OptBand: optimization-based confidence bands for functions to characterize time-to-event distributions

T. Chen¹ · S. Tracy^{2,3} · H. Uno³

Received: 15 August 2019 / Accepted: 19 March 2021 / Published online: 13 April 2021
© The Author(s) 2021

Abstract

Classical simultaneous confidence bands for survival functions (i.e., Hall–Wellner, equal precision, and empirical likelihood bands) are derived from transformations of the asymptotic Brownian nature of the Nelson–Aalen or Kaplan–Meier estimators. Due to the properties of Brownian motion, a theoretical derivation of the highest confidence density region cannot be obtained in closed form. Instead, we provide confidence bands derived from a related optimization problem with local time processes. These bands can be applied to the one-sample problem regarding both cumulative hazard and survival functions. In addition, we present a solution to the two-sample problem for testing differences in cumulative hazard functions. The finite sample performance of the proposed method is assessed by Monte Carlo simulation studies. The proposed bands are applied to clinical trial data to assess survival times for primary biliary cirrhosis patients treated with D-penicillamine.

Keywords Survival · Confidence band · Highest density region · Optimization

S. Tracy and T. Chen have contributed equally to the work

✉ S. Tracy
seasamgo@gmail.com

T. Chen
tomchen00@gamil.com

H. Uno
huno@ds.dfci.harvard.edu

¹ Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care, Boston, USA

² Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, USA

³ Department of Data Science, Dana-Faber Cancer Institute, Boston, USA

1 Introduction

For time-to-event outcomes in clinical studies, inference for the cumulative distribution and survival functions are of interest. The Kaplan–Meier (KM) estimator and its corresponding asymptotic variance are often used to construct pointwise confidence intervals. When interest lies in inference of the survival curve over an entire time interval, “simultaneous” confidence bands (CB) need to be constructed instead of a set of pointwise confidence intervals.

For continuous function $\varphi(t)$ which characterizes the event time distribution, we say two stochastic processes $\mathcal{L}(t)$ and $\mathcal{U}(t)$ are $(1 - \alpha)$ coverage level simultaneous CBs for $\varphi(t)$ on the interval $[t_L, t_U]$ if

$$\mathbb{P}(\mathcal{L}(t) \leq \varphi(t) \leq \mathcal{U}(t) \quad \forall t \in [t_L, t_U]) \geq 1 - \alpha.$$

Common choices for α are 0.01, 0.05, 0.10. We distinguish this definition from the standard pointwise CBs, which are constructed such that at each time point t , $\mathbb{P}(\mathcal{L}(t) \leq \varphi(t) \leq \mathcal{U}(t)) = 1 - \alpha$. Clearly, pointwise CBs are narrower than simultaneous CBs, but the former do not attain the nominal coverage level $(1 - \alpha)$ on the specified interval $[t_L, t_U]$. Two classical simultaneous CBs based on the asymptotics of the KM estimator were originally developed by Hall and Wellner (1984) and Nair (1984); they are now popularly referred to as the Hall–Wellner (HW) and equal precision (EP) bands, respectively. Both of these aforementioned papers derived a pivotal quantity for the estimated survival function and showed that this quantity weakly converges to a Brownian motion or Brownian bridge, which were then inverted to obtain simultaneous CBs for the survival function. Hollander et al. (1997) derived simultaneous CBs for survival and cumulative hazard functions based on the empirical likelihood (EL) approach for confidence intervals introduced by Thomas and Grunkemeier (1975). This technique has also been applied to quantile functions by Li et al. (1996), but was not applied to the two-sample problem. Furthermore, McKeague and Zhao (2002) proposed CBs for ratios of survival functions using EL techniques. They also extended their approaches to construct CBs for differences and ratios of linear functionals of the cumulative hazard functions (McKeague and Zhao 2005). The extended approach relies on simulation to estimate critical values. Moreover, EL CBs for survival functions are not readily accessible as they have not been implemented in standard computing software. Parzen et al. (1994) introduced a perturbation resampling method as a generalized approach in calculating critical values for transformed distributions and notably applied this to the two-sample problem (Parzen et al. 1997). Tian et al. (2011) derived bands which target the highest confidence density region (HCDR), but this approach requires the standard specifications and tuning procedures that accompany a Markov chain Monte Carlo process and can be computationally burdensome. Finally, Cui and Hannig (2019a) introduced a nonparametric fiducial approach to confidence bands, which has been shown to be robust and efficient in small samples; several auxiliary works (Cui and Hannig 2019b; Martin 2019) explored the implications of this work and how fiducial inference fits in the context of modern statistics.

In this paper, we propose a novel procedure to construct analytical simultaneous CBs $\{\mathfrak{L}(t), \mathfrak{U}(t)\}$ for $\varphi(t)$ which approximately target the HCDR. That is, such bands aim to minimize

$$\begin{aligned} & \mathbb{E} \left(\int_{t_L}^{t_U} \{\mathfrak{U}(t) - \mathfrak{L}(t)\} dt \right) \\ & \text{s.t. } \mathbb{P} \left(\mathfrak{L}(t) \leq \varphi(t) \leq \mathfrak{U}(t) \quad \forall t \in [t_L, t_U] \right) = 1 - \alpha. \end{aligned} \quad (1)$$

To accomplish the objective in Eq. (1), we utilize an analytical result from Kendall et al. (2007), which provides an approximate solution to a related optimization problem

$$\begin{aligned} & \min_u \left(\int_{t_L}^{t_U} u(t) dt \right) \\ & \text{s.t. } \mathbb{P} \left(-u(t) \leq W(\sigma^2(t)) \leq u(t) \quad \forall t \in [t_L, t_U] \right) = 1 - \alpha. \end{aligned} \quad (2)$$

The approximate solution to Eq. (2) is $u_\kappa^*(t) = \psi(\kappa\sigma^2(t))\sigma(t)$, where $W(\sigma^2(t))$ is a mean-zero Brownian motion with strictly-increasing variance function $\sigma^2(t)$, κ is a critical value related to the desired coverage level $1 - \alpha$, $\psi(x) = (-\mathcal{W}_{-1}(-x^2))^{1/2} \mathbb{I}(x \leq e^{-1/2})$, $\mathcal{W}_{-1}$ is the Lambert W function on the -1 branch, and $\mathbb{I}$ is the indicator function (see Appendix A for details of the derivation). The solution $u_\kappa^*(t)$ is approximate in the sense that it replaces the probability constraint in Eq. (2) with its local time portion from a Doob-Meyer decomposition. Nevertheless, Kendall et al. (2007) have demonstrated its success in practical contexts, and henceforth we coin our simultaneous CBs derived from their results as simply “OptBand”. To wit, our strategy is to form pivotal quantities that are asymptotically Brownian, apply the result by Kendall et al. (2007), and transform these pivotal quantities into CBs for $\varphi(t)$, which could be either the cumulative hazard function or survival function.

In practice, there are two challenges in adopting the work of Kendall et al. (2007). The first is the burden in calculating the critical value parameter κ . As shown in Appendix A, κ depends on t_L , t_U , $\sigma^2(t)$, and α , which are typically study-specific and depend on the clinical research question of interest. Secondly, how to apply this theory to construct OptBand for survival functions is not entirely clear. Non-linear pivotal quantity transformations do not necessarily preserve the HCDR property from the optimization problem in Eq. (2) to the optimization problem in Eq. (1). The pivotal quantity transformation for the cumulative hazard function happens to be exactly linear and therefore it is straightforward to derive OptBand for this case. The survival function, however, requires more sophisticated means. In this paper, we tackle these challenges. In Sect. 2, we derive a highly accurate functional approximation in calculating κ so that users may automatically determine κ for their specific problems and not have to be constrained to specific parameters found in standard critical value tables. We construct OptBand for the one-sample cumulative hazard function and two-sample difference in cumulative hazard functions in Sect. 3. OptBand for the one-sample survival function is derived in Sect. 4. We evaluate OptBand’s performance against performances of fiducial, empirical likelihood, equal precision, and Hall-Wellner bands in Sect. 5, and

illustrate with an application to clinical trial data concerning primary biliary cirrhosis in Sect. 6. The paper concludes with a discussion in Sect. 7.

2 Functional approximation of the critical value parameter

Akin to the critical values required for the HW or EP bands, OptBand's critical value, κ , is required to attain a desired $(1 - \alpha)$ coverage level. Such value is computed to satisfy

$$\mathbb{P}\left(|W(\sigma^2(t))| \leq \psi(\kappa\sigma^2(t)/\sigma^2(t_U))\sqrt{\sigma^2(t)} \quad \forall t \in [t_L, t_U]\right) = 1 - \alpha.$$

Kendall et al. (2007) performed a Monte Carlo simulation for the case $t_L = 0$, $t_U = 1$, $\alpha = 0.05$, $\sigma^2(t) = t$ (the standard Brownian motion on the unit interval) in approximating $\kappa \approx 0.105$. We provide an approximation procedure that generalizes their simulation.

Without loss of generality, we may consider minimizing the interval around $W(\sigma^2(t)/\sigma^2(t_U))$ for $t \in [t_L, t_U]$, since by properties of Brownian motion, $\sigma^{-1}(t_U)W(\sigma^2(t))$ and $W(\sigma^2(t)/\sigma^2(t_U))$ are equal in distribution. Hence, we only need to compute κ such that

$$\mathbb{P}(|W(s)| \leq \psi(\kappa s)s^{1/2} \quad \forall s \in [L, 1]) = 1 - \alpha. \quad (3)$$

Therefore, κ is now only a function of α and $L \stackrel{\text{def}}{=} \sigma^2(t_L)/\sigma^2(t_U) \in [0, 1]$. Through Monte Carlo simulations calculations using Eq. (3), we form an array of $(1 - \alpha)$ coverage values corresponding to (κ, L) . While this array is sufficient as a lookup table for the critical values, the functional relationship is accurately approximated with $1 - \alpha \approx 1 + (a + bL)\kappa + a\kappa^2$, or

$$\kappa \approx -\frac{(a + bL) + \{(a + bL)^2 - 4a[1 - (1 - \alpha)]\}^{1/2}}{2a}. \quad (4)$$

where $a = -0.4272$, $b = 0.2848$ (see Appendix A). As a measure of goodness of fit, the adjusted R^2 for Eq. (4) based on our Monte Carlo data is 0.9987. Furthermore, the resulting $(1 - \alpha)$ from the simulation ranges from 0.871 to 0.999, encompassing accurate interpolation for clinically relevant coverage levels of 0.90, 0.95, and 0.99. Applying Eq. (4) to the standard Brownian motion on $[0, 1]$, we calculate $\kappa = 0.106$, which is very close to the value specified in Kendall et al. (2007).

3 OptBand for the cumulative hazard function

3.1 One-sample problem

Let T following distribution F and C following distribution G denote failure times and censoring times, respectively. Let $\{T_i, C_i\}_{i=1}^n$ be i.i.d. copies of $\{T, C\}$. Throughout

the paper, we assume T and C are independent of each other. Our observed data are $(\tilde{T}_i, \Delta_i)_{i=1}^n$, where $\tilde{T}_i = \min(T_i, C_i)$ and $\Delta_i = \mathbb{I}\{T_i \leq C_i\}$. Let $\hat{H}(t)$ be the Nelson-Aalen estimator of the true cumulative hazard function $H(t)$. Restricting our interval of interest to $[t_L, t_U]$, standard asymptotic results dictate that $n^{1/2}\{\hat{H}(t) - H(t)\}/\sigma(t_U)$ weakly converges to $W(\sigma^2(t)/\sigma^2(t_U))$ (Fleming and Harrington 1991), where

$$\sigma^2(t) = \int_0^t \frac{dH(s)}{(1 - F(s))(1 - G(s))}.$$

OptBand's inference is based on the pivotal quantity

$$\left| n^{1/2} \frac{\hat{H}(t) - H(t)}{\sigma(t_U)} \right| \leq \psi \left(\kappa \frac{\sigma^2(t)}{\sigma^2(t_U)} \right) \frac{\sigma(t)}{\sigma(t_U)}.$$

Since we do not observe the true variance $\sigma^2(t)$, we replace it with an estimator. Throughout, we will use Greenwood's formula

$$\hat{\sigma}^2(t) = n \sum_{j: T_j \leq t} \frac{d_j}{n_j(n_j - d_j)},$$

where $d_i = \sum_{j: T_j = \tilde{T}_i} T_j \Delta_j$ is the number of failures at $\tilde{T}_i$, and $n_i = |\{j : T_j \geq \tilde{T}_i\}|$ is the number of individuals at risk at time $\tilde{T}_i$. Pivoting, the asymptotic $(1 - \alpha)$ level OptBand for $H(t)$ is $\{\mathcal{L}(t), \mathcal{U}(t)\} = \hat{H}(t) \pm c_{CH}(t)$, where

$$c_{CH}(t) = \psi \left(\kappa \frac{\hat{\sigma}^2(t)}{\hat{\sigma}^2(t_U)} \right) \frac{\hat{\sigma}(t)}{n^{1/2}}$$

and κ is computed using Eq. (4).

3.2 Two-sample problem

The two-sample problem considers CBs for $H_1(t) - H_2(t)$, where $H_1(t)$ and $H_2(t)$ are the cumulative hazard functions which give rise to data $(\tilde{T}_{1i}, \Delta_{1i})_{i=1}^{n_1}$ and $(\tilde{T}_{2i}, \Delta_{2i})_{i=1}^{n_2}$, respectively. The quantity

$$[\hat{H}_1(t) - \hat{H}_2(t)] - [H_1(t) - H_2(t)] = [\hat{H}_1(t) - H_1(t)] - [\hat{H}_2(t) - H_2(t)]$$

weakly converges to $W(n_1^{-1}\sigma_1^2(t) + n_2^{-1}\sigma_2^2(t))$. Applying the same strategy as before, the pivotal quantity must satisfy, with probability $(1 - \alpha)$, as follows:

$$\left| (n_1 + n_2)^{-1/2} \frac{[\hat{H}_1(t) - H_1(t)] - [\hat{H}_2(t) - H_2(t)]}{\sigma_p(t_U)} \right| \leq \psi \left(\kappa \frac{\sigma_p^2(t)}{\sigma_p^2(t_U)} \right) \frac{\sigma_p(t)}{\sigma_p(t_U)},$$

where $\sigma_p^2(t) = (n_1 + n_2) \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)$. Replacing $\sigma_p^2(t)$ with estimator $\widehat{\sigma}_p^2(t)$ and pivoting, we derive the $(1 - \alpha)$ level OptBand for $H_1(t) - H_2(t)$ as $\{\mathfrak{L}(t), \mathfrak{U}(t)\} = [\widehat{H}_1(t) - \widehat{H}_2(t)] \pm c_{CH}(t)$, where

$$c_{CH}(t) = \psi \left(\kappa \frac{\widehat{\sigma}_p^2(t)}{\widehat{\sigma}_p^2(t_U)} \right) \frac{\widehat{\sigma}_p(t)}{(n_1 + n_2)^{1/2}}.$$

As before, κ is computed using Eq. (4).

4 OptBand for the survival function

4.1 One-sample problem

Let $\widehat{S}(t)$ be either the Fleming-Harrington (Fleming and Harrington 1991) or Kaplan–Meier estimator for the true survival function $S(t)$. Standard asymptotic results show that both estimators satisfy that

$$n^{1/2} \frac{\widehat{S}(t) - S(t)}{S(t)} \text{ weakly converges to } W \left(\sigma^2(t) \right). \quad (5)$$

As noted in the Introduction, we cannot apply the same strategy with this pivotal quantity as with cumulative hazard functions as it would focus on minimizing the area between the bands of $\{\widehat{S}(t) - S(t)\}/S(t)$, not $\widehat{S}(t) - S(t)$. We overcome this problem by weighting the Eq. (5) by $S(t)$ and re-optimizing; that is, we target $\widetilde{u}(t) = S(t)u(t)$ so that $n^{1/2}|\widehat{S}(t) - S(t)| \leq \widetilde{u}(t)$ with the same optimization strategy from Kendall et al. (2007). As shown in Appendix B, the $(1 - \alpha)$ level OptBand for $S(t)$ is $\{\mathfrak{L}(t), \mathfrak{U}(t)\} = \widehat{S}(t) (1 \pm c_S(t))$, where

$$c_S(t) = \psi \left(\kappa \widehat{S}(t) \frac{\widehat{\sigma}^2(t)}{\widehat{\sigma}^2(t_U)} \right) \frac{\widehat{\sigma}(t)}{n^{1/2}}$$

and κ is calculated according to

$$\kappa = - \frac{\tilde{b} + \{\tilde{b}^2 - 4\tilde{a}\tilde{c}\}^{1/2}}{2\tilde{a}}$$

to obtain a coverage level of $1 - \alpha$, where

$$\begin{aligned} \tilde{a} &= a \bar{S}_{K-1}^2, \\ \tilde{b} &= \frac{b}{\widehat{\sigma}^2(t_U)} \left\{ \sum_{i=1}^{K-2} \bar{S}_i (\widehat{\sigma}^2(\xi_i) - \widehat{\sigma}^2(\xi_{i+1})) \right\} + \left(a + b \frac{\widehat{\sigma}^2(\xi_{K-1})}{\widehat{\sigma}^2(t_U)} \right) \bar{S}_{K-1}, \\ \tilde{c} &= \alpha, \end{aligned}$$

and $\bar{S}_i = (\widehat{S}(\xi_i) + \widehat{S}(\xi_{i+1}))/2$, where $\xi_1 \leq \xi_2 \leq \dots \leq \xi_K$ are the failure time points we observe in the sample and K is the total number of observed points (see Appendix B for derivation).

5 Simulation

We compare our proposed OptBand against FD-I (Cui and Hannig 2019a), EL (Hollander et al. 1997), EP (Nair 1984), and HW (Hall and Wellner 1984) bands. Hollander et al. (1997) proposed several types of EL bands; we compare against a bias-corrected EP-type EL band, which was shown to have the most consistently good performance. We generate T_i from a unit exponential distribution and C_i from an exponential distribution with rate parameters 0, 0.25, 1 and 9 to demonstrate 0%, 20%, 50% and 90% censored observations, respectively. These were done at sample sizes $n = 100, 500, 1000$ across $R = 2,000$ replicate simulations. We restrict to time intervals $\mathcal{T}_{a,b} = \{t : a \leq d(t) \leq b\}$, where $d(t) = \widehat{\sigma}^2(t)/[1 + \widehat{\sigma}^2(t)]$ and $(a, b) = (0.05, 0.95), (0.05, 0.8), (0.2, 0.8), (0.2, 0.95)$. This form of time restriction, as opposed to trimming off the first a and last $1 - b$ quantiles of times, was the formulation used by Nair (1984) and Hollander et al. (1997) when evaluating the validity of the coverage of their bands. The form of $d(t)$ is derived from a transformation used in Nair (1984) to yield EP bands. Critical values for EP bands were only established for $0.02 \leq a < b \leq 0.98$ (Klein and Moeschberger 2006), and both Nair (1984) and Hollander et al. (1997) recommended their methods to be used for $0.05 \leq a < b \leq 0.95$. Furthermore, we include time interval restrictions as part of the simulation scenarios to evaluate the regions where each method is noteworthy in producing narrower bands. We shall expand the restriction to $(0.02, 0.98)$ in the data application in Sect. 6 to demonstrate the adequacy of the selected methods when approaching the boundaries of the entire time interval. Finally, nominal confidence level is set at 95% for all scenarios.

Each replicate simulation produces estimated bands $\{\widehat{\mathcal{L}}_r(t), \widehat{\mathcal{U}}_r(t)\}$ for $r = 1, \dots, R$. We then compute the empirical coverage (EC) and average area between bands (AABB) as

$$EC = R^{-1} \sum_{r=1}^R \mathbb{I}(\widehat{\mathcal{L}}_r(t) \leq S(t) \leq \widehat{\mathcal{U}}_r(t) \quad \forall t \in \mathcal{T}_{a,b}) \text{ and}$$

$$AABB = R^{-1} \sum_{r=1}^R \int_{\mathcal{T}_{a,b}} \{\widehat{\mathcal{U}}_r(t) - \widehat{\mathcal{L}}_r(t)\} dt, \text{ respectively,}$$

where $S(t) = e^{-t}$ is the survival function for T_i . Note that $S(t)$ is decreasing and each $\{\widehat{\mathcal{L}}_r(t), \widehat{\mathcal{U}}_r(t)\}$ is a step function. Therefore, for each constant step $[t_i, t_{i+1}]$, EC is practically computed by checking $S(t_i) \leq \widehat{\mathcal{U}}_r(t_i)$ and $\widehat{\mathcal{L}}_r(t_{i+1}-) \leq S(t_{i+1})$, where $f(t-) = \lim_{x \nearrow t} f(x)$. These metrics for each of the five methods are reported in Table 1. Because HW tends to produce the largest AABB among the five methods,

all areas are normalized by the HW areas under the same scenario (e.g., EP AABB divided by HW AABB).

As exhibited in Table 1, and also originally noted by Nair (1984) and Hollander et al. (1997), the HW bands tend to be more conservative and have a larger area between bands than the other four methods do. For 0% censoring, OptBand, FD-I, EL, and EP attain nominal confidence level and are highly competitive with each other in terms of area between bands, with OptBand outperforming in terms of AABB in most restriction and sample size combinations under no censoring. In general, we see that EP bands are less competitive when the restriction is of the form $\mathcal{T}_{0.05,0.80}$ or $\mathcal{T}_{0.20,0.80}$. This is a well-known phenomenon, as EP bands are noted to have narrower tails than HW bands, but the advantage is diminished as bands are constructed on intervals bounded away from the tails. At 20% censoring, OptBand outperforms EL bands in most scenarios and at 50% censoring, OptBand generally outperforms whenever the restriction does not include the tails (e.g., $\mathcal{T}_{0.05,0.80}$ or $\mathcal{T}_{0.20,0.80}$) or under larger sample sizes. In general, EL bands outperform OptBand for smaller sample sizes, higher censoring, and restrictions including more of the tail ends of the survival curves. For the 20% and 50% censoring scenarios, OptBand and FD-I bands consistently produce at or nearly at nominal confidence level for all settings, with FD-I bands generally having less area between bands in smaller sample sizes and OptBand having less area in larger sample sizes. At 90% censoring, FD-I bands are conservative, while OptBand maintains nominal confidence level and produces less area between bands than FD-I does except in small sample sizes ($n = 100$) and with the restrictions of the form $\mathcal{T}_{0.05,0.80}$, $\mathcal{T}_{0.05,0.95}$. Based on these observations, when censoring is very high and sample size is small, we recommend the use of FD-I. For larger sample sizes, we recommend the use of OptBand.

6 Primary biliary cirrhosis data analysis

The Mayo Clinic trial in primary biliary cirrhosis (PBC) of the liver, described in Appendix D of Fleming and Harrington (1991) and accessible in the *survival* package in R (Therneau 2015), was a double-blind, randomized, placebo-controlled study conducted between 1974 and 1984 to evaluate the efficacy of the treatment drug D-penicillamine. The primary endpoint was death, but some patients received a liver transplant. A total of 312 patients were enrolled in the randomized trial, with 158 given placebo and 154 given treatment. This dataset had an additional 112 non-randomized patients who consented to have basic measurements recorded and to be followed for survival; we excluded these patients in our application. The mean follow-up time for the randomized subjects was 5.5 years with approximately 50% censoring across groups. The original analysis found no benefit of D-penicillamine over placebo. For illustrative purposes, we fit survival curves and confidence bands for the D-penicillamine group, the placebo group, and over both groups pooled together and consider a liver transplant to be a censoring event.

We construct OptBand, FD-I, EL, EP, and HW bands for time to endpoint on the D-penicillamine, placebo, and aggregated groups on the restricted interval $\mathcal{T}_{0.02,0.98}$; as stated under Sect. 5, this is the largest interval for which critical values for EL and

Table 1 Simulation results among the five methods comparing empirical coverage levels and average area between bands

p	r	n	(Empirical coverage, average area between bands)				
			OptBand	FD-I	EL	EP	HW
0	(.05, .95)	100	(.940, .854)	(.953, .886)	(.970, .882)	(.948, .895)	(.974, 1)
		500	(.942, .850)	(.949, .969)	(.967, .884)	(.955, .885)	(.959, 1)
		1000	(.942, .849)	(.940, .979)	(.964, .882)	(.952, .884)	(.952, 1)
	(.05, .8)	100	(.952, .961)	(.946, .927)	(.963, .989)	(.960, 1.01)	(.961, 1)
		500	(.952, .964)	(.949, .971)	(.947, 1.01)	(.958, 1.01)	(.958, 1)
		1000	(.949, .964)	(.949, .981)	(.948, 1.01)	(.949, 1.01)	(.958, 1)
	(.2, .95)	100	(.943, .853)	(.953, .889)	(.971, .863)	(.934, .876)	(.970, 1)
		500	(.950, .847)	(.952, .969)	(.965, .865)	(.963, .866)	(.961, 1)
		1000	(.952, .844)	(.957, .979)	(.967, .866)	(.961, .864)	(.966, 1)
	(.2, .8)	100	(.961, .973)	(.942, .935)	(.955, .976)	(.963, .996)	(.962, 1)
		500	(.940, .973)	(.945, .973)	(.955, .990)	(.949, .995)	(.953, 1)
		1000	(.948, .972)	(.946, .981)	(.958, .991)	(.951, .994)	(.952, 1)
.2	(.05, .95)	100	(.942, .856)	(.959, .824)	(.973, .864)	(.950, .888)	(.972, 1)
		500	(.945, .839)	(.947, .869)	(.959, .866)	(.944, .869)	(.955, 1)
		1000	(.950, .838)	(.954, .877)	(.956, .864)	(.964, .867)	(.962, 1)
	(.05, .8)	100	(.953, .964)	(.938, .914)	(.961, .986)	(.954, 1.01)	(.965, 1)
		500	(.962, .965)	(.958, .967)	(.964, 1.00)	(.962, 1.01)	(.966, 1)
		1000	(.952, .966)	(.952, .973)	(.965, 1.01)	(.960, 1.01)	(.960, 1)
	(.2, .95)	100	(.942, .855)	(.936, .819)	(.963, .850)	(.949, .868)	(.955, 1)
		500	(.944, .834)	(.955, .859)	(.947, .845)	(.957, .848)	(.965, 1)
		1000	(.950, .832)	(.941, .866)	(.951, .849)	(.950, .846)	(.950, 1)
	(.2, .8)	100	(.962, .978)	(.944, .918)	(.961, .973)	(.967, .996)	(.972, 1)
		500	(.944, .975)	(.938, .957)	(.959, .988)	(.942, .994)	(.945, 1)
		1000	(.951, .975)	(.954, .964)	(.954, .989)	(.957, .993)	(.956, 1)
.5	(.05, .95)	100	(.953, .896)	(.959, .800)	(.974, .858)	(.950, .902)	(.961, 1)
		500	(.949, .830)	(.950, .835)	(.954, .842)	(.950, .850)	(.958, 1)
		1000	(.956, .819)	(.951, .822)	(.959, .837)	(.951, .840)	(.947, 1)
	(.05, .8)	100	(.958, .976)	(.961, .932)	(.967, .979)	(.948, 1.01)	(.962, 1)
		500	(.951, .970)	(.944, .995)	(.956, .998)	(.958, 1.00)	(.956, 1)
		1000	(.960, .970)	(.953, 1.01)	(.958, 1.00)	(.949, 1.00)	(.959, 1)
	(.2, .95)	100	(.957, .896)	(.963, .784)	(.963, .847)	(.954, .881)	(.957, 1)
		500	(.954, .823)	(.949, .787)	(.956, .820)	(.964, .826)	(.968, 1)
		1000	(.954, .814)	(.946, .794)	(.953, .815)	(.952, .819)	(.952, 1)
	(.2, .8)	100	(.959, .995)	(.950, .927)	(.959, .972)	(.964, 1.00)	(.966, 1)
		500	(.957, .982)	(.949, .969)	(.947, .986)	(.954, .992)	(.952, 1)
		1000	(.957, .981)	(.954, .981)	(.951, .988)	(.958, .991)	(.961, 1)

Table 1 continued

p	r	n	(Empirical coverage, average area between bands)				
			OptBand	FD-I	EL	EP	HW
.9	(.05, .95)	100	(.908, .911)	(.983, 1.15)	(.955, .928)	(.860, .928)	(.940, 1)
		500	(.938, 1.03)	(.974, 1.49)	(.955, .967)	(.902, .955)	(.889, 1)
		1000	(.936, .988)	(.984, 1.57)	(.941, .939)	(.903, .923)	(.888, 1)
	(.05, .8)	100	(.908, .908)	(.988, 1.17)	(.952, .949)	(.859, .924)	(.948, 1)
		500	(.935, 1.04)	(.983, 1.39)	(.942, 1.01)	(.887, .995)	(.891, 1)
		1000	(.934, 1.01)	(.980, 1.36)	(.959, 1.01)	(.906, .994)	(.908, 1)
	(.2, .95)	100	(.954, .998)	(.982, 1.17)	(.957, .946)	(.925, .972)	(.960, 1)
		500	(.946, 1.04)	(.979, 1.42)	(.954, .967)	(.912, .945)	(.900, 1)
		1000	(.944, .991)	(.988, 1.46)	(.945, .928)	(.906, .909)	(.894, 1)
	(.2, .8)	100	(.944, 1.01)	(.993, 1.21)	(.955, .981)	(.919, .978)	(.950, 1)
		500	(.953, 1.08)	(.995, 1.34)	(.946, 1.04)	(.926, 1.02)	(.926, 1)
		1000	(.943, 1.04)	(.983, 1.25)	(.951, 1.02)	(.915, 1.01)	(.906, 1)

The average area between bands are all normalized by the HW areas. Here, p = censoring proportion, r = restriction interval T_r , n = sample size

EP are available. Figure 1a–f compare OptBand against each of FD-I, EL, EP, and HW bands for the survival function in the placebo group, D-penicillamine group, and the two groups pooled together, respectively. Table 2 displays the normalized area between bands for the aforementioned scenarios. We observe that EL provides the least area between bands for placebo and D-penicillamine groups separately, while OptBand provides the least area for the pooled group. This is completely in line with our simulation conclusions, as EL outperforms when censoring is high ($\sim 50\%$ or higher) and sample size is small (~ 100), which match the characteristics of the placebo or D-penicillamine groups separately. Combining the two groups, the sample size becomes sufficiently large such that optimality properties of OptBand become more pronounced.

We construct OptBand for the difference in cumulative hazard functions, as seen in Fig. 2, where Fig. 2a and b depict the 95% CBs for the placebo and treatment groups, respectively, while Fig. 2c depicts the difference between cumulative hazard functions of the two groups. The bands suggest a lack of evidence for a significant difference in cumulative hazard functions between the two groups at the 0.05 significance level.

7 Discussion

In this work, we present analytical CBs for the one-sample cumulative hazard function, two-sample difference in cumulative hazard functions, and one-sample survival functions which approximately minimize the area between bands and hence approximates the HCDR. Classical bands such as HW and EP require only a table of critical values and hence are fast to compute, albeit the larger area between bands. EL and fiducial methods, while possibly providing lower area, can be computationally intensive.

Fig. 1 Survival plots for placebo group (a, b), D-penicillamine group (c, d), and pooling both groups together (e, f). 95% OptBand, Fiducial (FD-I), Empirical Likelihood (EL), Equal Precision (EP), and Hall-Wellner bands plotted over all combinations. OptBand is present on all plots for easier comparison

Table 2 Area between bands, normalized by Hall-Wellner area, for each band type and each grouping of the PBC data

Group	p	n	Average area between bands				
			OptBand	FD-I	EL	EP	HW
Placebo	.535	158	.980	1.035	.964	.981	1
D-penicillamine	.563	154	.987	1.044	.971	.989	1
Overall	.558	312	.957	1.030	.960	.970	1

Here, p = censoring proportion, n = sample size

OptBand strikes a delicate balance between computational simplicity and roughly targeting the HCDR. Intuitively, one would expect a CB to take the form of an estimator $\hat{\theta}(t)$ plus or minus some variation $\alpha(t) \cdot \text{se}(\hat{\theta}(t))$. OptBand takes exactly this form, with $\alpha(\cdot)$ encompassing the $\psi(\cdot)$ function within. Hence, the $\psi(\cdot)$ function can be viewed as the appropriate weighting function that shapes the bands to have approximately minimal area.

Unfortunately, designing OptBand for the risk difference $S_2(t) - S_1(t)$ or ratio $S_2(t)/S_1(t)$, at least with our current framework derived from Kendall et al. (2007), remains an intractable problem. The difference

$$[S_2(t) - S_1(t)] - [\hat{S}_2(t) - \hat{S}_1(t)]$$

can be shown to converge weakly to $S_1(t)W(\sigma_1^2(t)) + S_2(t)W(\sigma_2^2(t))$, which cannot be further manipulated into a form befitting derivation of OptBand; HW and EP bands face a similar intractability for the difference of survival curves. Bands for the ratio face a different intractability. Since $H_1(t) - H_2(t) = \log(S_2(t)/S_1(t))$, we have that

$$\frac{\hat{S}_2(t)/\hat{S}_1(t) - S_2(t)/S_1(t)}{S_2(t)/S_1(t)}$$

weakly converges to $W(\sigma_1^2(t) + \sigma_2^2(t))$, and hence at first glance OptBand for the ratio could be derived in a similar manner to Optband for $S(t)$ in the one-sample problem. However, in the derivation for $S(t)$, we rely on the non-increasing monotonicity of $S(t)$ (see Appendix B for how this fact is utilized), but $S_2(t)/S_1(t)$ is no longer monotone. We hope to derive bands for the ratio, and in general, functions of $S_1(t)$ and $S_2(t)$, in future works.

Further work expanding OptBand to conditionally independent censoring scenarios would be interesting. Given treatment indicator $A \in \{0, 1\}$ and baseline covariates such that T and C are independent given A and X , one could fit a Cox model

$$h(t|A, X) = h_A(t)e^{\beta^T X}$$

where the proportional hazards assumption is imposed on X , but the baseline hazards do not need to be proportional between treatments. In testing the null hypothesis $H_0(t) = H_1(t)$ for all $t \in [t_L, t_U]$ vs the alternative $H_0(t) \neq H_1(t)$ for some $t \in$

Fig. 2 OptBand applied to the placebo (a) and then the D-penicillamine (b) group of the PBC trial data to compute 95% CBs for the cumulative hazard functions, as well as for the difference in cumulative hazard functions between the two groups (c)

$[t_L, t_U]$, OptBand can be extended to the Breslow estimators for $H_0(t)$ and $H_1(t)$. Using the fact that the Breslow estimators are asymptotically Brownian, we may compute confidence intervals for $H_1(t) - H_0(t)$ in the exact same manner as seen in Sect. 3.2.

These methods have been published for use in the R package *optband* on CRAN. We recommend the use of OptBand with some minor restrictions of the event times (e.g., $T_{0.05,0.95}$). Our experiences with OptBand show that it is quite robust to censoring, even at 90% censoring.

Acknowledgements This work was supported by the National Institutes of Health under Award Numbers T32 ES007142 and T32 CA009337. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

A Minimal-area bands from Kendall et al. (2007)

Below are the results specified by Kendall et al. (2007) with a bit more generality. Suppose we intend to find a continuous function u on $[t_L, t_U]$ that is a solution to the following optimization problem:

$$\min_u \left(\int_{t_L}^{t_U} u(t) dt \right) \quad \text{s.t.} \quad \mathbb{P}(-u(t) \leq W(\sigma^2(t)) \leq u(t) \quad \forall t \in [t_L, t_U]) = \gamma,$$

where $W(\sigma^2(t))$ is a zero-mean Weiner process with strictly increasing variance function $\sigma^2(t)$. We call γ the *coverage level*, and it is more commonly thought of as $\gamma = 1 - \alpha$, where α is the *significance level*. The objective to be minimized is clearly one-half the area between the bands, which is exactly the objective of finding the HCDR. We may consider the dual problem

$$\min_u 1 - \mathbb{P}(-u(t) \leq W(\sigma^2(t)) \leq u(t) \quad \forall t \in [t_L, t_U]) \quad \text{s.t.} \quad \left(\int_{t_L}^{t_U} u(t) dt \right) = \beta,$$

where β is a constant related to γ . Due to the intractability of the objective function, the key observation Kendall et al. noted was to consider a related problem based on local times:

$$\min_u \mathbb{E}[L^u[t_L, t_U] + L^{-u}[t_L, t_U]] \quad \text{s.t.} \quad \left(\int_{t_L}^{t_U} u(t) dt \right) = \beta,$$

where $L^u[t_L, t_U]$ is the local time accumulated by $W(\sigma^2(t))$ along curve u on the interval $[t_L, t_U]$. We can calculate

$$\begin{aligned} \mathbb{E}L^u[t_L, t_U] &= \mathbb{E} \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{t_L}^{t_U} \mathbb{I}\{u(t) \leq W(\sigma^2(t)) < u(t) + \epsilon\} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{t_L}^{t_U} \exp\left(-\frac{u(t)^2}{2\sigma^2(t)}\right) \frac{1}{\sigma(t)} dt \end{aligned}$$

and similarly for $\mathbb{E}L^{-u}[t_L, t_U]$. So we ultimately have

$$\min_u \frac{2}{\sqrt{2\pi}} \int_{t_L}^{t_U} \exp\left(-\frac{u(t)^2}{2\sigma^2(t)}\right) \frac{1}{\sqrt{\sigma^2(t)}} dt \quad \text{s.t.} \quad \left(\int_{t_L}^{t_U} u(t) dt\right) = \beta,$$

which can be solved through the Euler-Lagrange equation to obtain

$$u_{\kappa}^*(t) = \psi(\kappa\sigma^2(t))\sigma(t).$$

For all combinations of (κ, L) from $\kappa = 0.01, 0.02, \dots, 0.25$ and $L = 0.005, 0.010, \dots, 0.995$, we estimated $\gamma = 1 - \alpha$ with

$$\hat{\gamma}_{\kappa, L} = N^{-1} \sum_{n=1}^N \mathbb{I}\left(|\hat{W}_{\tau, n}(s)| \leq \psi(\kappa s)\sqrt{s} \quad \forall s = \left\{L, L + \frac{1}{\tau}, \dots, 1 - \frac{1}{\tau}, 1\right\}\right)$$

where $\hat{W}_{\tau, n}(t) = \frac{1}{\sqrt{\tau}} \sum_{k=1}^{\tau} \sqrt{t} Z_{k, n}$, with $Z_{k, n}$ being i.i.d. standard normal, and $t = \frac{1}{\tau}, \dots, 1$. This provides an approximation to the standard Brownian motion on the unit interval, with the step-size parameter τ controlling how granular the approximated process is and N being the number of Monte Carlo simulations. In our simulations, $N = 4 \times 10^5$ and $\tau = 10^5$. Note that this choice of τ partitions the interval $[L, 1]$ into intervals of equal size, given the sequence for L . We further justify our functional approximation with a heatmap in Fig. 3. The borders between the colors on the heatmap are smooth and closely match up with the γ contours defined from our functional approximation $\hat{\gamma}_{\kappa, L} = 1 + (a + bL)\kappa + a\kappa^2$ in Eq. (4) proposed in the main text. We chose this specific functional approximation based on empirical observations. We first noticed by setting $L = 0$, the coefficients of κ and κ^2 were nearly identical in the quadratic fit of κ and $\hat{\gamma}_{\kappa, 0}$. Next, we noticed that the coefficient of κ increased linearly and the coefficient of κ^2 remained nearly constant with respect to L , which is exactly described in the functional approximation above. Of course, one could include more parameters for marginal increases in accuracy, but in favor of parsimony, we selected this relatively simple two-parameter approximation.

B derivation of OptBand for the survival function

In order to exactly target the bands around $\hat{S}(t) - S(t)$, we may consider the optimization problem

$$\min_u \frac{2}{\sqrt{2\pi}} \int_{t_L}^{t_U} \exp\left(-\frac{u(t)^2}{2\sigma^2(t)}\right) \frac{1}{\sigma(t)} dt \quad \text{s.t.} \quad \left(\int_{t_L}^{t_U} S(t)u(t) dt\right) = \beta$$

which, through similar steps taken in Appendix A, gives us the solution

$$u^*(t) = \psi(\kappa S(t)\sigma^2(t))\sigma(t).$$

Fig. 3 Simulated coverage in terms of κ (vertical axis) and L (horizontal axis). This smooth relationship (a) is closely matched by a functional approximation (b)

Again, making the transformation $\sigma^2(t) \mapsto \sigma^2(t)/\sigma^2(t_U)$, we find κ so that the event

$$|W(s)| \leq \psi(\kappa S(t)s) \sqrt{s} \quad \forall s \in \left[\frac{\sigma^2(t_L)}{\sigma^2(t_U)}, 1 \right]$$

occurs with probability γ , where $t = [\sigma^2]^{-1} \{ \sigma^2(t_U)s \}$ and $[\sigma^2]^{-1}$ denotes the inverse variance function. We now face a problem where our κ is also a function of $S(t)$. We propose the computation scheme. Denote

$$\mathcal{E}(\eta(\cdot), \tau, \nu) = \left\{ \left| W \left(\frac{\sigma^2(t)}{\sigma^2(t_U)} \right) \right| \leq \psi \left(\eta(t) \frac{\sigma^2(t)}{\sigma^2(t_U)} \right) \frac{\sigma(t)}{\sigma(t_U)} \quad \forall t \in [\tau, \nu] \right\}.$$

Now if $\eta(\cdot) = \eta$ is a constant, we can approximate this quantity with Eq. (4):

$$\begin{aligned} \mathbb{P}(\mathcal{E}(\eta, \tau, \nu)) &= \mathbb{P} \left(\left| W \left(\frac{\sigma^2(t)}{\sigma^2(t_U)} \right) \right| \leq \psi \left(\eta \frac{\sigma^2(t)}{\sigma^2(t_U)} \right) \frac{\sigma(t)}{\sigma(t_U)} \quad \forall t \in [\tau, \nu] \right) \\ &\approx 1 + \left(a + b \frac{\sigma^2(\tau)}{\sigma^2(t_U)} \right) \left(\eta \frac{\sigma^2(t_U)}{\sigma^2(t_U)} \right) + a \left(\eta \frac{\sigma^2(t_U)}{\sigma^2(t_U)} \right)^2. \end{aligned}$$

In what follows, we will need the fact

$$\eta_1 \leq \eta_2 \implies \mathcal{E}(\eta_1, \tau, \nu) \supseteq \mathcal{E}(\eta_2, \tau, \nu) \quad (6)$$

which follows, since a larger η corresponds to tighter bands.

Let $t_L = \xi_1 \leq \xi_2 \leq \dots \leq \xi_K = t_U$ be some partition of the interval $[t_L, t_U]$, with the most natural choice being the observed event time points $\{t_i : t_L \leq t_i \leq t_U\}$.

Let $\bar{S}_i = (\widehat{S}(\xi_i) + \widehat{S}(\xi_{i+1}))/2$. The construction of OptBand for survival functions is based on two approximations: (1) approximate $S(t)$ for $t \in [\xi_i, \xi_{i+1}]$ with a midpoint Riemann sum $\bar{S}_i$, (2) approximate with Eq. (4) for each locally constant portion of $\bar{S}_i$. More specifically, the first approximation is

$$\gamma = \mathbb{P} \left(\bigcap_{i=1}^{K-1} \mathcal{E}(\kappa S(\cdot), \xi_i, \xi_{i+1}) \right) \approx \mathbb{P} \left(\bigcap_{i=1}^{K-1} \mathcal{E}(\kappa \bar{S}_i, \xi_i, \xi_{i+1}) \right) \quad (7)$$

with the approximation increasingly accurate as $K \rightarrow \infty$. In practice, there are no advantages in choosing $\{\xi_1, \dots, \xi_K\}$ to be anything else but the observed event times because only at these values does the estimated survival function $\widehat{S}(\cdot)$ change, and therefore the Riemann sum cannot become more accurate than that.

The interpretation of $\bigcap_{i=1}^{K-1} \mathcal{E}(\kappa \bar{S}_i, \xi_i, \xi_{i+1})$ is the event that our Brownian motion falls within our bands on interval $[\xi_i, \xi_{i+1}]$ for each i . Equivalently, since $\kappa \bar{S}_i$ is non-increasing, hence $\mathcal{E}(\kappa \bar{S}_i, \xi_{i+1}, \xi_K)$ are tighter than $\mathcal{E}(\kappa \bar{S}_{i+1}, \xi_{i+1}, \xi_K)$ by Eq. (6), we can imagine first having to fall within $\mathcal{E}(\kappa \bar{S}_1, \xi_1, \xi_K)$, then next having to fall within $\mathcal{E}(\kappa \bar{S}_2, \xi_2, \xi_K)$, etc, leading to the equivalence

$$\bigcap_{i=1}^{K-1} \mathcal{E}(\kappa \bar{S}_i, \xi_i, \xi_{i+1}) = \bigcup_{i=1}^{K-1} \mathcal{E}(\kappa \bar{S}_i, \xi_i, \xi_K).$$

Continuing Eq. (7) and defining $\mathcal{E}(\kappa \bar{S}_0, \xi_0, \xi_K) = \emptyset$, we have

$$\begin{aligned} \gamma &\approx \mathbb{P} \left(\bigcup_{i=1}^{K-1} \mathcal{E}(\kappa \bar{S}_i, \xi_i, \xi_K) \right) \\ &= \mathbb{P} \left(\bigcup_{i=1}^{K-1} [\mathcal{E}(\kappa \bar{S}_i, \xi_i, \xi_K) \setminus \mathcal{E}(\kappa \bar{S}_{i-1}, \xi_i, \xi_K)] \right) \\ &= \sum_{i=1}^{K-1} \{ \mathbb{P}(\mathcal{E}(\kappa \bar{S}_i, \xi_i, \xi_K)) - \mathbb{P}(\mathcal{E}(\kappa \bar{S}_{i-1}, \xi_i, \xi_K)) \} \\ &\approx \frac{\kappa b}{\sigma^2(t_U)} \left\{ \sum_{i=1}^{K-2} \bar{S}_i (\sigma^2(\xi_i) - \sigma^2(\xi_{i+1})) \right\} + 1 + \left(a + b \frac{\sigma^2(\xi_{K-1})}{\sigma^2(t_U)} \right) (\kappa \bar{S}_{K-1}) \\ &\quad + a (\kappa \bar{S}_{K-1})^2 \end{aligned}$$

where the last $\approx$ comes from continued application of Eq. (4). This is a quadratic function in κ , and hence we can solve for κ analytically. We also estimate $\sigma^2(t)$ with $\widehat{\sigma}^2(t)$, as before. Finally, we pivot to get a 100 $\gamma\%$ CI:

$$\begin{aligned} \gamma &\approx \mathbb{P} \left(\left| n^{1/2} \frac{\widehat{S}(t) - S(t)}{\widehat{S}(t) \widehat{\sigma}(t_U)} \right| \leq \psi \left(\kappa \widehat{S}(t) \frac{\widehat{\sigma}^2(t)}{\widehat{\sigma}^2(t_U)} \right) \frac{\widehat{\sigma}(t)}{\widehat{\sigma}(t_U)} \quad \forall t \in [t_L, t_U] \right) \\ &= \mathbb{P} \left(\widehat{S}(t) (1 - c_S(t)) \leq S(t) \leq \widehat{S}(t) (1 + c_S(t)) \quad \forall t \in [t_L, t_U] \right), \end{aligned}$$

where $c_S(t) = \psi \left(\kappa \widehat{S}(t) \frac{\widehat{\sigma}^2(t)}{\widehat{\sigma}^2(t_U)} \right) \frac{\widehat{\sigma}^2(t)}{n^{1/2}}$.

References

- Cui Y, Hannig J (2019a) Nonparametric generalized fiducial inference for survival functions under censoring. *Biometrika* 106(3):501–518
- Cui Y, Hannig J (2019b) Rejoinder: nonparametric generalized fiducial inference for survival functions under censoring. *Biometrika* 106(3):527–531
- Emerson S (2009) Small sample performance and calibration of the Empirical Likelihood method. PhD thesis, Stanford University
- Fleming TR, Harrington DP (1991) Counting processes and survival analysis. Wiley, New York
- Hall WJ, Wellner JA (1984) Confidence bands for a survival curve from censored data. *Biometrika* 67(1):133
- Hollander M, McKeague IW, Yang J (1997) Likelihood ratio-based confidence bands for survival functions. *J Am Stat Assoc* 92(437):215–226
- Kendall WS, Marin JM, Robert CP (2007) Confidence bands for Brownian motion and applications to Monte Carlo simulation. *Stat Comput* 17(1):1–10
- Klein JP, Moeschberger ML (2006) Survival analysis: techniques for censored and truncated data. Springer, Berlin
- Martin R (2019) Discussion: nonparametric generalized fiducial inference for survival functions under censoring. *Biometrika* 106(3):519–522
- McKeague IW, Zhao Y (2002) Simultaneous confidence bands for ratios of survival functions via empirical likelihood. *Stat Prob Lett* 60:405–415
- McKeague IW, Zhao Y (2005) Comparing distribution functions via empirical likelihood. *Int J Biostat* 1(1):5
- Li G, Hollander M, McKeague IW, Yang J (1996) Nonparametric likelihood ratio confidence bands for quantile functions from incomplete survival data. *Ann Stat* 24(2):628–640
- Nair VN (1984) Confidence bands for survival functions with censored data: a comparative study. *Technometrics* 26(3):265–275
- Parzen MI, Wei LJ, Ying Z (1994) A resampling method based on pivotal estimating functions. *Biometrika* 81(2):341–350
- Parzen MI, Wei LJ, Ying Z (1997) Simultaneous confidence intervals for the difference of two survival functions. *Scand J Stat* 24(3):309–314
- Therneau T (2015) The survival package (version 2.38). The Comprehensive R Archive Network Retrieved from <https://cran.r-project.org/web/packages/survival/>
- Thomas David R, Grunkemeier Gary L (1975) Confidence interval estimation of survival probabilities for censored data. *J Am Stat Assoc* 70(352):865–871
- Tian L, Wang R, Cai T, Wei LJ (2011) The highest confidence density region and its usage for joint inferences about constrained parameters. *Biometrics* 67(2):604–610

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.