Ann Inst Stat Math (2019) 71:213–246
https://doi.org/10.1007/s10463-017-0634-9

Bootstrapping the Kaplan–Meier estimator on the whole line

Dennis Dobler[1]

Received: 27 June 2016 / Revised: 18 October 2017 / Published online: 13 January 2018

Abstract This article is concerned with proving the consistency of Efron's bootstrap for the Kaplan–Meier estimator on the whole support of a survival function. While previous works address the asymptotic Gaussianity of the Kaplan–Meier estimator without restricting time, we enable the construction of bootstrap-based time-simultaneous confidence bands for the whole survival function. Other practical applications include bootstrap-based confidence bands for the mean residual lifetime function or the Lorenz curve as well as confidence intervals for the Gini index. Theoretical results are complemented with a simulation study and a real data example which result in statistical recommendations.

Keywords Counting process · Right censoring · Resampling · Efron's bootstrap · Mean residual lifetime · Lorenz curve · Gini index

1 Introduction

This article reconsiders Efron's classical bootstrap of Kaplan–Meier estimators; cf. Efron (1981). It is well known that drawing with replacement directly from the original observations consisting of (event time, censoring indicator) reproduces the correct covariance structure; see, for example, Akritas (1986), Lo and Singh (1986), Horvath and Yandell (1987) or van der Vaart and Wellner (1996) for an application in empirical processes. Let $T : \Omega \to (0, \tau)$ be a continuously distributed random survival time with survival function given by $S(t) = 1 - F(t) = P(T > t)$. For conceptual convenience we mainly refer to T as a random *survival time*, although other interpretations are

✉ Dennis Dobler
dennis.dobler@uni-ulm.de

1 Institute of Statistics, Ulm University, Helmholtzstr. 20, 89081 Ulm, Germany

also reasonable; see the examples below. In the previously mentioned articles the typical assumption $S(\tau) > 0$ is met for mathematical convenience in proving weak convergence of estimators of S on the Skorohod space $D[0, \tau]$ and because most studies involve a rather strict censoring mechanism: after a pre-specified end of study time each individual without an observed event is considered as right-censored. Thus, it is often not possible to draw inference on functionals of the whole survival function.

Some functionals, however, indeed require the possibility to observe arbitrarily large survival times. For instance, consider the *mean residual lifetime* function

$$t \longmapsto g(t) = \mathbb{E}[T - t \mid T > t] = \frac{1}{S(t)} \int_t^\tau S(u)\mathrm{d}u; \tag{1}$$

see, for example, Meilijson (1972), Remark 3.3 in Gill (1983), and Stute and Wang (1993). This function describes the expected remaining lifetime given the survival until a point of time $t > 0$. Another, econometric example of a functional of whole survival curves is the *Lorenz curve*

$$p \longmapsto L(p) = \mu^{-1} \int_0^p F^{-1}(t)\mathrm{d}t = \mu^{-1} \int_0^{F^{-1}(p)} s\mathrm{d}F(s), \tag{2}$$

where $F^{-1}(t) = \inf\{u \geqslant 0 : F(u) \geqslant t\}$ is the left-continuous generalized inverse of F and $\mu = \int_0^\tau t\mathrm{d}F(t)$ is its mean. With the interpretation of T being the income of a random individual in a population, this function L obviously represents the total income of the lowest pth fraction of all incomes; cf. Gastwirth (1971). A closely related quantity is the *Gini index*

$$\mathcal{G} = \frac{\int_0^1 (u - L(u))\mathrm{d}u}{\int_0^1 u\mathrm{d}u} \in [0, 1] \tag{3}$$

as a measure of uniformity of all incomes within a population; see, for example, Tse (2006). The value $\mathcal{G} = 0$ represents perfect equality of all incomes, whereas $\mathcal{G} = 1$ describes the other extreme: only one person gains everything and the rest nothing.

All quantities (1), (2) and (3) are statistical functionals of the *whole* survival function S. First analyzing S only on a subset of its support results inevitably in an alternation of the above functionals in a second step. And this affects the interpretation of the above quantities. In order to circumvent such problems, estimating the whole survival function is the obvious solution: Henceforth, denote by $\tau = \inf\{t \geqslant 0 : S(t) = 0\} \in (0, \infty)$ the support's right end point. Wang (1987) and Stute and Wang (1993) showed the uniform consistency of the *Kaplan–Meier* or *product-limit estimator* $(\widehat{S}(t))_{t\in[0,\tau]}$ for $(S(t))_{t\in[0,\tau]}$, and Gill (1983) and Ying (1989) proved its weak convergence on the Skorohod space $D[0, \tau]$. For robust statistical inference procedures concerning the above functionals of S it is thus necessary to extend well-known bootstrap results for the Kaplan–Meier estimator to the whole Skorohod space $D[0, \tau]$. After presenting this primary result we deduce inference procedures for the quantities (1)–(3).

This article is organized as follows. Section 2 introduces all required estimators, recapitulates previous weak convergence results for the Kaplan–Meier estimator on

$D[0, \tau]$, and provides handy results for checking all main assumptions. The main theorems on weak convergence of the bootstrap Kaplan–Meier estimator are presented in Sect. 3, including a consistency theorem for a bootstrap variance function estimator. Section 4 deduces inference procedures for (1)–(3). The coverage probabilities of various confidence bands for the whole survival function in case of small- to medium-sized samples are assessed in an extensive simulation study in Sect. 5. These are complemented with simulations on the coverage of linear and log-transformed confidence bands for the mean residual lifetime function. In Sect. 6, these results are illustrated and revalued in the light of a real data-set of male lung cancer patients. The final Sect. 7 gives a discussion on future research possibilities. All proofs are given in "Appendix." Most of this article's results originate from the Ulm University PhD thesis of Dobler (2016); cf. Chapters 6 and 7 therein.

2 Preliminary results

Let $T_1, \ldots, T_n : \Omega \to (0, \infty), n \in \mathbb{N}$, be independent survival times with continuous survival functions $S(t) = 1 - F(t) = P(T_1 > t)$ and cumulative hazard function $A(t) = \int_0^t \alpha(u)\mathrm{d}u = -\int_0^t (\mathrm{d}S)/S_- = -\log S(t)$. Independent thereof, let $C_1, \ldots, C_n : \Omega \to (0, \infty)$ be i.i.d. (censoring) random variables with (possibly discontinuous) survival function $G(t) = P(C_1 > t)$ such that the observable data consist of all $1 \leqslant i \leqslant n$ pairs $(X_i, \delta_i) := (T_i \wedge C_i, \mathbf{1}\{X_i = T_i\})$. Here, $\mathbf{1}\{\cdot\}$ is the indicator function and $a \wedge b$ denotes the minimum of a and b. Thus, the survival function of X_1 is $H = S \cdot G$. The Kaplan–Meier estimator is defined by $\widehat{S}_n(t) = \prod_{i:X_{i:n} \leqslant t}(1 - \frac{\delta_{[i:n]}}{n-i+1})$, where $(X_{1:n}, \ldots, X_{n:n})$ is the order statistic of $(X_1, \ldots, X_n)$ and $(\delta_{[1:n]}, \ldots, \delta_{[n:n]})$ are their concomitant censoring indicators. Throughout, we assume that

$$-\int_0^\tau \frac{\mathrm{d}S}{G_-} < \infty \tag{4}$$

which restricts the magnitude of censoring to a reasonable level. Here and throughout, the minus subscript indicates the left-continuous version of right-continuous functions. For instance Gill (1983), Ying (1989) and Akritas and Brunner (1997) require Condition (4) for an analysis of the large sample properties of Kaplan–Meier estimators on the whole support $[0, \tau]$. Thereof, it is utilized in Gill (1983) for a vanishing upper bound in Lenglart's inequality. Obviously, the above condition implies that $[0, \tau]$ is contained in the support of G; see also Allignol et al. (2014) for a similar condition in a non-Markov illness–death model, reduced to a competing risks problem.

Denote by $\widehat{T}_n := X_{n:n}$ the largest observed event or censoring time and let, for a function $t \mapsto f(t)$, the function $f^{\widehat{T}_n}$ be its stopped version, i.e., $f^{\widehat{T}_n}(t) = f(t \wedge \widehat{T}_n)$. The monotone function $t \mapsto \sigma^2(t) = \int_0^t (\mathrm{d}A)/H_-$ is the asymptotic variance function of the related Nelson-Aalen estimator for A and reappears in the asymptotic covariance function of $\widehat{S}_n$. Throughout, all convergences (in distribution, probability, or almost surely) are understood to hold as $n \to \infty$ and convergence in distribution and convergence in probability are denoted by $\xrightarrow{d}$ and $\xrightarrow{p}$, respectively. The present theory relies on the following weak convergence results for the Kaplan–Meier process $\widehat{S}_n$ of S. More precisely, both assertions are corollaries from the cited theorems.

Lemma 1 *Let B denote a Brownian motion on $[0, \infty)$ and suppose (4) holds.*

(a) Theorem 1.2(i) of (Gill 1983): *On $D[0, \tau]$ we have* $\sqrt{n}(\widehat{S}_n - S)^{\widehat{T}_n} \xrightarrow{d} W := S \cdot (B \circ \sigma^2)$,

(b) Part of Theorem 2 in (Ying 1989): *On $D[0, \tau]$ we have* $\sqrt{n}(\widehat{S}_n - S) \xrightarrow{d} W = S \cdot (B \circ \sigma^2)$.

Denote by $\widehat{A}_n(t) = \sum_{i: X_{i:n} \leqslant t} \frac{\delta_{[i:n]}}{n-i+1}$ the Nelson-Aalen estimator for the cumulative hazard function $A(t)$ and by $\widehat{G}_n$ the Kaplan–Meier estimator for the censoring survival function G. Note that $\widehat{H}_n = \widehat{G}_n \widehat{S}_n$ holds for the empirical survival function of H since, almost surely (a.s.), no survival time equals a censoring time: $T_i \neq C_j$ a.s. for all i, j. The asymptotic covariance function Γ of W in Lemma 1 and a natural estimator $\widehat{\Gamma}_n$ are given by

$$\Gamma(u, v) = S(u) \left(\int_0^{u \wedge v} \frac{\mathrm{d}A}{H_-} \right) S(v) \quad \text{and} \quad \widehat{\Gamma}_n(u, v) = \widehat{S}_n(u) \left(\int_0^{u \wedge v} \frac{\mathrm{d}\widehat{A}_n}{\widehat{H}_{n-}} \right) \widehat{S}_n(v).$$

The following lemma is helpful for an assessment of Condition (4) and for studentizations.

Lemma 2 (a) *For all $t \in [0, \tau]$ it holds that*

$$- \int_t^\tau \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}} \xrightarrow{p} - \int_t^\tau \frac{\mathrm{d}S}{G_-} \leqslant \infty.$$

In case the right-hand side is infinite, the convergence is even almost surely.
(b) *In case of* (4) *we have*

$$\sup_{(u,v) \in [0,\tau]^2} |\widehat{\Gamma}_n(u, v) - \Gamma(u, v)| \xrightarrow{p} 0.$$

Lemma 1(a) may be applied in the following way for assessing Condition (4): If the integral on the left-hand side does not seem to tend to infinity with increasing n for a particular realization of the data, then this is an indication for the validity of Condition (4).

3 Main results

The limit distribution of the Kaplan–Meier process in Lemma 1 shall be assessed via bootstrapping. To this end, we independently draw n times with replacement from $(X_1, \delta_1), \ldots, (X_n, \delta_n)$ and denote the thus obtained bootstrap sample by

$$(X_1^*, \delta_1^*), \ldots, (X_n^*, \delta_n^*).$$

Throughout, denote by Γ_n^*, S_n^*, etc., the obvious estimators but based on the bootstrap sample. Note that this requires a discontinuous extension of the above quantities. The following theorem is the basis of all later inference methods.

Theorem 1 *Let B denote a Brownian motion on $[0, \tau]$ and suppose that (4) holds. Then, we have, conditionally on $X_1, \delta_1, X_2, \delta_2, \ldots$ and as $n \to \infty$,*

$$\sqrt{n}(S_n^* - \widehat{S}_n) \xrightarrow{d} W = S \cdot (B \circ \sigma^2)$$

on $D[0, \tau]$ in probability.

Many statistical applications involve a consistent variance estimator, e.g., Hall–Wellner or equal precision confidence bands for S; cf. Andersen et al. (1993), p. 266. In order to asymptotically reproduce the same limit on the bootstrap side, the uniform consistency of a bootstrapped variance estimator (defined on the whole support $[0, \tau]^2$ of the covariance function) needs to be verified. To this end, introduce the bootstrap version of $\widehat{\Gamma}_n$, that is,

$$\Gamma_n^*(u, v) = S_n^*(u) \left(\int_0^{u \wedge v} \frac{\mathrm{d}A_n^*}{H_{n-}^*} \right) S_n^*(v).$$

For all $\varepsilon > 0$, its uniform consistency (here and below always meaning conditional convergence in probability given $X_1, \delta_1, X_2, \delta_2, \ldots$ in probability) over all points $(u, v) \in [0, \tau]^2 \setminus [\tau - \varepsilon, \tau]^2$ is an immediate consequence of Theorem 1 in combination with the continuous mapping theorem: Write the absolute value of the integral part minus its estimated counterpart as

$$\begin{aligned} &\left| \int_0^{u \wedge v} \frac{(\widehat{H}_{n-} - H_{n-}^*)\mathrm{d}A_n^* - H_{n-}^*\mathrm{d}(\widehat{A}_n - A_n^*)}{H_{n-}^* \widehat{H}_{n-}} \right| \\ &\leqslant \frac{\sup_{(0,u\wedge v)} |\widehat{H}_n - H_n^*|}{H_n^*((u \wedge v)-)\widehat{H}_n((u \wedge v)-)} A_n^*(u \wedge v) + \left| \int_0^{u \wedge v} \frac{\mathrm{d}(\widehat{A}_n - A_n^*)}{\widehat{H}_{n-}} \right|. \end{aligned}$$

The first term is asymptotically negligible due to Pòlya's theorem and the second term becomes small due to the continuous mapping theorem applied to the integral functional and the logarithm functional. Here the restriction to $[0, \tau]^2 \setminus [\tau - \varepsilon, \tau]^2$ simplified the calculations since all denominators are asymptotically bounded away from zero.

For uniform consistency on the whole rectangle $[0, \tau]^2$, however, similar arguments as for the bootstrapped Kaplan–Meier process on $[0, \tau]$ are required. Compared to (4), we postulate a slightly more restrictive censoring condition.

Lemma 3 *Suppose that*

$$-\int_0^{\tau} \frac{\mathrm{d}S}{G_-} - \int_0^{\tau} \frac{S\mathrm{d}S}{G_-^2} < \infty. \tag{5}$$

Then, we have the following conditional uniform consistency given $X_1, \delta_1, X_2, \delta_2, \ldots$ in probability:

$$\sup_{(u,v)\in[0,\tau]^2} |\Gamma_n^*(u,v) - \widehat{\Gamma}_n(u,v)| \xrightarrow{p} 0 \quad \text{in probability as } n \to \infty. \tag{6}$$

4 Applications

Apart from time-simultaneous confidence bands for the whole survival curve S with asymptotically exact coverage probability, applications of Theorem 1 concern confidence intervals for the mean residual lifetime $g(t) = \mathbb{E}[T - t \mid T > t]$ on compact subintervals $[t_1, t_2] \subset [0, \tau)$ as well as confidence regions for the Lorenz curve L and the Gini index G. To this end, we apply the functional delta method (e.g., Andersen et al. (1993), Theorem II.8.1) which in turn requires the Hadamard differentiability of all involved statistical functionals.

Confidence bands for the mean residual lifetime function

Let $0 \leqslant t_1 \leqslant t_2$ and introduce the space $C[t_1, \tau]$ of continuous functions on $[t_1, \tau]$ equipped with the supremum norm as well as the subset

$$\widetilde{C}[t_1, t_2] = \left\{ f \in C[t_1, \tau] : \inf_{s\in[t_1,t_2]} |f(s)| > 0 \right\} \subset C[t_1, \tau]$$

containing all continuous functions having a positive distance to the constant zero function on the interval $[t_1, t_2]$. Similarly, let

$$\widetilde{D}[t_1, t_2] = \left\{ f \in D[t_1, \tau] : \inf_{s\in[t_1,t_2]} |f(s)| > 0, \sup_{s\in[t_1,\tau]} |f(s)| < \infty \right\} \subset D[t_1, \tau]$$

be the extension of $\widetilde{C}[t_1, t_2]$ to possibly discontinuous, bounded càdlàg functions. For the notion of Hadamard differentiability tangentially to subsets of $D[t_1, \tau]$, see Definition II.8.2, Theorem II.8.2 and Lemma II.8.3 in Andersen et al. (1993), p. 111f. The following lemma allows the applicability of the functional delta method for the mean residual lifetime function.

Lemma 4 *Let $\tau < \infty$ and $[t_1, t_2] \subset [0, \tau)$ be a compact interval. Then,*

$$\psi : \widetilde{D}[t_1, t_2] \to D[t_1, t_2], \quad \theta(\cdot) \mapsto \frac{1}{\theta(\cdot)} \int_{\cdot}^{\tau} \theta(s)\mathrm{d}s$$

is Hadamard-differentiable at each $\theta \in \widetilde{C}[t_1, t_2]$ tangentially to $C^2[t_1, \tau]$ with continuous linear derivative $\mathrm{d}\psi(\theta) \cdot h \in D[t_1, t_2]$ given by

$$(\mathrm{d}\psi(\theta) \cdot h)(s) := \frac{1}{\theta(s)} \int_s^{\tau} h(u)\mathrm{d}u - h(s) \int_s^{\tau} \frac{\theta(u)}{\theta^2(s)}\mathrm{d}u.$$

As pointed out in Gill (1989) or Andersen et al. (1993), p. 110, the functional delta method is established on the functional space $D[t_1, \tau]$ (or subsets thereof) equipped with the supremum norm. However, in case of limiting processes with continuous sample paths, "weak convergence in the sense of the [Skorohod] metric and in the sense of the supremum norm are exactly equivalent" (Andersen et al. 1993). See also Problem 7 in Pollard (1984), p. 137. The convergence result of Theorem 1 combined with the functional ψ of Lemma 4 constitutes the following weak convergence.

Lemma 5 *Suppose that* (4) *holds. On the Skorohod space* $D[t_1, t_2]$ *we then have*

$$\sqrt{n}\left(\int_{\cdot}^{\tau} \frac{\widehat{S}_n(u)}{\widehat{S}_n(\cdot)} \mathrm{d}u - \int_{\cdot}^{\tau} \frac{S(u)}{S(\cdot)} \mathrm{d}u\right) \xrightarrow{d} U$$

and, given $X_1, \delta_1, X_2, \delta_2, \ldots,$

$$\sqrt{n}\left(\int_{\cdot}^{\tau} \frac{S_n^*(u)}{S_n^*(\cdot)} \mathrm{d}u - \int_{\cdot}^{\tau} \frac{\widehat{S}_n(u)}{\widehat{S}_n(\cdot)} \mathrm{d}u\right) \xrightarrow{d} U$$

in outer probability. The Gaussian process U *has a.s. continuous sample paths, mean zero and covariance function*

$$(r, s) \mapsto \int_{r \vee s}^{\tau} \int_{r \vee s}^{\tau} \frac{\Gamma(u, v)}{S(r)S(s)} \mathrm{d}u \mathrm{d}v - \sigma^2(r \vee s) g(r) g(s),$$

where $g(t) = \mathbb{E}[T_1 - t \mid T_1 > t] = \int_t^{\tau} \frac{S(u)}{S(t)} \mathrm{d}u$ *is again the mean residual lifetime function and* $a \vee b$ *denotes the maximum of* a *and* b.

The previous lemma in combination with the continuous mapping theorem almost immediately gives rise to the construction of asymptotically valid confidence regions for the mean residual lifetime function. According to the functional delta method we may first apply, e.g., a log-transformation to ensure that only positive values are included in the confidence regions; cf. Sect. IV.1.3 in Andersen et al. (1993), p. 208ff. For ease of presentation, only the linear regions are stated below.

Theorem 2 *Let* $0 \leqslant t_1 \leqslant t_2 < \tau$. *Choose any* $\alpha \in (0, 1)$ *and suppose that* (4) *holds. An asymptotic two-sided* $(1-\alpha)$*-confidence band for the mean residual lifetime function* $(\mathbb{E}[T_1 - t \mid T_1 > t])_{t \in [t_1, t_2]}$ *is given by*

$$\left[\int_t^{\tau} \frac{\widehat{S}_n(u)}{\widehat{S}_n(t)} \mathrm{d}u - \frac{q_{n_1,n_2}^{MRLT}}{\sqrt{n}}, \int_t^{\tau} \frac{\widehat{S}_n(u)}{\widehat{S}_n(t)} \mathrm{d}u + \frac{q_{n_1,n_2}^{MRLT}}{\sqrt{n}}\right]_{t \in [t_1, t_2]}$$

where q_{n_1,n_2}^{MRLT} *is the* $(1-\alpha)$*-quantile of the conditional law given all observations* $(X_1, \delta_1), \ldots, (X_n, \delta_n)$ *of*

$$\sqrt{n} \sup_{t \in [t_1, t_2]} \left| \int_t^{\tau} \frac{S_n^*(u)}{S_n^*(t)} \mathrm{d}u - \int_t^{\tau} \frac{\widehat{S}_n(u)}{\widehat{S}_n(t)} \mathrm{d}u \right|.$$

Remark 1 (a) Instead of using a transformation as indicated above Theorem 2, one could also employ a studentization using $\widehat{\Gamma}_n$ and Γ_n^*. Plugging these and consistent estimators for the other unknown quantities into the asymptotic variance representation yields consistent variance estimators for the statistic of interest. This yields a Gaussian process with asymptotic variance 1 at all points of time for the mean residual lifetime estimates.
(b) In practice, the construction of confidence bands for the mean residual lifetime function requires to choose t_2 depending on the data: else, too large choices of t_2 might result in $\widehat{S}_n(t_2) = 0$, in which case the above estimator would not be well defined. Following the suggestion of a referee, an asymptotically exact ad hoc solution would be replacing the estimated mean residual lifetime with 0 in this case.

Confidence regions for the Lorenz curve and the Gini index

As estimators for the Lorenz curve and the Gini index we consider the plug-in estimates

$$\widehat{L}_n(p) = \frac{1}{\widehat{\mu}_n} \int_0^p (1 - \widehat{S}_n(t))^{-1} \mathrm{d}t \quad \text{and} \quad \widehat{\mathcal{G}}_n = \frac{\int_0^1 (u - \widehat{L}_n(u)) \mathrm{d}u}{\int_0^1 u \mathrm{d}u},$$

where $\widehat{\mu}_n = \int_0^\tau s \mathrm{d}\widehat{S}_n(s)$. The restricted and unscaled Lorenz curve estimator under independent right censoring has already been bootstrapped by Horvath and Yandell (1987). Furthermore, Tse (2006) discussed the large sample properties of the above Lorenz curve estimator (even under left-truncation) and also of the normalized estimated Gini index

$$\sqrt{n}(\widehat{\mathcal{G}}_n - \mathcal{G}) = \sqrt{n} \left(\frac{\int_0^1 (u - \widehat{L}_n(u)) \mathrm{d}u}{\int_0^1 u \mathrm{d}u} - \frac{\int_0^1 (u - L(u)) \mathrm{d}u}{\int_0^1 u \mathrm{d}u} \right)$$
$$= 2\sqrt{n} \int_0^1 (L(u) - \widehat{L}_n(u)) \mathrm{d}u.$$

Again equip all subsequent function spaces with the supremum norm. Let $D_\uparrow[0, \tau] \subset D[0, \tau]$ be the set of all distribution functions on $[0, \tau]$ with no atom in 0, and let $D_-[0, \tau]$ be the set of all càglàd functions on $[0, \tau]$. First, we consider the normalized estimated Lorenz curve, i.e., the process $W_n : \Omega \to [0, 1]$ given by

$$W_n(p) = \sqrt{n} \left(\frac{1}{\widehat{\mu}_n} \int_0^p (1 - \widehat{S}_n)^{-1}(s) \mathrm{d}s - \frac{1}{\mu} \int_0^p (1 - S)^{-1}(s) \mathrm{d}s \right)$$
$$= \sqrt{n}(\widehat{\mu}_n^{-1} \cdot (\Phi \circ \Psi \circ (1 - \widehat{S}_n))(p) - \mu^{-1} \cdot (\Phi \circ \Psi \circ (1 - S))(p)).$$

Here the functionals Φ and Ψ are

$$\Phi : D_-[0, 1] \mapsto C[0, 1], \quad h \mapsto \left(p \mapsto \int_0^p h(s) \mathrm{d}s \right), \quad \text{and}$$
$$\Psi : D_\uparrow[0, \tau] \mapsto D_-[0, 1], \quad k \mapsto k^{-1} \quad \text{(the left-continuous generalized inverse).}$$

Suppose that S is continuously differentiable on its support with strictly positive derivative f, bounded away from zero. The Hadamard differentiability of Ψ at $(1-S)$ tangentially to $C[0,\tau]$ then holds according to Lemma 3.9.23 in van der Vaart and Wellner (1996), p. 386. Its derivative map is given by $\alpha \mapsto -\frac{\alpha}{f} \circ (1-S)^{-1}$. The other functional Φ is obviously Hadamard-differentiable at $S^{-1} \in C[0,1]$ tangentially to $C[0,1]$ since Φ itself is linear and the domain of integration is bounded. Next,

$$\sqrt{n}\left(\frac{1}{\widehat{\mu}} - \frac{1}{\mu}\right) = \sqrt{n}(\Upsilon(\widehat{g}_n(0)) - \Upsilon(g(0)))$$

where $\Upsilon : (0,\infty) \to (0,\infty), r \mapsto \frac{1}{r}$, $g(0) = \mathbb{E}[T - a \mid T > 0] = \mathbb{E}[T]$ is the mean residual lifetime function at 0 and $\widehat{g}_n(0)$ its estimated counterpart. Clearly, Υ is (Hadamard-)differentiable and the required Hadamard differentiability of $(1-S) \mapsto g(0)$ follows immediately from Lemma 4. Finally, the multiplication functional is also Hadamard-differentiable. All in all, we conclude that $W_n = \sqrt{n}(\Xi(\widehat{S}_n) - \Xi(S))$ for a functional $\Xi : D[0,\tau] \to C[0,1]$ which is Hadamard-differentiable at S tangentially to $C[0,\tau]$. Theorem 1 in combination with the functional δ-method (for the bootstrap) immediately implies that W_n and W_n^* both converge in (conditional) distribution to the same continuous Gaussian process (in outer probability given $X_1, \delta_1, X_2, \delta_2, \ldots$). Time-simultaneous inference procedures for the Lorenz curve such as tests for equality and confidence bands are constructed straightforwardly.

Finally, the normalized estimated Gini index allows the representation

$$\sqrt{n}(\widehat{\mathcal{G}}_n - \mathcal{G}) = 2\sqrt{n}(\{\Phi \circ \Xi\}(1) \circ \widehat{S}_n - \{\Phi \circ \Xi\}(1) \circ S)$$

of which $\{\Phi \circ \Xi\}(1)$ is again Hadamard-differentiable at S tangentially to $C[0,\tau]$. Hence, confidence intervals for $\mathcal{G}$ with bootstrap-based quantiles are constructed in the same way as before.

5 Small sample performances

In this section, we evaluate the behavior of survival and mean residual lifetime curve confidence bands in terms of their coverage probabilities for small- to medium-sized samples. All subsequent simulation results have been obtained using R version 3.2.3 (R Development Core Team 2016).

5.1 Comparison of various confidence bands for S

We assess the performance of the present bootstrap method by comparing the coverage probabilities of several possible confidence bands for the whole survival function S with nominal coverage $1-\alpha \in (0,1)$. Even though transformations such as $x \mapsto \log(-\log x)$ are known to improve the bands' coverage probabilities for small samples in case of restricted time spans (cf. p. 266 in Andersen et al. 1993), we decided to only consider untransformed bands since such transformations are not suitable for the boundary values $S(0) = 1$ and $S(\tau) = 0$.

For the construction of the bands, we note that Theorem 1 also implies the validity of convergences similar to those in Theorem 1.2 of Gill (1983): Define the slightly modified Nelson-Aalen variance estimators

$$\widehat{C}_n(t) = \int_0^t \frac{\mathrm{d}\widehat{A}_n(u)}{\widehat{H}_n(u)} \quad \text{and} \quad C_n^*(t) = \int_0^t \frac{\mathrm{d}A_n^*(u)}{H_n^*(u)}$$

and denote $K(t) = 1-(1+\Gamma(t,t))^{-1}$, $\widehat{K}_n(t) = 1-(1+\widehat{C}_n(t))^{-1}$, and $K_n^*(t) = 1-(1+C_n^*(t))^{-1}$. Then, Condition (4) implies that, conditionally on $X_1, \delta_1, X_2, \delta_2, \ldots$,

$$\sqrt{n}(S_n^* - \widehat{S}_n)\frac{1-K_n^*}{S_n^*} \xrightarrow{d} B^0 \circ K \quad \text{and} \quad \sqrt{n}(S_n^* - \widehat{S}_n)\frac{1-\widehat{K}_n}{\widehat{S}_n} \xrightarrow{d} B^0 \circ K$$

on $D[0,\tau]$ in probability, where B^0 is a standard Brownian bridge.

Following the convention of Gill (1983), we let $(1-\widehat{K}_n)/\widehat{S}_n$ and $(1-K_n^*)/S_n^*$ by the latest defined value in case of $0/0$. The convergences in the previous display hold due to $(1-K)/S$, $(1-\widehat{K}_n)/\widehat{S}_n$ and $(1-K_n^*)/S_n^*$ being contained in $[0,1]$ and non-increasing (cf. equation 1.2 in Gill 1983), and the latter two are pointwise consistent for the first function. But pointwise consistency, in combination with monotonicity and boundedness, implies uniform consistency. Therefore, Slutzky's theorem yields the above convergences.

It follows that the following confidence bands for the whole function S have asymptotic coverage probability $1-\alpha$:

$$\widehat{S}_n \pm \frac{q^0}{\sqrt{n}}\frac{\widehat{S}_n}{1-\widehat{K}_n}, \quad \widehat{S}_n \pm \frac{\widehat{q}_n^0}{\sqrt{n}}\frac{\widehat{S}_n}{1-\widehat{K}_n}, \quad \widehat{S}_n \pm \frac{q_n^{0*}}{\sqrt{n}}\frac{\widehat{S}_n}{1-\widehat{K}_n}, \quad \text{and} \quad \widehat{S}_n \pm \frac{q_n^*}{\sqrt{n}},$$

where q^0 denotes the $(1-\alpha)$-quantile of the distribution of $\sup_{u\in[0,1]}|B^0(u)|$, $\widehat{q}_n^0$ denotes the $(1-\alpha)$-quantile of the conditional distribution of $\sup_{[0,\tau]}|\sqrt{n}(S_n^* - \widehat{S}_n)\frac{1-\widehat{K}_n}{\widehat{S}_n}|$ given $X_1, \delta_1, X_2, \delta_2, \ldots$, and similarly, q_n^{0*} and q_n^* are obtained from $\sup_{[0,\tau]}|\sqrt{n}(S_n^* - \widehat{S}_n)\frac{1-K_n^*}{S_n^*}|$ and $\sup_{[0,\tau]}|\sqrt{n}(S_n^* - \widehat{S}_n)|$, respectively. Due to $1-\widehat{K}_n = 1/(1+\widehat{C}_n)$, the first three of the above confidence bands are of Hall–Wellner type; cf. p. 266 in Andersen et al. (1993). Therefore, we denote the four bands by HW, $\widehat{HW}$, HW^* and BM^*, respectively, "BM" standing for Brownian motion.

Furthermore, we compare the above confidence bands with the asymptotic and bootstrapped equal precision bands, cf. p. 266 in Andersen et al. (1993),

$$\widehat{S}_n \pm \frac{\tilde{q}^0}{\sqrt{n}}\sqrt{\widehat{\Gamma}_n}, \quad \widehat{S}_n \pm \frac{\widehat{q}_n^{EP}}{\sqrt{n}}\sqrt{\widehat{\Gamma}_n}, \quad \widehat{S}_n \pm \frac{q_n^{EP*}}{\sqrt{n}}\sqrt{\widehat{\Gamma}_n},$$

respectively, which we denote by EP, $\widehat{EP}$, EP^*. Here, $\tilde{q}^0$ denotes the $(1-\alpha)$-quantile of the distribution of $\sup_{u\in[0,1]}|B^0(u)(u(1-u))^{-1/2}|$ and $\widehat{q}_n^{EP}$ and q_n^{EP*} are obtained from $\sup_{[0,\tau]}|\sqrt{n}(S_n^* - \widehat{S}_n)/\widehat{\Gamma}_n^{1/2}|$ and $\sup_{[0,\tau]}|\sqrt{n}(S_n^* - \widehat{S}_n)/\Gamma_n^{*1/2}|$, resulting in pointwise studentizations of the (bootstrapped) Kaplan–Meier estimators. Similarly as above, we use the convention that $\Gamma_n^*(t)$ is the first nonzero value in case of $S_n^*(t) = 1$ and the latest nonzero value in case of $S_n^*(t) = 0$.

Note that, for various nominal levels α, the quantiles q^0 and $\tilde{q}^0$ are tabulated; see, for example, Appendix C in Klein and Moeschberger (2003) and the R package *km.ci*. Scaling by the empirical counterparts of $(1-K)/S$ and $1/\sqrt{\Gamma}$ yielded the above pivotal limit distributions, i.e., they do not depend on unknown quantities. The pivotality holds due to the surjectivity of K onto [0, 1] implying that $\sup_{[0,\tau]} |B^0 \circ K| = \sup_{[0,1]} |B^0|$. This feature typically leads to a good small sample performance.

Throughout the following simulations, we chose $\alpha = 5\%$, i.e., we create 95% confidence bands for S. We consider these simulation setups:

1. Exponential distributions $S(t) = \exp(-\lambda t)$ and $G(t) = \exp(-\mu t)$ for which the stronger Condition (5) is satisfied if $0 < \mu < \lambda < \infty$. In particular, we chose $(\mu, \lambda) = (1, 2), (1, 1.5), (1, 1.2)$. The resulting censoring probabilities $P(C < T) = \mu/(\lambda + \mu)$ are 33.33%, 40%, and 45.45%. We refer to these setups as (i), (ii) and (iii), respectively.
2. Gompertz distribution $S(t) = \exp(-\eta(\exp(bt) - 1))$ and exponential distribution $G(t) = \exp(-\mu t)$ for which Condition (5) is satisfied for all possible parameters $\eta, b, \mu > 0$. In particular, we chose $(\mu, \eta, b) = (1/3, 1, 1), (1/3, 0.3, 1), (1/3, 0.02, 2)$, yielding censoring rates of approximately 17.19%, 31.68% and 42.43%, respectively. These were found via simulation of 1,000,000 independent individuals. We also refer to these setups as (i), (ii) and (iii), respectively.

For each setup, we created 10,000 confidence intervals and each bootstrap-based quantile has been obtained from 1,999 bootstrap replicates. Compared to EP and HW, respectively, the confidence bands $\widehat{EP}$ and $\widehat{HW}$ yielded overall comparable but often worse coverage probabilities. Following the reasoning of the referee, this bad behavior of $\widehat{EP}$ and $\widehat{HW}$ could be due to not taking the randomness of estimating the standard deviation into account: Given $X_1, \delta_1, X_2, \delta_2, \ldots$, the studentizations of $\sqrt{n}(S_n^* - \widehat{S}_n)$ via $(1 - \widehat{K}_n)/\widehat{S}_n$ and $\widehat{\Gamma}_n$ are no random quantities, in contrast to the same studentizations for $\sqrt{n}(\widehat{S}_n - S)$. Therefore, the simulation results for $\widehat{EP}$ and $\widehat{HW}$ are not shown here.

The remaining results are shown in Tables 1, 2 and 3. The coverage probabilities of the equal precision bands EP are much too low in all considered setups. Appar-

Table 1 Simulated coverage probabilities (in %) of asymptotic 95% confidence bands for the exponential (left) and the Gompertz (right) survival function S in each setup (i), i.e., under light censoring

n	HW	HW^*	BM^*	EP	EP^*	n	HW	HW^*	BM^*	EP	EP^*
30	93.3	90.4	93.3	79.6	93.8	30	93.4	88.8	92.1	77.4	93.6
50	94.0	91.7	93.5	78.7	94.0	50	94.6	91.4	93.1	76.2	93.6
100	94.9	93.0	93.8	76.5	93.9	100	94.9	92.9	93.4	73.4	94.1
150	94.7	93.3	93.5	76.1	93.7	150	95.0	93.4	93.4	73.0	93.6
200	95.1	93.8	94.1	77.1	94.1	200	95.5	94.2	94.4	73.4	93.9
500	95.1	94.0	94.3	75.2	94.2	500	95.1	94.2	94.2	72.7	94.0
1000	95.0	94.3	94.5	74.7	93.8	1000	95.1	94.5	94.4	71.0	93.5

Table 2 Simulated coverage probabilities (in %) of asymptotic 95% confidence bands for the exponential (left) and the Gompertz (right) survival function S in each setup (ii), i.e., under medium censoring

n	HW	HW^*	BM^*	EP	EP^*	n	HW	HW^*	BM^*	EP	EP^*
30	91.3	90.0	91.8	78.5	93.4	30	93.3	89.4	92.8	78.4	93.5
50	93.3	91.8	92.9	78.5	93.7	50	94.5	91.2	93.1	76.0	93.7
100	94.5	93.0	93.6	79.1	94.2	100	95.1	92.8	93.3	74.7	94.1
150	94.4	92.9	93.4	77.2	93.4	150	95.4	93.4	93.8	74.5	94.1
200	94.7	93.2	94.2	76.2	93.7	200	95.2	93.8	94.0	73.0	93.9
500	95.2	94.2	94.2	76.0	94.1	500	94.9	94.1	94.2	72.3	94.0
1000	94.9	94.2	94.6	75.4	93.9	1000	95.4	94.6	94.3	70.9	93.1

Table 3 Simulated coverage probabilities (in %) of asymptotic 95% confidence bands for the exponential (left) and the Gompertz (right) survival function S in each setup (iii), i.e., under strong censoring

n	HW	HW^*	BM^*	EP	EP^*	n	HW	HW^*	BM^*	EP	EP^*
30	89.2	89.7	90.0	78.4	93.8	30	92.6	89.0	92.7	77.6	93.2
50	91.9	91.6	91.5	77.8	93.4	50	94.0	91.1	93.2	74.8	93.3
100	94.2	93.1	92.3	76.7	93.2	100	94.8	92.3	93.3	75.1	93.7
150	94.5	93.1	92.0	76.5	93.6	150	95.0	93.3	93.8	73.3	93.5
200	94.8	93.6	93.1	77.1	93.3	200	95.3	93.7	93.9	72.3	93.7
500	95.4	94.4	93.4	75.4	93.8	500	95.1	94.2	94.3	72.0	93.8
1000	95.0	94.2	93.9	74.6	93.6	1000	95.4	94.7	94.6	70.8	93.8

ently, the quantiles $\tilde{q}^0$ do not adequately correspond to the large sample behavior of $\sup_{[0,\tau]} \sqrt{n}|\widehat{S}_n - S|/\widehat{\Gamma}_n^{1/2}$. Indeed, it is not obvious how the weak convergence thereof may be verified. Therefore, it surprises to see that the quantile based on the bootstrapped Kaplan–Meier estimator and the bootstrapped estimated variance function Γ_n^* apparently succeeds in mimicking the correct asymptotic distribution: all simulated coverage probabilities of the EP^* band are between 93.2 and 94.4%, even for the smallest sample size $n = 30$ and under strong censoring. Even though these coverage probabilities are a bit too low but else satisfying, this outcome has to be treated with caution as no theoretic justification is available yet. For example, it is not even apparent from the simulations that the coverage probabilities will finally converge to 95%.

In contrast, the asymptotically exact confidence bands HW^* yield slightly too small coverage probabilities which seem to converge quite slowly from below to their limit of 95%. Typically, this gap is reduced by applications of suitable transformations such as $\log - \log$ which, unfortunately, are not available as explained above. Similar results are obtained in other distribution setups (results not shown).

Finally, the most accurate simulated coverage probabilities among those confidence bands, for which asymptotic correctness has been verified, are achieved by HW and BM^* with a slight preference for the Hall–Wellner band based on the asymptotic quantile. This is a bit surprising when comparing these results with the simulation study by Akritas (1986): He found the (time-restricted version of) HW to be too conservative and the corresponding HW^* to be quite accurate for quite small sample sizes $n = 25, 50$. We found the same relation between those bands, even though all these bands lost a few percentage points of coverage probability in comparison with the simulations by Akritas (1986) when creating confidence bands for all time points. Therefore, the former conservative band HW appears to be the most accurate in our simulations. The band HW approaches the nominal coverage level satisfactorily for $n \geqslant 100$, even under strong censoring. For small sample sizes, the simulated coverage probabilities of HW and BM^* are comparable, but the convergence seems to be slower for BM^* as n grows. Therefore, based on these simulation outcomes, one should choose the HW band for a confidence region for the whole survival function S. However, we will learn in Sect. 6 below why the BM^* band might, in fact, be the preferable one.

5.2 Confidence bands for the mean residual lifetime function

Complementing the previous section's results, we here demonstrate the performance of bootstrap-based 95% confidence bands for the mean residual lifetime function which involve non-pivotal limit distributions. That is, applications of the bootstrap are essentially necessary and cannot be circumvented by means of tabulated quantiles. For practical reasons, the estimated mean residual lifetime is only integrated up to the largest observed event time. Otherwise, an integration of the eventually constant Kaplan–Meier estimate beyond the largest observation and up to an unknown end point τ, which would require an arbitrary choice by the statistician, could result in a bias and hence in a bad performance of the confidence band. We used the same number of iterations as in the previous subsection and chose the following setup underlying the simulations:

1. The same distributions as in setup 1 (Exponential distributions) of the previous section with confidence bands along the time points $[t_1, t_2] = [0, 1], [0, 1.2], [0, 1.4]$, where the length of the time interval increases with the censoring intensity. All other time span-censoring combinations have been considered in "Appendix D." The survival probabilities at the right boundaries of these time intervals are approximately 13.53%, 16.53% and 18.64% respectively. Due to memorylessness of the exponential distribution, the true mean residual lifetime functions are constant, $\psi(S) \equiv \text{const}$.
2. The same distributions as in setup 2 (Gompertz distributions) of the previous section with confidence bands along the time points $[t_1, t_2] = [0, 1.2], [0, 2], [0, 2.3]$, respectively. The survival probabilities at the right boundaries of these time intervals are approximately 9.83, 14.71 and 13.95%, respectively. The resulting mean residual lifetime functions are strictly decreasing and convex.

The respective scenarios in each setup are numbered as (i), (ii), (iii).

We simulated both the untransformed, i.e., linear, and the log-transformed confidence bands for $\psi(S)$. Considering the simulated coverage probabilities of confidence bands for the mean residual lifetime functions for the exponential distribution setup in Table 4 (left), we see, first, that the application of the log-transformation has a bad effect on the confidence bands: For small sample sizes, the coverage probabilities of the log-transformed bands are far too low, whereas it may even exceed 95% for larger sample sizes as seen in the second column. This effect is even more pronounced in the Gompertz distribution setup with simulation results presented in Table 4 (right) where the coverage probabilities increase up to 98%.

In contrast, the linear confidence bands show much more acceptable coverage probabilities, even for small samples. Indeed, all of these are contained in the interval [92.6%, 97.3%] in all considered setups. For the exponential setups (i) and (ii), Table 4 (left) suggests convergence to 95% from below, whereas the linear bands appear to be slightly too wide in case of the underlying Gompertz distribution; see Table 4 (right). This may be due to the smaller survival probability at the right boundaries t_2 of the considered time intervals.

All in all, we advise using the linear confidence band for the mean residual lifetime function in combination with the bootstrap. As suggested from the conducted simulations, the sample sizes need not be very large in order to ensure a sufficient accurateness of the developed confidence bands, for example, the case $n \geqslant 150$ yields confidence bands with coverage probabilities close to the nominal level even under stronger censoring. The impact of the choice of the time span $[t_1, t_2]$ on the coverage probabilities is assessed in "Appendix D." There it is seen that the coverage probabilities can deteriorate if the survival probabilities at the intervals' right boundary are extremely small.

6 Application to the lung cancer data-set

In order to illustrate the advantages and drawbacks of the confidence bands empirically assessed in the previous section, we consider the lung cancer data-set *lung* which is freely available in the R package *survival* (Therneau and Lumley 2017). The data consist of, among others, possibly right-censored survival times (in days) of patients with advanced lung cancer from the North Central Cancer Treatment Group. In the original study, several additional scores concerning performance of daily activities and other covariates have been measured for fitting Cox models; cf. Loprinzi et al. (1994). In this section, however, we focus on a nonparametric, patient-averaged confidence bands-based analysis of the survival curve and the mean residual lifetime function using the methods developed above.

Out of this data-set, we analyze the subsample of male patients which consists of $n = 138$ individuals. Thereof, $26/138 \approx 18.84\%$ are right-censored and the death time has been recorded for the rest. Therefore, judging from the simulation results of the previous section for the case of light censoring and $n = 150$, we expect to find confidence bands for the true survival function with coverage probabilities between 93.5 and 95%. Similarly, the linear confidence band for the mean residual lifetime function will presumably have coverage probability between 95 and 97%. The corresponding log-transformed bands will possibly be too conservative with coverage probabilities between 96% and 98%. For matching the data-set to the present theory, we broke ties in the survival times by randomly adding small perturbations, but still respecting the convention that censorings occur after events. The more restrictive censoring condition 5 seems to be satisfied as

Table 4 Simulated coverage probabilities (in %) of asymptotic 95% confidence bands for the mean residual lifetime function for distribution setups 1 (left) and 2 (right)

Setup	Exponential survival distribution						Setup	Gompertz survival distribution					
	(i)		(ii)		(iii)			(i)		(ii)		(iii)	
n	Linear	Log	Linear	Log	Linear	Log	n	Linear	Log	Linear	Log	Linear	Log
30	94.4	61.7	93.1	56.9	92.6	51.0	30	96.6	84.0	96.1	87.5	95.6	84.5
50	94.6	79.9	93.5	74.8	93.0	70.3	50	97.0	94.9	96.6	96.5	96.1	95.2
100	95.3	94.2	94.0	91.9	93.5	88.1	100	97.3	98.6	96.7	98.3	95.8	98.5
150	95.8	96.3	94.7	95.6	94.2	94.2	150	96.7	98.2	96.8	97.6	96.1	98.2
200	95.7	96.6	94.9	95.8	94.5	95.3	200	96.7	97.7	96.6	97.5	95.8	97.6
500	95.6	96.3	95.5	95.8	95.1	95.5	500	95.5	96.0	95.6	95.8	95.2	95.8
1000	95.7	95.6	95.4	95.5	95.4	95.5	1000	95.3	95.5	95.1	95.3	95.0	95.3

$$-\int_0^\tau \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}} - \int_0^\tau \frac{\widehat{S}_n \mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}^2} \approx 1.81$$

is still a very small number.

Figure 1 shows the plots of all survival confidence bands with nominal level 95% under consideration. On day 700 there were only $8/138 \approx 5.60\%$ of all patients still at risk. At first, we pay attention to the survival probability bands. Thereof, all three Hall–Wellner bands look similar, but the band based on the asymptotic quantile, HW, is slightly wider, which might reflect the reason for the better accurateness in the simulation study of Sect. 5. Apart from that, the Hall–Wellner bands exhibit a quite unfavorable behavior: At early time points, the bands are unnecessarily broad, while, at late time points, they are far too narrow. The latter is a real problem as, due to the last two occurring times being censorings, the estimated final survival probability is 3.57%, i.e., slightly greater than zero. However, the Hall–Wellner bands suggest a definitive certainty concerning the Kaplan–Meier estimations, but the estimators and bands are constant from this point on. Hence, the Hall–Wellner bands cannot contain the real survival function which eventually drops down to zero.

On the other hand, the linear confidence band avoids this obstacle by maintaining the same width for the whole time. Therefore, it also provides a width for early time points which is much more reasonable than the early widths of the Hall–Wellner bands that completely cover [0, 1]. Even though the rates of convergence of BM^*'s coverage probabilities toward the nominal level were slower than those of HW, the BM^* band is, in this example, the confidence region of choice due to its reasonable shape and its asymptotic exactness. Furthermore, due to the quite light censoring in the data-set, the above simulation study suggests that BM^* has coverage probabilities of approximately 93.5%.

Finally, we consider the equal precision bands which provide the perhaps most attractive shapes: narrower, but not too narrow at early and late time points and wider in between where the estimation uncertainty is higher. The discrepancy between EP^*, EP and $\widehat{EP}$ nicely illustrate the undercoverage of the latter two, which had also been observed in the simulation study of Sect. 5. The same simulations indicate that EP^* should have coverage probabilities close to the nominal level. However, due to the ignorance of its asymptotic correctness, one should still choose the BM^* band instead.

Let us turn our attention to the 95% confidence bands for the mean residual lifetime function. Figure 2 shows the linear and log-transformed confidence bands on the time interval [0, 700] days. Apart from the ragged shape of the function estimator, the mean residual lifetime function seems to be decreasing almost linearly from 325 days at the beginning to 194 days at day 700. Especially for early time points, we see that the log-transformed band is much wider which is presumably the reason for the unnecessary high coverage probability in comparison with the linear bands; see again the simulation study in Sect. 5. At later time points, the lower bounds of both types of bands even cross such that the log-transformed band is overall slightly elevated. All in all, we suggest using linear simultaneous confidence bands for the mean residual lifetime function which have a reasonably narrow shape and coverage probabilities close to the nominal level.

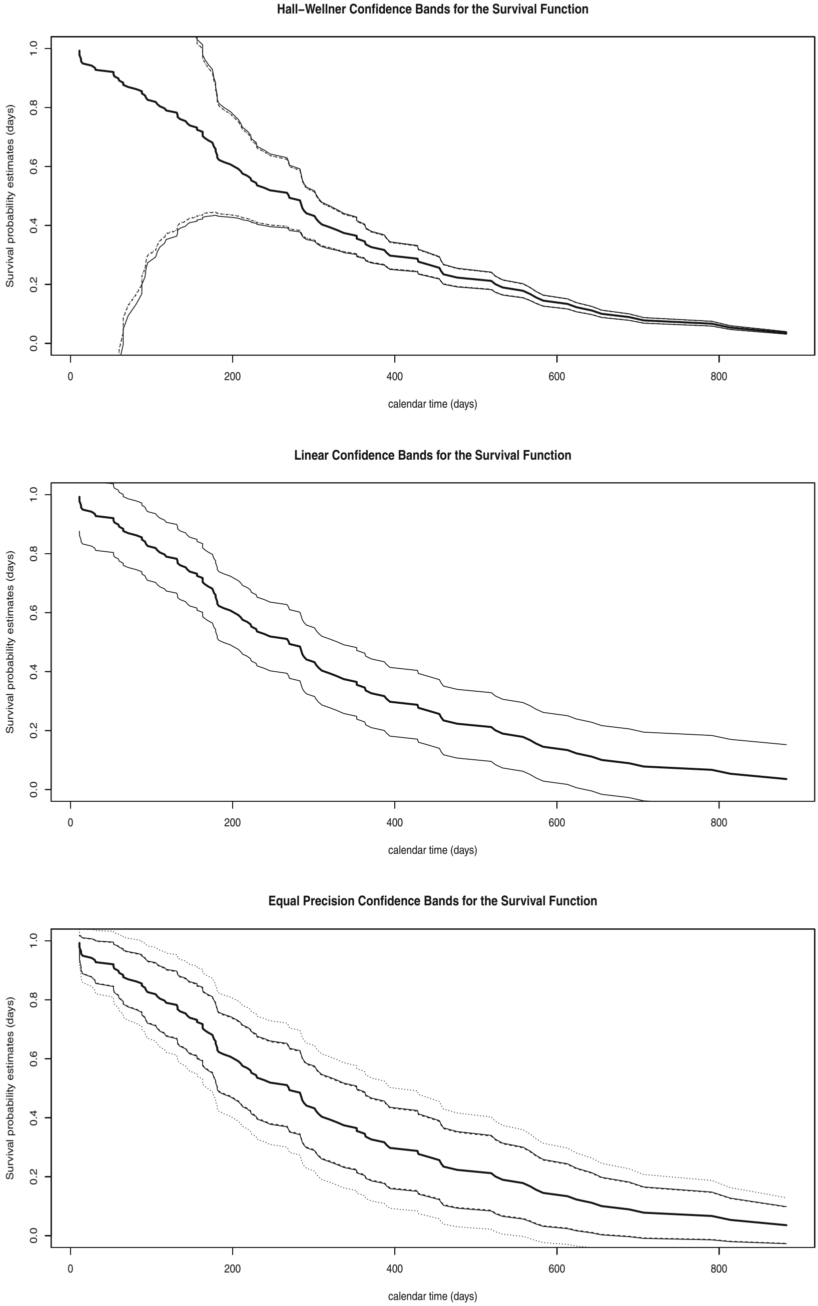

Fig. 1 Confidence bands of Hall–Wellner, linear and equal precision bands for the survival function, respectively. The straight bands (—) correspond to HW, BM^* and EP, the dashed bands (- -) to $\widehat{HW}$ and $\widehat{EP}$ and the dotted bands ($\cdots$) to HW^* and EP^*. In the upper panel, $\widehat{HW}$ and HW^* are virtually the same band. In the lower panel, the EP and $\widehat{EP}$ bands are almost the same. The thick straight lines (**—**) show the Kaplan–Meier estimator

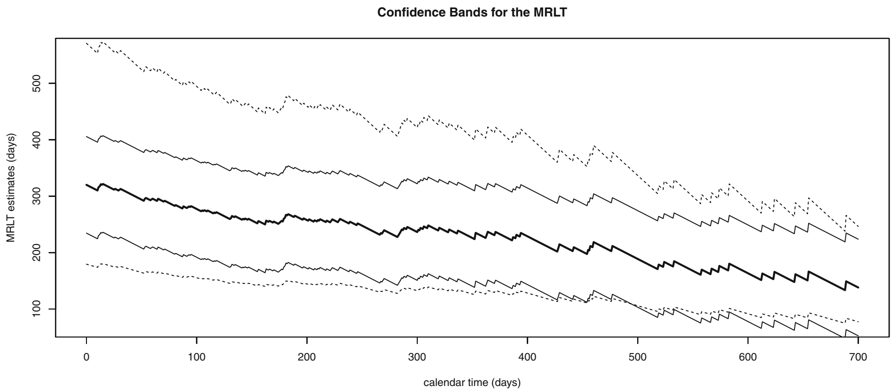

Fig. 2 Linear (—) and log-transformed (- -) confidence bands for the mean residual lifetime function. The thick straight line (**—**) shows the estimator for the mean residual lifetime curve

7 Discussion

In this article we established consistency of the bootstrap for Kaplan–Meier estimators on the whole support of the estimated survival function. By means of the functional delta method, this conditional weak convergence is transferred to Hadamard-differentiable functionals such as the mean residual lifetime, the Lorenz curve, or the Gini index. Further applications include the expected length of stay in the transient state (e.g., Grand and Putter 2016) or the probability of concordance in a two-sample problem (e.g., Pocock et al. 2012; Dobler and Pauly 2017).

Based on the empirical results of Sect. 5 for the mean residual lifetime function, we saw that apart from already good coverage probabilities in smaller samples of the linear confidence bands, the log-transformation in general yields undesirable deteriorations in terms of coverage probabilities. Furthermore, such a transformation is not available for the construction of confidence bands for the whole survival functions which partially resulted in slightly too narrow confidence bands. Here, on the one hand, the particular choice of the Hall–Wellner band appeared to yield reliable confidence bands in terms of coverage, even for small sample sizes and under strong censoring. On the other hand, the lung cancer data example discussed in Sect. 6 suggests a quite unattractive shape of the Hall–Wellner bands: Far too wide in the beginning and far too narrow in the end. Especially if the latest few observations consist entirely of censorings, the Kaplan–Meier estimate will remain strictly greater than zero, but the widths of the Hall–Wellner bands vanish. Even though the considered bootstrap-based equal precision bands revealed good empirical coverage rates, this type of confidence band should not be used as long as its asymptotic accurateness has not been verified theoretically. Therefore, we suggest using the linear, bootstrap-based confidence band for the whole survival function which provides a reasonable shape and also acceptable, though perhaps a bit too low coverage rates for smaller sample sizes.

The presently analyzed bootstrap consistency on the whole support may also be extended to more general inhomogeneous Markovian multistate models. Based on the martingale representation of Aalen-Johansen estimators for transition probability matrices (e.g., Andersen et al. 1993, p. 289), one could try to generalize the results of Gill (1983) to this setting. Here the notion of the 'largest event times' requires special attention as these may differ for different types of transitions. A reasonable first step toward such a generalization would be an analysis of competing risks setups where the support of each cumulative incidence function provides a natural domain to investigate weak convergences on. Once weak convergence of the estimators on the whole support is verified, martingale arguments similar to those of Akritas (1986) and Gill (1983) may be employed in order to obtain such (now conditional) weak convergences for the resampled Aalen-Johansen estimator using a variant of Efron's bootstrap. In more general Markovian multistate models we could independently draw with replacement from the sample that contains all individual *trajectories* rather than single observed transitions in order to not corrupt the dependencies within each individual; see for example Tattar and Vaman (2012) for a similar suggestion. Applications of this theory could include inference on more refined variants of the probability of concordance or the expected length of stay. Considering a progressive disease in a two-sample situation, for instance, one could compare the probability that an individual of group one remains longer in a less severe disease state than an individual of group two. Accurate inference procedures for the mean residual lifetime in a state of disability given any state at present time offer another kind of application.

Acknowledgements The author would like to thank Markus Pauly (Ulm University) for helpful discussions. Furthermore, the discussion with a referee and his / her suggestions to present numerical results and to illustrate the methods with the help of real data have helped to improve this manuscript significantly.

Appendix

Denote by $\mathbf{X}$ the σ-algebra generated by all observations $X_1, \delta_1, X_2, \delta_2, \dots$. Some of the following proofs (in "Appendix A") rely on the ideas of Gill (1983). In order to also apply (variants of) his lemmata in our bootstrap context, "Appendix B" below contains all required results. 'Tightness' in the support's right boundary τ for the bootstrapped Kaplan–Meier estimator is essentially shown via a bootstrap version of the approximation theorem for truncated estimators as in Theorem 3.2 in Billingsley (1999); cf. "Appendix C." Define by $Y(u) = n\widehat{H}_{n-}(u)$ the process counting the number of individuals at risk of dying, and by $Y^*(u)$ its bootstrap version.

A Proofs

Proof of Lemma 2 Proof of (a): This part of the lemma is proven by showing the convergence of the integral from 0 to τ in probability. Since, by the continuous mapping theorem and the boundedness away from zero of $\frac{1}{G}$ on $[0, t]$, the convergence $-\int_0^t \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}} \xrightarrow{p} -\int_0^t \frac{\mathrm{d}S}{G_-}$ holds as $n \to \infty$, the assertion then follows from an application of the continuous mapping theorem to the difference functional.

Let $t < \tau$ and suppose (4) holds. Letting $t \uparrow \tau$, the integral $-\int_0^t \frac{\mathrm{d}S}{G_-}$ converges toward $-\int_0^\tau \frac{\mathrm{d}S}{G_-} < \infty$. It remains to apply Theorem 3.2 of Billingsley (1999) to the distance $\rho(-\int_0^t \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}}, -\int_0^\tau \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}}) = |-\int_0^\tau \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}} + \int_0^t \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}}|$ for a verification of the assertion for $t = 0$. Thus, we show that for all $\varepsilon > 0$,

$$\lim_{t\uparrow\tau} \limsup_{n\to\infty} P\left(-\int_t^\tau \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}} > \varepsilon\right) = 0.$$

Let $\widehat{T}_n = X_{n:n}$ again be the largest observation among $X_1, \ldots, X_n$ and define, for any $\beta > 0$,

$$B_\beta := \left\{\widehat{S}_n(s) \leqslant \beta^{-1} S(s) \text{ and } \widehat{H}_n(s-) \geqslant \beta H(s-) \text{ for all } s \in [0, \widehat{T}_n]\right\}.$$

By Lemmata 6 and 7, the probability $p_\beta := 1 - P(B_\beta) \leqslant \beta + \frac{e}{\beta}\exp(-1/\beta)$ is arbitrary small for sufficiently small $\beta > 0$. Hence, by Theorem 1.1 of Stute and Wang (1993) (applied for the concluding convergence),

$$\begin{aligned} P\left(-\int_t^\tau \frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}} > \varepsilon\right) &= P\left(-\int_t^\tau \frac{\widehat{S}_{n-}\mathrm{d}\widehat{S}_n}{\widehat{H}_{n-}} > \varepsilon\right) \\ &\leqslant P\left(-\beta^{-2}\int_t^\tau \frac{S_-\mathrm{d}\widehat{S}_n}{H_-} > \varepsilon\right) + p_\beta \\ &= P\left(-\beta^{-2}\int_t^\tau \frac{\mathrm{d}\widehat{S}_n}{G_-} > \varepsilon\right) + p_\beta \to P\left(-\beta^{-2}\int_t^\tau \frac{\mathrm{d}S}{G_-} > \varepsilon\right) + p_\beta. \end{aligned}$$

For large $t < \tau$ and by the continuity of S, the far right-hand side of the previous display equals p_β.

Suppose now that (4) is violated. By the Glivenko–Cantelli theorems in Stute and Wang (1993) for Kaplan–Meier estimators of continuous survival functions and by letting G be continuous w.l.o.g. (distributing the atoms of G uniformly on small intervals with no mass of S, without affecting the integral), the integral over $(0, t]$ converges almost surely to $-\int_0^t (\mathrm{d}S)/G_-$ for each $t < \tau$ by the continuous mapping theorem. But this integral is arbitrarily large for sufficiently large $t \uparrow \tau$. Hence, the stated a.s. convergence follows.

Proof of (b): First note that the uniform convergences in probability in Theorems IV.3.1 and IV.3.2 of Andersen et al. (1993), p. 261ff., yield, for any $\varepsilon > 0$,

$$\sup_{(u,v)\in[0,\tau-\varepsilon]^2} |\widehat{\Gamma}_n(u, v) - \Gamma(u, v)| \xrightarrow{p} 0 \quad \text{as } n \to \infty.$$

Further, the dominated convergence theorem and $S_-\mathrm{d}A = -\mathrm{d}S$ show that

$$\Gamma(u, v) = -\int_0^\tau \mathbf{1}\{w \leqslant u \wedge v\} \frac{S(u)S(v)}{S(w-)S(w-)} \frac{\mathrm{d}S(w)}{G(w-)} \to -\int_0^\tau 0 \frac{\mathrm{d}S}{G_-} = 0$$

as $u, v \to \tau$. Hence, it remains to verify the remaining Condition (3.8) of Theorem 3.2 in Billingsley (1999) in order to conclude this proof. That is, for each positive δ we show

$$\lim_{u,v\to\tau}\limsup_{n\to\infty} P\left(\sup_{(u,v)\in[0,\tau]^2} |\widehat{\Gamma}_n(u,v)-\widehat{\Gamma}_n(\tau,\tau)| \geqslant \delta\right) = 0.$$

To this end, rewrite $\widehat{\Gamma}_n(u,v)-\widehat{\Gamma}_n(\tau,\tau)$ as

$$\int_{u\wedge v}^{\tau} \frac{\widehat{S}_n(\tau)\widehat{S}_n(\tau)}{\widehat{S}_n(w-)\widehat{S}_n(w-)} \frac{\mathrm{d}\widehat{S}_n(w)}{\widehat{G}_n(w-)} + \int_0^{u\wedge v} \frac{\mathrm{d}\widehat{S}_n}{\widehat{S}_{n-}^2\widehat{G}_{n-}} (\widehat{S}_n^2(\tau)-\widehat{S}_n(u)\widehat{S}_n(v)).$$

The left-hand integral is bounded in absolute value by $-\int_{u\wedge v}^{\tau}\frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}}$ which goes to $-\int_{u\wedge v}^{\tau}\frac{\mathrm{d}S}{G_-}$ in probability as $n\to\infty$ by (a). For large u,v this is arbitrarily small.

The remaining integral is bounded in absolute value by

$$-\int_0^{u\wedge v} \frac{\widehat{S}_n(u)\widehat{S}_n(v)}{\widehat{S}_{n-}^2}\frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}} = -n^2\int_0^{u\wedge v}\frac{\widehat{S}_n(u)\widehat{S}_n(v)}{Y^2}\widehat{G}_{n-}\mathrm{d}\widehat{S}_n.$$

By Lemmata 6 and 7 this integral is bounded from above by a constant times

$$-\int_0^{u\wedge v}\frac{S(\widehat{T}_n\wedge u)S(\widehat{T}_n\wedge v)}{H_-^2}G_-\mathrm{d}\widehat{S}_n = -\int_0^{u\wedge v}\frac{S(\widehat{T}_n\wedge u)S(\widehat{T}_n\wedge v)}{S_-^2G_-}\mathrm{d}\widehat{S}_n$$

on a set with arbitrarily high probability. For sufficiently large n we also have $\widehat{T}_n > u\wedge v$ with arbitrarily high probability. Next, Theorem 1.1 in Stute and Wang (1993) yields

$$-\int_0^{u\wedge v}\frac{S(u)S(v)}{S_-^2G_-}\mathrm{d}\widehat{S}_n \xrightarrow{p} -\int_0^{u\wedge v}\frac{S(u)S(v)}{S_-^2G_-}\mathrm{d}S \quad \text{as } n\to\infty.$$

As above the dominated convergence theorem shows the negligibility of this integral as $u,v\to\tau$. □

Proof of Theorem 1 For the proof of weak convergence of the bootstrapped Kaplan–Meier estimator on each Skorohod space $D[0,t]$, $t<\tau$, see, for example, Akritas (1986), Lo and Singh (1986) or Horvath and Yandell (1987). By defining these processes as constant functions after t, the convergences equivalently hold on $D[0,\tau]$. This takes care of Condition (a) in Lemma 9, while (c) is obviously fulfilled by the continuity of the limit Gaussian process.

To close the indicated gap for the bootstrapped Kaplan–Meier process on the whole support $[0,\tau]$, it remains to analyze Condition (b). This is first verified for the truncated process by following the strategy of Gill (1983) while applying the martingale theory of Akritas (1986) for the bootstrapped counting processes. Thus, the truncation technique of Lemma 9 shows the convergence in distribution of the truncated process. Finally, the negligibility of the remainder term is shown similarly as in Ying (1989).

We will make use of the fact that our martingales, stopped at arbitrary stopping times, retain the martingale property; cf. Andersen et al. (1993), p. 70, for sufficient conditions on this matter. Similarly to the largest event or censoring time $\widehat{T}_n$, introduce the largest bootstrap time $T_n^* = \max_{i=1,\dots,n} X_i^*$, being an integrable stopping time with respect to the filtration of Akritas (1986) who used Theorem 3.1.1 of Gill (1980): Hence, we choose the filtration given by

$$\mathcal{F}_t := \{X_i, \delta_i, \delta_i^* \mathbf{1}\{X_i^* \leqslant t\}, X_i^* \mathbf{1}\{X_i^* \leqslant t\} : i = 1, \dots, n\}, \quad 0 \leqslant t \leqslant \tau;$$

see also Gill (1980), p. 26, for a similar minimal filtration. Note that we did not include the indicators $\mathbf{1}\{X_i^* \leqslant t\}$ into the filtration since their values are already determined by all the $X_i^* \mathbf{1}\{X_i^* \leqslant t\}$: According to our assumptions, $X_i^* > 0$ a.s. for all $i = 1, \dots, n$.

We would first like to verify condition (b) in Lemma 9 for the stopped bootstrap Kaplan–Meier process. That is, for each $\varepsilon > 0$ and an arbitrary subsequence $(n') \subset (n)$ there is another subsequence $(n'') \subset (n')$ such that

$$\begin{aligned}
&\lim_{t \uparrow \tau} \limsup_{n'' \to \infty} P\left(\sup_{t \leqslant s < T_n^*} \sqrt{n''} |(S_n^* - \widehat{S}_n)(s) - (S_n^* - \widehat{S}_n)(t)| > \varepsilon \mid \mathbf{X} \right) \\
&\leqslant \lim_{t \uparrow \tau} \limsup_{n'' \to \infty} P\left(\sup_{t \leqslant s < \widehat{T}_n} \sqrt{n''} |(S_n^* - \widehat{S}_n)(s \wedge T_n^*) - (S_n^* - \widehat{S}_n)(t \wedge T_n^*)| > \varepsilon \mid \mathbf{X} \right) \\
&= 0 \quad \text{a.s.}
\end{aligned} \tag{7}$$

for all $\varepsilon > 0$. Here $\sigma(\mathbf{X}) = \mathcal{F}_0$ summarizes the collected data. Due to the boundedness away from zero, i.e., $\inf_{t \leqslant s < \widehat{T}_n} \widehat{S}_n(s) > 0$, we may rewrite the bootstrap process

$$\sqrt{n}(S_n^* - \widehat{S}_n)(s) = \sqrt{n}\Big(\frac{S_n^*(s)}{\widehat{S}_n(s)} - 1 \Big) \widehat{S}_n(s)$$

for each $s \in [t, \widehat{T}_n)$ of which the bracket term is a square integrable martingale; see Akritas (1986) again. Hence, the term $\sqrt{n}(S_n^* - \widehat{S}_n)(s)$ in (7) equals

$$M_n^*(s) \widehat{S}_n(s \wedge T_n^*) := \sqrt{n} \left(\frac{S_n^*(s \wedge T_n^*)}{\widehat{S}_n(s \wedge T_n^*)} - 1 \right) \widehat{S}_n(s \wedge T_n^*), \tag{8}$$

whereof $(M_n^*(s))_{s \in [0, \widehat{T}_n)}$ is again a square integrable martingale. Indeed, its predictable variation process evaluated at the stopping time $s = T_n^*$ is finite (having the sufficient condition of Andersen et al. (1993), p. 70, for a stopped martingale to be a square integrable martingale in mind): The predictable variation is given by

$$s \mapsto \langle M_n^* \rangle(s) = \int_0^{s \wedge T_n^*} \left(\frac{S_{n-}^*}{\widehat{S}_n} \right)^2 \frac{(1 - \Delta \widehat{A}_n) \mathrm{d} \widehat{A}_n}{H_{n-}^*},$$

where H_n^* is the empirical survival function of $X_1^*, \dots, X_n^*$ and Δf denotes the increment process $s \mapsto f(s+) - f(s-)$ of a monotone function f. The supremum in (7) is bounded by

$$\sup_{t \leqslant s < \widehat{T}_n} |M_n^*(s) - M_n^*(t)| \widehat{S}_n(s \wedge T_n^*) + \sup_{t \leqslant s < \widehat{T}_n} |M_n^*(t)| |\widehat{S}_n(s \wedge T_n^*) - \widehat{S}_n(t \wedge T_n^*)|$$

of which the right-hand term is not greater than $|M_n^*(t)|\widehat{S}_n(t\wedge T_n^*)$. By the convergence in distribution of the bootstrapped Kaplan–Meier estimator on each $D[0,\widetilde{\tau}]$, $\widetilde{\tau}<\tau$, we have convergence in conditional distribution of $M_n^*(t)\widehat{S}_n(t\wedge T_n^*)$ given $\mathbf{X}$ toward $N(0, S^2(t)\Gamma(t,t))$ in probability. Hence,

$$\lim_{n''\to\infty} P(|M_n^*(t)\widehat{S}_n(t\wedge T_n^*)| > \varepsilon/2 \mid \mathbf{X}) \to 1 - N(0,\Gamma(t,t))(-\varepsilon/2,\varepsilon/2)$$

almost surely along subsequences (n'') of arbitrary subsequences $(n')\subset(n)$. Since the variance of the normal distribution in the previous display goes to zero as $t\uparrow\tau$, cf. (2.4) in Gill (1983), the above probability vanishes as $t\uparrow\tau$.

By Lemma 8, the remainder $\sup_{t\leqslant s<\widehat{T}_n}|M_n^*(s)-M_n^*(t)|\widehat{S}_n(s\wedge T_n^*)$ is not greater than

$$2\sup_{t\leqslant s<\widehat{T}_n}\Big|\int_t^s \widehat{S}_n(u)\mathrm{d}M_n^*(u)\Big|. \tag{9}$$

Since, given $\mathbf{X}$, $\widehat{S}_n$ is a bounded and predictable process, this integral is a square integrable martingale on $[t,\widehat{T}_n)$. We proceed as in Gill (1983) by applying Lenglart's inequality, cf. Sect. II.5.2 in Andersen et al. (1993): For each $\eta>0$ we have

$$\begin{aligned} &P\Big(\sup_{t\leqslant s<T_n^*\wedge\tau}\Big|\int_t^s \widehat{S}_n\mathrm{d}M_n^*\Big| > \varepsilon \,\Big|\, \mathbf{X}\Big) \\ &\quad\leqslant \frac{\eta}{\varepsilon^2} + P\Big(\Big|\int_t^{\tau\wedge T_n^*} S_{n-}^{*2}\frac{(1-\Delta\widehat{A}_n)\mathrm{d}\widehat{A}_n}{H_{n-}^*}\Big| > \eta \,\Big|\, \mathbf{X}\Big). \end{aligned} \tag{10}$$

We intersect the event on the right-hand side of (10) with $B_{H,n,\beta}^* := \{H_n^*(s-)\geqslant \beta\widehat{H}_n(s-)$ for all $s\in[t,T_n^*]\}$ and also with $B_{S,n,\beta}^* := \{S_n^*(s)\leqslant\beta^{-1}\widehat{S}_n(s)$ for all $s\in[t,T_n^*]\}$. According to Lemmata 6 and 7, the conditional probabilities of these events are at least $1-\exp(1-1/\beta)/\beta$ and $1-\beta$, respectively, for any $\beta\in(0,1)$. Thus, (10) is less than or equal to

$$\frac{\eta}{\varepsilon^2}+\beta+\frac{\exp(1-1/\beta)}{\beta}+\mathbf{1}\Big\{\beta^{-3}\Big|\int_t^{\tau\wedge\widehat{T}_n}\widehat{S}_{n-}^2\frac{(1-\Delta\widehat{A}_n)\mathrm{d}\widehat{A}_n}{\widehat{H}_{n-}}\Big|>\eta\Big\}. \tag{11}$$

In order to show the almost sure negligibility of the indicator function as $n\to\infty$ and then $t\uparrow\tau$, we analyze the corresponding convergence of the integral. Since $-\mathrm{d}\widehat{S}_n=\widehat{S}_{n-}\mathrm{d}\widehat{A}_n$, the integral is less than or equal to

$$-\int_t^\tau\frac{\widehat{S}_{n-}\mathrm{d}\widehat{S}_n}{\widehat{H}_{n-}} = -\int_t^\tau\frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}}.$$

Lemma 2 implies that for each subsequence $(n')\subset(n)$ there is another subsequence $(n'')\subset(n')$ such that $-\int_t^\tau\frac{\mathrm{d}\widehat{S}_n}{\widehat{G}_{n-}}\to-\int_t^\tau\frac{\mathrm{d}S}{G_-}$ a.s. for all $t\in[0,\tau]\cap\mathbb{Q}$ along (n''). Due to $P(Z_1\in\mathbb{Q})=0$, the same convergence holds for all $t\leqslant\tau$. Letting now $t\uparrow\tau$

shows that the indicator function in (11) vanishes almost surely in limit superior along (n''). The remaining terms are arbitrarily small for sufficiently small $\eta, \beta > 0$. Hence, all conditions of Lemma 9 are met and the assertion follows for the stopped process

$$(\mathbf{1}\{s < T_n^*\}\sqrt{n}(S_n^*(s) - \widehat{S}_n(s)) + \mathbf{1}\{s \geqslant T_n^*\}\sqrt{n}(S_n^*(T_n^*-) - \widehat{S}_n(T_n^*-)))_{s\in[0,\tau]}.$$

Finally, we show the asymptotic negligibility of

$$\sup_{T_n^* \leqslant s \leqslant \tau} \sqrt{n}|S_n^*(s) - \widehat{S}_n(s)| \leqslant \sup_{T_n^* \leqslant s \leqslant \tau} \sqrt{n}(S_n^*(s) + \widehat{S}_n(s)) \\ = \sqrt{n}S_n^*(T_n^*) + \sqrt{n}\widehat{S}_n(T_n^*);$$

cf. Ying (1989) for similar considerations. Again by Lemma 6, we have for any $\varepsilon > 0$, $\beta \in (0, 1)$ that

$$P(\sqrt{n}S_n^*(T_n^*) + \sqrt{n}\widehat{S}_n(T_n^*) > \varepsilon \mid \mathbf{X}) \\ \leqslant P(\sqrt{n}S_n^*(T_n^*) > \varepsilon/2 \mid \mathbf{X}) + P(\sqrt{n}\widehat{S}_n(T_n^*) > \varepsilon/2 \mid \mathbf{X}) \\ \leqslant P(\sqrt{n}\widehat{S}_n(T_n^*) > \beta\varepsilon/2 \mid \mathbf{X}) + P(\sqrt{n}\widehat{S}_n(T_n^*) > \varepsilon/2 \mid \mathbf{X}) + \beta.$$

Define the generalized inverse $\widehat{S}_n^{-1}(u) := \inf\{s \leqslant \tau : \widehat{S}_n(s) \geqslant u\}$. The independence of the bootstrap drawings as well as arguments of quantile transformations yields

$$P(\sqrt{n}\widehat{S}_n(T_n^*) > \varepsilon \mid \mathbf{X}) = P(X_1^* < \widehat{S}_n^{-1}(\varepsilon/\sqrt{n}) \mid \mathbf{X})^n \\ = \Big[1 - \frac{1}{n}|\{i : X_i \geqslant \widehat{S}_n^{-1}(\varepsilon/\sqrt{n})\}|\Big]^n.$$

The cardinality in the display goes to infinity in probability, and hence almost surely along subsequences. Indeed, for any constant $C > 0$,

$$P(|\{i : X_i \geqslant \widehat{S}_n^{-1}(\varepsilon/\sqrt{n})\}| \geqslant C) \\ = P(|\{i : \widehat{S}_n(X_i) \geqslant \varepsilon/\sqrt{n}\}| \geqslant C) \\ \geqslant P(|\{i : \widehat{H}_n(X_i) \geqslant \varepsilon/\sqrt{n}\}| \geqslant C) \\ = P\Big(\Big|\Big\{i : \frac{i-1}{n} \geqslant \frac{\varepsilon}{\sqrt{n}}\Big\}\Big| \geqslant C\Big) \\ = \mathbf{1}\Big\{\Big|\Big\{i : \frac{i-1}{n} \geqslant \frac{\varepsilon}{\sqrt{n}}\Big\}\Big| \geqslant C\Big\} = \mathbf{1}\{|\{\lceil\varepsilon\sqrt{n}\rceil + 1, \ldots, n\}| \geqslant C\}.$$

Clearly, this indicator function goes to 1 as $n \to \infty$. □

Proof of Lemma 3 For the most part, we follow the lines of the above proof of Lemma 2 by verifying Condition (3.8) of Theorem 3.2 in Billingsley (1999). To point out the major difference to the previous proof, we consider

$$-\int_{u\wedge v}^{\tau} \frac{\mathrm{d}S_n^*}{G_{n-}^*} = \int_{u\wedge v}^{\tau} \frac{S_{n-}^*\mathrm{d}A_n^*}{G_{n-}^*} = \int_{u\wedge v}^{\tau} \frac{S_{n-}^*}{G_{n-}^*} J^*\mathrm{d}(A_n^* - \widehat{A}_n) + \int_{u\wedge v}^{\tau} \frac{S_{n-}^*}{G_{n-}^*} J^*\mathrm{d}\widehat{A}_n,$$

where $J^*(u) = \mathbf{1}\{Y^*(u) > 0\}$. The arguments of Akritas (1986) show that

$$\int_{u\wedge v}^{\cdot} \frac{S^*_{n-}}{G^*_{n-}} J^* \mathrm{d}(A^*_n - \widehat{A}_n)$$

is a square integrable martingale with predictable variation process given by

$$t \longmapsto \int_{u\wedge v}^{t} \frac{S^{*2}_{n-}}{G^{*2}_{n-}} \frac{J^*}{Y^*} (1 - \Delta \widehat{A}_n) \mathrm{d}\widehat{A}_n.$$

After writing $S^*_n G^*_n = H^*_n$, a twofold application of Lemmata 6 and 7 (at first to the bootstrap quantities S^*_n and H^*_n, then to the Kaplan–Meier estimators $\widehat{S}_n$ and $\widehat{H}_n$) shows that the predictable variation in the previous display is bounded from above by

$$-\beta^{-9} \int_{u\wedge v}^{t} \frac{S^3_-}{H^2_-} \mathrm{d}\widehat{S}_n = -\beta^{-9} \int_{u\wedge v}^{t} \frac{S \mathrm{d}\widehat{S}_n}{G^2_-}$$

on a set with arbitrarily large probability depending on $\beta \in (0, 1)$. Here we also used that $\widehat{S}_{n-} \mathrm{d}\widehat{A}_n = \mathrm{d}\widehat{S}_n$. Due to (5), Theorem 1.1 of Stute and Wang (1993) yields

$$-\int_{u\wedge v}^{t} \frac{S \mathrm{d}\widehat{S}_n}{G^2_-} \xrightarrow{p} -\int_{u\wedge v}^{t} \frac{S \mathrm{d}S}{G^2_-} < \infty$$

and hence the asymptotic negligibility of the predictable variation process in probability. By Rebolledo's theorem (Theorem II.5.1 in Andersen et al. 1993, p. 83), $\int_{u\wedge v}^{\tau} \frac{S^*_{n-}}{G^*_{n-}} J^* \mathrm{d}(A^*_n - \widehat{A}_n)$ hence goes to zero in conditional probability. The remaining integral $\int_{u\wedge v}^{\tau} \frac{S^*_{n-}}{G^*_{n-}} J^* \mathrm{d}\widehat{A}_n$ is treated similarly with Lemmata 6 and 7 and Theorem 1.1 of Stute and Wang (1993) yielding a bound in terms of $\int_{u\wedge v}^{\tau} \frac{\mathrm{d}S}{G_-}$. This is arbitrarily small for sufficiently large $u, v < \tau$. □

Proof of Lemma 4 Let the function spaces $D[t_1, \tau]$ and $\widetilde{D}[t_1, t_2]$ be equipped with the supremum norm. For some sequences $t_n \downarrow 0$ and $h_n \to h$ in $D[t_1, \tau]$ such that $\theta + t_n h_n \in \widetilde{D}[t_1, t_2]$, consider the supremum distance

$$\sup_{s\in[t_1,t_2]} \left| \frac{1}{t_n} [\psi(\theta + t_n h_n)(s) - \psi(\theta)(s)] - (\mathrm{d}\psi(\theta) \cdot h)(s) \right|. \tag{12}$$

The proof is concluded if (12) goes to zero. For an easier access the expression in the previous display is first analyzed for each fixed $s \in [t_1, t_2]$:

$$\frac{1}{t_n} [\psi(\theta + t_n h_n)(s) - \psi(\theta)(s)] - (\mathrm{d}\psi(\theta) \cdot h)(s)$$
$$= \frac{1}{t_n} \frac{1}{\theta(s) + t_n h_n(s)} \frac{1}{\theta(s)}$$

$$
\begin{aligned}
&\times\left[\theta(s)\int_s^\tau(\theta(u)+t_nh_n(u))\mathrm{d}u-(\theta(s)+t_nh_n(s))\int_s^\tau\theta(u)\mathrm{d}u\right]\\
&-\frac{1}{\theta(s)}\int_s^\tau h(u)\mathrm{d}u+h(s)\int_s^\tau\frac{\theta(u)}{\theta^2(s)}\mathrm{d}u\\
=&\frac{1}{\theta(s)+t_nh_n(s)}\int_s^\tau h_n(u)\mathrm{d}u-\frac{h_n(s)}{\theta(s)+t_nh_n(s)}\frac{1}{\theta(s)}\int_s^\tau\theta(u)\mathrm{d}u\\
&-\frac{1}{\theta(s)}\int_s^\tau h_n(u)\mathrm{d}u+h_n(s)\int_s^\tau\frac{\theta(u)}{\theta^2(s)}\mathrm{d}u\\
&-\frac{1}{\theta(s)}\int_s^\tau(h(u)-h_n(u))\mathrm{d}u-(h(s)-h_n(s))\int_s^\tau\frac{\theta(u)}{\theta^2(s)}\mathrm{d}u\\
=&-\int_s^\tau h_n(u)\mathrm{d}u\frac{t_nh_n(s)}{[\theta(s)+t_nh_n(s)]\theta(s)}+h_n(s)\int_s^\tau\frac{\theta(u)}{\theta^2(s)}\mathrm{d}u\frac{t_nh_n(s)}{\theta(s)+t_nh_n(s)}\\
&-\frac{1}{\theta(s)}\int_s^\tau(h(u)-h_n(u))\mathrm{d}u-(h(s)-h_n(s))\int_s^\tau\frac{\theta(u)}{\theta^2(s)}\mathrm{d}u. \qquad (13)
\end{aligned}
$$

For large n, each denominator is bounded away from zero: To see this, denote $\varepsilon := \inf_{s\in[t_1,t_2]}|\theta(s)|$ and $C := \sup_{s\in[t_1,t_2]}|h(u)|$. Thus,

$$
\sup_{s\in[t_1,t_2]}|h_n(s)| \leqslant \sup_{s\in[t_1,t_2]}|h_n(s)-h(u)| + \sup_{s\in[t_1,t_2]}|h(s)| \leqslant \varepsilon + C
$$

for each n large enough. It follows that, for each such n additionally satisfying $t_n \leqslant \varepsilon(2\varepsilon+2C)^{-1}$, the denominators are bounded away from zero, in particular, $\inf_{s\in[t_1,t_2]}|\theta(s)+t_nh_n(s)| \geqslant \varepsilon/2$. Thus, taking the suprema over $s\in[t_1,t_2]$, the first two terms in (13) become arbitrarily small by letting t_n be sufficiently small. The remaining two terms converge to zero since $\sup_{s\in[t_1,t_2]}|h_n(s)-h(s)| \to 0$ and $\sup_{u\in[t_1,\tau]}|\theta(u)| < \infty$. Note here that

$$
\int_{t_1}^\tau |h(u)-h_n(u)|\mathrm{d}u \leqslant \sup_{u\in[t_1,\tau]}|h(u)-h_n(u)|(\tau-t_1) \to 0.
$$

□

Proof of Lemma 5 The convergences are immediate consequences of the functional delta method, Theorem 1 and the bootstrap version of the delta method; cf. Sect. 3.9 in van der Vaart and Wellner (1996). Simply note that all considered survival functions are elements of $D_{<\infty} \cap \widetilde{D}[t_1, t_2]$ (on increasing sets with probability tending to one) and that the survival function of the lifetimes is assumed continuous and bounded away from zero on compact subsets of $[0, \tau)$. Further, there is a version of the limit Gaussian processes with almost surely continuous sample paths.

For the representation of the variance of the limit distribution in part (a) we refer to van der Vaart and Wellner (1996), p. 383 and 397. The asymptotic covariance structure in part (b) is easily calculated using Fubini's theorem—for its applicability note that the variances $\Gamma(r, r)$ of the limit process W of the Kaplan–Meier estimator exist at all points of time $r \in [0, \tau]$. Thus, since W is a zero-mean process, we have for any $0 \leqslant r \leqslant s < \tau$,

$$
\begin{aligned}
&cov\left(\int_r^\tau \frac{W(u)}{S(r)}\mathrm{d}u - \int_r^\tau \frac{W(r)S(u)}{S^2(r)}\mathrm{d}u, \int_s^\tau \frac{W(v)}{S(s)}\mathrm{d}v - \int_s^\tau \frac{W(s)S(v)}{S^2(s)}\mathrm{d}v\right)\\
&= \int_r^\tau \int_s^\tau \left[\Gamma(u,v) - \frac{S(u)}{S(r)}\Gamma(r,v) - \frac{S(v)}{S(s)}\Gamma(s,u) + \frac{S(u)S(v)}{S(r)S(s)}\Gamma(r,s)\right]\frac{\mathrm{d}u\mathrm{d}v}{S(r)S(s)}.
\end{aligned}
$$

Inserting the definition $\Gamma(r,s) = S(r)S(s)\sigma^2(r\wedge s)$ and splitting the first integral into $\int_r^\tau = \int_r^s + \int_s^\tau$ yields that the last display equals

$$
\begin{aligned}
&\int_r^\tau \int_s^\tau \frac{S(u)S(v)}{S(r)S(s)}[\sigma^2(u\wedge v) - \sigma^2(r\wedge v) - \sigma^2(s\wedge u) + \sigma^2(r\wedge s)]\mathrm{d}u\mathrm{d}v\\
&= \int_s^\tau \int_s^\tau \frac{S(u)S(v)}{S(r)S(s)}[\sigma^2(u\wedge v) - \sigma^2(r) - \sigma^2(s) + \sigma^2(r)]\mathrm{d}u\mathrm{d}v\\
&\quad + \int_r^s \int_s^\tau \frac{S(u)S(v)}{S(r)S(s)}[\sigma^2(u) - \sigma^2(r) - \sigma^2(u) + \sigma^2(r)]\mathrm{d}u\mathrm{d}v\\
&= \int_s^\tau \int_s^\tau \frac{\Gamma(u,v)}{S(r)S(s)}\mathrm{d}u\mathrm{d}v - \sigma^2(r\vee s)g(r)g(s).
\end{aligned}
$$

□

Proof of Theorem 2 The theorem follows from Lemma 5 combined with the continuous mapping theorem applied to the supremum functional $D[t_1,t_2] \to \mathbb{R}$, $f \mapsto \sup_{t\in[t_1,t_2]}|f(t)|$ which is continuous on $C[t_1,t_2]$. For the connection between the consistency of a bootstrap distribution of a real statistic and the consistency of the corresponding tests (and the equivalent formulation in terms of confidence regions), see Lemma 1 in Janssen and Pauls (2003). □

B Adaptations of Gill's (1983) Lemmata

Abbreviate again the sigma algebra containing all the information of the original sample as $\mathbf{X} := \sigma(X_i, \delta_i : i = 1, \ldots, n)$. The proofs in "Appendix A" rely on bootstrap versions of Lemmata 2.6, 2.7 and 2.9 in Gill (1983). Since those are stated under the assumption of a continuous distribution function S, but ties in the bootstrap sample are inevitable, these lemmata need a slight extension. For completeness, parts (a) of the following two Lemmata correspond to the original Lemmata 2.6 and 2.7 in Gill (1983).

Lemma 6 *(Extension of Lemma 2.6 in Gill 1983) For any* $\beta \in (0,1)$,

(a) $P(\widehat{S}_n(t) \leqslant \beta^{-1}S(t)$ *for all* $t \leqslant \widehat{T}_n) \geqslant 1-\beta$,
(b) $P(S_n^*(t) \leqslant \beta^{-1}\widehat{S}_n(t)$ *for all* $t \leqslant T_n^* \mid \boldsymbol{X}) \geqslant 1-\beta$ *almost surely.*

Proof of (b). All equalities and inequalities concerning conditional expectations are understood as to hold almost surely. As in the proof of Theorem 1, $(S_n^*(t\wedge T_n^*)/\widehat{S}_n(t\wedge T_n^*))_{t\in[0,\widehat{T}_n)}$ defines a right-continuous martingale for each fixed n and for almost every given sample $\mathbf{X}$. Hence, Doob's L_1-inequality (e.g., Revuz and Yor 1999, Theorem 1.7 in Chapter II) yields for each $\beta \in (0,1)$

$$P\left(\sup_{t\in[0,\widehat{T}_n)} S_n^*(t\wedge T_n^*)/\widehat{S}_n(t\wedge T_n^*) \geqslant \beta^{-1} \mid \mathbf{X}\right)$$
$$\leqslant \beta \sup_{t\in[0,\widehat{T}_n)} E(S_n^*(t\wedge T_n^*)/\widehat{S}_n(t\wedge T_n^*) \mid \mathbf{X})$$
$$= \beta E(S_n^*(0)/\widehat{S}_n(0) \mid \mathbf{X}) = \beta.$$

This implies $P(S_n^* \leqslant \beta^{-1}\widehat{S}_n \text{ on } [0, T_n^*) \mid \mathbf{X}) \geqslant 1-\beta$. It remains to extend this result to the interval's endpoint. If the observation corresponding to T_n^* is uncensored, we have $0 = S_n^*(T_n^*) \leqslant \beta^{-1}\widehat{S}_n(T_n^*)$. Else, the event of interest $\{S_n^* \leqslant \beta^{-1}\widehat{S}_n \text{ on } [0, T_n^*)\}$ (given $\mathbf{X}$) implies that

$$S_n^*(T_n^*) = S_n^*(T_n^*-) \leqslant \beta^{-1}\widehat{S}_n(T_n^*-) = \beta^{-1}\widehat{S}_n(T_n^*).$$

Thus, for given $\mathbf{X}$, $\{S_n^*(T_n^*) \leqslant \beta\widehat{S}_n(T_n^*)\} \subset \{S_n^* \leqslant \beta\widehat{S}_n \text{ on } [0, T_n^*)\}$. □

Lemma 7 *(Extension of Lemma 2.7 in Gill 1983) For any* $\beta \in (0,1)$*,*

(a) $P(\widehat{H}_n(t-) \geqslant \beta H(t-)$ *for all* $t \leqslant \widehat{T}_n) \geqslant 1 - \frac{e}{\beta}\exp(-1/\beta)$,

(b) $P(H_n^*(t-) \geqslant \beta \widehat{H}_n(t-)$ *for all* $t \leqslant T_n^* \mid \boldsymbol{X}) \geqslant 1 - \frac{e}{\beta}\exp(-1/\beta)$ *almost surely.*

Proof of (a). As pointed out by Gill (1983), the assertion follows from the inequality for the uniform distribution in Remark 1(ii) of Wellner (1978). By using quantile transformations, his inequality can be shown to hold for random variables having an arbitrary, even discontinuous distribution function.

Proof of (b). Fix $X_i(\omega), \delta_i(\omega), i = 1, \ldots, n$. Since H in part (a) is allowed to have discontinuities, (b) follows from (a) for each ω. □

Let $a, b \in D[0,\tau]$ be two (stochastic) jump processes, i.e., processes being constant between two discontinuities. If b has bounded variation, we define the integral of a with respect to b via

$$\int_0^s a\mathrm{d}b = \sum a(t)\Delta b(t), \quad s \in (0,\tau],$$

where the sum is over all discontinuities of b inside the interval $(0, s]$. If a has bounded variation, we define the above integral via integration by parts: $\int_0^s a\mathrm{d}b = a(s)b(s) - a(0)b(0) - \int_0^s b_-\mathrm{d}a$.

Lemma 8 *(Adaptation of Lemma 2.9 in Gill 1983) Let* $h \in D[0,\tau]$ *be a nonnegative and non-increasing jump process such that* $h(0) = 1$ *and let* $Z \in D[0,\tau]$ *be a jump process which is zero at time zero. Then, for all* $t \leqslant \tau$*,*

$$\sup_{s\in[0,t]} h(s)|Z(s)| \leqslant 2 \sup_{s\in[0,t]} \left|\int_0^s h(u)\mathrm{d}Z(u)\right|.$$

Proof The original proof of Lemma 2.9 in Gill (1983) still applies for the most part with the assumptions of this lemma. For the sake of completeness, we present the whole proof.

Let $U(t) = \int_0^t h(s)\mathrm{d}Z(s)$ with a $t \leqslant \tau$ such that $h(t) > 0$. Then,

$$Z(t) = \int_0^t \frac{\mathrm{d}U(s)}{h(s)} = \frac{U(t)}{h(t)} - \int_0^t U(s-)\mathrm{d}\Big(\frac{1}{h(s)}\Big)$$
$$= \int_0^t (U(t) - U(s-))\mathrm{d}\Big(\frac{1}{h(s)}\Big) + \frac{U(t)}{h(0)}.$$

Thus, following the lines of the original proof,

$$|h(t)Z(t)| \leqslant \Big|\int_0^t (U(t) - U(s-))\mathrm{d}\Big(\frac{h(t)}{h(s)}\Big)\Big| + |U(t)|h(t)$$
$$\leqslant 2 \sup_{0<s\leqslant t} |U(s)|\Big(1 - \frac{h(t)}{h(0)}\Big) + \sup_{0<s\leqslant t} |U(s)|h(t) \leqslant 2 \sup_{0<s\leqslant t} |U(s)|.$$

□

C Bootstrap version of the truncation technique for weak convergence

The following lemma is a conditional variant of Theorem 3.2 in Billingsley (1999). Let ρ be the modified Skorohod metric J_1 on $D[0,\tau]$ as in Billingsley (1999), i.e., $\rho(f,g) = \inf_{\lambda\in\Lambda}(\|\lambda\|^o \vee \sup_{t\in[0,\tau]} |f(t) - g(\lambda(t))|)$, where Λ is the collection of non-decreasing functions onto $[0,\tau]$ and

$$\|\lambda\|^o = \sup_{s\neq t}\left|\log \frac{\lambda(s) - \lambda(t)}{s - t}\right|.$$

For an application in the proof of Theorem 1, note that $\rho(f,g) \leqslant \sup_{t\in[0,\tau]} |f(t) - g(t)|$.

Lemma 9 *Let $X : (\Omega, \mathcal{A}, P) \to (D[0,\tau], \rho)$ be a stochastic process and let the sequences of stochastic processes X_{un} and X_n satisfy the following convergences given a σ-algebra $\mathcal{C}$:*

(a) $X_{un} \xrightarrow{d} Z_u$ given $\mathcal{C}$ in probability as $n \to \infty$ for every fixed u,
(b) $Z_u \xrightarrow{d} X$ given $\mathcal{C}$ in probability as $u \to \infty$,
(c) for all $\varepsilon > 0$ and for each subsequence $(n') \subset (n)$ there exists another subsequence $(n'') \subset (n')$ such that

$$\lim_{u\to\infty} \limsup_{n''\to\infty} P(\rho(X_{un''}, X_{n''}) > \varepsilon \mid \mathcal{C}) = 0 \quad \textit{almost surely.}$$

Then, $X_n \xrightarrow{d} X$ given $\mathcal{C}$ in probability as $n \to \infty$.

Proof Choose a sequence $\varepsilon_m \downarrow 0$. Let $(n') \subset (n)$ be an arbitrary subsequence and choose subsequences $(n''(\varepsilon_m)) \subset (n')$ and $(u') \subset (u)$ such that (a) and (b) hold almost surely and also such that (c) holds along these subsequences. Replace $(n''(\varepsilon_m))$ by their diagonal sequence (n'') ensuring (c) simultaneously for all ε_m. Let $F \subset D[0,\tau]$ be a closed subset and let $F_{\varepsilon_m} = \{f \in D[0,\tau] : \rho(f,F) \leqslant \varepsilon_m\}$ be its closed ε_m-enlargement. We proceed as in the proof of Theorem 3.2 in Billingsley (1999), whereas all inequalities now hold almost surely.

$$P(X_{n''} \in F \mid \mathcal{C}) \leqslant P(X_{u'n''} \in F_{\varepsilon_m} \mid \mathcal{C}) + P(\rho(X_{u'n''}, X_{n''}) > \varepsilon_m \mid \mathcal{C}).$$

The Portmanteau theorem in combination with (a) yields

$$\limsup_{n''\to\infty} P(X_{n''} \in F \mid \mathcal{C}) \leqslant P(Z_{u'} \in F_{\varepsilon_m} \mid \mathcal{C}) + \limsup_{n''\to\infty} P(\rho(X_{u'n''}, X_{n''}) > \varepsilon_m \mid \mathcal{C}).$$

Condition (b) and another application of the Portmanteau theorem imply that

$$\limsup_{n''\to\infty} P(X_{n''} \in F \mid \mathcal{C}) \leqslant P(X \in F_{\varepsilon_m} \mid \mathcal{C}).$$

Let $m \to \infty$ to deduce $\limsup_{n''\to\infty} P(X_{n''} \in F \mid \mathcal{C}) \leqslant P(X \in F \mid \mathcal{C})$ almost surely. Thus, a final application of Portmanteau theorem as well as the subsequence principle leads to the conclusion that $X_n \xrightarrow{d} X$ given $\mathcal{C}$ in probability. □

D Further simulation results for mean residual lifetime confidence bands

As suggested by a referee, we now evaluate the influence of the choice of the underlying time span $[t_1, t_2]$ on the coverage probabilities of the confidence bands for the mean residual lifetime function. To this end, we again carried out the simulations of Sect. 5.2 while allowing for any combination of the above-chosen time intervals in all setups, i.e., with underlying exponential and Gompertz distributions with any of the three scenarios (i)–(iii). From these results, we conclude that the choice of time span does not have a great impact on the linear bands' coverage probabilities. In the exponential setup (Table 5), these empirical probabilities vary only slightly, whereas broader intervals $[t_1, t_2]$ appear to make the bands more conservative in the case of an underlying Gompertz distribution (Table 6).

The behavior of the log-transformed confidence bands is quite surprising: The empirical coverage probabilities deteriorate considerably for small sample sizes in the case of increased interval lengths $t_2 - t_1$: They may drop by several dozens of percentage points (see the Exponential setup (i) and $n \in \{30, 50\}$) or even from 94.9 to 0.3% (see the Gompertz setup (i) and $n = 50$) (Table 6).

We will find the reason for the sometimes good, sometimes dramatic behavior of the confidence bands in the eventual survival probabilities; Table 7 shows the survival probabilities at the right boundaries of the time intervals for each considered setup. Even in the case of a final survival probability of about 6% and sample sizes $n \geqslant 500$, the linear confidence bands have satisfactory empirical coverage probabilities within 95.2 and 96.2%. On the other hand, the reason for the bad performance of both types

Table 5 Simulated coverage probabilities of asymptotic 95% confidence bands for the mean residual lifetime function under light (left), medium (middle), and strong (right) censoring for distribution setup 1

Time span	[0, 1]		[0, 1.2]		[0, 1.4]		[0, 1]		[0, 1.2]		[0, 1.4]		[0, 1]		[0, 1.2]		[0, 1.4]	
n	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log
30	94.4	61.7	94.8	40.8	94.6	25.1	93.0	74.9	93.1	56.9	93.2	40.0	92.7	81.9	92.5	67.0	92.6	51.0
50	94.6	79.9	94.4	58.5	95.0	38.6	94.2	89.2	93.5	74.8	94.1	58.1	93.3	91.5	93.2	83.6	93.0	70.3
100	95.3	94.2	94.4	81.5	95.0	62.1	95.0	95.9	94.0	91.9	94.3	81.4	94.2	95.2	94.0	94.1	93.5	88.1
150	95.8	96.3	95.2	91.2	94.7	76.3	95.2	96.3	94.7	95.6	94.6	90.7	94.7	95.4	94.8	95.8	94.2	94.2
200	95.7	96.6	95.6	94.5	94.8	84.4	95.4	95.9	94.9	95.8	94.5	93.9	95.2	95.5	94.6	95.5	94.5	95.3
500	95.6	96.3	95.7	96.4	95.3	96.0	95.9	95.7	95.5	95.8	95.5	96.2	95.1	95.1	95.0	95.0	95.1	95.5
1000	95.7	95.6	95.2	95.8	95.3	96.2	95.5	95.5	95.4	95.5	95.2	95.5	94.8	94.7	95.0	95.0	95.4	95.5

Table 6 Simulated coverage probabilities of asymptotic 95% confidence bands for the mean residual lifetime function under light (left), medium (middle), and strong (right) censoring for distribution setup 2

Time span	[0, 1.2]		[0, 2.0]		[0, 2.3]		[0, 1.2]		[0, 2.0]		[0, 2.3]		[0, 1.2]		[0, 2.0]		[0, 2.3]	
n	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log	Linear	Log
30	96.6	84.0	96.6	2.4	96.7	0.2	94.9	96.9	96.1	87.5	96.3	58.3	92.6	96.6	94.9	98.4	95.6	84.5
50	97.0	94.9	97.7	4.0	97.7	0.3	95.1	96.8	96.6	96.5	97.1	75.8	92.8	95.7	94.5	98.2	96.1	95.2
100	97.3	98.6	98.1	8.1	98.0	0.5	94.6	95.6	96.7	98.3	97.0	94.3	93.7	95.1	94.9	96.8	95.8	98.5
150	96.7	98.2	98.1	11.9	98.2	0.9	94.9	95.6	96.8	97.6	96.9	97.6	93.6	95.0	94.6	96.4	96.1	98.2
200	96.7	97.7	98.2	15.1	98.3	1.2	94.9	95.2	96.6	97.5	97.1	98.2	94.0	95.0	94.6	96.0	95.8	97.6
500	95.5	96.0	98.5	33.4	98.2	3.2	94.6	94.7	95.6	95.8	96.2	97.1	94.4	94.8	94.5	94.8	95.2	95.8
1000	95.3	95.5	98.6	55.7	98.5	5.4	94.7	95.3	95.1	95.3	96.0	96.1	94.7	94.8	94.7	95.3	95.0	95.3

Table 7 Survival probabilities (in %) at the right boundary of the time interval

Distribution \ Setup	(i)			(ii)			(iii)		
End time	1.0	1.2	1.4	1.0	1.2	1.4	1.0	1.2	1.4
Exponential	13.53	9.07	6.08	22.31	16.53	12.25	30.12	23.69	18.64
End time	1.2	2.0	2.3	1.2	2.0	2.3	1.2	2.0	2.3
Gompertz	9.83	0.17	0.01	49.86	14.71	6.77	81.84	34.23	13.95

of confidence bands in setup (i) with an underlying Gompertz distribution is the very small survival probability at the right boundary of the time interval: With survival probabilities between 0.01 and 0.17%, one cannot clearly see the asymptotic exactness of the confidence bands, even with sample sizes as big as $n = 1000$. The linear bands tend to be quite conservative, whereas the log-transformed bands are very liberal. Much larger samples are necessary to get reliable confidence bands in this extreme setup. These results also indicate that the asymptotic behavior breaks down whenever one seeks to derive confidence bands for the mean residual lifetime function on its whole support.

These observations again lead us to the conclusion that the linear confidence bands appear to be the best choice for the mean residual lifetime function. Depending on the length of the underlying time interval, it may be slightly conservative. But this improves with an increasing sample size if the survival probability at the right boundary is not excessively small.

References

Akritas, M. G. (1986). Bootstrapping the Kaplan–Meier Estimator. *Journal of the American Statistical Association*, *81*(396), 1032–1038.

Akritas, M. G., Brunner, E. (1997). Nonparametric methods for factorial designs with censored data. *Journal of the American Statistical Association*, *92*(438), 568–576.

Allignol, A., Beyersmann, J., Gerds, T., Latouche, A. (2014). A competing risks approach for nonparametric estimation of transition probabilities in a non-Markov illness-death model. *Lifetime Data Analysis*, *20*(4), 495–513.

Andersen, P. K., Borgan, Ø., Gill, R. D., Keiding, N. (1993). *Statistical models based on counting processes*. New York: Springer.

Billingsley, P. (1999). *Convergence of probability measures* (2nd ed.). New York: Wiley.

Dobler, D. (2016). *Nonparametric inference procedures for multi-state Markovian models with applications to incomplete life science data*. PhD thesis Universität Ulm, Deutschland.

Dobler, D., Pauly, M. (2017). Bootstrap- and permutation-based inference for the Mann–Whitney effect for right-censored and tied data. *TEST* early view:1–20. https://doi.org/10.1007/s11749-017-0565-z.

Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, *76*(374), 312–319.

Gastwirth, J. L. (1971). A general definition of the Lorenz curve. *Econometrica*, *39*(6), 1037–1039.

Gill, R. D. (1980). *Censoring and stochastic integrals*. Mathematical Centre Tracts 124, Amsterdam: Mathematisch Centrum.

Gill, R. D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *The Annals of Statistics*, *11*(1), 49–58.

Gill, R. D. (1989). Non- and semi-parametric maximum likelihood estimators and the von mises method (part 1) [with discussion and reply]. *Scandinavian Journal of Statistics*, *16*(2), 97–128.

Grand, M. K., Putter, H. (2016). Regression models for expected length of stay. *Statistics in Medicine*, *35*(7), 1178–1192.
Horvath, L., Yandell, B. (1987). Convergence rates for the bootstrapped product-limit process. *The Annals of Statistics*, *15*(3), 1155–1173.
Janssen, A., Pauls, T. (2003). How do bootstrap and permutation tests work? *The Annals of Statistics*, *31*(3), 768–806.
Klein, J. P., Moeschberger, M. L. (2003). *Survival analysis*: *Techniques for censored and truncated data* (2nd ed.). New York: Springer.
Lo, S.-H., Singh, K. (1986). The product-limit estimator and the bootstrap: Some asymptotic representations. *Probability Theory and Related Fields*, *71*(3), 455–465.
Loprinzi, C. L., Laurie, J. A., Wieand, H. S., Krook, J. E., Novotny, P. J., Kugler, J. W., et al. (1994). Prospective evaluation of prognostic variables from patient-completed questionnaires. *Journal of Clinical Oncology*, *12*(3), 601–607.
Meilijson, I. (1972). Limiting properties of the mean residual lifetime function. *The Annals of Mathematical Statistics*, *43*(1), 354–357.
Pocock, S. J., Ariti, C. A., Collier, T. J., Wang, D. (2012). The win ratio: A new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *European Heart Journal*, *33*(2), 176–182.
Pollard, D. (1984). *Convergence of stochastic processes*. New York: Springer.
R Development Core Team. (2016). *R*: *A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. http://www.R-project.org.
Revuz, D., Yor, M. (1999). *Continuous martingales and Brownian motion* (3rd ed.). Berlin: Springer.
Stute, W., Wang, J.-L. (1993). The strong law under random censorship. *The Annals of Statistics*, *21*(3), 1591–1607.
Tattar, P., Vaman, H. (2012). Extension of the Harrington–Fleming tests to multistate models. *Sankhya B*, *74*(1), 1–14.
Therneau, T. M., Lumley, T. (2017). *A package for survival analysis in S*. http://CRAN.R-project.org/package=survival. version 2.41-3.
Tse, S.-M. (2006). Lorenz curve for truncated and censored data. *Annals of the Institute of Statistical Mathematics*, *58*(4), 675–686.
van der Vaart, A. W., Wellner, J. (1996). *Weak convergence and empirical processes*. New York: Springer.
Wang, J.-G. (1987). A note on the uniform consistency of the Kaplan–Meier estimator. *The Annals of Statistics*, *15*(3), 1313–1316.
Wellner, J. A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, *45*(1), 73–88.
Ying, Z. (1989). A note on the asymptotic properties of the product-limit estimator on the whole line. *Statistics & Probability Letters*, *7*(4), 311–314.