ABSTRACT

LIKELIHOOD RATIO CONFIDENCE BANDS FOR SURVIVAL AND RELATED FUNCTIONS FROM CENSORED TWO-SAMPLE LOCATION-SCALE FAMILIES

by
Kristian Nestor

Two-sample location-scale refers to a situation in which a pair of random variables are linearly related to a base random variable that has mean 0 and variance 1. Using a formulation that leverages the location-scale structure, a plug-in likelihood ratio (LR) method is employed to obtain improved simultaneous confidence bands (SCBs) for the base survival function. The plug-ins are the estimated means and standard deviations of the two samples. Because of the plug-ins, the scaled log LR is now indexed by estimated parameters, making the large-sample study challenging. Using empirical process theory, the indexing by estimated parameters is addressed. The large-sample distribution of the scaled log LR is in an intractable form, which makes computation of the critical values difficult. Two bootstrap schemes are deployed to obtain the thresholds for SCB construction. Two algorithms are proposed for computing the SCBs. The SCBs are shown to have correct empirical coverage and the relative reduction in the average enclosed areas over the nonparametric SCBs is considerable.

The proposed approach advances the computation of LR SCBs for the difference or the ratio of survival functions, and vertical quantile comparison functions. It is expected that the SCBs would be considerably tighter and hence more informative than those obtained through the nonparametric approach. These will be the focus of continuing research and will be reported as soon as they are completed.

LIKELIHOOD RATIO CONFIDENCE BANDS FOR SURVIVAL AND RELATED FUNCTIONS FROM CENSORED TWO-SAMPLE LOCATION-SCALE FAMILIES

by
Kristian Nestor

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology and
Rutgers, The State University of New Jersey—Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Mathematical Sciences

Department of Statistics

May 2026

APPROVAL PAGE

LIKELIHOOD RATIO CONFIDENCE BANDS FOR SURVIVAL AND RELATED FUNCTIONS FROM CENSORED TWO-SAMPLE LOCATION-SCALE FAMILIES

Kristian Nestor

Dr. Sundaraman Subramanian, Dissertation Advisor — Date
Associate Professor, NJIT

Dr. Sunil Dhar, Committee Member — Date
Professor, NJIT

Dr. Antai Wang, Committee Member — Date
Associate Professor, NJIT

Dr. Wenge Guo, Committee Member — Date
Associate Professor, NJIT

Dr. Sudhakar Rao, Committee Member — Date
Director of Biostatistics, Johnson & Johnson

BIOGRAPHICAL SKETCH

Author: Kristian Nestor

Degree: Doctor of Philosophy

Date: May 2026

Undergraduate and Graduate Education:

- Doctor of Philosophy in Mathematical Sciences
  New Jersey Institute of Technology, Newark, NJ, 2026
- Master of Science in Applied Statistics,
  New Jersey Institute of Technology, Newark, NJ, 2022
- Bachelor of Science, Appplied Mathematical and Physical Sciences
  National Technical University of Athens, Athens, Greece, 2020

Major: Mathematical Sciences

All models are wrong, but some are usefull.

George E. P. Box

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This research is concerned with a study of the effectiveness of some survival analysis techniques that will be introduced and will be tailored to bridge the gap in estimation when there are additional imposed structures. It is well understood that the methods that operate under fully nonparametric specifications would require finer adaptations to take advantage of additional structures, leading to more efficient methods. It is the aim of this research to bring into sharper view the existence of improved likelihood ratio (LR) tailored simultaneous confidence bands (SCBs) for survival functions, difference of survival functions, ratio of survival functions, vertical quantile comparison (VQC) functions, and other important measures, under the framework of two-sample location-scale (LS).

Two-sample LS models have the property that two random variables, when standardized, have the same distribution, which we shall call the base distribution. Adapting the LR approach by incorporating the LS structure into the estimations would lead to improved procedures for censored data. In this research, the overarching goal is to obtain improved SCBs for a wide variety of parameters built on survival or quantile functions, their composition, difference or ratio.

Let S be a survival function. Thomas and Grunkemeier (1975) devised an LR method of constructing pointwise confidence intervals (PCIs) for $S(t)$ from censored data. It is a test-based approach that inverts the LR to obtain PCIs and works as follows. The nonparametric likelihood, expressible as a function of the unknown $S(t)$, is parameterized via discrete hazards. The maximum likelihood estimates (MLEs) of the discrete hazards, when plugged into a product integral representation, yields the nonparametric maximum likelihood estimator (NPMLE), the time-honored Kaplan–Meier (KM) estimator of the survival function, see, for example, Kalbfleisch and Prentice (2002). The KM estimator is plugged into the nonparametric likelihood, yielding a baseline measure that constitutes the denominator of the LR. For the numerator of the LR, one substitutes the maximizer of the likelihood subject to a constraint that the survival function at a specified time point (say t_0) would be a specified number (say p, where $0 < p < 1$). That is, $S(t_0) = p$ is the constraint, which is the null hypothesis. Any number of discrete survival functions supported on the uncensored times may be out there. A subset of those would satisfy the null hypothesis. The Lagrange Multiplier (LM) method permits obtaining the survival function among this subset that maximizes the likelihood. The MLEs of the discrete hazards in this case depend on the unknown LM and, hence, the maximizing survival function and, in turn, the numerator of the LR will also be a function of the LM. The null hypothesis would be accepted for high values of the LR (close to 1). To produce a string of accepted numbers p, however, an uncountable number of null hypotheses will have to be tested.

There is a way out of this difficulty. The log of the LR scaled by the factor -2 is, in its final compact form, a function of the data and the LM. It is 0 when the LM is

0 and is decreasing for negative values and increasing for positive values. The insight that it can be approximated by a chi-squared distribution with 1 degree of freedom is an important one. There are two LM values for which the scaled log LR equals the upper α quantile of a chi-squared distribution with 1 degree of freedom. These are the lower and upper limits for the range of LM values for which the null hypothesis would not be rejected. It follows that, plugging the two values of the LM into the product integral form for the survival function, one obtains two limits within which bracket the true survival function at t_0 will be included with $100(1-\alpha)\%$ confidence. The Brent's algorithm can be used to obtain the aforementioned two LM values.

Li (1996) derived the asymptotic distribution of the scaled log LR. It is, as Thomas and Grunkemeier (1975) concluded, chi-squared with 1 degree of freedom. Li (1996) gave a rigorous justification for heuristic arguments in the literature concerning Taylor expansion of the log LR. He derived large-sample rates for the LM estimate, which were crucial for the Taylor expansion of the scaled log LR leading to its asymptotic distribution. Li's (1996) formal approach identified a comprehensive road map for deriving SCBs for survival functions and functionals thereof, as well as SCBs for a host of related measures such as a difference or a ratio of two survival functions.

Hollander, McKeague, and Yang (1997), henceforth HMY, made a fundamental contribution that guides the construction of SCBs. They extended the approach of Thomas and Grunkemeier (1975) to initiate the provision of SCBs for survival functions. The survival function at t replaced the number p and the scaled log LR was framed as a function of S and its evaluating point t. The critical value for computing the two LM values were obtained from the supremum of a Brownian bridge process. Computation of their SCBs, over regions excluding the lower end point of zero would, however, require reliance on special-purpose tables, see Subramanian (2016). The bootstrap, therefore, would offer a feasible alternative.

HMY's technique was further applied to derive SCBs for quantile functions (Li, Hollander, McKeague, and Yang, 1996). McKeague and Zhao (2002) proposed SCBs for the ratio of survival functions. McKeague and Zhao (2005) proposed further extensions and derived LR SCBs for differences and ratios of linear functionals of cumulative hazard functions. McKeague and Zhao (2006) applied the LR method and proposed width-scaled SCBs.They were derived by scaling the LR PCIs, using a data-driven inflation factor, around the NPMLE of the survival function, that is, the KM estimator. The approach improved the stability of the SCBs, especially in the tails.

For LR SCBs of $D_0(t) = S_1(t) - S_2(t)$, the difference of two survival functions, a strikingly novel contribution merits deeper consideration. Treating $S_2(t)$ as a nuisance parameter η, Shen and He (2005) framed the maximization in terms $S_1(t) = \eta + D_0(t)$ and $S_2(t) = \eta$. They derived the log LR which is a function of two Lagrange multipliers and proved the existence of an η that maximized the LR. The algorithm for computing the SCBs for the true difference $D(t)$ was not given, however.

Semiparametric random censorship models (SRCMs), proposed by Dikta (1998) permit improved LR SCBs around a semiparametric survival function estimator. When the model for the conditional non-censoring probability given the event time is specified correctly, the semiparametric survival function estimator is asymptotically more efficient than the KM estimator (Dikta, 1998). This fact has been exploited to derive SRCM LR SCBs for survival functions (Subramanian, 2016), and the difference of survival functions (Ahmed and Subramanian, 2016).

Chen, Tracy, and Uno (2021) constructed SCBs for survival functions and difference of hazard functions that were derived from a related optimization problem with local time processes. Their "OptBand" formulated the problem as an optimization task to find SCBs that minimize the enclosed area between them for a given coverage level.

Hall and Wellner (1984) and Nair (1984) derived two classical SCBs using the weak convergence of a pivotal quantity to a Brownian motion or Brownian bridge, which they then inverted to obtain SCBs for survival functions. Parzen, Wei, and Ying (1994) introduced a perturbation resampling method to compute critical values; Parzen, Wei, and Ying (1997) applied it to the two-sample problem.

We propose two sample LS-based SCBs for survival functions, their difference, ratio, as well as SCBs for vertical quantile comparison functions. We show that nonparametric SCBs can be effectively adapted to the LS framework. The proposed methods exhibit superior performance in terms of some evaluation metrics such as estimated average enclosed areas and estimated average widths. For each task we provide a comprehensive algorithm that describes how threshold values for inverting the LR are obtained and how the SCBs are computed.

CHAPTER 2

SIMULTANEOUS CONFIDENCE BANDS FOR SURVIVAL FUNCTIONS

2.1 The Location–Scale (LS) Framework

Suppose that

$$X_i = \mu_i + \sigma_i Z, \quad i = 1, 2,$$

where Z is a random variable with distribution function F, mean 0, and variance 1. We refer to Z as having the *baseline distribution* F (for example, Z may be standard normal). The parameters μ_i and σ_i denote the mean and standard deviation of X_i, respectively. Hence,

$$F_i(x) = F\left(\frac{x - \mu_i}{\sigma_i}\right), \quad i = 1, 2. \tag{2.1}$$

The distributions F_1 and F_2 are said to belong to an LS family with baseline distribution function F. LS models are widely used in medical research, where outcome distributions are critical for evaluating treatment efficacy and for supporting clinical decision-making. Differences in disease progression, for instance, often appear as shifts and/or scales in health-related outcome distributions such as blood pressure, time to an event, or biomarker levels.

LS models are particularly useful in survival analysis, as they naturally accommodate heterogeneity between patient groups. In the context of accelerated failure time (AFT) models, the log-survival times are assumed to follow an LS family (e.g., log-normal or log-logistic), allowing clinicians to interpret treatment effects as multiplicative changes in median survival [17].

2.2 Censoring and Estimation under the LS Model

Let X_{ij} ($j = 1, 2, \ldots, n_i$; $i = 1, 2$) denote the failure time of the j-th individual in the i-th sample, and let C_{ij} be the corresponding censoring time. We observe the pairs $(\tilde{Z}_{ij}, \tilde{\delta}_{ij})$, where

$$\tilde{Z}_{ij} = \min(X_{ij}, C_{ij}), \qquad \tilde{\delta}_{ij} = \mathbb{1}(X_{ij} \leq C_{ij}), \quad j = 1, \ldots, n_i;\ i = 1, 2.$$

Let $\hat{S}_i$ be the KM estimator of S_i, the survival function of X_i. Let $\Delta\hat{S}_i(t)$ be the jump of $\hat{S}_i$ at t. The estimated means and variances are

$$\hat{\mu}_i \;=\; -\sum_{j=1}^{n_i} \tilde{\delta}_{ij} \tilde{Z}_{ij}\, \Delta\hat{S}_i\left(\tilde{Z}_{ij}\right); \qquad \hat{\psi}_i \;=\; -\sum_{j=1}^{n_i} \tilde{\delta}_{ij} \tilde{Z}_{ij}^2\, \Delta\hat{S}_i\left(\tilde{Z}_{ij}\right); \qquad \hat{\sigma}_i^2 \;=\; \hat{\psi}_i - \hat{\mu}_i^2.$$

Let the notional standardized variable U_{ij}^0 and its concomitant censoring V_{ij}^0 be

$$U_{ij}^0 = \frac{X_{ij} - \mu_i}{\sigma_i} \sim F, \qquad V_{ij}^0 = \frac{C_{ij} - \mu_i}{\sigma_i}.$$

The proposed research leverages the LS structure to pool the two samples from which the KM estimator of $S(t) = 1 - F(t)$ is computed. That pooling does not create a conflict in producing the correct estimator is borne from the following argument, when the means and standard deviations are known.

Define the notional variables $Z^0_{ij} = \min(U^0_{ij}, V^0_{ij})$ and $\delta^0_{ij} = I(U^0_{ij} \le V^0_{ij})$. Let $\bar{G}_i(t)$ be the survival function of the concomitant censoring variable $V^0_i = (C_i - \mu_i)/\sigma_i$. Let W^0 be the pooled minimum and let ξ be the group membership indicator with $\mathbb{P}(\xi = 0) = \alpha$. We continue to let δ^0 be the (concomitant) censoring indicator. Then

$$
\begin{aligned}
P(W^0 > t) &= P(W^0 > t|\xi = 0)\alpha + P(W^0 > t|\xi = 1)(1 - \alpha) \\
&= P(Z^0_1 > t|\xi = 0)\alpha + P(Z^0_2 > t|\xi = 1)(1 - \alpha) \\
&= S(t)\left\{\alpha\bar{G}_1(t) + (1 - \alpha)\bar{G}_2(t)\right\}. \qquad (2.2)
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
\mathbb{P}(W^0 \le t, \delta^0 = 1) &= \mathbb{P}(W^0 \le t, \delta^0 = 1|\xi = 0)\alpha + \mathbb{P}(W^0 \le t, \delta^0 = 1|\xi = 1)(1 - \alpha) \\
&= \alpha\int_{-\infty}^{t} \bar{G}_1(s)\,dF(s) + (1 - \alpha)\int_{-\infty}^{t} \bar{G}_2(s)\,dF(s). \qquad (2.3)
\end{aligned}
$$

We then have that

$$
\int_{-\infty}^{t} \frac{d\mathbb{P}(W^0 \le s, \delta^0 = 1)}{\mathbb{P}(W^0 > s)} = \int_{-\infty}^{t} \frac{(\alpha\bar{G}_1(s) + (1 - \alpha)\bar{G}_2(s))\,dF(s)}{S(s)(\alpha\bar{G}_1(s) + (1 - \alpha)\bar{G}_2(s))} = \Lambda(t),
$$

where $\Lambda(t)$ is the cumulative hazard function associated with F. These calculations indicate that $\Lambda(t)$ and $S(t)$ can be estimated by the (notional) Nelson–Aalen estimator and the KM estimator respectively from the pooled standardized notional data $\{W^0_j, j = 1, \ldots, n\}$, where $n = n_1 + n_2$.

To enable transition from notional to the actual, let U_{ij} and V_{ij} be the estimated counterparts of U^0_{ij} and V^0_{ij} respectively, defined by

$$
U_{ij} = \frac{X_{ij} - \hat{\mu}_i}{\hat{\sigma}_i}, \qquad V_{ij} = \frac{C_{ij} - \hat{\mu}_i}{\hat{\sigma}_i},
$$

Let $W_j, j = 1, \ldots, n$ be the pooled minimum of U_{ij} and V_{ij}, and let δ be the concomitant censoring indicator. On the basis of $W_1, \ldots, W_n$, form the counting process $N_{\hat{\theta}}(t) = \sum_{j=1}^{n} I(W_j \le t, \delta_j = 1)$ and the "at-risk" process $Y_{\hat{\theta}}(t) = \sum_{j=1}^{n} I(W_j \ge t)$. Note that $\hat{\theta}$ is embedded in the W_j's. Write $\Delta N_{\hat{\theta}}(t) = N_{\hat{\theta}}(t) - N_{\hat{\theta}}(t-)$. The KM estimator from the pooled data is

$$
\hat{S}_{\hat{\theta}}(t) = \prod_{s \le t}\left(1 - \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)}\right). \qquad (2.4)
$$

Note that $\hat{S}_\theta(s) \equiv \hat{S}(s)$, the KM estimator based on the notional data $W^0_j, j = 1, \ldots, n$. From Major and Rejtö (1988) and the functional delta method (Theorem II.8.1. of Andersen et al. (1993), it follows that

$$
n^{1/2}(\hat{S}(t) - S(t)) = -S(t)\, n^{-1/2}\sum_{l=1}^{n} I^{(1)}_l(t) + o_{\mathbb{P}}(1), \qquad (2.5)
$$

where

$$I_l^{(1)}(t) \quad = \quad \frac{I(W_l^0 \leq t, \delta_l^0 = 1)}{\mathbb{P}(W^0 > W_l^0)} - \int_0^t \frac{I(W_l^0 > y)d\Lambda(y)}{\mathbb{P}(W^0 > y)}. \tag{2.6}$$

Let f be the density of F. Nestor and Subramanian (2025) derived a large-sample representation for $\hat{S}_{\hat{\theta}}(t) - S(t)$. We review their results here. For cadlag functions $\alpha(s)$ and $\beta(s)$, let

$$\mathcal{D}_t(\alpha, \beta) \quad = \quad \int_{-\infty}^t d\,\alpha(s)/\beta(s).$$

Let $C_{G_i}(t) = \mathcal{D}_t(\Lambda_{G_i}, 1 - H_i)$, where Λ_{G_i} is the cumulative hazard function associated with the censoring distribution function G_i. Then

$$n^{1/2}(\hat{\mu}_i - \mu_i) \quad = \quad \kappa_i^{-1/2} \left(n_i^{-1/2} \sum_{j=1}^{n_i} I_{ij}^{(2)} \right) + o_{\mathbb{P}}(1), \tag{2.7}$$

$$n^{1/2}(\hat{\sigma}_i - \sigma_i) \quad = \quad \frac{\kappa_i^{-1/2}}{2\sigma_i} \left\{ n_i^{-1/2} \sum_{j=1}^{n_i} \left(I_{ij}^{(3)} - 2\mu_i\, I_{ij}^{(2)} \right) \right\} + o_{\mathbb{P}}(1), \tag{2.8}$$

where

$$\begin{aligned} I_{ij}^{(2)} \quad &= \quad \left\{ \tilde{Z}_{ij}\beta_{i0}(\tilde{Z}_{ij})\tilde{\delta}_{ij} - \mu_i \right\} + \frac{1 - \tilde{\delta}_{ij}}{1 - H_i(\tilde{Z}_{ij})} \int_{\tilde{Z}_{ij}}^{\infty} w\, dF_i(w) \\ &- \quad \int_{-\infty}^{\infty} C_{G_i}(\tilde{Z}_{ij} \wedge w) w\, dF_i(w), \end{aligned} \tag{2.9}$$

$$\begin{aligned} I_{ij}^{(3)} \quad &= \quad \left\{ \tilde{Z}_{ij}^2\beta_{i0}(\tilde{Z}_{ij})\tilde{\delta}_{ij} - \psi_i \right\} + \frac{1 - \tilde{\delta}_{ij}}{1 - H_i(\tilde{Z}_{ij})} \int_{\tilde{Z}_{ij}}^{\infty} w^2\, dF_i(w) \\ &- \quad \int_{-\infty}^{\infty} C_{G_i}(\tilde{Z}_{ij} \wedge w) w^2\, dF_i(w). \end{aligned} \tag{2.10}$$

Note that $\mathbb{E}(I_{ij}^{(2)}) = 0$ and $\mathbb{E}(I_{ij}^{(3)}) = 0$ for $j = 1, \ldots, n_i$ and $i = 1, 2$.

From their Proposition 1, $\hat{A}(t) := \hat{S}_{\hat{\boldsymbol{\theta}}}(t) - S(t)$ admits the large sample representation

$$\begin{aligned} \hat{A}(t) \quad &= \quad \hat{S}(t) - S(t) + \frac{\alpha}{\sigma_1} f(t) \left\{ (\hat{\mu}_1 - \mu_1) + t(\hat{\sigma}_1 - \sigma_1) \right\} \\ &\quad + \frac{1-\alpha}{\sigma_2} f(t) \left\{ (\hat{\mu}_2 - \mu_2) + t(\hat{\sigma}_2 - \sigma_2) \right\} + o_{\mathbb{P}}(n^{-1/2}). \end{aligned} \tag{2.11}$$

The resulting large-sample representation for $\hat{A}(t)$, obtained by plugging representations given by Eq. (2.5), Eq. (2.7) and Eq. (2.8) into Eq. (2.11) yields a weak Gaussian limit with complicated covariance structure.

2.3 SCBs for individual survival functions under two-sample LS

As will be shown below, the estimated standardized pooled data permit the construction of SCBs for $S(t)$. In turn, SCBs for $S_i(t)$ can be obtained from evaluations at time points shifted and scaled from the support points of $\hat{S}_{\hat{\theta}}(t)$. We describe the details now.

Let $D(t) = \max_{s\le t}(\Delta N_{\hat{\theta}}(s) - Y_{\hat{\theta}}(s))$. Let Γ be the space of all survival functions on $(-\infty,\infty)$, with first and second moments 0 and 1 respectively. For $0<p<1$, the nonparametric LR, when θ is known, is (Thomas and Grunkemeier, 1975)

$$R(p,t) = \frac{\sup\{L(\gamma) : \gamma(t) = p, \gamma\in\Gamma\}}{\sup\{L(K) : K\in\Gamma\}}. \tag{2.12}$$

Replacing p by $K(t)$, where $K\in\Gamma$, the formula for the plug-in scaled log LR is exactly as in HMY, but all the functions will now be indexded by $\hat{\theta}$, the estimated θ:

$$\begin{aligned}\mathcal{L}(K(t),t) &:= -2\log R(K(t),t)\\ &= -2\sum_{s\le t}\Bigg[(Y_{\hat{\theta}}(s)-\Delta N_{\hat{\theta}}(s))\log\left(1+\frac{\hat{\lambda}_{\hat{\theta}}(t)}{Y_{\hat{\theta}}(s)-\Delta N_{\hat{\theta}}(s)}\right)\\ &\qquad - Y_{\hat{\theta}}(s)\log\left(1+\frac{\hat{\lambda}_{\hat{\theta}}(t)}{Y_{\hat{\theta}}(s)}\right)\Bigg],\end{aligned} \tag{2.13}$$

where $\hat{\lambda}_{\hat{\theta}}(t)$ satisfies

$$\prod_{s\le t}\left(1-\frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)+\hat{\lambda}_{\hat{\theta}}(t)}\right) = K(t). \tag{2.14}$$

Eqs. (2.13) and (2.14) differ from their counterparts in HMY, because all quantities are now indexed by estimated means and standard deviations. If those were replaced by their true values one obtains the HMY counterparts. In particular, $\hat{\lambda}_{\hat{\theta}}(t)$ is a perturbed version of $\hat{\lambda}$ in HMY. Introduce the factor

$$\hat{\sigma}^2_{\hat{\theta}}(t) \;=\; n\sum_{s\le t}\frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)(Y_{\hat{\theta}}(s)-\Delta N_{\hat{\theta}}(s))}. \tag{2.15}$$

The following theorem gives a large-sample representation for $\mathcal{L}(S(t),t)$.

Theorem 1. *With $\hat{A}(t)$ defined by Eq. (2.11), the scaled log LR admits the large-sample representation*

$$\mathcal{L}(S(t),t) = \frac{1}{\hat{\sigma}^2_{\hat{\theta}}(t)S^2(t)}\left\{n^{1/2}\hat{A}(t)\right\}^2 + o_{\mathbb{P}}(1), \tag{2.16}$$

uniformly for $t\in[t_1,t_2]\subset(0,\tau_H]$.

The proof of Theorem 1 is given in the Appendix. Our perturbed version of the HMY scaled log LR requires empirical process theory to "smooth out" the perturbation; see Subsection 2.2 where it is given in detail. What emerges from the analysis is a weak limit for $n^{1/2}(\hat{S}_{\hat{\theta}}(t) - S(t)) \equiv n^{1/2}\hat{A}(t)$ that is the sum of the HMY Brownian bridge process and additional terms due to the plug-ins for μ_i and σ_i that caused the perturbation. The asymptotic version of the supremum over t of the scaled $\mathcal{L}(S(t), t)$, therefore, is the supremum over t of an intractable limiting process, which makes it difficult to obtain thresholds needed for computing the SCBs. We employ the bootstrap to obtain the critical values. Two algorithms, Algorithm 1 and Algorithm 2, are provided below. These differ in the way the resampling is done.

Algorithm 1: Bootstrap LR SCBs for the base survival function I

Data: Sample X_{ij}, C_{ij}, failure and censoring times, respectively $i = 1, 2$, $j = 1, 2, \cdots, n_i$

Result: Simultaneous confidence bands $\mathrm{SCB}_i(t)$, $i = 1, 2$

Step 1: Estimate location and scale parameters;
for $i \in \{1, 2\}$ **do**
  Obtain $\tilde{Z}_{ij} = \min(X_{ij}, C_{ij})$, $\tilde{\delta}_{ij} = \mathbb{1}(X_{ij} \le C_{ij})$;
  Compute the KM estimator $\hat{S}_i(t)$ from $(\tilde{Z}_{ij}, \tilde{\delta}_{ij})$;
  Compute the jumps $\Delta\hat{S}_i(\tilde{Z}_{ij})$;
  Estimate $\hat{\mu}_i = -\sum_j^{n_i} \tilde{\delta}_{ij}\tilde{Z}_{ij}\Delta\hat{S}_i(\tilde{Z}_{ij})$, $\hat{\psi}_i = -\sum_j^{n_i} \tilde{\delta}_{ij}\tilde{Z}_{ij}^2\Delta\hat{S}_i(\tilde{Z}_{ij})$, $\hat{\sigma}_i^2 = \hat{\psi}_i - \hat{\mu}_i^2$;

Step 2: Standardize and pool samples;
For each sample $i = 1, 2$, set $U_{ij} = (X_{ij} - \hat{\mu}_i)/\hat{\sigma}_i$ and $V_{ij} = (C_{ij} - \hat{\mu}_i)/\hat{\sigma}_i$;
Obtain pooled failure and censor time each of size $n = n_1 + n_2$
$\boldsymbol{U} = (U_{11}, U_{12}, \cdots, U_{1n_1}, U_{21}, U_{22}, \cdots, U_{2n_2})^\top$,
$\boldsymbol{V} = (V_{11}, V_{12}, \cdots, V_{1n_1}, V_{21}, V_{22}, \cdots, V_{2n_2})^\top$;
Form pooled data: $W_j = \min(\boldsymbol{U_j}, \boldsymbol{V_j})$, $\delta_j = \mathbb{1}(\boldsymbol{U_j} \le \boldsymbol{V_j})$), $j = 1, 2, \cdots, n$;
Compute the pooled KM estimator $\hat{S}_{\hat{\theta}}(t)$ from (W_j, δ_j), where $\hat{\theta} = (\hat{\mu}_1, \hat{\sigma}_1, \hat{\mu}_2, \hat{\sigma}_2)$;
Store ordered event (failure) times $T_{(j)}$, $\hat{S}_{\hat{\theta}}\left(T_{(j)}\right)$, $Y_{\hat{\theta}}\left(T_{(j)}\right)$, $\Delta N_{\hat{\theta}}\left(T_{(j)}\right)$, $j = 1, 2, \cdots, k$;
Compute $w_n(t) := \frac{\hat{\sigma}^2_{\hat{\theta}}(t)}{1+\hat{\sigma}^2_{\hat{\theta}}(t)}$ where $\hat{\sigma}^2_{\hat{\theta}}(t)$ is defined by Eq. (2.15) ;

Step 3: Bootstrap critical value computation;
for $b = 1, \ldots, B$ **do**
  Sample indices with replacement from pooled data to form bootstrap sample $(Z^{*(b)}, \delta^{*(b)})$;
  Compute $\hat{S}^{*(b)}_{\hat{\theta}}(t)$, $Y^{*(b)}_{\hat{\theta}}(t)$, and $\Delta N^{*(b)}_{\hat{\theta}}(t)$;
  Solve for $\hat{\lambda}^{*(b)}(t)$ from $\prod_{s \le t}\left(1 - \frac{\Delta N^{*(b)}_{\hat{\theta}}(s)}{Y^{*(b)}_{\hat{\theta}}(s)+\lambda}\right) = \hat{S}_{\hat{\theta}}(t)$;
  Evaluate test statistic $\mathcal{L}^{*(b)}(t) = -2w_n(t)\sum_{s\le t}\left[(Y^{*(b)}_{\hat{\theta}}(s) - \Delta N^{*(b)}_{\hat{\theta}}(s))\log\left(1 + \frac{\lambda}{Y^{*(b)}_{\hat{\theta}}(s)-\Delta N^{*(b)}_{\hat{\theta}}(s)}\right) - Y^{*(b)}_{\hat{\theta}}(s)\log\left(1 + \frac{\lambda}{Y^{*(b)}_{\hat{\theta}}(s)}\right)\right]$;
  Store $\max_t \mathcal{L}^{*(b)}(t)$;
Compute the 95^{th} percentile of $\{\max_t \mathcal{L}^{*(b)}(t)\}_{b=1}^B$ as $q_{\mathrm{LR}_{\hat{\theta}}}$;

Step 4: Construct simultaneous confidence bands;
for *each* $t \in [T_\alpha, T_{1-\alpha}]$, *where* T_α *is the* $\alpha - th$ *percentiles of* $\left\{T_{(j)}\right\}_{j=1}^n$ **do**
  Solve $\mathcal{L}(S(t), t) = q_{\mathrm{LR}_{\hat{\theta}}}/w_n(t)$ for $\lambda_L(t)$ and $\lambda_U(t)$ using Brent's method;
  Compute $\hat{S}_L(t) = \prod_{s\le t}\left(1 - \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)+\lambda_L(t)}\right)$, $\hat{S}_U(t) = \prod_{s\le t}\left(1 - \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)+\lambda_U(t)}\right)$;
return $\mathrm{SCB}_i(t) = \left[\hat{S}_L\left(\frac{t-\hat{\mu}_i}{\hat{\sigma}_i}\right), \hat{S}_U\left(\frac{t-\hat{\mu}_i}{\hat{\sigma}_i}\right)\right]$;

Algorithm 2: Bootstrap LR SCBs for the base survival function II

Data: Sample X_{ij}, C_{ij}, failure and censor times, respectively $i = 1, 2$, $j = 1, 2, \cdots, n_i$

Result: Simultaneous confidence bands $\text{SCB}_i(t)$, $i = 1, 2$

Step 1: Estimate location and scale parameters;
for $i \in \{1, 2\}$ **do**
  Obtain $\tilde{Z}_{ij} = \min(X_{ij}, C_{ij})$, $\tilde{\delta}_{ij} = \mathbb{1}(X_{ij} \le C_{ij})$;
  Compute Kaplan–Meier estimator $\hat{S}_i(t)$ from $(\tilde{Z}_{ij}, \tilde{\delta}_{ij})$;
  Compute jumps $\Delta\hat{S}_i(t_j)$ at uncensored times;
  Estimate $\hat{\mu}_i = -\sum_j^{n_i} \tilde{\delta}_{ij}\tilde{Z}_{ij}\Delta\hat{S}_i(\tilde{Z}_{ij})$, $\hat{\psi}_i = -\sum_j^{n_i} \tilde{\delta}_{ij}\tilde{Z}_{ij}^2\Delta\hat{S}_i(\tilde{Z}_{ij})$, $\hat{\sigma}_i^2 = \hat{\psi}_i - \hat{\mu}_i^2$;

Step 2: Standardize and pool samples;
For each sample $i = 1, 2$, set $U_{ij} = (X_{ij} - \hat{\mu}_i)/\hat{\sigma}_i$ and $V_{ij} = (C_{ij} - \hat{\mu}_i)/\hat{\sigma}_i$;
Obtain pooled failure and censor time each of size $n = n_1 + n_2$
$\boldsymbol{U} = (U_{11}, U_{12}, \cdots, U_{1n_1}, U_{21}, U_{22}, \cdots, U_{2n_2})^\top$,
$\boldsymbol{V} = (V_{11}, V_{12}, \cdots, V_{1n_1}, V_{21}, V_{22}, \cdots, V_{2n_2})^\top$;
Form pooled data: $W_j = \min(\boldsymbol{U_j}, \boldsymbol{V_j})$, $\delta_j = \mathbb{1}(\boldsymbol{U_j} \le \boldsymbol{V_j})$), $j = 1, 2, \cdots, n$;
Compute pooled KM estimator $\hat{S}_{\hat{\theta}}(t)$ from (W_j, δ_j);
Store ordered failure times $T_{(j)}$, $\hat{S}_{\hat{\theta}}\left(T_{(j)}\right)$, $Y_{\hat{\theta}}\left(T_{(j)}\right)$, $\Delta N_{\hat{\theta}}\left(T_{(j)}\right)$, $j = 1, 2, \cdots, k$;
Compute $w_n(t) := \frac{\hat{\sigma}^2_{\hat{\theta}}(t)}{1+\hat{\sigma}^2_{\hat{\theta}}(t)}$ where $\hat{\sigma}^2_{\hat{\theta}}(t)$ is defined on (2.15) ;

Step 3: Bootstrap critical value computation;
for $b = 1, \ldots, B$ **do**
  for $i \in \{1, 2\}$ **do**
    Sample indiced from original data to obtain bootstrap samples $\left(X_{ij}^*, C_{ij}^*\right)$;
    Obtain $\tilde{Z}_{ij}^* = \min\left(X_{ij}^*, C_{ij}^*\right)$, $\tilde{\delta}_{ij}^* = \mathbb{1}\left(X_{ij}^* \le C_{ij}^*\right)$;
    Compute Kaplan–Meier estimator $\hat{S}_i^*(t)$ from $(\tilde{Z}_{ij}^*, \tilde{\delta}_{ij}^*)$;
    Compute jumps $\Delta\hat{S}_i^*(t_j)$ at uncensored times;
    Estimate $\hat{\mu}_i^* = -\sum_j^{n_i} \tilde{\delta}_{ij}^*\tilde{Z}_{ij}^*\Delta\hat{S}_i^*(\tilde{Z}_{ij})$, $\hat{\psi}_i^* = -\sum_j^{n_i} \tilde{\delta}_{ij}^*\tilde{Z}_{ij}^{2*}\Delta\hat{S}_i^*(\tilde{Z}_{ij}^*)$, $\hat{\sigma}_i^{2*} = \hat{\psi}_i^* - \hat{\mu}_i^{*2}$;
  Compute $\hat{S}_{\hat{\theta}}^{*(b)}(t)$, $Y_{\hat{\theta}}^{*(b)}(t)$, and $\Delta N_{\hat{\theta}}^{*(b)}(t)$;
  Solve for $\hat{\lambda}^{*(b)}(t)$ from $\prod_{s\le t}\left(1 - \frac{\Delta N_{\hat{\theta}}^{*(b)}(s)}{Y_{\hat{\theta}}^{*(b)}(s)+\lambda}\right) = \hat{S}_{\hat{\theta}}(t)$;
  Evaluate test statistic $\mathcal{L}^{*(b)}(t) = -2w_n(t)\sum_{s\le t}\left[(Y_{\hat{\theta}}^{*(b)}(s) - \Delta N_{\hat{\theta}}^{*(b)}(s))\log\left(1 + \frac{\lambda}{Y_{\hat{\theta}}^{*(b)}(s)-\Delta N_{\hat{\theta}}^{*(b)}(s)}\right) - Y_{\hat{\theta}}^{*(b)}(s)\log\left(1 + \frac{\lambda}{Y_{\hat{\theta}}^{*(b)}(s)}\right)\right]$;
  Store $\max_t \mathcal{L}^{*(b)}(t)$;
Compute the 95^{th} percentile of $\{\max_t \mathcal{L}^{*(b)}(t)\}_{b=1}^B$ as $q_{\text{LR}_{\hat{\theta}}}$;

Step 4: Construct simultaneous confidence bands;
for *each* $t \in [T_\alpha, T_{1-\alpha}]$, *where* T_α *is the* $\alpha - th$ *percentiles of* $\left\{T_{(j)}\right\}_{j=1}^n$ **do**
  Solve $\mathcal{L}(S(t), t) = q_{\text{LR}_{\hat{\theta}}}/w_n(t)$ for $\lambda_L(t)$ and $\lambda_U(t)$ using Brent's method;
  Compute $\hat{S}_L(t) = \prod_{s\le t}\left(1 - \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)+\lambda_L(t)}\right)$, $\hat{S}_U(t) = \prod_{s\le t}\left(1 - \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)+\lambda_U(t)}\right)$;
return $\text{SCB}_i(t) = \left[\hat{S}_L\left(\frac{t-\hat{\mu}_i}{\hat{\sigma}_i}\right), \hat{S}_U\left(\frac{t-\hat{\mu}_i}{\hat{\sigma}_i}\right)\right]$;

Remark As mentioned in the Introduction, for each uncensored time point t in the pooled ordered sample, $\mathcal{L}(S(t), t)$,the scaled log LR, is set equal to the critical value obtained by the bootstrap scaled by the reciprocal of a weight function $w_n(t)$. From Eq. (2.13), $\mathcal{L}(S(t), t)$, considered as a function of λ, is 0 at 0, decreasing for $\lambda < 0$ and increasing for $\lambda > 0$. It follows that there are two values of λ where $\mathcal{L}(S(t), t)$ equals the bootstrap critical value. These can be obtained by the Brent's method of root finding. The two values of λ are substituted into the LHS of Eq. (2.14) to yield a bracket for $S(t)$.

CHAPTER 3

SIMULATIONS

In this section, we report the results of various simulation studies we conducted to evaluate the finite-sample performance of the LS-based SCBs. The goal is to examine how well the theoretical properties derived in the previous section hold in practice and to compare the performance of our method with the HMY SCBs. We considered three scenarios.

Scenario 1 (Normal family) The failure times were generated from the normal family of distributions with means μ_i and standard deviations σ_i, $i = 1, 2$. The censoring times were also normally distributed, with means $\mu_{c_1} = \mu_1 + 3$ and $\mu_{c_2} = \mu_2 + 1$, and standard deviations $\sigma_{c_1} = \sigma_1 + 3$ and $\sigma_{c_2} = \sigma_2 + 1$. These choices yield censoring rates of approximately 23% and 42% for the two samples respectively.

Scenario 2 (Exponential family) The base distribution was exponential with rate $\lambda = 3$. These were scaled by σ_i and shifted by μ_i to obtain the failure times for the two samples. Likewise scaling the base exponential by σ_i and shifting thereafter by $\mu_1 + 3$ and μ_2 respectively, the censoring times were obtained. These settings produced censoring rates of approximately 7% for sample 1 and 50% for sample 2.

Scenario 3 (Extreme value distribution (EVD)) The failure times follow shifted and scaled extreme value distributions with location parameters μ_i and scale parameters σ_i, $i = 1, 2$. The censoring times are generated from the same family of distributions, with locations $\mu_1 + 2$ and $\mu_2 + 1$, and scales $\sigma_1 + 3$ and $\sigma_2 + 2$. The resulting censoring rates are approximately 25% and 37% for samples 1 and 2 respectively.

We first followed Algorithm 1 to obtain the SCBs for all the three scenarios described above. Critical values were estimated from $B = 1000$ bootstrap samples. The empirical coverage probability (ECP) is the proportion of $M = 1000$ replications that yielded SCBs which included the true survival curve over the entire interval $[t_1, t_2]$, where in each case t_1 and t_2 represent the 10-th and the 90-th percentile, respectively. The ECPs for the various scenarios and various censoring rates (CRs) are summarized in Table 3.1.

Table 3.1: Empirical coverage probabilities (ECPs) of proposed 95% SCBs

Baseline distribution	$n_1 = n_2$	$\theta_1 = (-3, 1)^\top$	$\theta_2 = (5, 2)^\top$
	25	0.968	0.966
Gaussian	50	0.951	0.953
	100	0.948	0.948
CR		23%	42%
	25	0.940	0.942
Exp($\lambda = 3$)	50	0.948	0.945
	100	0.950	0.949
CR		7%	45%
	25	0.930	0.930
Extreme Value	50	0.946	0.948
	100	0.950	0.951
CR		25%	37%

The estimated average enclosed area (EAEA) is defined as

$$\text{EAEA} = \frac{1}{M} \sum_{i=1}^{M} \left\{ \sum_{j: x_j \in [t_1, t_2]} l_j \Delta x_j \right\},$$

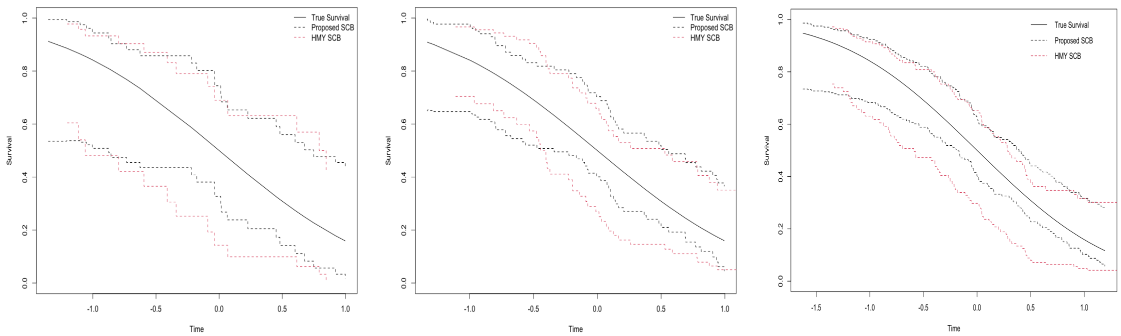
where l_j is the width of the band at the uncensored time $x_j, \Delta x_j = x_{j+1} - x_j$. The percentage reduction in EAEA of the proposed relative to the HMY SCBs are reported in Table 3.2.

Table 3.2: Percent reduction in EAEA of proposed relative to the HMY SCBs

Baseline distribution	$n_1 = n_2$	$\theta_1 = (-3, 1)^\top$	$\theta_2 = (5, 2)^\top$
	25	28.00	80.55
Gaussian	50	20.63	80.53
	100	22.46	81.11
CR		23%	42%
	25	23.90	72.24
Exp($\lambda = 3$)	50	25.83	70.45
	100	31.25	71.66
CR		10%	45%
	25	28.57	79.01
Extreme Value	50	26.16	80.07
	100	28.01	80.55
CR		25%	37%

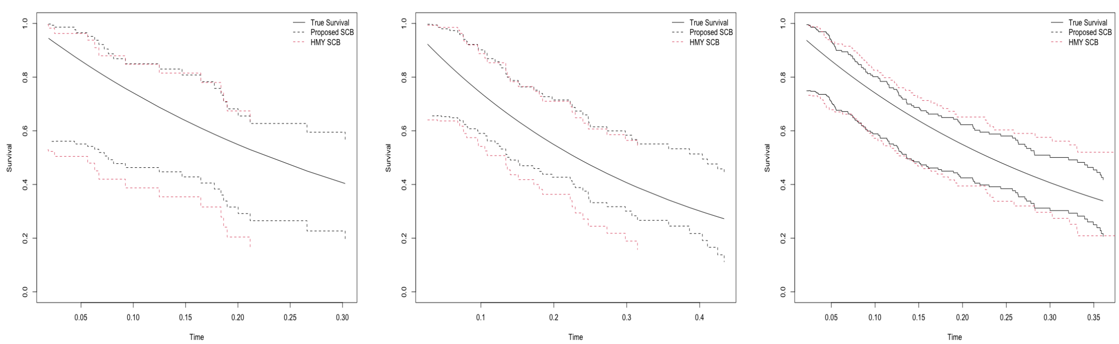
On the basis of Tables 1 and 2, it is clear that leveraging the LS framework yields narrower SCBs, with correct nominal 95% coverage. The reduction in the areas captured by the proposed SCBs is significant and can be up to 80%. Therefore, it is very conceivable that such improvements would manifest in the case of SCBs based on two samples as well. Those will be the focus of our continuing research.

In Figs 3.1 – 3.3, we show some sample SCBs obtained using the two algorithms.



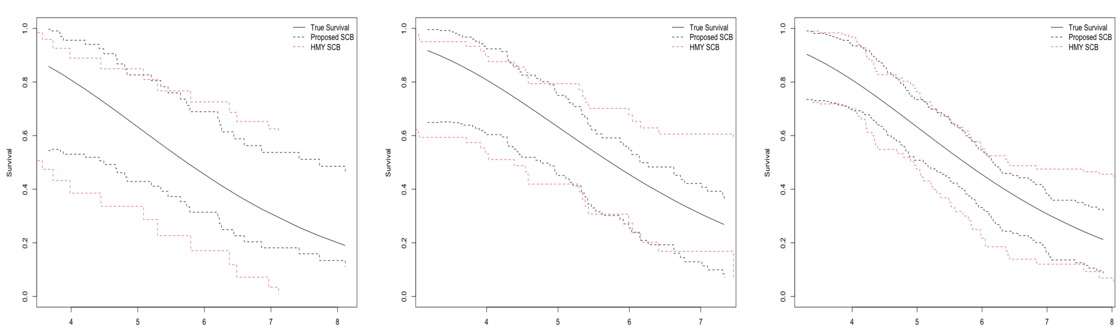
(a) $n = 25$ (b) $n = 50$ (c) $n = 100$

Figure 3.1: Sample SCBs for $N(0, 1)$ with 10% CR using Algorithm 1



(a) $n = 25$ (b) $n = 50$ (c) $n = 100$

Figure 3.2: Sample SCBs for Exponential($\lambda = 3$) with 30% CR using Algorithm 1



(a) $n = 25$ (b) $n = 50$ (c) $n = 100$

Figure 3.3: Sample SCBs for EVD(5, 2) with 45% CR using Algorithm 2

In Table 3.3 and Table 3.4, we present ECPs and EAEAs obtained using Algorithm 2.

Table 3.3: ECPs of proposed 95% SCBs for different CRs using Algorithm 2

Baseline distribution	$n_1 = n_2$	$\theta_1 = (-3, 1)^\top$	$\theta_2 = (5, 2)^\top$
	25	0.956	0.965
Gaussian	50	0.945	0.940
	100	0.949	0.951
CR		23%	42%
	25	0.940	0.942
Exp($\lambda = 3$)	50	0.945	0.943
	100	0.955	0.954
CR		7%	45%
	25	0.938	0.930
Extreme Value	50	0.945	0.955
	100	0.956	0.951
CR		25%	37%

Table 3.4: Percent reduction in EAEA relative to the HMY SCBs using Algorithm 2

Baseline distribution	$n_1 = n_2$	$\theta_1 = (-3, 1)^\top$	$\theta_2 = (5, 2)^\top$
	25	21.06	78.67
Gaussian	50	21.61	80.65
	100	22.82	88.06
CR		23%	42%
	25	25.55	71.05
Exp($\lambda = 3$)	50	32.48	71.78
	100	31.78	72.15
CR		10%	45%
	25	28.95	79.18
Extreme Value	50	27.32	80.46
	100	28.15	73.59
CR		25%	37%

In addition, we considered the performance of the proposed SCBs when the LS assumption was violated. For the first sample, the failure times were of the form $X_1 = \mu_1 + \sigma_1 Y$ where $Y = \sqrt{1-\Delta^2} Z_1 + \Delta |Z_2|$, where $\Delta \in [0,1]$ and Z_1 and Z_2 are two independent copies of the standard normal distribution. The failure times for the second sample were generated from the normal distribution with mean μ_2 and standard deviation σ_2^2. The censoring times were also normally distributed, with means $\mu_{c_1} = \mu_1 + 3$ and $\mu_{c_2} = \mu_2 + 1$, and standard deviations $\sigma_{c_1} = \sigma_1 + 3$ and $\sigma_{c_2} = \sigma_2 + 1$. These choices yield censoring rates of approximately 23% and 24% for the two samples respectively. For $\Delta = 0.05$ and $\Delta = 0.1$, X_1 does not belong to an LS family. In fact, there is significant misspecification. The proposed approach gave ECPs close to the nominal level of 95% for small sample sizes (25 and 50) but the ECPs degraded for sample size 100.

Table 3.5: ECPs of proposed 95% SCBs for mispecification study using Algorithm 2

Baseline distribution	$\boldsymbol{n_1 = n_2}$	Δ 0.05	0.1
	25	0.952	0.946
Skewed Normal	50	0.938	0.930
	100	0.934	0.916
CR		23%	24%

Table 3.6: Percent reduction in EAEA of proposed relative to the HMY SCBs using Algorithm 2

Baseline distribution	$\boldsymbol{n_1 = n_2}$	Δ 0.05	0.1
	25	-4.76	-4.98
Skewed Normal	50	-1.80	-1.82
	100	-1.45	-1.55
CR		23%	24%

CHAPTER 4

CONTINUING RESEARCH

The research proposed will be continued to obtain SCBs for the ratio and difference of two survival functions. The procedure is briefly outlined in the following subsections.

4.1 SCBs for the ratio of survival functions

Consider the LS framework $W_j = \min(U_j, V_j)$ and $\delta_j = \mathbb{1}(U_j \le V_j)$, where U_j and V_j, $j = 1, 2, \cdots, n$, are the pooled standardized failure and censored times respectively. For any two survival functions $\tilde{S}_1$ and $\tilde{S}_2$ having membership in an LS family with base survival function $\tilde{S}$, note that

$$\tilde{\varphi}(t) \;:=\; \frac{\tilde{S}_1(t)}{\tilde{S}_2(t)} \;=\; \tilde{S}\left(\frac{t-\tilde{\mu}_1}{\tilde{\sigma}_1}\right)\Big/\tilde{S}\left(\frac{t-\tilde{\mu}_2}{\tilde{\sigma}_2}\right). \tag{4.1}$$

With focus on the given pair S_1 and S_2 having base distribution S, the true ratio is

$$\varphi(t) \;=\; \frac{S_1(t)}{S_2(t)} \;=\; S\left(\frac{t-\mu_1}{\sigma_1}\right)\Big/S\left(\frac{t-\mu_2}{\sigma_2}\right).$$

To build SCBs for $\varphi(t)$, the LR when $\tilde{\theta}$ is known is (cf. McKeague and Zhao, 2002)

$$R(\tilde{\varphi}(t), t) = \frac{\sup\left\{L(\tilde{S}) : \tilde{S} \in \Gamma \text{ and Eq. (4.1) holds }\right\}}{\sup\left\{L(K) : K \in \Gamma\right\}.} \tag{4.2}$$

The survival function that maximizes the constrained likelihood in the numerator of Eq. (4.2) is discrete, supported on the uncensored time points in the pooled data and is parameterized through discrete hazards $\lambda_j, j = 1, \ldots, k$, one for each uncensored point (see, for example, Kalbfleisch and Prentice, 2002). The pooled data is arranged in a sequence of uncensored points from the smallest to the largest, with each t_j, a distinct uncensored time point, having its associated d_j, "the number of events (deaths)" and r_j, the number who are at risk at the time point "just before" t_j. For the moment assume that θ is known. With $\lambda_\theta \equiv \lambda$ without subscript denoting the LM, the objective function for the numerator of Eq. (4.2) is

$$\begin{aligned} H(\lambda_j, \lambda) = {} & \sum_{j=1}^{k} \left\{ d_j \log \lambda_j + (r_j - d_j) \log(1-\lambda_j) \right\} \\ & + \lambda \left(\sum_{j: u_j \le \frac{t-\mu_1}{\sigma_1}} \log(1-\lambda_j) - \sum_{j: u_j \le \frac{t-\mu_2}{\sigma_2}} \log(1-\lambda_j) - \log \varphi(t) \right) \end{aligned}$$

It follows that

$$
\begin{aligned}
H(\lambda_j, \lambda) &= \sum_{j=1}^{k} \{d_j \log \lambda_j + (r_j - d_j) \log (1 - \lambda_j)\} \\
&\quad + \lambda \left(\sum_{j:\min\left(\frac{t-\mu_1}{\sigma_1}, \frac{t-\mu_2}{\sigma_2}\right) \leq u_j \leq \max\left(\frac{t-\mu_1}{\sigma_1}, \frac{t-\mu_2}{\sigma_2}\right)} \log (1 - \lambda_j) - \log \varphi(t) \right)
\end{aligned}
$$

Let $B(t) = \#\{j : U_j \leq t\}$. Let $u_i = (t - \mu_i)/\sigma_i$, $i = 1, 2$. Write $a(t) = \min (B(u_1(t)), B(u_2(t))$ and $b(t) = \max (B(u_1(t)), B(u_2(t)))$. Working as in McKeague and Zhao (2002), with the estimated standardized and pooled data, it can be shown that the plug-in scaled log LR is

$$
\begin{aligned}
\mathcal{L}(\tilde{\varphi}(t), t) = -2 \sum_{a(t) \leq s \leq b(t)} \Bigg[& (Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s)) \log \left(1 + \frac{\lambda_{\hat{\theta}}(t)}{Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s)} \right) \\
& - Y_{\hat{\theta}}(s) \log \left(1 + \frac{\lambda_{\hat{\theta}}(t)}{Y_{\hat{\theta}}(s)} \right) \Bigg] \qquad (4.3)
\end{aligned}
$$

where $\lambda_{\hat{\theta}}(t)$ satisfies

$$
\prod_{a(t) \leq s \leq b(t)} \left(1 - \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s) + \lambda_{\hat{\theta}}(t)} \right) = \varphi(t). \qquad (4.4)
$$

Interestingly, Eqs. (4.3) and (4.4) are generalized versions of Eqs. (2.13) and (2.14). If $a(t) = 0$ and $b(t) = t$, then Eqs. (2.13) and (2.14) ensue. As in Eqs. (2.13) and (2.14), all quantities are indexed by estimated means and standard deviations. If those were replaced by their true values one obtains the generalized HMY counterparts. In particular, $\hat{\lambda}_{\hat{\theta}}(t)$ is a perturbed version of $\hat{\lambda}_\theta \equiv \hat{\lambda}$.

4.2 Confidence Bands for the Difference of Survival Functions

Consider the LS framework $W_j = \min (U_j, V_j)$ and $\delta_j = \mathbb{1}(U_j \leq V_j)$, where U_j and V_j, $j = 1, 2, \cdots, n$, are the pooled standardized failure and censored times respectively. Let $B(t) = \#\{j : U_j \leq t\}$. Simultaneous confidence bands for the difference $D_0(t) = S_1(t) - S_2(t)$ of survival functions will be obtained by considering the LR (Shen and He, 2006)

$$
R(S(t), t) = \frac{\sup \{L(S_1, S_2) : S_1(t) = D_0(t) + \eta(t), S_2(t) = \eta(t), S_1, S_2 \in \Theta\}}{L(\hat{S}_{\hat{\theta}})} \qquad (4.5)
$$

Using the Lagrange multiplier, we want to optimize the function

$$
\begin{aligned}
H(\lambda_j, \lambda_\theta^{(1)}, \lambda_\theta^{(2)}) = & \sum_{j=1}^{n} \{(r_j - d_j)\log(1-\lambda_j) + dj \log(\lambda_j)\} \\
& + \lambda_\theta^{(1)} \left(\sum_{j=1}^{B(u_1(t))} \log(1-\lambda_j) - \log(D_0 + \eta) \right) \\
& + \lambda_\theta^{(2)} \left(\sum_{j=1}^{B(u_2(t))} \log(1-\lambda_j) - \log(\eta) \right)
\end{aligned}
$$

An iterative solution of η, the nuisance parameter, will be obtained and, in turn, of the two LMs. The algorithm for SCBs will flow from these estimates and the bootstrap.

4.3 Numerical results

Extensive studies that will illustrate the power of the proposed methods' ability to capture the correct coverage probability will be undertaken. Furthermore the percent reduction in the EAEAs relative to the nonparametric methods will be obtained. Studies that will monitor the price of possible misspecification will be carried out. Finally the methods will be illustrated using publicly available data sets from biomedical and other public health studies.

4.4 Validity of the bootstrap

That the bootstrap produces correct thresholds for computing the SCBs will be proved. This is the asymptotic validity of the bootstrap.

APPENDIX A

APPENDIX

Lemma 1. *Suppose S_0 is continuous and that for $[t_1, t_2] \subset (0, \tau_H]$,*

$$\Psi(t_1) := -\log S(t_1) > \epsilon$$

for some $\epsilon > 0$. Then

$$\left\|\hat{\lambda}_{\hat{\theta}}\right\|_{t_1}^{t_2} = o\left((n \log n)^{1/2}\right) \text{ a.s.} \tag{A.1}$$

Proof. We follow the exact steps in the proof of lemma 2.2 of Li (1996) to get $-\log S(t) = \Psi(t) \geq \hat{\Psi}_{\hat{\theta}}(t)$ and

$$\left|\hat{\lambda}_{\hat{\theta}}\right| \leq \frac{n\left(\Psi(t) - \hat{\Psi}_{\hat{\theta}}(t)\right)}{\Psi(t)}. \tag{A.2}$$

Nestor and Subramanian (2025) obtained a large-sample representation for $\hat{A}(s) = \hat{F}_{\hat{\theta}}(s) - F(s)$, see their Eq. (2.6). Specifically, $\|\hat{A}\|_{t_1}^{t_2} = o((n/\log n)^{-1/2})$ a.s. because each centered expression on the RHS of their Eq. (2.6) admits that rate, uniformly in $t \in [t_1, t_2]$.

The Duhamel equation applied to $\hat{S}_{\hat{\theta}} - S$ (see p. 1535, Gill and Johansen, 1990) gives

$$\frac{\hat{S}_{\hat{\theta}}(t) - S(t)}{S(t)} = -\int_{(0,t]} \frac{\hat{S}_{\hat{\theta}}(s-)}{S(s)} d\{\hat{\Psi}_{\hat{\theta}}(s) - \Psi(s)\}.$$

It follows that

$$\begin{aligned}
\hat{\Psi}_{\hat{\theta}}(t) - \Psi(t) &= -\frac{\hat{S}_{\hat{\theta}}(t) - S(t)}{S(t)} - \int_{(0,t]} \left(\frac{\hat{S}_{\hat{\theta}}(s-)}{S(s)} - 1\right) d\hat{\Psi}_{\hat{\theta}}(s) + \int_{(0,t]} \left(\frac{\hat{S}_{\hat{\theta}}(s-)}{S(s)} - 1\right) d\Psi(s) \\
&\leq \frac{\|\hat{S}_{\hat{\theta}} - S\|_{t_1}^{t_2}}{S(t_2)} \left(1 + \hat{\Psi}_{\hat{\theta}}(t_2) + \Psi(t_2)\right).
\end{aligned}$$

Therefore, $\|\hat{\Psi}_{\hat{\theta}} - \Psi\|_{t_1}^{t_2} = o((n/\log n)^{-1/2})$ a.s. Inequality (A.2) then implies Eq. (A.1).

When $\hat{\lambda}_{\hat{\theta},0}(t) > 0$ for $t \in [t_1, t_2]$, we follow the steps leading to Eq. (2.13) on page 101 of Li (1996) to get

$$\log \hat{S}_{\hat{\theta}}(t) + \hat{\Psi}_{\hat{\theta}}(t) - \log S(t) \leq \hat{\Psi}_{\hat{\theta}}(t) \left(\frac{n}{n + |\hat{\lambda}_{\hat{\theta}}|}\right).$$

The strong uniform consistency of $\hat{\Psi}_{\hat{\theta}}$ and $\hat{S}_{\hat{\theta}}$ is clear from Eq. (A.6) of Nestor and Subramanian (2025). Following the steps in the proof of Lemma 1 of Subramanian (2016) leads us to conclude that for large n

$$|\hat{\lambda}_{\hat{\theta}}| \;\leq\; \frac{n\Big(\log S_0(t) - \log \hat{S}_{\hat{\theta}}(t)\Big)}{\hat{\Psi}_{\hat{\theta}}(t) + \log \hat{S}_{\hat{\theta}}(t) - \log S(t)} \;\leq\; \frac{n^{1/2}\left\|n^{1/2}\left(\log S(t) - \log \hat{S}_{\hat{\theta}}(t)\right)\right\|_{t_1}^{t_2}}{\Psi(t_1) - \epsilon}.$$

A Taylor expansion and the strong uniform consistency rate for $\hat{S}_{\hat{\theta}}$ implies Eq. (A.1), completing the proof.

□

Lemma 2. *Under the conditions of Lemma 1, uniformly for $t \in [t_1, t_2] \subset [0, \tau_H]$,*

$$\hat{\lambda}_{\hat{\theta}}(t) = -\frac{n}{\hat{\sigma}^2_{\hat{\theta}}}\left(\log \hat{S}_{\hat{\theta}}(t) - \log S(t)\right) + o_{\mathbb{P}}(\log N).$$

Proof. From Eq. (2.14), with $K(t) = S(t)$, the LM estimate $\hat{\lambda}_{\hat{\theta}}(t)$ satisfies

$$f(\hat{\lambda}_{\hat{\theta}}(t)) = \log S(t), \quad \text{where } f(\lambda) := \sum_{s\leq t} \log\left(1 - \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s) + \lambda}\right).$$

Taylor expansion of $f(\hat{\lambda}_{\hat{\theta}}(t))$ about 0 gives

$$f(\hat{\lambda}_{\hat{\theta}}(t)) = f(0) + \hat{\lambda}_{\hat{\theta}}(t) f'(0) + \frac{(\hat{\lambda}_{\hat{\theta}}(t))^2}{2} f''(\eta(t)),$$

where $|\eta(t)| \leq |\hat{\lambda}_{\hat{\theta}}(t)|$. Note that $f(0) = \log \hat{S}_{\hat{\theta}}(t)$. Note also that $f'(0) = (\hat{\sigma}_{\hat{\theta}}(t))^2/n$, where

$$\hat{\sigma}_{\hat{\theta}}(t) = n \sum_{s\leq t} \frac{\Delta N_{\hat{\theta}}(s)}{Y_{\hat{\theta}}(s)\,(Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s))} \tag{A.3}$$

Then

$$\log S_0(t) = \log \hat{S}_{\hat{\theta}}(t) + \hat{\lambda}_{\hat{\theta}}(t)\frac{(\hat{\sigma}_{\hat{\theta}}(t))^2}{N} + \frac{(\hat{\lambda}_{\hat{\theta}}(t))^2}{2} f''(\eta(t)).$$

Hence,

$$\hat{\lambda}_{\hat{\theta}}(t) \;=\; -\frac{n}{(\hat{\sigma}_{\hat{\theta}}(t))^2}\left\{\left(\log \hat{S}_{\hat{\theta}}(t) - \log S(t)\right) + \frac{(\hat{\lambda}_{\hat{\theta}}(t))^2}{2} f''(\eta(t))\right\}. \tag{A.4}$$

It remains to show $f''(\eta(t)) = O_{\mathbb{P}}(n^{-2})$ uniformly for $t \in [t_1, t_2]$, and then apply Lemma (2). Note that

$$f''(\eta) = -\sum_{s\leq t} \frac{\Delta N_{\hat{\theta}}(s)(2[Y_{\hat{\theta}}(s) + \eta(t)] - \Delta N_{\hat{\theta}}(s))}{(Y_{\hat{\theta}}(s) + \eta - \Delta N_{\hat{\theta}}(s))^2 (Y_{\hat{\theta}}(s) + \eta)^2}.$$

Note that we consider only $[t_1, t_2] \subset (0, \tau_H]$, where τ_H is such that $1 - H(\tau_H) > 0$. Choose $\epsilon > 0$ so that $\epsilon < 1 - H(\tau_H) < 1 - H(\tau_H-)$. Since $n^{-1}\left(Y_{\hat{\theta}}(\tau_H) - \Delta N_{\hat{\theta}}(\tau_H)\right) \xrightarrow{\text{a.s.}} 1 - H(\tau_H)$, we conclude that, for large n,

$$\frac{1}{n}\left[Y_{\hat{\theta}}(\tau_H) - \Delta N_{\hat{\theta}}(\tau_H)\right] \;>\; \epsilon. \tag{A.5}$$

For $s \leq t \leq \tau_H$, we have that $Y_{\hat{\theta}}(t) - \Delta N_{\hat{\theta}}(t) \geq Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s)$. Therefore

$$Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s) \;\geq\; Y_{\hat{\theta}}(\tau_H) - \Delta N_{\hat{\theta}}(\tau_H). \tag{A.6}$$

Furthermore, from $|\eta(t)/n| < \epsilon$ for all $t \in [t_1, t_2]$ when n is large. It follows from inequalities (A.5) & (A.6) that

$$Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s) + \eta(t) > Y_{\hat{\theta}}(\tau_H) - \Delta N_{\hat{\theta}}(\tau_H) - n\epsilon > 0.$$

Because $\sum_{s\leq t} \Delta N_{\hat{\theta}}(s) \leq n$, we have

$$|f''(\eta(t))| \leq \frac{n\left(2\left(n + \sup_{t\in[t_1,t_2]} |\eta(t)|\right) + n\right)}{\left(Y_{\hat{\theta}}(\tau_H) - \Delta N_{\hat{\theta}}(\tau_H) - n\epsilon\right)^2 \left(Y_{\hat{\theta}}(\tau_H) - n\epsilon\right)^2} = O_{\mathbb{P}}(n^{-2}).$$

By Lemma (1), therefore, the dominant part of the remainder term in Eq. (A.4), $N(\hat{\lambda}_{\hat{\theta}}(t))^2 f''(\eta(t))$, is $o_{\mathbb{P}}(n^2 \log n) O_{\mathbb{P}}(n^{-2}) = o_{\mathbb{P}}(\log n)$ which completes the proof. □

Proof of Theorem 1

Recall from Eq. (2.13) that $\mathcal{L}(S(t), t)$, the scaled log LR, is

$$-2\sum_{s\leq t}\left\{(Y(s) - \Delta N(s))\log\left(1 + \frac{\hat{\lambda}_{\hat{\boldsymbol{\theta}}}(t)}{Y(s) - \Delta N(s)}\right) - Y(s)\log\left(1 + \frac{\hat{\lambda}_{\hat{\boldsymbol{\theta}}}(t)}{Y(s)}\right)\right\},$$

where $\hat{\lambda}_{\hat{\theta}}(t)$ satisfies Eq. (2.14) with $K(t)$ replaced by $S(t)$. Following Li (1996), apply $\log(1 + x) = x - x^2/2 + x^3/3 - x^4/4 + O(x^5)$ as $x \to \infty$, to obtain

$$\begin{aligned}
-2\log\mathcal{L}(S(t), t) &= \hat{\lambda}^2_{\hat{\boldsymbol{\theta}}}(t)\sum_{s\leq t}\left\{\frac{1}{Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s)} - \frac{1}{Y_{\hat{\theta}}(s)}\right\} \\
&\quad - \frac{2}{3}\hat{\lambda}^3_{\hat{\boldsymbol{\theta}}}(t)\sum_{s\leq t}\left\{\left(\frac{1}{Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s)}\right)^2 - \left(\frac{1}{Y_{\hat{\theta}}(s)}\right)^2\right\} \\
&\quad + \frac{1}{2}\hat{\lambda}^4_{\hat{\boldsymbol{\theta}}}(t)\sum_{s\leq t}\left\{\left(\frac{1}{Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s)}\right)^3 - \left(\frac{1}{Y_{\hat{\theta}}(s)}\right)^3\right\} + O_{\mathbb{P}}\left(n^{-1/2}\right) \\
&:= A_1(t) + A_2(t) + A_3(t) + o_{\mathbb{P}}(1)
\end{aligned} \tag{A.7}$$

Applying Lemma 2, we have

$$
\begin{aligned}
A_1(t) &= \frac{\hat{\sigma}^2_{\hat{\theta}}(t)}{n}\,\hat{\lambda}^2_{\hat{\theta}}(t) \\
&= \frac{\hat{\sigma}^2_{\hat{\theta}}(t)}{n}\left\{-\frac{n}{\hat{\sigma}^2_{\hat{\theta}}(t)}\left(\log \hat{S}_{\hat{\theta}}(t) - \log S(t)\right) + o_{\mathbb{P}}(\log n)\right\}^2 \\
&= \frac{1}{\hat{\sigma}^2_{\hat{\theta}}(t)}\left\{n^{1/2}\left(\log \hat{S}_{\hat{\theta}}(t) - \log S(t)\right)\right\}^2 + o_{\mathbb{P}}\left(\frac{\log n}{n^{1/2}}\right),
\end{aligned}
$$

which leads to the RHS of Eq. (2.16). We now show that $A_2(t)$ and $A_3(t)$ are asymptotically negligible.

As in Li (1996), we have

$$
\begin{aligned}
||A_2(t)||_{t_1}^{t_2} &\leq \frac{2}{3}\left(||\lambda_{\hat{\boldsymbol{\theta}}}(t)||_{t_1}^{t_2}\right)^3\left(\frac{1}{Y_{\hat{\theta}}(t_2) - \Delta N_{\hat{\theta}}(t_2)} + \frac{1}{Y_{\hat{\theta}}(t_2)}\right) \\
&\quad \times \sum_{s \leq t_2}\left(\frac{1}{Y_{\hat{\theta}}(s) - \Delta N_{\hat{\theta}}(s)} + \frac{1}{Y_{\hat{\theta}}(s)}\right).
\end{aligned}
$$

Apply Lemma 1 to obtain

$$
\begin{aligned}
||A_2(t)||_{t_1}^{t_2} &= o_{\mathbb{P}}\left((n \log n)^{3/2}\right) O_{\mathbb{P}}\left(n^{-1}\right)\frac{\hat{\sigma}^2_{\hat{\theta}}(t_2)}{n} \\
&= o_{\mathbb{P}}\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right) \\
&= o_{\mathbb{P}}(1).
\end{aligned}
$$

The summation term of $A_3(t)$ can be bounded above by

$$
\left(\frac{1}{\left(Y_{\hat{\theta}}(t_2) - \Delta N_{\hat{\theta}}(t_2)\right)^2} + \frac{1}{Y_{\hat{\theta}}(t_2)\left(Y_{\hat{\theta}}(t_2) - \Delta N_{\hat{\theta}}(t_2)\right)} + \frac{1}{Y^2(t_2)}\right)\frac{\hat{\sigma}^2_{\hat{\theta}}(t_2)}{n}
$$

which is $O_{\mathbb{P}}\left(\left(\frac{1}{n}\right)^2\right) O_{\mathbb{P}}\left(\frac{1}{n}\right)$. By Lemma 1, it follows that

$$
\begin{aligned}
||A_3(t)||_{t_1}^{t_2} &= o_{\mathbb{P}}\left((n \log n)^2\right) O_{\mathbb{P}}\left(n^{-2}\right) O_{\mathbb{P}}(n^{-1}) \\
&= o_{\mathbb{P}}\left(\frac{(\log n)^2}{n}\right) \\
&= o_{\mathbb{P}}(1).
\end{aligned}
$$

This completes the proof of Theorem 1. □

REFERENCES

[1] N. Ahmed and S. Subramanian. 'Semiparametric simultaneous confidence bands for the difference of survival functions'. In: *Lifetime Data Analysis* (2015).

[2] M.G. Akritas. 'Bootstrapping the Kaplan–Meier estimator'. In: *Journal of the American Statistical Association* 81 (1986), pp. 1032–1038.

[3] R. Berk and D. Jones. 'Goodness-of-fit test statistics that dominate the Kolmogorov statistics'. In: *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* 47 (1979), pp. 47–59.

[4] N.E. Breslow and J.J. Crowley. 'A large sample study of the life table and product limit estimates under random censorship'. In: *Annals of Statistics* 2 (1973), pp. 437–453.

[5] C.F. Chung. *Formulas for probabilities associated with Wiener and Brownian bridge processes*. Tech. rep. 79. Laboratory for Research in Statistics and Probability, Carleton University, Ottawa, Canada, 1986.

[6] C.F. Chung. *Wiener pack: a subroutine package for computing probabilities associated with Wiener and Brownian bridge processes*. Tech. rep. 87-12. Geological Survey of Canada, 1987.

[7] G. Claeskens and N.L. Hjort. *Model Selection and Model Averaging*. Cambridge University Press, 2008.

[8] G. Dikta. 'Asymptotically efficient estimation under semi-parametric random censorship models'. In: *Journal of Multivariate Analysis* 124 (2014), pp. 10–24.

[9] G. Dikta. 'On semiparametric random censorship models'. In: *Journal of Statistical Planning and Inference* 66 (1998), pp. 253–279.

[10] G. Dikta, M. Kvesic, and C. Schmidt. 'Bootstrap approximations in model checks for binary data'. In: *Journal of the American Statistical Association* 101 (2006), pp. 521–530.

[11] B. Efron. 'Censored data and the bootstrap'. In: *Journal of the American Statistical Association* 76 (1981), pp. 312–319.

[12] J.H.J. Einmahl and I.W. McKeague. 'Confidence tubes for multiple quantile plots via empirical likelihood'. In: *Annals of Statistics* 27 (1999), pp. 1348–1367.

[13] R.D. Gill and S. Johansen. 'A survey of product-integration with a view toward application in survival analysis'. In: *Annals of Statistics* 18 (1990), pp. 1501–1555.

[14] W. Hall and J.A. Wellner. 'Confidence bands for a survival curve from censored data'. In: *Biometrika* 67 (1980), pp. 133–143.

[15] M. Hollander, I.W. McKeague, and J. Yang. 'Likelihood ratio-based confidence bands for survival functions'. In: *Journal of the American Statistical Association* 92 (1997), pp. 215–226.

[16] L. Jager and J.A. Wellner. *A new goodness of fit test: The reversed Berk–Jones statistic*. Tech. rep. 443. Department of Statistics, University of Washington, 2005.

[17] John D. Kalbfleisch and Ross L. Prentice. *The Statistical Analysis of Failure Time Data*. 2nd. Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley-Interscience, 2002. ISBN: 9780471363576.

[18] R. Koenker and J. Yoon. 'Parametric links for binary choice models: A Fisherian–Bayesian colloquy'. In: *Journal of Econometrics* 152 (2009), pp. 120–130.

[19] G. Li. 'On nonparametric likelihood ratio estimation of survival probabilities for censored data'. In: *Statistics and Probability Letters* 25 (1995), pp. 95–104.

[20] G. Li et al. 'Nonparametric likelihood ratio confidence bands for quantile functions from incomplete survival data'. In: *Annals of Statistics* 23 (1996), pp. 628–640.

[21] D. Matthews. 'Exact nonparametric confidence bands for the survivor function'. In: *International Journal of Biostatistics* 9 (2013), pp. 185–204.

[22] P. McCullagh and J.A. Nelder. *Generalized Linear Models*. 2nd ed. Chapman Hall, 1989.

[23] I.W. McKeague and Y. Zhao. 'Comparing distribution functions via empirical likelihood'. In: *International Journal of Biostatistics* 1 (2005), Article 5.

[24] Ian W. McKeague and Yuehua Zhao. 'Simultaneous Confidence Bands for Ratios of Survival Functions via Empirical Likelihood'. In: *Statistics & Probability Letters* 60.4 (2002), pp. 405–415.

[25] S. Mondal and S. Subramanian. 'Model assisted Cox regression'. In: *Journal of Multivariate Analysis* 123 (2014), pp. 281–303.

[26] S. Mondal and S. Subramanian. 'Simultaneous confidence bands for Cox regression from semiparametric random censorship'. In: *Lifetime Data Analysis* 22 (2016), pp. 122–144.

[27] B.J.T. Morgan and D.M. Smith. 'A note on Wadley's problem with overdispersion'. In: *Applied Statistics* 41 (1992), pp. 349–354.

[28] V. Nair. 'Confidence bands for survival functions with censored data: a comparative study'. In: *Technometrics* 14 (1984), pp. 265–275.

[29] A. Owen. 'Empirical likelihood ratio confidence intervals for a single functional'. In: *Biometrika* 75 (1988), pp. 265–275.

[30] A. Owen. 'Empirical likelihood ratio confidence regions'. In: *Annals of Statistics* 18 (1990), pp. 90–120.

[31] A. Owen. 'Nonparametric confidence bands for a distribution function'. In: *Journal of the American Statistical Association* 90 (1995), pp. 516–521.

[32] M.I. Parzen, L.J. Wei, and Z. Ying. 'Simultaneous confidence intervals for the difference of two survival functions'. In: *Scandinavian Journal of Statistics* 24 (1997), pp. 309–314.

[33] W.H. Press et al. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. 2nd ed. Cambridge University Press, 1992.

[34] Junshan Shen and Shuyuan He. 'Empirical likelihood for the difference of two survival functions under right censorship'. In: *Statistics & Probability Letters* 76.2 (2006), pp. 169–181.

[35] W. Stute. 'Nonparametric model checks for regression'. In: *Annals of Statistics* 25 (1997), pp. 613–641.

[36] W. Stute. 'Strong consistency of the MLE under random censoring'. In: *Merika* 39 (1992), pp. 257–267.

[37] S. Subramanian. 'Bootstrap likelihood ratio confidence bands for survival function under random censorship and its semiparametric extension'. In: *Journal of Multivariate Analysis* 147 (2016), pp. 51–81.

[38] S. Subramanian. 'Model-based likelihood ratio confidence intervals for survival functions'. In: *Statistics and Probability Letters* 82 (2012), pp. 626–635.

[39] S. Subramanian. 'The missing censoring-indicator model of random censorship'. In: *Advances in Survival Analysis, Handbook of Statistics, Vol. 23*. Ed. by N. Balakrishnan and C.R. Rao. 2004, pp. 123–141.

[40] S. Subramanian and D. Bean. 'Hazard function estimation from homogeneous right censored data with missing censoring indicators'. In: *Statistical Methodology* 5 (2008), pp. 515–527.

[41] S. Subramanian and P. Zhang. 'Model-based confidence bands for survival functions'. In: *Journal of Statistical Planning and Inference* 143 (2013), pp. 1355–1374.

[42] D.R. Thomas and G.L. Grunkemeier. 'Confidence interval estimation of survival probabilities for censored data'. In: *Journal of the American Statistical Association* 70 (1975), pp. 865–871.

[43] A.A. Tsiatis, M. Davidian, and B. McNeney. 'Multiple imputation methods for testing treatment differences in survival distributions with missing cause of failure'. In: *Biometrika* 89 (2002), pp. 238–244.

[44] M. Yuan. 'Semiparametric censorship model with covariates'. In: *TEST* 14 (2005), pp. 1–26.

[45] M.J. Zhang and J. Klein. 'Confidence bands for the difference of two survival curves under proportional hazards model'. In: *Lifetime Data Analysis* 7 (2001), pp. 243–254.