r \leq s \leq k-1$ with $s - r = k - m$, define

$$\begin{aligned}
S^{\dagger}_{r:s}(x) &= \inf_{y \leq x} \frac{\hat{S}_{r:s}(y)}{\bar{S}_{r:s+1}(y)} \bar{S}_{r:s+1}(x), &\quad S^{\ddagger}_{r:s}(x) &= \sup_{y \leq x} \frac{\hat{S}_{r:s}(y)}{\bar{S}_{r-1:s}(y)} \bar{S}_{r-1:s}(x), \\
S^{\ddagger\dagger}_{r:s}(x) &= \inf_{y \leq x} \frac{S^{\ddagger}_{r:s}(y)}{\bar{S}_{r:s+1}(y)} \bar{S}_{r:s+1}(x), &\quad S^{\dagger\ddagger}_{r:s}(x) &= \sup_{y \leq x} \frac{S^{\dagger}_{r:s}(y)}{\bar{S}_{r-1:s}(y)} \bar{S}_{r-1:s}(x),
\end{aligned} \tag{2.2}$$

and

$$\bar{S}_{r:s} = [S^{\ddagger\dagger}_{r:s} + S^{\dagger\ddagger}_{r:s}]/2, \quad 2 \leq r \leq s \leq k-1. \tag{2.3}$$

Note that $S^{\dagger}_{r:s}$ ($S^{\ddagger}_{r:s}$) is not necessarily uniformly stochastically larger (smaller) than $\bar{S}_{r-1:s}$ ($\bar{S}_{r:s+1}$). However, a second projection, as in the last line of (2.2) works, i.e., both SFs $S^{\ddagger\dagger}_{r:s}$ and $S^{\dagger\ddagger}_{r:s}$ obey the order restrictions,

$$\bar{S}_{r-1:s} \leq_{uso} S^{\ddagger\dagger}_{r:s} \leq_{uso} \bar{S}_{r+1:s}, \quad \text{and} \quad \bar{S}_{r-1:s} \leq_{uso} S^{\dagger\ddagger}_{r:s} \leq_{uso} \bar{S}_{r+1:s},$$

as shown by the following lemma.

Lemma 2.1. *Let S_1, S_2 and S_3 be SFs.*

(i) Assume $S_1 \leq_{uso} S_3$ and $S_2 \leq_{uso} S_3$. Then $S_2^{\ddagger} \leq_{uso} S_3$, where

$$S_2^{\ddagger}(x) = \sup_{y \leq x} \frac{S_2(y)}{S_1(y)} S_1(x).$$

(ii) Assume $S_1 \leq_{uso} S_2$ and $S_1 \leq_{uso} S_3$. Then $S_2^{\dagger} \geq_{uso} S_1$, where

$$S_2^{\dagger}(x) = \inf_{y \leq x} \frac{S_2(y)}{S_3(y)} S_3(x).$$

Proof. We prove only (i) since the proof of (ii) is similar. By assumption the functions $\theta_1 = S_1/S_3$ and $\theta_2 = S_2/S_3$ are nonincreasing. We claim that

$$\theta(x) := \frac{S_2^{\ddagger}(x)}{S_3(x)} = \sup_{y \leq x} \frac{S_2(y)S_3(y)}{S_1(y)S_3(y)} \frac{S_1(x)}{S_3(x)} = \sup_{y \leq x} \frac{\theta_2(y)}{\theta_1(y)} \theta_1(x)$$

is nonincreasing. Assume that $x \leq \tilde{x}$. Then $\theta_1(x) \geq \theta_1(\tilde{x})$, and we can write

$$\theta(\tilde{x}) = \frac{\theta_1(\tilde{x})}{\theta_1(x)} \max\left\{ \sup_{x < y \leq \tilde{x}} \frac{\theta_2(y)}{\theta_1(y)} \theta_1(x), \theta(x) \right\}.$$

If the maximum is attained in $\theta(x)$, we are done. Otherwise,

$$\sup_{x<y\leq\tilde{x}}\frac{\theta_2(y)}{\theta_1(y)}\geq\sup_{y\leq x}\frac{\theta_2(y)}{\theta_1(y)},$$

and hence

$$\theta(\tilde{x})=\sup_{x<y\leq\tilde{x}}\frac{\theta_2(y)}{\theta_1(y)}\theta_1(\tilde{x})\leq\frac{\theta_2(x)}{\theta_1(\tilde{x})}\theta_1(\tilde{x})=\theta_2(x).$$

Since $\theta(x)\geq\theta_2(x)$, this implies the claim. □

Considering the original projections by Rojo and Samaniego (1993) and Mukerjee (1996) given at (1.2), one might wonder why there is never a second summand in our construction. This second term vanishes since the support of the survival function in the denominator contains the one of the numerator.

We apply part (i) of the lemma with $S_1=\bar{S}_{r-1:s}, S_2=S^{\dagger}_{r:s}$ and $S_3=\bar{S}_{r:s+1}$ and part (ii) with $S_1=\bar{S}_{r-1:s}, S_2=S^{\ddagger}_{r:s}$ and $S_3=\bar{S}_{r:s+1}$. Unfortunately, $S^{\dagger\ddagger}_{r:s}$ is generally not equal to $S^{\ddagger\dagger}_{r:s}$. We could use either or any convex combination due to transitivity of USO. Having no prior preference, we choose their arithmetic average. Our pyramidal construction goes from the top to the bottom until we reach $S^*_i=\bar{S}_{i:i}$, which satisfy the order constraints $S^*_1\leq_{uso}S^*_2\leq_{uso}\cdots\leq_{uso}S^*_k$ by construction. Next we show that these estimators are consistent. The argument follows the same sequential pattern as the construction of the estimators.

Lemma 2.2. *Assume that $\hat{U}$ and $\hat{V}$ are SFs that estimate the SFs U and V, respectively. Assume $\|\hat{U}-U\|\to 0$ a.s., and $\|\hat{V}-V\|\to 0$ a.s., where $\|\cdot\|$ denotes the sup-norm.*

(i) Assume $U\leq_{uso}V$ and let U^* be defined by

$$U^*(x)=\inf_{y\leq x}\frac{\hat{U}(y)}{\hat{V}(y)}\hat{V}(x)I(\hat{V}(x)>0).$$

Then $\|U^*-U\|\to 0$ a.s.

(ii) Assume $U\geq_{uso}V$ and let U^{**} be defined by

$$U^{**}(x)=\sup_{y\leq x}\frac{\hat{U}(y)}{\hat{V}(y)}\hat{V}(x)I(\hat{V}(x)>0)+\hat{U}(x)I(\hat{V}(x)=0).$$

Then $\|U^{**}-U\|\to 0$ a.s.

Proof. The proof uses the following result (Rojo and Samaniego, 1993, Lemmas 1 and 2): If f and g are bounded functions on $[0,x]$, then

$$|\inf_{y\leq x}f(y)-\inf_{y\leq x}g(y)|\leq\sup_{y\leq x}|f(y)-g(y)|, \tag{2.4}$$

$$|\sup_{y\leq x}f(y)-\sup_{y\leq x}g(y)|\leq\sup_{y\leq x}|f(y)-g(y)|. \tag{2.5}$$

For the proof of (i), consider $E=\{\omega:\|\hat{U}-U\|\to 0,\|\hat{V}-V\|\to 0\}$ with $P(E)=0$. For $\hat{V}(x)>0$, we have

$$\begin{aligned}
|U^*(x)-U(x)|&=\left|\inf_{y\leq x}\frac{\hat{U}(y)}{\hat{V}(y)}\hat{V}(x)-\inf_{y\leq x}\frac{U(y)}{V(y)}V(x)\right|\\
&\leq\sup_{y\leq x}\left|\frac{\hat{U}(y)}{\hat{V}(y)}\hat{V}(x)-\frac{U(y)}{V(y)}V(x)\right|\\
&\leq\sup_{y\leq x}\frac{\hat{V}(x)}{\hat{V}(y)}|\hat{U}(y)-U(y)|+\sup_{y\leq x}\frac{U(y)}{\hat{V}(y)}|\hat{V}(x)-V(x)|\\
&\quad+\sup_{y\leq x}\frac{V(x)U(y)}{\hat{V}(y)V(y)}|V(y)-\hat{V}(y)|\\
&\leq\sup_{y\leq x}|\hat{U}(y)-U(y)|+2\sup_{y\leq x}\frac{U(y)}{\hat{V}(y)}\sup_{y\leq x}|\hat{V}(y)-V(y)|\\
&\leq\sup_{y\leq x}|\hat{U}(y)-U(y)|+2\sup_{y\leq x}\frac{V(y)}{\hat{V}(y)}\sup_{y\leq x}|\hat{V}(y)-V(y)|,
\end{aligned}\tag{2.6}$$

where we have used (2.4) for the first inequality, the triangular inequality in the next, the facts that $\hat{V}(x)/\hat{V}(y)\leq 1$ and $V(x)/V(y)\leq 1$ for all $y\leq x$, $|\hat{V}(x)-V(x)|\leq\sup_{y\leq x}|\hat{V}(y)-V(y)|$, and an elementary inequality about supremum of products in the next, and $U(y)\leq V(y)$ for all y in the last.

Uniform consistency of U^* on $[0,\infty)$ is equivalent to uniform consistency on $[0,x]$ for all $x\geq 0$. Suppose an arbitrary $\epsilon>0$ is given. First assume there exists $\tau>0$ such that $0<V(\tau)<\epsilon$. On E we know that $\|\hat{V}-V\|<V(\tau)/2$ and $\|\hat{U}-U\|<\epsilon$ for a sufficiently large sample size and

$$\frac{V(y)}{\hat{V}(y)}\leq\frac{V(y)}{V(y)-V(\tau)/2}\leq\frac{V(\tau)}{V(\tau)-V(\tau)/2}=2,$$

using $V(y)/[V(y)-V(\tau)/2]$ is increasing in $0 \le y \le \tau$. From (2.6) we have

$$\sup_{y \le \tau} |U^*(y) - U(y)| < 5\epsilon \quad \text{on } E. \tag{2.7}$$

If there is no τ such that $0 < V(\tau) < \epsilon$, then there must be a jump $\ge \epsilon$ at the right endpoint of V. In this case, we can use ϵ in place of $V(\tau)$ in the proof of the first case to get (2.7), using the fact that $V(y) \ge \epsilon$ for all y where $V(y) > 0$. Since $\epsilon > 0$ is arbitrary, this completes the proof of (i).

For the proof of (ii), by using the triangular inequality, we note that (recall our notation at the beginning of Section 2 about omitting an indicator function)

$$\begin{aligned} |U^{**}(x) - U(x)| \le & \left| \sup_{y \le x} \frac{\hat{U}(y)}{\hat{V}(y)} \hat{V}(x) - \sup_{y \le x} \frac{U(y)}{V(y)} V(x) \right| \\ & + |\hat{U}(x) I(\hat{V}(x) = 0) - U(x) I(V(x) = 0)|. \end{aligned}$$

The proof of uniform convergence of the first term using (2.5) is similar to that in (i); that of the second term follows from standard probability results. □

The estimators S_i^* are strongly and uniformly consistent by the following Theorem 2.1, and the construction ensures that they satisfy the desired order constraints $S_1^* \le_{uso} S_2^* \le_{uso} \cdots \le_{uso} S_k^*$.

Theorem 2.1. *The estimator $S_i^* = \bar{S}_{i:i}$ in (2.1)–(2.3) is strongly and uniformly consistent for S_i for $1 \le i \le k$ under the USO in (1.4) and the relative sample size assumption (1.5).*

Proof. Under assumption (1.5), $\hat{S}_{r:s}$ is strongly and uniformly consistent for $S_{r:s}$ for all $1 \le r \le s \le k$ from standard probability theory. The conclusion of the theorem then follows simply by applying Lemma 2.2(i) and (ii) alternately to the adjacent pairs

$$\bar{S}_{1:k-m+1} \le_{uso} \bar{S}_{1:k-m+2} \le_{uso} \bar{S}_{2:k-m+2} \le_{uso} \bar{S}_{2:k-m+3} \le_{uso} \cdots \le_{uso} \bar{S}_{m-1:k} \le_{uso} \bar{S}_{m:k}$$

sequentially for $m = 2, 3, \ldots, k-1$. For example, for $m = 2$, we let $U = S_{1:k-1}$ and $V = S_{1:k}$ in Lemma 2.2(i) for consistency of $\bar{S}_{1:k-1}$ and $U = S_{2:k}$ and $V = S_{1:k}$ in Lemma 2.2 (ii) for consistency of $\bar{S}_{2:k}$. For $m = 3$, consistency of $\bar{S}_{1:k-2}$ and $\bar{S}_{3:k}$ follow similarly. For $\bar{S}_{2:k-1}$, we first get consistency of $S^{\dagger}_{2:k-1}$ using $\hat{U} = \hat{S}_{2:k-1}$ and $\hat{V} = \bar{S}_{2:k}$ in Lemma 2.2(i), and then of $S^{\dagger\ddagger}_{2:k-1}$ using $\hat{U} = S^{\dagger}_{2:k-1}$ and $\hat{V} = \bar{S}_{1:k-1}$ in Lemma 2.2 (ii). Consistency of $S^{\ddagger\dagger}_{2:k-1}$, and hence of $\bar{S}_{2:k-1}$, follow similarly. □

In the case of random right censoring of the life distributions, we assume that the entire set of life distributions and censoring distributions are jointly independent. We use the KM estimator S_i^{KM} of S_i when there are no order restrictions. It is well known that the KM estimator S_i^{KM} is strongly and uniformly consistent on $[0, b]$ for all $b < b_i$, where b_i is the minimum of the right endpoints of the supports of the ith life distribution and the corresponding censoring distribution. We consider the estimation under USO only on $[0, b]$ for some $b < \min\{b_i \mid i = 1, \ldots, k\}$.

Now consider the uncensored case with the empirical being generalized by a step function, with possibly unequal jumps, with all the SFs having the support $[0, b]$. Then the estimation and consistency will follow exactly as described above using the empiricals. Thus, the estimators replacing the empiricals by the KM estimators will have the desired consistency on $[0, b]$.

3. Errors in El Barmi and Mukerjee (2016)

For the k-sample problem considered above with the assumption of zero mass at the origin for all populations, let $0 < t_1 < t_2 < \cdots < t_c$ be the distinct observation points of the combined sample, and define $t_0 = 0$. Dykstra et al. (1991) showed that the USO in (1.4) is equivalent to $\theta_i = S_i(t_j)/S_i(t_{j-1})$ is increasing in i for each t_j and $S_i(t) = \prod_{t_j \le t} \theta_i(t_j)$. The empirical estimate of $\theta_i(t_j)$ is $\hat{\theta}_i(t_j) = \hat{S}_i(t_j)/\hat{S}_i(t_{j-1})$. Then they showed that the NPMLE $(\theta_1^*(t_j), \ldots, \theta_k^*(t_j))$ is the isotonic regression of $(\hat{\theta}_1(t_j), \ldots, \hat{\theta}_k(t_j))$ with the weight vector $(n_1 \hat{S}_1(t_{j-1}), \ldots, n_k \hat{S}_k(t_{j-1}))$ for all j, performed independently. El Barmi and Mukerjee (2016) considered essentially the NPMLEs, but, instead of assigning the weight $n_i \hat{S}_i(t_{j-1})$ at t_j, which is the number of items in the ith population at risk at time t_j^-, they assigned the weight $n_i S_i^*(t_{j-1})$ at t_j, arguing that this is the updated "effective" number at risk at time t_j^- using the past information. This forces the estimation to be constructed sequentially. For $k = 2$, Mukerjee (1996) showed that this estimator is the same as the inf/sup definition of his 2-sample estimator in (1.3). The estimator turned out to have nice properties in simulations, but two crucial mistakes were made in proving consistency. Unable to find any direct methods, they looked at the pyramidal structure of $\{S_{r:s}\}$ by their orderings as shown in the last section, and suggested a roundabout proof. Consider the case $k = 3$. They defined the estimators of all of the $\{S_{r:s}\}$ the same way as in Section 2, except for that of $S_{2:2} = S_2$. The estimators of $S_{1:1}, S_{1:2}, S_{2:3}, S_{1:3}$ and $S_{3:3}$ are all consistent as shown in Section 2 since they are estimated under no order restrictions or just one. Letting $\theta_{(\cdot)}(t_j) = S_{(\cdot)}(t_j)/S_{(\cdot)}(t_{j-1})$, and noting that $\theta_{12}(t_j) \le \theta_2(t_j) \le \theta_{23}(t_j)$ They claimed that their original estimator of $S_2(t)$ is given by

$$S_2^*(t) = \prod_{t_j \le t} \bar{\theta}_2(t_j), \quad \text{where} \quad \bar{\theta}_2(t_j) = \bar{\theta}_{1:2}(t_j) \vee \hat{\theta}_2(t_j) \wedge \bar{\theta}_{2:3}(t_j). \tag{3.1}$$

Once $\bar{\theta}_{1:2}, \bar{\theta}_{2:3}$ and $\hat{\theta}_2$ have been computed, one can compute $S_2^*(t_j)$ for all j simultaneously. However, in the original estimation process, it is not possible to compute $S_2^*(t_j)$ without first computing $S_2^*(t_{j-1})$ except for some trivial cases. Thus, S_2^* in (3.1) cannot be the original estimator.

Moreover, El Barmi and Mukerjee (2016) claimed that $S_2^{\dagger}$, defined by

$$S_2^{\dagger}(t) = \prod_{t_j \le t} [\hat{\theta}_2(t_j) \wedge \bar{\theta}_{23}(t_j)], \tag{3.2}$$

is consistent, quoting the consistency of the 2-sample estimator of Mukerjee (1996) that was shown to be equivalent to the original estimator for $k = 2$. However, this equivalence occurred when $\theta_{(\cdot)}(t_j)$ was defined to be $S_{(\cdot)}(t_j)/S^{*}_{(\cdot)}(t_{j-1})$, not $S_{(\cdot)}(t_j)/\hat{S}_{(\cdot)}(t_{j-1})$. In fact, (3.2) defines precisely the NPMLE given in Dykstra et al. (1991) for $k = 2$, and Mukerjee (1996) showed that the estimators are typically inconsistent when all the SFs are continuous. El Barmi and Mukerjee apologize for such serious mistakes.

We conclude in remarking that the computational cost of the estimators suggested in this paper grows at rate k^2 by the nature of the pyramidal structure in Section 2 and therefore estimation will be costly if we are given a large number k of classes. The estimators originally given in El Barmi and Mukerjee (2016) are more attractive in this regard since we can compute them by sequentially isotonizing certain thresholds at all distinct observations and hence the computational cost grows only linearly with the sample size. However, the problem of consistency remains unresolved.

Acknowledgments

Sebastian Arnold and Johanna Ziegel gratefully acknowledge funding from Swiss National Science Foundation.

Appendix A. Supplementary material

The supplementary material to this article contains a short simulation study and the characterization of some alternative consistent estimators which are slightly simpler but less optimal than the estimators we presented in Section 2. We also show weak convergence of the restricted estimators in a special case and analyze a real data example taken from Dykstra et al. (1991) which gives a graphical comparison of the restricted and unrestricted estimators of the survival times for patients with carcinoma of the oropharynx. The data is found in Table 1 of Dykstra et al. (1991).

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2024.110045.

References

Dykstra, R., Kochar, S., Robertson, T., 1991. Statistical inference for uniform stochastic ordering in several populations. Ann. Statist. 19, 870–888.
El Barmi, H., Mukerjee, H., 2016. Consistent estimation of survival functions under uniform stochastic ordering; The k-sample case. J. Multivariate Anal. 144, 99–109.
Mukerjee, H., 1996. Estimation of survival functions under uniform stochastic ordering. J. Amer. Statist. Assoc. 91, 1684–1689.
Rojo, J., 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., Samaniego, F.J., 1993. On estimating a survival curve subject to a uniform stochastic ordering constraint. J. Amer. Statist. Assoc. 88, 566–572.
Shaked, M., Shanthikumar, J.G., 2007. Stochastic Orders. Springer, New York.