Statistica Sinica **27** (2017), 645-664
doi:http://dx.doi.org/10.5705/ss.202015.0042

TESTING FOR UNIFORM STOCHASTIC ORDERING VIA EMPIRICAL LIKELIHOOD UNDER RIGHT CENSORING

Hammou El Barmi

Baruch College, City University of New York

Abstract: Empirical likelihood based tests for the presence of uniform stochastic ordering (or hazard rate ordering) among two univariate distributions functions (DFs) are developed when the data are right censored in the one- and two-sample cases. The proposed test statistics are formed by taking the supremum of some functional of localized empirical likelihood test statistics. The null asymptotic distributions of these test statistics are distribution-free and have simple representations in terms of a standard Brownian motion. Simulations show that the tests we propose outperform, in terms of power, the one sided log-rank test at many distributions. The stochastic ordering case is shown to be a special case of our procedure. We illustrate our theoretical results with an example.

Key words and phrases: Empirical likelihood, stochastic ordering, uniform stochastic ordering.

1. Introduction

The concept of ordering univariate random variables according to their DFs is an important area of statistics and applied probability. Many orders are discussed extensively in the literature and they include in increasing order of strength: stochastic ordering, uniform stochastic ordering, and likelihood ratio ordering. Such orderings arise in the biomedical sciences and reliability engineering. For a broad overview of these and other stochastic orders, see Shaked and Shanthikumar (2006).

Let X_1 and X_2 be two nonnegative random variables with DFs F_1 and F_2, respectively. F_1 is said to be *stochastically smaller* than F_2 ($F_1 \preceq_{SO} F_2$) if $\bar{F}_1 \leq \bar{F}_2$ where $\bar{F}_i = 1 - F_i$ is the survival function (SF) corresponding to F_i. On the other hand F_1 is said to be *uniformly stochastically smaller* than F_2 ($F_1 \preceq_{USO} F_2$) if the ratio $\bar{F}_1(t)/\bar{F}_2(t)$ is nonincreasing for t in $(-\infty, F_2^{-1}(1))$. When F_1 and F_2 are absolutely continuous with hazard rates λ_1 and λ_2, respectively, uniform stochastic ordering is equivalent to $\lambda_1(t) \geq \lambda_2(t)$ and hence, uniform stochastic ordering is sometimes called hazard rate ordering. This ordering is also equivalent to

$$P[X_1 > t + s | X_1 > t] \leq P[X_2 > t + s | X_2 > t], \quad \text{for all}\, s \geq 0, t.$$

That is, the conditional distribution of X_1, given that $X_1 > t$, is stochastically smaller than that corresponding to X_2. For this reason, uniform stochastic ordering is a more useful concept than the usual stochastic ordering, in applications. As Dykstra, Kochar and Robertson (1991) point out, uniform stochastic ordering is especially of interest when the populations correspond to survival times for different medical treatments. Even if the corresponding survival times are stochastically ordered initially, they may not be when the patients are examined at a later time. However, if they are uniformly stochastically ordered, then clearly one treatment is better than the others.

Estimation of DFs under uniform stochastic ordering has been considered in the literature. Dykstra, Kochar and Robertson (1991) derived the nonparametric maximum likelihood estimators (NPMLE) of k uniformly stochastically ordered DFs when $k \geq 2$. They also studied the likelihood ratio test (LRT) for equality of multinomial distributions against the alternative that they are uniformly stochastically ordered. In the one-sample case, Rojo and Samaniego (1991) obtained the NPMLE of the DF F_1 when $F_1 \preceq_{USO} F_2$ with F_2 known, continuous, and strictly increasing, and showed that it is not consistent. Mukerjee (1996) showed that this is also the case in the two-sample problem. Alternative estimators that are consistent have since been developed in the one- and the two-sample cases. For more on this, see Mukerjee (1996) and Rojo and Samaniego (1993). The weak convergence of these new estimators was studied by Arcones and Samaniego (2000). Recently, El Barmi and Mukerjee (2015) developed consistent estimators in the $k-$sample case, and El Barmi and McKeague (2016) developed EL-based tests for the presence of this ordering in the uncensored case.

This paper develops EL-based one- and two-sample tests for the presence of uniform stochastic ordering under right censoring. The EL method was originally used by Owen (1988, 1990) to construct confidence regions for parameters defined by general classes of estimating equations. The EL approach has been extended to many areas in statistics, references can be found in Owen (2001). Einmahl and McKeague (2003) developed a localized version of EL to allow nonparametric hypothesis testing. The EL approach was extended in the uncensored case to testing for the presence of stochastic ordering and uniform stochastic ordering in the $k-$sample case in El Barmi and McKeague (2013) and El Barmi and McKeague (2016), respectively. Chang (2014) extended their test for stochastic ordering to the censored case, and in the present paper we do the same for uniform stochastic ordering.

In Section 2, EL-based tests are developed for the presence of uniform stochastic ordering in the one- and the two-sample cases, and in Section 3 we show that a test for stochastic ordering is obtained as a special case of our test for uniform stochastic. Chang (2014) also treated the stochastic ordering problem and our results improve on hers. In Section 4 we give the results of simulations that compare our tests in Section 2 with the one sided log-rank test, and in Section 5 we discuss an example. Some concluding remarks are given in Section 6. Throughout the paper, $\stackrel{d}{\to}$, $\stackrel{P}{\to}$ and $\stackrel{w}{\Rightarrow}$ are used to denote convergence in distribution, convergence in probability, and weak convergence, respectively. The proofs are relegated to the Appendix.

2. Uniform Stochastic Ordering

In this section we confine attention to uniform stochastic ordering. We first introduce the one-sample case and then extend the results to the two-sample case. To derive the asymptotic distributions of our test statistics, we appeal to some elegant results obtained by Præstgaard and Huang (1996) who developed the asymptotic theory for the NPMLE of two stochastically ordered DFs.

2.1. One-sample case

Suppose $X_1, X_2, \ldots, X_n$ are independent lifetimes with a common but unknown DF F and SF $\bar{F}$, and independent hereof, $C_1, C_2, \ldots, C_n$ is a random sample of censoring times with DF G and SF $\bar{G}$. What we observe are the n independent and identically distributed pairs $\{(Z_i, \delta_i) \equiv (\min\{X_i, C_i\}, I[X_i \le C_i]), i = 1, \ldots, n\}$. We assume the setting as Dykstra (1982) in that complete observations (which we call deaths) occur on a subset of the times $T_1 < T_2 < \ldots < T_m$, and let $T_0 = 0$ and $T_{m+1} = \infty$ for convenience. Let

$$d_i = \text{number of complete observations at } T_i;$$
$$l_i = \text{ number of observations censored in } [T_i, T_{i+1});$$
$$n_i = \sum_{j=i}^{n} (d_j + l_j) = \text{the number of observations at risk at time } T_i;$$
$$\theta_i = \frac{\bar{F}(T_i)}{\bar{F}(T_{i-1})}, i = 1, 2, \ldots, m.$$

In addition, we let $n_0 = n, d_0 = 0$ and $n_{m+1} = 0$. We note that

$$\bar{F}(T_j) = \prod_{i=1}^{j} \theta_i, j = 1, \ldots, m. \tag{2.1}$$

Suppose we want to test

$$H_0 : F = F_0 \quad \text{against} \quad H_1 : F \prec_{USO} F_0,$$

where F_0 is a known DF and $\prec_{USO}$ denotes $\preceq_{USO}$ with equality excluded. The DF F_0 may, for instance, be some reference DF or a DF available from demographical studies.

Let $\pi(t)=\bar F(t)\bar G(t)$ denote the probability of remaining under study at time t. The approach we adapt is based on testing a family of "local" hypotheses of the form

$$H_0^{s,t}: \frac{\bar F(t)}{\bar F_0(t)}=\frac{\bar F(s)}{\bar F_0(s)} \quad \text{versus} \quad H_1^{s,t}: \frac{\bar F(t)}{\bar F_0(t)}<\frac{\bar F(s)}{\bar F_0(s)}$$

for fixed $0\le s<t\le b$ for some b with $\pi(b)>0$. The test statistic for H_0 against H_1 is some functional of the test statistics corresponding to these local hypotheses. The local EL procedure rejects $H_0^{s,t}$ for small values of

$$\mathcal{R}(s,t)=\frac{\sup\left\{L(\bar F):\phi=\phi_0\right\}}{\sup\left\{L(\bar F):\phi\le\phi_0\right\}}, \tag{2.2}$$

where $\phi=\phi(F;s,t)\equiv\bar F(t)/\bar F(s), \phi_0=\phi(F_0;s,t)\equiv\bar F_0(t)/\bar F_0(s)$ and

$$\begin{aligned}
L(\bar F)&=\prod_{\text{uncensored}}[\bar F(Z_i)-\bar F(Z_{i-1})]\prod_{\text{censored}}\bar F(Z_i)\\
&=\prod_{i=1}^m[\bar F(T_i)-\bar F(T_{i-1})]^{d_i}\prod_{i=0}^m[\bar F(T_i)]^{n_i-n_{i+1}-d_i}\\
&=\prod_{i=1}^m\left[1-\frac{\bar F(T_i)}{\bar F(T_{i-1})}\right]^{d_i}\left[\frac{\bar F(T_i)}{\bar F(T_{i-1})}\right]^{n_i-d_i}\\
&=\prod_{i=1}^m\theta_i^{n_i-d_i}(1-\theta_i)^{d_i}.
\end{aligned}$$

We have suppressed the argument (s,t) in ϕ for simplicity. The suprema in (2.2) are taken over DFs F that are supported by the data points and by convention we tale $\sup\emptyset=0$ and $0/0=1$. Kaplan and Meier (1958) showed that, under no restrictions, the NPMLE of $\bar F$ is the maximizer of $L(\bar F)$ and is given by

$$\hat{\bar F}(t)=\prod_{\{T_i\le t\}}\hat\theta_i,$$

where

$$\hat\theta_i=\frac{n_i-d_i}{n_i},\ i=1,2,\ldots,m, \tag{2.3}$$

and $\prod_\emptyset\equiv 1$.

By (2.1), the parameter of interest $\phi=\prod_{i=n(s)+1}^{n(t)}\theta_i$, where for $u>0, n(u)\equiv\sum_{i=1}^n I[T_i\le u]$ is the number of distinct uncensored observations in the time interval $[0,u]$. Write

$$\begin{aligned}
L(\bar F)=&\left\{\prod_{i=1}^{n(s)}\theta_i^{n_i-d_i}(1-\theta_i)^{d_i}\right\}\times\left\{\prod_{i=n(s)+1}^{n(t)}\theta_i^{n_i-d_i}(1-\theta_i)^{d_i}\right\}\\
&\times\left\{\prod_{i=n(t)+1}^{n}\theta_i^{n_i-d_i}(1-\theta_i)^{d_i}\right\}.
\end{aligned}$$

The three terms in braces in $L(\bar{F})$ can be maximized individually under $H_j^{s,t}$, $j = 0, 1$. Since the constraints on ϕ under both hypotheses have no effect on the first and the third factors, these two factors make no contribution to $\mathcal{R}(s,t)$ and hence they cancel out. The remaining term

$$\prod_{i=n(s)+1}^{n(t)} \theta_i^{n_i-d_i}(1-\theta_i)^{d_i} \tag{2.4}$$

is then maximized under $H_0^{s,t}$ and $H_1^{s,t}$. Let $\psi_i = \log(\theta_i)$. In terms of this new parametrization, to compute $\mathcal{R}(s,t)$, it suffices to maximize

$$\sum_{i=n(s)+1}^{n(t)} \{(n_i-d_i)\psi_i + d_i \log(1-e^{\psi_i})\} \tag{2.5}$$

subject to the constraint $\sum_{i=n(s)+1}^{n(t)} \psi_i = \log(\phi_0)$ under $H_0^{s,t}$ or $\sum_{i=n(s)+1}^{n(t)} \psi_i \leq \log(\phi_0)$ under $H_1^{s,t}$. A simple Lagrangian argument shows that, under $H_0^{s,t}$, (2.5) is maximized by a vector whose ith component is

$$\bar{\psi}_i \equiv \log(\overline{\theta}_i) = \log\left(\frac{n_i+\beta_n(s,t)-d_i}{n_i+\beta_n(s,t)}\right), \ i = n(s)+1, n(s)+2, \ldots, n(t),$$

and $\beta_n(s,t)$ is the unique solution for β in the equation

$$H(\beta) \equiv \prod_{i=n(s)+1}^{n(t)} \frac{n_i+\beta-d_i}{n_i+\beta} - \phi_0 = 0. \tag{2.6}$$

Let $(\hat{\psi}_{n(s)+1}, \hat{\psi}_{n(s)+2}, \ldots, \hat{\psi}_{n(t)}) = (\log(\hat{\theta}_{n(s)+1}), \log(\hat{\theta}_{n(s)+2}), \ldots, \log(\hat{\theta}_{n(t)}))$ where $\hat{\theta}_i$ is defined in (2.3). Clearly, under no constraints, (2.4) is maximized by $(\hat{\theta}_{n(s)+1}, \hat{\theta}_{n(s)+2}, \ldots, \hat{\theta}_{n(t)})$. So when $\sum_{i=n(s)+1}^{n(t)} \hat{\psi}_i \leq \log(\phi_0)$, $(\hat{\psi}_{n(s)+1}, \hat{\psi}_{n(s)+1}, \ldots, \hat{\psi}_{n(t)})$ maximizes (2.5) under $H_1^{s,t}$ and this occurs if and only if $\beta_n(s,t) \geq 0$ where $\beta_n(s,t)$ is defined as before. On the other hand, if $\sum_{i=n(s)+1}^{n(t)} \hat{\psi}_i > \log(\phi_0)$, strict concavity of the objective function in (2.5) implies that the maximum occurs at $(\bar{\psi}_{n(s)+1}, \bar{\psi}_{n(s)+1}, \ldots, \bar{\psi}_{n(t)})$. These steps together imply that

$$\mathcal{R}(s,t) = \begin{cases} \frac{\prod_{i=n(s)+1}^{n(t)} \overline{\theta}_i^{n_i-d_i}(1-\overline{\theta}_i)^{d_i}}{\prod_{i=n(s)+1}^{n(t)} \hat{\theta}_i^{n_i-d_i}(1-\hat{\theta}_i)^{d_i}}, & \beta_n(s,t) \geq 0, \\ 1, & \text{otherwise.} \end{cases}$$

To derive the asymptotic distribution of the test statistic that we propose for testing H_0 against H_1, we assume that the DF F has density f on $]0, b[$ with a bounded, continuous, and strictly positive hazard rate $\lambda(t) = f(t)/\bar{F}(t)$ on $[0, b]$. Define

$$c(t) = \int_0^t \frac{\lambda(u)du}{\pi(u)}$$

and note that $c(b) < \infty$ since $\pi(b) > 0$. For $(s,t) \in [0,b]^2$, let

$$B_n(s,t) = \frac{(c(t) - c(s))\beta_n(s,t)}{\sqrt{n}}.$$

This process has been studied extensively in Præstgaard and Huang (1996). They showed in particular that, under H_0 and the assumptions we made about F, this process converges weakly to a Gaussian process on $[0,b]^2$, specifically,

$$B_n \overset{w}{\Longrightarrow} U \tag{2.7}$$

in $\ell^\infty[0,b]^2$ where $U(s,t) = W(c(t)) - W(c(s))$ and W is a standard Brownian motion. It can be shown using a Taylor expansion of $\log(1+y)$ about $y = 0$ that, under H_0 (see proof of Theorem 1),

$$\begin{aligned} -2\log(\mathcal{R}(s,t)) &= [\beta_n(s,t)^+]^2 \sum_{i=1}^n \frac{d_i}{n_i(n_i - d_i)} I(s < T_i \le t) \\ &= (\hat{c}(t) - \hat{c}(s)) \frac{[\beta_n^+(s,t)]^2}{n} + o_p(1), \end{aligned}$$

where $a^+ = \max(a, 0)$ and

$$\hat{c}(t) = n \sum_{T_i \le t} \frac{d_i}{n_i(n_i - d_i)}. \tag{2.8}$$

Since $\hat{c}(t)$ is a consistent estimator of $c(t)$,

$$-2\log(\mathcal{R}(s,t)) \overset{d}{\to} \frac{[U^+(s,t)]^2}{c(t) - c(s)} \equiv \frac{[(W(c(t)) - W(c(s)))^+]^2}{c(t) - c(s)} \overset{d}{=} Z^{+2}$$

using (2.7), Slutsky's, and Continuous Mapping Theorems, where $Z \sim N(0,1)$. To test H_0 against H_1 we propose to use

$$\mathcal{T}_n = \sup_{0 \le s < t \le b} \sqrt{\frac{-2(\hat{c}(t) - \hat{c}(s)) \log(\mathcal{R}(s,t))}{\hat{c}(b)}}.$$

An appealing aspect of this test is that it is an asymptotically similar test under H_0, its null asymptotic distribution does not depend on the underlying DF F or the censoring DF G.

Theorem 1. *Under H_0, if $\pi(b) > 0$, then*

$$\mathcal{T}_n \overset{d}{\to} \sup_{0 \le u \le 1} |W(u)|,$$

where W is a standard Brownian motion.

The distribution of $\sup_{0\le u\le 1}|W(u)|$ has been extensively studied in the literature (see Billingsley (1968, p.79)). In particular,

$$P(\mathcal{T}_n > t) \to P(\sup_{0\le u\le 1}|W(u) > t) = 1 - \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\left\{-\frac{\pi^2(2k+1)^2}{8t^2}\right\} \quad (2.9)$$

as $n\to\infty$. Hall and Wellner (1979) have shown that the approximation of the right hand side of (2.9) by $4(1-\Phi(t))$ gives a three decimal place accuracy for $t > 1.4$. Here Φ is the standard normal DF. Using this approximation, the asymptotic 0.90, 0.95 and 0.99 quantiles of $\mathcal{T}_n$ are 1.96, 2.241 and 2.807.

Remark. Clearly a similar test for $H_0: F = F_0$ against $H_1: F_0 \prec_{USO} F$ can be developed using the same technique. In this case

$$-2\log(\mathcal{R}(s,t)) = -2\sum_{i=n(s)+1}^{n(t)}\left[(n_i - d_i)\log\left(\frac{\bar{\theta}_i}{\hat{\theta}_i}\right) + d_i\log\left(\frac{1-\bar{\theta}_i}{1-\hat{\theta}_i}\right)\right] I[\beta_n(s,t)\le 0],$$

where $\beta_n(s,t)$ is as defined in (2.6). The proposed test statistic is again $\mathcal{T}_n$ defined above, and its null asymptotic distribution is given in Theorem 1.

2.2. Two-sample case

In this section we assume that we have sets of right censored survival data from two populations. Specifically, we assume that the standard setting of right censoring holds in that we only observe

$$(Z_{ij}, \delta_{ij}) \equiv (\min\{X_{ij}, C_{ij}\}, I[X_{ij}\le C_{ij}]), \quad i = 1, 2, \quad j = 1, 2, \ldots, n_i,$$

where $(X_{i1}, X_{i2}, \ldots, X_{in_i})$ and $(C_{1i}, C_{i2}, \ldots, C_{in_i})$ are independent nonnegative random samples and $X_{ij}(C_{ij})$ has DF $F_i(G_i)$ and SF $\bar{F}_i(\bar{G}_i)$.

Let $\pi_i(t) = \bar{F}_i(t)\,\bar{G}_i(t)$ denote the probability of remaining under study at time t in the ith population. Henceforth we consider the setting as in Dykstra (1982) and assume that complete observations occur on a subset of the times $T_1 < T_2 < \ldots < T_m$, and let $T_0 = 0$ and $T_{m+1} = \infty$ for convenience. Let $n = n_1 + n_2$ and assume that $\lim_{n\to\infty} n_1/n \to \alpha$ where $0 < \alpha < 1$. Define, for $j = 1, \ldots, m$,

d_{ij} = number of complete observations from the ith sample at T_j;

l_{ij} = number of observations censored from the ith sample at $[T_j, T_{j+1})$;

$$n_{ij} = \sum_{l=j}^{n_i}(d_{il} + l_{il})$$
= the number of observations at risk from the ith sample at time T_j

and $n_{i0} = n_i$ and $d_{i0} = 0$. In addition, define

$$\theta_{ij} = \frac{\bar{F}_i(T_j)}{\bar{F}_i(T_{j-1})}, j = 1, 2, \ldots, m,$$

and note that

$$\bar{F}_i(T_j) = \prod_{l=1}^{j} \theta_{il}, \quad i = 1, 2.$$

We develop an EL-based test for $H_0 : F_1 = F_2$ versus $H_1 : F_1 \prec_{USO} F_2$. We again test a family of "local" null hypotheses of the form

$$H_0^{s,t} : \frac{\bar{F}_1(t)}{\bar{F}_2(t)} = \frac{\bar{F}_1(s)}{\bar{F}_2(s)} \quad \text{versus} \quad H_1^{s,t} : \frac{\bar{F}_1(t)}{\bar{F}_2(t)} < \frac{\bar{F}_1(s)}{\bar{F}_2(s)}$$

for a given (s, t) where $0 \le s < t \le b$ for some b with $\pi_i(b) > 0, i = 1, 2$. The local EL procedure rejects $H_0^{s,t}$ for small values of

$$\mathcal{R}(s,t) = \frac{\sup\left\{L(\bar{F}_1)L(\bar{F}_2) : \phi_1 = \phi_2\right\}}{\sup\left\{L(\bar{F}_1)L(\bar{F}_2) : \phi_1 \le \phi_2\right\}},$$

where $\phi_i = \phi(F_i; s, t) \equiv \bar{F}_i(t)/\bar{F}_i(s)$ and

$$L(\bar{F}_i) = \prod_{j=1}^{m} \theta_{ij}^{n_{ij}-d_{ij}} (1 - \theta_{ij})^{d_{ij}}, i = 1, 2.$$

Clearly, under no constraints, $L(\bar{F}_i)$ achieves its maximum value at a vector whose jth component is

$$\hat{\theta}_{ij} = \frac{(n_{ij} - d_{ij})}{n_{ij}}, \quad j = 1, 2, \ldots, n_i \quad \text{and} \quad i = 1, 2. \tag{2.10}$$

Write

$$L(\bar{F}_i) = \Big\{ \prod_{j=1}^{n_i(s)} \theta_{ij}^{n_{ij}-d_{ij}} (1 - \theta_{ij})^{d_{ij}} \Big\} \times \Big\{ \prod_{j=n(s)+1}^{n(t)} \theta_{ij}^{n_{ij}-d_{ij}} (1 - \theta_{ij})^{d_{ij}} \Big\}$$
$$\times \Big\{ \prod_{j=n(t)+1}^{n} \theta_{ij}^{n_{ij}-d_{ij}} (1 - \theta_{ij})^{d_{ij}} \Big\},$$

where $n(u)$ is defined as before. Here $H_0^{s,t} : \phi_1 = \phi_2$ is equivalent to $\prod_{j=n(s)+1}^{n(t)} \theta_{1j}$ $= \prod_{j=n(s)+1}^{n(t)} \theta_{2j}$, while $H_1^{s,t} : \phi_1 < \phi_2$ is equivalent to $\prod_{j=n(s)+1}^{n(t)} \theta_{1j} < \prod_{j=n(s)+1}^{n(t)}$ θ_{2j}. As the constraints for the first and the third terms of $L(\bar{F}_i), i = 1, 2$, are the same under both hypotheses, these terms make no contribution to $\mathcal{R}(s,t)$ and cancel out. The remaining term

$$\prod_{i=1}^{2}\prod_{j=n(s)+1}^{n(t)} \theta_{ij}^{n_{ij}-d_{ij}}(1-\theta_{ij})^{d_{ij}} \tag{2.11}$$

is then maximized under $H_0^{s,t}$ and $H_1^{s,t}$. Take $\psi_{ij} = \log(\theta_{ij})$ and, in terms of this new parametrization, to compute $\mathcal{R}(s,t)$, it suffices to maximize

$$\sum_{i=1}^{2}\sum_{j=n(s)+1}^{n(t)} \{(n_{ij}-d_{ij})\psi_{ij} + d_{ij}\log(1-e^{\psi_{ij}})\}, \tag{2.12}$$

subject to the constraint $\sum_{i=1}^{2}\sum_{j=n(s)+1}^{n(t)}(-1)^{i-1}\psi_{ij} = 0$ under $H_0^{s,t}$ or $\sum_{i=1}^{2}\sum_{j=n(s)+1}^{n(t)}(-1)^{i-1}\psi_{ij} \le 0$ under $H_1^{s,t}$. Again, a Lagrangian argument shows that under $H_0^{s,t}$, (2.12) is maximized by $\{\bar\psi_{ij}, j = n(s)+1, n(s)+2, \ldots, n(t), i = 1, 2\}$, where

$$\bar\psi_{1j} \equiv \log(\bar\theta_{1j}) = \log\left(\frac{n_{1j}+\beta_n(s,t)-d_{1j}}{n_{1j}+\beta_n(s,t)}\right),$$
$$\bar\psi_{2j} \equiv \log(\bar\theta_{2j}) = \log\left(\frac{n_{2j}-\beta_n(s,t)-d_{2j}}{n_{2j}-\beta_n(s,t)}\right)$$

and $\beta_n(s,t)$ is the unique solution for β in the equation

$$H(\beta) \equiv \prod_{j=n(s)+1}^{n(t)} \frac{n_{1j}+\beta-d_1j}{n_{1j}+\beta} - \prod_{j=n(s)+1}^{n(t)} \frac{n_{2j}-\beta-d_2j}{n_{2j}-\beta} = 0. \tag{2.13}$$

For $i = 1, 2$, let $(\hat\psi_{i,n(s)+1}, \hat\psi_{i,n(s)+2}, \ldots, \hat\psi_{i,n(t)}) = (\log(\hat\theta_{i,n(s)+1}), \log(\hat\theta_{i,n(s)+2}), \ldots, \log(\hat\theta_{i,n(t)}))$ where $\hat\theta_{ij}$ is defined in (2.10). Under no constraints, (2.11) is maximized by $\{\hat\theta_{ij}, j = n(s)+1, n(s)+2, \ldots, n(t), i = 1, 2\}$. As a result, when $\sum_{i=1}^{2}\sum_{j=n(s)+1}^{n(t)}(-1)^{i-1}\hat\psi_{ij} \le 0$, $\{\hat\psi_{ij}, j = n(s)+1, n(s)+2, \ldots, n(t), i = 1, 2\}$ maximize (2.12) under $H_1^{s,t}$ and this occurs if and only if $\beta_n(s,t) \ge 0$ where $\beta_n(s,t)$ is defined in (2.13). Otherwise, strict concavity of the objective function in (2.12) implies that the maximum occurs at $\{\bar\psi_{ij}, j = n(s)+1, n(s)+2, \ldots, n(t), i = 1, 2\}$. Consequently, we have

$$\mathcal{R}(s,t) = \begin{cases} \dfrac{\prod_{i=1}^{2}\prod_{j=n(s)+1}^{n(t)} \bar\theta_{ij}^{n_{ij}-d_{ij}}(1-\bar\theta_{ij})^{d_{ij}}}{\prod_{i=1}^{2}\prod_{j=n(s)+1}^{n(t)} \hat\theta_{ij}^{n_{ij}-d_{ij}}(1-\hat\theta_{ij})^{d_{ij}}}, & \beta_n(s,t) \ge 0, \\ 1, & \text{otherwise.} \end{cases}$$

To derive the null asymptotic distribution of the test statistic that we propose, we rely on results in Præstgaard and Huang (1996). To this end we assume that under H_0, the DF, F, has density f on $]0, b[$ with a bounded, continuous and strictly positive hazard rate $\lambda(t) = f(t)/\bar F(t)$ on $[0, b]$. Define

$$c_i(t) = \int_0^t \frac{\lambda(u)du}{\pi_i(u)}$$

and note that $c_i(b) < \infty$. Let

$$c(t) = \alpha^{-1}c_1(t) + (1-\alpha)^{-1}c_2(t) \text{ and } B_n(s,t) = (c(t) - c(s))\frac{\beta_n(s,t)}{\sqrt{n}}.$$

Under these assumptions, Præstgaard and Huang (1996) showed that

$$B_n \overset{w}{\Longrightarrow} U \tag{2.14}$$

on $\ell^\infty[0,b]^2$ where $U(t,s) = V(t) - V(s)$,

$$V(t) = \alpha^{-1/2}W_1(c_1(t)) + (1-\alpha)^{-1/2}W_2(c_2(t))$$

for all t and W_1 and W_2 are independent standard Brownian motions. Using a Taylor expansion of $\log(1+y)$ about $y = 0$, we get under H_0 (see proof of Theorem 2)

$$-2\log(\mathcal{R}(s,t)) = (\hat{c}(t) - \hat{c}(s))\frac{[\beta_n^+(s,t))]^2}{n} + o_P(1),$$

where

$$\hat{c}(t) = \frac{n}{n_1}\hat{c}_1(t) + \frac{n}{n_2}\hat{c}_2(t),$$
$$\hat{c}_i(t) = n_i \sum_{T_j \le t} \frac{d_{ij}}{n_{ij}(n_{ij} - d_{ij})}, \quad i = 1,2. \tag{2.15}$$

Since $\hat{c}(t)$ is a consistent estimator of $c(t)$, (2.14) and the Continuous Mapping Theorem imply that

$$-2\log(\mathcal{R}(s,t)) \overset{d}{\to} \frac{[(V(t) - V(s)^+]^2}{c(t) - c(s)} \overset{d}{=} Z^{+2},$$

where $Z \sim N(0,1)$. Analogously to the one-sample case, we propose

$$\mathcal{T}_n = \sup_{0 \le s < t < b} \sqrt{\frac{-2(\hat{c}(t) - \hat{c}(s))\log(\mathcal{R}(s,t))}{\hat{c}(b)}}$$

as a test for H_0 against H_1.

Theorem 2. *Under H_0, if $\pi_i(b) > 0, i = 1,2$,*

$$\mathcal{T}_n \overset{d}{\to} \sup_{0 \le u \le 1} |W(u)|,$$

where W is a standard Brownian motion.

3. Stochastic Ordering

This section discusses how to extend our tests to test $H_0 : F = F_0$ against $H_1 : F \prec_{SO} F_0$ where F_0 is a (pre-)specified DF in the one-sample and $H_0 : F_1 = F_2$ against $H_1 : F_1 \prec_{SO} F_2$ in the two-sample case. Here $\prec_{SO}$ is used to denote $\preceq_{SO}$ with equality excluded. We show that the stochastic ordering problem is a special case of the uniform stochastic ordering problem. Assuming the settings of Sections 2.1 and 2.2 hold in the one- and the two-sample cases, respectively, the results are easily modified to handle this situation since $\bar{F}_1(0) = \bar{F}_2(0) = 1$,

$$\bar{F}_1(t) \le \bar{F}_2(t) \Longleftrightarrow \phi_1(0,t) = \frac{\bar{F}_1(t)}{\bar{F}_1(0)} \le \phi_2(0,t) = \frac{\bar{F}_2(t)}{\bar{F}_2(0)}.$$

As a consequence, by considering only "local" hypotheses of the form $H_0^{0,t}$ and $H_1^{0,t}, t > 0$, or equivalently, by setting $s = 0$ throughout the previous section, the test $\mathcal{T}_n$ reduces to a test for the presence of the classical stochastic ordering. In this case, the test statistic, now denoted $\mathcal{S}_n$, is

$$\mathcal{S}_n = \sup_{0<t<b} \sqrt{\frac{-2\hat{c}(t)\log(\mathcal{R}(0,t))}{\hat{c}(b)}},$$

where $\mathcal{R}(s,t)$ and $\hat{c}$ are defined in Section 2.1 for the one-sample and in Section 2.2 in the two sample case. A careful inspection of the proofs of Theorems 1 and 2 shows that, under H_0,

$$\mathcal{S}_n \xrightarrow{d} \sup_{0\le t\le 1} W(t) \stackrel{d}{=} |N(0,1)| \tag{3.1}$$

since $W(0) = 0$ in the proof of Theorem 1 and $V(0) = 0$ in the proof of Theorem 2, with probability 1. The last equality follows from Billingsley (1968).

4. Simulations

In this section we report on a simulation study to compare the performance of our test $\mathcal{T}_n$ with the standard log-rank test denoted $\mathcal{R}_n$. In each simulation run, 10,000 data sets were used to approximate the rejection probabilities at a nominal level of $\alpha = 0.05$; the critical value for $\mathcal{T}_n$ was taken to be 2.241. We take b to be a number less than the minimum of the largest uncensored observation(s) from the one-(two-) sample(s).

In the one-sample case, the known DF, F_0, is a Weibull distribution with shape and scale parameters equal to 2 and $\sqrt{2}$, respectively. The data were simulated from a DF, F, with hazard rate $\lambda(t) = tI(0 < t < 1) + atI(t \ge 1)$ for various choices of $a \ge 1$. The censoring DF was taken to be Weibull with shape

Table 1. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 10% censoring.

	$n=30$		$n=50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.055	0.059	0.053	0.046
1.20	0.184	0.140	0.209	0.174
1.40	0.272	0.239	0.347	0.335
1.60	0.393	0.354	0.546	0.514
1.80	0.541	0.467	0.696	0.674
2.00	0.670	0.583	0.801	0.797
2.20	0.761	0.684	0.905	0.882
2.40	0.845	0.759	0.935	0.930
2.60	0.890	0.817	0.976	0.966
2.80	0.936	0.866	0.983	0.780
3.00	0.960	0.902	0.996	0.990

parameter 2, and scale parameter specified in a way that produces 10% , 25% or 50% censoring. The results are given Tables 1, 2, and 3 and they show that the new tests outperform, in terms of power, the one-sample one-sided log-rank test $\mathcal{R}_n$.

In the two-sample case, we considered two scenarios. In the first, we took F_1 to be the uniform(0, 1) and F_2 to be the uniform(0, a), for various choices of $a \geq 1$. The censoring DF in each sample was taken to be exponential with scale parameter specified in a way that produces 10%, 25% or 50% censoring. The results are given in Tables 4, 5, and 6, and they indicate that the test $\mathcal{T}_n$ has greater power than the one-sided log-rank test $\mathcal{R}_n$. In the second scenario we specified each DF, F_j, in terms of its hazard function $\lambda_j(t)$, $t \geq 0$, setting $\lambda_1(t) = tI(0 < t < 1) + atI(t \geq 1)$ and $\lambda_2(t) = t$, for various choices of $a \geq 1$. The censoring distribution in each sample is taken to be Weibull with shape parameter 2, and scale parameter specified in a way that produces 10%, 25% or 50% censoring. The results are given in Tables 7, 8, and 9. The results show that the log-rank test $\mathcal{R}_n$ has slightly greater power than the proposed test $\mathcal{T}_n$ when close to the null hypothesis and that $\mathcal{T}_n$ has substantially greater power otherwise. The reason for this could be that, close to the null hypothesis, the hazard rates are almost proportional and it is well known that the log-rank test performs well when the hazard rates are proportional.

5. Example

To illustrate the results discussed in earlier sections, we consider Data Set II from Kalbfleisch and Prentice (1980). The data consists of survival times for

Table 2. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 25% censoring.

	$n = 30$		$n = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.058	0.064	0.055	0.059
1.20	0.168	0.122	0.196	0.146
1.40	0.273	0.194	0.343	0.251
1.60	0.337	0.271	0.464	0.381
1.80	0.474	0.341	0.639	0.497
2.00	0.575	0.434	0.749	0.621
2.20	0.666	0.507	0.864	0.709
2.40	0.764	0.565	0.908	0.782
2.60	0.851	0.618	0.938	0.839
2.80	0.887	0.677	0.970	0.877
3.00	0.933	0.718	0.985	0.911

Table 3. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 50% censoring.

	$n = 30$		$n = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.058	0.056	0.055	0.046
1.20	0.184	0.097	0.209	0.174
1.40	0.272	0.149	0.347	0.335
1.60	0.393	0.224	0.546	0.514
1.80	0.541	0.299	0.696	0.674
2.00	0.670	0.347	0.801	0.797
2.20	0.761	0.413	0.905	0.882
2.40	0.845	0.506	0.935	0.930
2.60	0.890	0.576	0.976	0.966
2.80	0.936	0.590	0.983	0.780
3.00	0.960	0.610	0.996	0.990

patients with carcinoma of the oropharynx and several covariates. These patients had been diagnosed with squamous carcinoma of the oropharynx and were classified by the degree to which the regional lymph nodes were affected by this disease into four populations. Because lymph node deterioration is an indication of the seriousness of the carcinoma, one would expect the corresponding four distributions to be stochastically ordered. Dykstra, Kochar and Robertson (1991) used the LRT they developed for the presence of uniform stochastic ordering in the multinomial case to analyze these data. They grouped it into seven arbitrarily chosen groups (intervals) and obtained a p-value of .04 in their LRT. Wang (1996) used the same grouping , but deleting all censored observations and

Table 4. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 10% censoring.

	$n_1 = n_2 = 30$		$n_1 = n_2 = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.045	0.047	0.046	0.052
1.05	0.117	0.113	0.199	0.196
1.10	0.233	0.228	0.494	0.447
1.15	0.353	0.380	0.758	0.636
1.20	0.522	0.506	0.923	0.834
1.25	0.645	0.623	0.98	0.926
1.30	0.758	0.724	0.996	0.982
0.35	0.811	0.785	1	0.987
1.40	0.851	0.811	1	0.997

Table 5. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 25% censoring.

	$n_1 = n_2 = 30$		$n_1 = n_2 = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.046	0.052	0.055	0.054
1.05	0.113	0.108	0.144	0.134
1.10	0.210	0.190	0.369	0.328
1.15	0.467	0.296	0.613	0.503
1.20	0.549	0.381	0.835	0.708
1.25	0.642	0.479	0.934	0.856
1.30	0.739	0.585	0.974	0.917
1.35	0.822	0.658	0.997	0.982
1.40	0.859	0.722	0.999	0.995

Table 6. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 50% censoring.

	$n_1 = n_2 = 30$		$n_1 = n_2 = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.058	0.054	0.055	0.046
1.05	0.092	0.088	0.123	0.104
1.10	0.139	0.124	0.217	0.189
1.15	0.226	0.169	0.341	0.277
1.20	0.326	0.226	0.515	0.395
1.25	0.378	0.286	0.618	0.573
1.30	0.468	0.338	0.765	0.694
1.35	0.553	0.399	0.849	0.790
1.40	0.632	0.452	0.925	0.895

Table 7. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 10% censoring.

	$n_1 = n_2 = 30$		$n_1 = n_2 = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.044	0.056	0.044	0.048
1.20	0.098	0.097	0.099	0.108
1.40	0.167	0.149	0.207	0.216
1.60	0.236	0.224	0.312	0.335
1.80	0.292	0.299	0.400	0.415
2.00	0.360	0.347	0.526	0.515
2.20	0.461	0.413	0.625	0.608
2.40	0.530	0.506	0.741	0.693
2.60	0.591	0.576	0.768	0.745
2.80	0.652	0.590	0.841	0.780
3.00	0.707	0.610	0.899	0.810

Table 8. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 25% censoring.

	$n_1 = n_2 = 30$		$n_1 = n_2 = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.044	0.058	0.046	0.052
1.20	0.092	0.088	0.108	0.105
1.40	0.137	0.124	0.183	0.180
1.60	0.184	0.176	0.268	0.250
1.80	0.254	0.222	0.346	0.335
2.00	0.317	0.305	0.443	0.420
2.20	0.395	0.336	0.555	0.465
2.40	0.440	0.363	0.627	0.538
2.60	0.491	0.425	0.712	0.594
2.80	0.577	0.445	0.769	0.655
3.00	0.623	0.502	0.819	0.692

deleting the seventh group that had heavy censoring, and applied the LRT he developed for the presence of stochastic ordering alternative in the discrete case. He also pooled the populations 0, 1, and 2, and tested this against population 3 to obtain a p-value of .091. In applying the tests described in Sections 2 and 3, we used the same pooling as Wang (1996). The Kaplan-Meier estimators of the corresponding SFs are given in Figure 1. In practice, one can take b to be any number less than the minimum of the largest uncensored observations from the two-samples. In this case we chose b to be 1,564 days. To test for the presence of stochastic ordering, the p-value based on (3.1) is 0.016, providing evidence that the distribution of population 3 dominates that of the others pooled together.

Table 9. Power comparison of $\mathcal{T}_n$ and $\mathcal{R}_n$ (log-rank), 50% censoring.

	$n_1 = n_2 = 30$		$n_1 = n_2 = 50$	
a	$\mathcal{T}_n$	$\mathcal{R}_n$	$\mathcal{T}_n$	$\mathcal{R}_n$
1.00	0.039	0.061	0.049	0.050
1.20	0.075	0.073	0.078	0.073
1.40	0.101	0.094	0.127	0.109
1.60	0.116	0.101	0.169	0.160
1.80	0.148	0.126	0.266	0.187
2.00	0.193	0.160	0.313	0.209
2.20	0.234	0.175	0.361	0.257
2.40	0.261	0.185	0.437	0.276
2.60	0.305	0.210	0.483	0.306
2.80	0.344	0.251	0.527	0.339
3.00	0.433	0.238	0.612	0.340

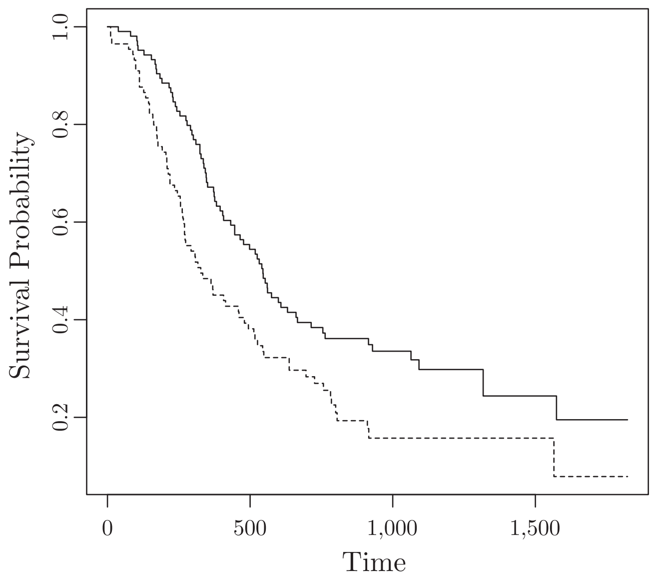

Figure 1. Unrestricted Kaplan Meier estimators of F_1 (solid) and F_2 (dotted).

This conclusion is also confirmed by the test in El Barmi and Mukerjee (2005) (p-value=0.024). In applying $\mathcal{T}_n$ for the presence of uniform stochastic ordering, the p-value is 0.033 based on the approximation in Hall and Wellner (1979). The log rank test in this case yielded a p-value of 0.031 confirming the presence of uniform stochastic ordering.

6. Concluding Remarks

In this paper we have developed an EL approach to the problem of nonparametrically testing for the presence of uniform stochastic ordering in the one- and two-sample cases under right censoring. We obtained a test for the presence of stochastic ordering as a special case. The null asymptotic distributions of the proposed tests are distribution-free. In addition, these tests are computationally efficient to implement and could be used with massive data sets because they do not rely on the bootstrap or any other simulation technique. We also carried out simulations to check finite sample performances and, we applied the tests to an example.

Acknowledgements

Special thanks are due to the Editor, an associate editor, two referees, and Hari Mukrjee for their constructive comments that have resulted in a much improved paper. This work was partially supported by City University of New York through a PSC-CUNY grant.

Appendix

Proof of Theorem 1. We have

$$-2\log(\mathcal{R}(s,t)) \equiv -2\sum_{i=n(s)+1}^{n(t)} \left[(n_i-d_i)\log\left(\frac{\bar{\theta}_i}{\hat{\theta}_i}\right)+d_i\log\left(\frac{1-\bar{\theta}_i}{1-\hat{\theta}_i}\right)\right] I[\beta_n(s,t)\geq 0]$$
$$= -2\sum_{i=n(s)+1}^{n(t)} \left[(n_i-d_i)\log\left(1+\frac{\beta_n^+(s,t)}{n_i-d_i}\right)-n_i\log\left(1+\frac{\beta_n^+(s,t)}{n_i}\right)\right].$$

Let $Y_n(b)$ be the number of observations at risk at b,

$$Y_n(b)=\sum_{i=1}^{n} I[Z_i\geq b].$$

Since $Y_n(b)\leq n_i-d_i$ for all i and $Y_n(b)/n\to\pi(b)>0$, we have

$$0\leq \max_{1\leq i\leq m}\max\left\{\frac{\beta_n^+(s,t)}{n_i-d_i},\frac{\beta_n^+(s,t)}{n_i}\right\}\leq\frac{\beta_n^+(s,t)}{Y_n(b)}=O_p(n^{-1/2})$$

as $n\to\infty$ by (2.7). This implies that $\beta_n^+(s,t)/(n_i-d_i)=O_p(n^{-1/2})$ and $\beta_n^+(s,t)/n_i=O_p(n^{-1/2})$, uniformly in $i=1,2,\ldots,m$, and since $n(t)-n(s)=O(n)$, we have using $\log(1+y)=y-y^2/2+y^3/3+y^4/4+O(y^5)$ as $y\to 0$ and the same arguments as those in Li (1995),

$$
\begin{aligned}
-2\log(\mathcal{R}(s,t)) &= [\beta_n^+(s,t)]^2 \sum_{i=n(s)+1}^{n(t)} \frac{d_i}{n_i(n_i-d_i)} + O_p(n^{-1/2}) \\
&= (\hat{c}(t)-\hat{c}(s))\frac{[\beta_n^+(s,t)]^2}{n} + O_p(n^{-1/2})
\end{aligned}
$$

uniformly on $[0,b]^2$ where $\hat{c}(t)$ is defined in (2.8). Since $\hat{c}(t)$ is a uniformly strongly consistent estimator of $c(t)$, (2.7), the Continuous Mapping Theorem and Slutsky's Theorem imply that the process $\{-2\log(\mathcal{R}(s,t)),\ (s,t)\in[0,b]^2\}$ converges weakly to the process $\{[U^+(s,t)]^2 \equiv [W(c(t))-W(c(s))]^{+2}, (s,t)\in[0,b]^2\}$. Thus,

$$
\begin{aligned}
\mathcal{T}_n \equiv& \sup_{0<s<t<b} \sqrt{\frac{-2(\hat{c}(t)-\hat{c}(s))\log(\mathcal{R}(s,t))}{\hat{c}(b)}} \\
\xrightarrow{d}& \sup_{0\le s\le t\le b} \frac{[W(c(t))-W(c(s))]^+}{\sqrt{c(b)}} \\
=& \sup_{0\le s\le t\le b} \frac{W(c(t))-W(c(s))}{\sqrt{c(b)}} \\
=& \sup_{0\le u\le 1} \sup_{0\le v\le u} [W(u)-W(v)] \\
=& \sup_{0\le u\le 1} [W(u) - \inf_{0\le v\le u} W(v)] \\
\stackrel{d}{=}& \sup_{0\le u\le 1} |W(u)|,
\end{aligned}
$$

where the last equality follows from Levy (1948).

Proof of Theorem 2. Using (2.14), $\log(1+y) = y - y^2/2 + y^3/3 + y^4/4 + O(y^5)$ as $y \to 0$, and the same arguments as those in Li (1995)

$$
\begin{aligned}
&-2\log(\mathcal{R}(s,t)) \\
&\equiv -2\sum_{i=1}^{2}\sum_{j=n(s)+1}^{n(t)}\left[(n_{ij}-d_{ij})\log\left(\frac{\bar{\theta}_{ij}}{\hat{\theta}_{ij}}\right)+d_i\log\left(\frac{1-\bar{\theta}_{ij}}{1-\hat{\theta}_{ij}}\right)\right]I[\beta_n(s,t)\geq 0] \\
&= -2\sum_{i=1}^{2}\sum_{j=n(s)+1}^{n(t)}(n_{ij}-d_i)\log\left(1+(-1)^{i-1}\frac{\beta_n^+(s,t)}{n_{ij}-d_{ij}}\right) \\
&+2\sum_{i=1}^{2}\sum_{j=n(s)+1}^{n(t)}n_{ij}\log\left(1+(-1)^{i-1}\frac{\beta_n^+(s,t)}{n_{ij}}\right) \\
&= \sum_{i=1}^{2}\left[\sum_{j=n(s)+1}^{n(t)}\frac{d_{ij}}{n_{ij}(n_{ij}-d_{ij})}[\beta_n(s,t)^+]^2+O_p(n^{-1/2})\right] \\
&= (\hat{c}(t)-\hat{c}(s))\frac{[\beta_n(s,t)^+]^2}{n}+O_p(n^{-1/2})
\end{aligned}
$$

uniformly on $[0,b]^2$ where $\hat{c}(\cdot)$ is as defined in (2.15). Having established this, it follows using the Continuous Mapping Theorem that

$$
\begin{aligned}
\mathcal{T}_n &\xrightarrow{d} \sup_{0\leq s\leq t\leq b}\frac{(V(t)-V(s))^+}{\sqrt{c(b)}} \\
&\xrightarrow{d} \sup_{0\leq s\leq t\leq b}\frac{[V(t)-V(s)]^+}{\sqrt{c(b)}} \\
&= \sup_{0\leq s\leq t\leq b}\frac{V(t)-V(s)}{\sqrt{c(b)}} \\
&\stackrel{d}{=} \sup_{0\leq u\leq 1}\sup_{0\leq v\leq u\leq 1}[W(u)-W(v)] \\
&= \sup_{0\leq u\leq 1}[W(u)-\inf_{0\leq v\leq u}W(v)] \\
&\stackrel{d}{=} \sup_{0\leq u\leq 1}|W(u)|,
\end{aligned}
$$

where W is a standard Brownian motion.

References

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**, 116-1150.

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

Chang, H. (2014). Empirical likelihood tests for stochastic ordering based on censored and biased data. Ph.D. Thesis, Columbia University.

Dykstra, R. L. (1982). Maximum likelihood estimation of the survival functions of stochastically ordered random variables. *J. Amer. Statist. Assoc.* **77**, 621-628.

Dykstra, R., Kochar, S. and Robertson, T. (1991). Statistical inference for uniform stochastic ordering in several populations. *Ann. Statist.* **19**, 870-888.

El Barmi, H. and McKeague, I. W. (2013). Empirical likelihood based tests for stochastic ordering. *Bernoulli* **19**, 295-307.

El Barmi, H. and McKeague, I. W. (2016). Testing for uniform stochastic ordering via empirical likelihood. To appear in *Ann. Inst. Statist. Math.*

El Barmi, H. and Mukerjee, H. (2005). Inference under a stochastic ordering constraint: the k-sample case. *J. Amer. Statist. Assoc.* **100**, 252-261.

El Barmi, H. and Mukerjee, H. (2015). Consistent estimation of survival functions under uniform stochastic ordering; the k-sample case. *J. Multivariate Anal.* **104**, 99-109.

Einmahl, J. H. J. and McKeague, I. W. (2003). Empirical likelihood based hypothesis testing. *Bernoulli* **9**, 267-290.

Hall, W. J. and Wellner, J. A. (1979). Estimation of a mean residual life. Unpublished manuscript.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. Wiley, New York.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 475-488.

Levy, P. (1948). *Processus Stochastiques et Mouvement Brownien.* Gauthier-Villars, Paris.

Li, G. (1995). On nonparamteric likelihood estimation of survival probabilities for censored data. *Statist. Probab. Lett.* **25**, 95-104.

Mukerjee, H. (1996). Estimation of survival functions under uniform stochastic ordering. *J. Amer. Statist. Assoc.* **91**, 1684-1689.

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237-249.

Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90-120.

Owen, A. B. (2001). *Empirical Likelihood*. Chapman and Hall, London.

Præstgaard and Huang (1996). Asymptotic theory for nonparametric estimation of survival curves under order restrictions. *Ann. Statist.* **24**, 1679-1716.

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.

Shaked, M. and Shanthikumar, G. J. (2006). *Stochastic Orders*. Springer, New York.

Wang, Y. (1996). A likelihood ratio test against stochastic ordering in several populations. *J. Amer. Statist. Assoc.* **79**, 686-691.

Department of Statistics and Computer Information Systems, Baruch College, City University of New York, USA.

E-mail: hammou.elbarmi@baruch.cuny.edu

(Received February 2015; accepted March 2016)