Consistency of bootstrap procedures for the nonparametric assessment of noninferiority with random censorship

Gudrun Freitag[1] and Axel Munk
Institut für Mathematische Stochastik
Georg-August-Universität Göttingen,
Germany

Abstract

In this paper we consider general Hadamard differentiable functionals $\phi(\Lambda_R, \Lambda_T)$ of the cumulative hazard functions of two samples of randomly right censored data, which can be used for the nonparametric assessment of noninferiority. We prove the consistency of various bootstrap procedures as suggested in Freitag et al. [1] for the practical implementation of tests for this problem.

Keywords: Noninferiority, Relative risk, Odds ratio, Hazard ratio, Hadamard differentiability, Weak convergence, Bootstrap

1 Introduction

We consider the problem of showing that a new, *test* treatment T is at most irrelevantly inferior to an established standard, *reference* treatment R, with respect to the corresponding survival probabilities. As to the motivation and specific methodological concerns related to noninferiority trials, we refer to the special issue of *Statistics in Medicine* on this topic in 2003 [2] and to a number of (draft) guidelines of the Committee for Proprietary Medicinal Products (CPMP) and the International Conference of Harmonization (ICH), such as [3]-[7]. A survey on existing methods for the assessment of noninferiority with censored data can be found in Freitag [8].

This paper deals with the approach suggested in Freitag et al. [1], where the comparison of the two samples is performed nonparametrically, using discrepancy functionals of the two underlying cumulative hazard functions, such as the cumulative hazard ratio or the cumulative odds ratio. These can be assessed over a whole time interval within the follow-up period of a clinical trial.

For this we assume the setting of two samples of failure times, which are subject to simple random right censoring. However, other censoring mechanisms could be treated as well (cf. the discussion of this issue in Freitag et al. [1]). We denote the cumulative distribution function (cdf) of a failure time under the reference and test treatment by F_R and F_T, respectively. Throughout the following we assume that the F_k are continuous cdfs with Lebesgue densities f_k, survivor functions $S_k = 1 - F_k$, and continuous hazard functions $\lambda_k = f_k/S_k$, $k = R, T$.

Let $T_{ki} \geq 0$, $i = 1, ..., n_k$, be independent and identically distributed (i.i.d.) failure times according to F_k, $k = R, T$. Further, let the corresponding censoring times $U_{ki} \geq 0$ be distributed

[1]Address for correspondence: Gudrun Freitag, Institut für Mathematische Stochastik, Georg-August-Universität Göttingen, Maschmühlenweg 8-10, 37073 Göttingen, Germany. Tel: +49/551/3913510, Fax: +49/551/3913505. Email: freitag@math.uni-goettingen.de.

according to G_k, where U_{ki} is independent of T_{ki}, $i = 1, ..., n_k$, $k = R, T$. The observations consist of pairs (X_{ki}, δ_{ki}), where $X_{ki} = T_{ki} \wedge U_{ki} \stackrel{\text{def}}{=} \min\{T_{ki}, U_{ki}\}$ are the observed failure times and $\delta_{ki} = I_{\{T_{ki} \leq U_{ki}\}}$ are the associated observable censoring indicators, $i = 1, ..., n_k$, $k = R, T$. For estimating the underlying cumulative hazard functions Λ_k and distribution functions F_k, we will use the standard nonparametric Nelson-Aalen and Kaplan-Meier estimators, respectively, which shall be denoted by $\hat{\Lambda}_k$ and $\hat{F}_k$, $k = R, T$.

For comparing the two samples of failure times nonparametrically, one can consider the difference of the survival functions, the cumulative odds ratio or the ratio of the survival functions, or the ratio of the cumulative hazard functions (cf. Remark 2.3). We will deal with these in a unified framework, using a general discrepancy functional $\phi = \phi(\Lambda_R, \Lambda_T)$ of the underlying cumulative hazard functions. Note that the survivor function S_k can be written as $S_k(t) = \prod_0^t [1 - d\Lambda_k(s)]$, where $\prod_0^t$ denotes the product integral (cf. Andersen et al. [9]). An analogous relation holds for the estimated survivor function $\hat{S}_k = 1 - \hat{F}_k$, i.e. $\hat{S}_k(t) = \prod_0^t [1 - d\hat{\Lambda}_k(s)]$.

Using these notations, the difference functional can be written as

$$d(t) \stackrel{\text{def}}{=} F_T(t) - F_R(t) = \prod_0^t [1 - d\Lambda_R(s)] - \prod_0^t [1 - d\Lambda_T(s)] \stackrel{\text{def}}{=} \phi_d(\Lambda_R, \Lambda_T)(t). \tag{1}$$

Here the subscript d in $\phi_d(\Lambda_R, \Lambda_T)$ indicates the concrete discrepancy functional $d(\cdot)$.

Freitag et al. [1] argue that for the nonparametric assessment of noninferiority of the test treatment as compared to the reference treatment, it is of interest to assess such a discrepancy over a whole time interval, instead of only at a single time point. Thus, in terms of a general discrepancy functional ϕ, we suggest hypotheses of the form

$$H_\phi : \phi(\Lambda_R, \Lambda_T)(t) \geq \Delta_\phi \text{ for some } t \in [\tau_0, \tau_1] \quad vs. \quad K_\phi : \phi(\Lambda_R, \Lambda_T)(t) < \Delta_\phi \, \forall t \in [\tau_0, \tau_1], \tag{2}$$

where Δ_ϕ is a fixed irrelevance bound which has to be defined in advance. For absolute discrepancy measures (such as the difference) this will be a small positive value, whereas for relative discrepancy measures such as the cumulative odds ratio it will be a value slightly larger than 1. For testing the above hypotheses we will use the nonparametric estimator $\hat{\phi}(\Lambda_R, \Lambda_T) = \phi(\hat{\Lambda}_R, \hat{\Lambda}_T)$.

In the following we will show the (weak) consistency of the bootstrap based methods suggested in Freitag et al. [1].

2 Weak convergence

An essential prerequisite is the weak convergence (denoted by $\stackrel{\mathcal{D}}{\Longrightarrow}$) of the two underlying cumulative hazard processes, $n_k^{1/2}(\hat{\Lambda}_k - \Lambda_k)$, $k = R, T$, in the space $\mathbf{D}[\tau_0, \tau_1]$ of càdlàg functions on $[\tau_0, \tau_1]$, which will be considered as equipped with the supremum norm $\|\cdot\|_\infty$ and the sigma field of open balls (cf. e.g. Shorack & Wellner [10] or Gill [11]).

Assumption 2.1 Suppose $\Lambda_k(t) < \infty$ for $t \leq \tau_1$, and $1 - (1 - F_k(\tau_1))(1 - G_k(\tau_1)) < 1$ for $k = R, T$.

Under Assumption 2.1, it is well known that, in $\mathbf{D}[\tau_0, \tau_1]$,

$$n_k^{1/2}(\hat{\Lambda}_k - \Lambda_k) \stackrel{\mathcal{D}}{\Longrightarrow} \mathbb{Z}_k, \quad k = R, T, \tag{3}$$

where $\mathbb{Z}_k$ is a mean zero Gaussian process with $\mathbb{Z}_k(0) = 0$ and covariance structure $\mathbf{COV}(\mathbb{Z}_k(s_1), \mathbb{Z}_k(s_2)) = D_k(s_1 \wedge s_2)$, with

$$D_k(t) = \int_0^t \frac{\lambda_k(s)}{(1 - F_k(s))(1 - G_k(s))} ds \tag{4}$$

(cf. Theorem IV.1.2 in Andersen et al. [9]).

The following result on the weak convergence of the discrepancy process follows immediately from the functional delta method for Hadamard differentiable functionals (cf. e.g. Gill [11]).

Lemma 2.2 *Suppose that the functional $\phi : \mathbf{D}[\tau_0, \tau_1] \times \mathbf{D}[\tau_0, \tau_1] \to \mathbf{D}[\tau_0, \tau_1]$ is Hadamard differentiable at (Λ_R, Λ_T) with derivative $d\phi_{(\Lambda_R, \Lambda_T)}$. Further, assume that Assumption 2.1 holds. Then, if*

$$n_R, n_T \to \infty, \; n_T/(n_R + n_T) \to \rho \in (0, 1), \tag{5}$$

we have that, in $\mathbf{D}[\tau_0, \tau_1]$,

$$\sqrt{\kappa_n}(\hat{\phi}(\Lambda_R, \Lambda_T) - \phi(\Lambda_R, \Lambda_T)) \stackrel{\mathcal{D}}{\Longrightarrow} d\phi_{(\Lambda_R, \Lambda_T)}(\sqrt{\rho}\,\mathbb{Z}_R, \sqrt{1-\rho}\,\mathbb{Z}_T) \stackrel{def}{=} \mathbb{W}_\phi, \tag{6}$$

where $\kappa_n \stackrel{def}{=} n_R n_T/(n_R + n_T)$ and $\mathbb{Z}_k$ is defined in (3), $k = R, T$.

Thus, the limiting random element $\mathbb{W}_\phi$ is a mean zero Gaussian process, and at each $t \in [\tau_0, \tau_1]$ we have asymptotic normality of $\mathbb{W}_\phi(t)$ with expectation zero and variance $\sigma^2_\phi(t)$, which depends on the underlying failure time and censoring distributions.

Remark 2.3 In the following we consider several special cases of the functional ϕ.

1. In case of the difference functional d from (1) the Hadamard differentiability is given immediately, using Assumption 2.1, the continuity of Λ_k, $k = R, T$, and Lemma 3.9.30 of van der Vaart & Wellner [12], which gives the Hadamard differentiability of the product integral. The covariance structure of the limiting random element $\mathbb{W}_d$ is given by

   $$\mathbf{COV}(\mathbb{W}_d(s), \mathbb{W}_d(t)) = (1 - \rho)S_T(s)S_T(t)D_T(s \wedge t) - \rho S_R(s)S_R(t)D_R(s \wedge t),$$

   where D_k, $k = R, T$, and ρ are defined in (4) and (5), respectively. Here the independence of the two processes $\mathbb{Z}_k$, $k = R, T$, was used.

2. For the relative risk functional

   $$r(t) \stackrel{\text{def}}{=} F_T(t)/F_R(t) = \frac{1 - \prod_0^t [1 - d\Lambda_T(s)]}{1 - \prod_0^t [1 - d\Lambda_R(s)]} \stackrel{\text{def}}{=} \phi_r(\Lambda_R, \Lambda_T)(t),$$

   the additional assumption $F_R(t) > 0$, $t \in [\tau_0, \tau_1]$, is required. The covariance structure of the corresponding limiting random element $\mathbb{W}_r$ is then given by

   $$\mathbf{COV}(\mathbb{W}_r(s), \mathbb{W}_r(t)) = (1 - \rho)\frac{S_T(s)S_T(t)}{F_R(s)F_R(t)}D_T(s \wedge t) - \rho\frac{F_T(s)S_R(s)F_T(t)S_R(t)}{F_R^2(s)F_R^2(t)}D_R(s \wedge t).$$

3. For the cumulative odds ratio functional,

$$o(t) \stackrel{\text{def}}{=} \frac{F_T(t)(1-F_R(t))}{(1-F_T(t))F_R(t)} = \frac{\mathop{\mathcal{P}}\limits_0^t [1-d\Lambda_T(s)](1-\mathop{\mathcal{P}}\limits_0^t [1-d\Lambda_R(s)])}{(1-\mathop{\mathcal{P}}\limits_0^t [1-d\Lambda_T(s)])\mathop{\mathcal{P}}\limits_0^t [1-d\Lambda_R(s)]} \stackrel{\text{def}}{=} \phi_o(\Lambda_R,\Lambda_T)(t),$$

we need the additional assumption $F_R(t)(1-F_T(t)) > 0,\ t \in [\tau_0,\tau_1]$. The covariance structure of the resulting limiting random element $\mathbb{W}_o$ is then given by

$$\begin{aligned} \mathbf{COV}(\mathbb{W}_o(s),\mathbb{W}_o(t)) &= (1-\rho)\,\frac{S_R(s)S_R(t)}{S_T(s)S_T(t)F_R(s)F_R(t)}\,D_T(s\wedge t) \\ &\quad -\rho\ \frac{F_T(s)F_T(t)S_R(s)S_R(t)}{S_T(s)S_T(t)F_R^2(s)F_R^2(t)}\,D_R(s\wedge t). \end{aligned}$$

4. For the cumulative hazard ratio,

$$h(t) \stackrel{\text{def}}{=} \Lambda_T(t)/\Lambda_R(t) \stackrel{\text{def}}{=} \phi_h(\Lambda_R,\Lambda_T)(t),$$

it has to be assumed in addition that $\Lambda_R(t) > 0,\ t \in [\tau_0,\tau_1]$. The covariance structure of the limiting random element $\mathbb{W}_h$ is then given by

$$\mathbf{COV}(\mathbb{W}_h(s),\mathbb{W}_h(t)) = (1-\rho)\frac{1}{F_R(s)F_R(t)}D_T(s\wedge t) - \rho\frac{F_T(s)F_T(t)}{F_R^2(s)F_R^2(t)}D_R(s\wedge t).$$

2.1 Suprema

The hypotheses in (2) suggest to base a test on $\sup_{t\in[\tau_0,\tau_1]}\hat\phi(t)$, which has be pursued in Freitag et al. [1]. In order to proof the consistency of such a test, we use results by Raghavachari [13] on the convergence in distribution of Kolmogorov-Smirnov type statistics under the alternative. Following the lines of the proof given therein, we can derive the convergence in distribution of our general supremum statistics.

Lemma 2.4 *Under the assumptions of Lemma 2.2 it follows that*

$$\sqrt{\kappa_n}\left(\sup_{t\in[\tau_0,\tau_1]}\hat\phi(t) - \sup_{t\in[\tau_0,\tau_1]}\phi(t)\right) \xrightarrow{\mathcal{D}} \sup_{t\in K_\phi^+}\mathbb{W}_\phi(t), \tag{7}$$

with $\mathbb{W}_\phi$ *from (6) and* $K_\phi^+ \stackrel{def}{=} \{t\in[\tau_0,\tau_1] : \sup_{s\in[\tau_0,\tau_1]}\phi(s) = \phi(t)\}$.

The proof of Lemma 2.4 essentially mimics the proof of Theorem 1 in Raghavachari [13]. However, the details will be required in the proof of the subsequent Theorem 3.2, hence we give the complete proof here. Moreover, instead of a result on the empirical process (cf. Lemma 1 in [13]), a more general result on weakly convergent processes is required in the following.

Proof:
Let

$$\lambda^+ \;=\; \sup_{t\in[\tau_0,\tau_1]}\phi(t),$$

$$
\begin{aligned}
D_n^+ &= \sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t), \\
Z_n^+ &= \sup_{t\in K_\phi^+} (\hat{\phi}(t) - \phi(t)), \\
D_n^{++} &= D_n^+ - \lambda^+ - Z_n^+.
\end{aligned}
$$

Thus, we have to show that $\sqrt{\kappa_n} D_n^{++} \xrightarrow{P} 0$ as $n_R, n_T \to \infty$, since the limiting distribution of $\sqrt{\kappa_n} Z_n^+$ is given by $\sup_{t\in K_\phi^+} \mathbb{W}_\phi(t)$ (cf. Lemma 2.2).

As $\phi(t)$ is a continuous, bounded function on $[\tau_0, \tau_1]$, the set K_ϕ^+ is a (non-empty) compact subset of $[\tau_0, \tau_1]$. From the week convergence result of Lemma 2.2 and Theorem V.1.3 in Pollard [14] we have that for every $\varepsilon, \eta > 0$, there exists a partition $\tau_0 = t_0 < t_1 < ... < t_l = \tau_1$ such that

$$
\limsup_{n_R,n_T\to\infty} P\Big(\sqrt{\kappa_n} \max_i \sup_{t\in J_i} |\hat{\phi}(t) - \phi(t) - \hat{\phi}(t_i) + \phi(t_i)| > \varepsilon\Big) < \eta, \tag{8}
$$

where $J_i = [t_i, t_{i-1})$ for $i = 0, 1, ..., l-1$. Choose those J_i for which there exists a $t \in J_i \cap K_\phi^+$, and denote these by $\tilde{J}_1, ..., \tilde{J}_p$ and $\tilde{t}_i, ..., \tilde{t}_p$, respectively ($p \le l$). Thus, $K_\phi^+ \subset \cup_{i=1}^p \tilde{J}_i \stackrel{\text{def}}{=} M_p$. Let $M_p^c = [\tau_0, \tau_1] \setminus M_p$. Then we have

$$
\begin{aligned}
P\Big\{\sqrt{\kappa_n} D_n^{++} > \varepsilon\Big\} &= P\Big\{\sqrt{\kappa_n}(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t) - \lambda^+ - Z_n^+) > \varepsilon\Big\} \\
&\le P\Big\{\sqrt{\kappa_n}(\sup_{t\in M_p} \hat{\phi}(t) - \lambda^+ - Z_n^+) > \varepsilon\Big\} \\
&+ P\Big\{\sqrt{\kappa_n}(\sup_{t\in M_p^c} \hat{\phi}(t) - \lambda^+ - Z_n^+) > \varepsilon\Big\}.
\end{aligned}
$$

From (8) it follows, for sufficiently large n_R, n_T,

$$
\begin{aligned}
& P\Big\{\sqrt{\kappa_n}(\sup_{t\in M_p} \hat{\phi}(t) - \lambda^+ - Z_n^+) > \varepsilon\Big\} \\
\le\ & P\Big\{\sqrt{\kappa_n}(\max_i \sup_{t\in\tilde{J}_i} (\hat{\phi}(t) - \phi(t) + \phi(t)) - \lambda^+ - Z_n^+) > \varepsilon\Big\} \\
\le\ & P\Big\{\sqrt{\kappa_n}(\max_i \sup_{t\in\tilde{J}_i} (\hat{\phi}(t) - \phi(t) + \phi(t)) - \phi(\tilde{t}_i) - \hat{\phi}(t_i) + \phi(t_i)) > \varepsilon\Big\} \\
\le\ & P\Big\{\sqrt{\kappa_n} \max_i \sup_{t\in\tilde{J}_i} |\hat{\phi}(t) - \phi(t) - \hat{\phi}(t_i) + \phi(t_i)| > \varepsilon\Big\} < \eta.
\end{aligned}
$$

Further, on M_p^c, $\phi(t)$ is bounded above by a number ρ with $0 < \rho < \lambda^+$. This follows from the continuity of $\phi(t)$, compactness of K_ϕ^+ and $K_\phi^+ \subset M_p$.

Let $\nu > 0$ such that $\nu < \lambda^+ - \rho$. By the Glivenko-Cantelli Theorem we have, for sufficiently large n_R, n_T,

$$
\sup_{t\in M_p^c} \hat{\phi}(t) - \lambda^+ < \sup_{t\in M_p^c} \hat{\phi}(t) - \lambda^+ + \nu < \rho - \lambda^+ + \nu < 0
$$

with probability one. Thus,

$$
\lim_{n_R,n_T\to\infty} P\Big\{\sup_{t\in M_p^c} \sqrt{\kappa_n}(\hat{\phi}(t) - \lambda^+ - Z_n^+) > \varepsilon\Big\} = 0.
$$

Hence,

$$\limsup_{n_R, n_T \to \infty} P\Big\{\sqrt{\kappa_n} D_n^{++} > \varepsilon\Big\} < \eta.$$

Since η was chosen arbitrary, $P\{\sqrt{\kappa_n} D_n^{++} > \varepsilon\} \to 0$ as $n_R, n_T \to \infty$. $\square$

3 Bootstrap approximations

The methods proposed in Freitag et al. [1] are based on bootstrapping. In the case of randomly right censored data, the so-called simple method suggested by Efron [15] can be used. According to Theorem 2.1 in Akritas [16], this method yields consistent estimators of the underlying distribution functions in the following sense.

Theorem 3.1 *For sample $k \in \{R, T\}$, a bootstrap sample $(X_{kj}^*, \delta_{kj}^*)$, $j = 1, ..., m_k$, is drawn from the pairs (X_{ki}, δ_{ki}), $i = 1, ..., n_k$. The corresponding Kaplan-Meier estimator $\hat{F}_k^*$ is then calculated. Then we have, as $m_k, n_k \to \infty$,*

$$m_k^{1/2}(\hat{F}_k^* - \hat{F}_k) \overset{\mathcal{D}}{\Longrightarrow} \mathbb{X}_k,$$

in $\mathbf{D}[\tau_0, \tau_1]$, for almost all sample sequences (X_{ki}, δ_{ki}), $i = 1, ..., n_k$. Here $\mathbb{X}_k$ denotes the limiting random element of the product-limit process $n_k^{1/2}(\hat{F}_k - F_k)$, and we have $\mathbb{X}_k = \mathbb{Z}_k / S_k$ on $[\tau_0, \tau_1]$ (cf. (3) and Th.IV.3.2 in Andersen et al. [9]).

An analogous result can be proved in the same way (via martingale representations; cf. [16]) for the bootstrapped Nelson-Aalen processes. This yields, as $m_k, n_k \to \infty$,

$$m_k^{1/2}(\hat{\Lambda}_k^* - \hat{\Lambda}_k) \overset{\mathcal{D}}{\Longrightarrow} \mathbb{Z}_k,$$

in $\mathbf{D}[\tau_0, \tau_1]$, for almost all sample sequences (X_{ki}, δ_{ki}), $i = 1, ..., n_k$.

Plugging the bootstrap estimators $\hat{\Lambda}_k^*$, $k = R, T$, into the definition of $\phi(t)$ yields a bootstrap estimator $\hat{\phi}^*(t)$. In Freitag et al. [1] two methods are proposed to solve the testing problem (2), the pointwise and the supremum approach. Both methods require the *weak consistency* of the bootstrap approximations they use. This is defined via the convergence of the sequences of distributions of the random elements under consideration as follows.

Let $\mathcal{L}[X]$ be the distribution of a random variable in $\mathbb{R}$, and suppose $\{X_n\} \subset \mathbb{R}$ with $X_n \xrightarrow{\mathcal{D}} X$ in $\mathbb{R}$. Further, let $\{X_m^*\}$ be a bootstrap version of $\{X_n\}$, and denote the conditional distribution, given $\{X_n\}$, by $\mathcal{L}^*$. Then the sequence $\{\mathcal{L}^*[X_m^*]\}$ is called weakly consistent for $\{\mathcal{L}[X_n]\}$, if $\rho_P(\mathcal{L}^*[X_m^*], \mathcal{L}[X]) \xrightarrow{P} 0$ for $m, n \to \infty$, where ρ_P is the Prohorov metric (cf. Gill [11], p.113).

3.1 Pointwise approach

Here an upper $(1-\alpha)$-confidence bound $\phi(t)^{(1-\alpha)}$ for $\phi(t)$ is constructed for each $t \in [\tau_0, \tau_1]$. Then the pointwise test of (2) consists of rejecting H_ϕ if $\phi(t)^{(1-\alpha)} < \Delta_\phi$ *for all* $t \in [\tau_0, \tau_1]$. We suggest using bootstrap confidence bounds for calculating $\phi(t)^{(1-\alpha)}$. Thus, it has to be assured that the bootstrapped discrepancy process at t, $\sqrt{\kappa_n}(\hat{\phi}^*(t) - \hat{\phi}(t))$ is weakly consistent for the distribution of $\mathbb{W}_\phi(t)$, which is normal with mean zero and and variance $\sigma_\phi^2(t)$ (cf. Lemma 2.2). Note that here the bootstrap sample sizes have been chosen as $m_k = n_k$, $k = R, T$. The required result follows immediately under the assumptions of Lemma 2.2 and Theorem 3.1, applying the functional delta method for the bootstrap as given in Theorem 5 in Gill [11]. Based on this, there are several possibilities to construct $\phi(t)^{(1-\alpha)}$, $t \in [\tau_0, \tau_1]$ (cf. Efron & Tibshirani [17] or Shao & Tu [18]). In Freitag et al. [1] the percentile and the bias-corrected accelerated percentile methods were applied.

3.2 Supremum approach

Here the aim is to calculate an upper $(1-\alpha)$-confidence bound $\phi^{(1-\alpha)}$ directly for the supremum of the discrepancy process, i.e. for $\sup_{t\in[\tau_0,\tau_1]} \phi(t)$. Then H_ϕ can be rejected if $\phi^{(1-\alpha)} < \Delta_\phi$. Thus, it has to be shown that the bootstrap approximation of the supremum functional is weakly consistent for the limiting random element in (7). Note that, in contrast to the case $\Delta_\phi = 0$, there arise technical difficulties due to the unknown set K_ϕ^+. These can be circumvented by choosing the bootstrap sample sizes *smaller* than the original sample sizes, as stated in the following Theorem.

Theorem 3.2 *Suppose that the assumptions of Lemma 2.2 are satisfied. Further, assume that m_k is chosen such that $m_k = o(n_k)$ and $m_T/(m_R + m_T) \to \rho$ as $m_k, n_k \to \infty$, $k = R, T$. Then it follows that the bootstrap approximation*

$$\sqrt{\kappa_m}\left(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}^*(t) - \sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t)\right)$$

is weakly consistent for the distribution of the limiting random element $\sup_{t\in K_\phi^+} \mathbb{W}_\phi(t)$ from Lemma 2.4.

Proof:
As in the proof of Theorem 5 in Gill [11] (but now focussing on the cumulative hazard functions instead of the distribution functions), application of the almost sure construction by Skorohod, Dudley and Wichura (cf. Shorack & Wellner [10]) yields sequences $\hat{\Lambda}'_k \overset{\mathcal{D}}{=} \hat{\Lambda}_k$ with $n_k^{1/2}(\hat{\Lambda}'_k - \Lambda_k) \xrightarrow{\|\cdot\|_\infty} \mathbb{Z}'_k$ a.s., $n_k \to \infty$, $\mathbb{Z}'_k \overset{\mathcal{D}}{=} \mathbb{Z}_k$, $k = R, T$ (cf. (3)). Then it follows for the bootstrapped Nelson-Aalen estimators, $\hat{\Lambda}_k^*$, that

$$m_k^{1/2}(\hat{\Lambda}_k^* - \hat{\Lambda}'_k) \overset{\mathcal{D}}{\Longrightarrow} \mathbb{Z}_k \quad a.s., \text{ as } n_k, m_k \to \infty,\ k = R, T.$$

Another application of the almost sure construction theorem yields for the given sequences $\hat{\Lambda}'_k$ new sequences $\hat{\Lambda}_k^{*\prime} \overset{\mathcal{D}}{=} \hat{\Lambda}_k^*$ with $m_k^{1/2}(\hat{\Lambda}_k^{*\prime} - \hat{\Lambda}'_k) \xrightarrow{\|\cdot\|_\infty} \mathbb{Z}_k^*$ a.s., $n_k, m_k \to \infty$, where $\mathbb{Z}_k^* \overset{\mathcal{D}}{=} \mathbb{Z}_k$, $k = R, T$. Using the assumption $m_k = o(n_k)$, $k = R, T$, yields, given $\hat{\Lambda}_k$, $k = R, T$,

$$\begin{aligned} m_k^{1/2}(\hat{\Lambda}_k^{*\prime} - \Lambda_k) &= m_k^{1/2}(\hat{\Lambda}_k^{*\prime} - \hat{\Lambda}'_k) + m_k^{1/2}(\hat{\Lambda}'_k - \Lambda_k) \\ &= m_k^{1/2}(\hat{\Lambda}_k^{*\prime} - \hat{\Lambda}'_k) + \sqrt{\frac{m_k}{n_k}} n_k^{1/2}(\hat{\Lambda}'_k - \Lambda_k) \\ &\xrightarrow{\|\cdot\|_\infty} \mathbb{Z}_k^* \overset{\mathcal{D}}{=} \mathbb{Z}_k \qquad a.s., \text{ as } n_k, m_k \to \infty,\ k = R, T. \end{aligned}$$

Let $\hat{\phi}^{*\prime}$ be the discrepancy functional applied to $\hat{\Lambda}_k^{*\prime}$, $k = R, T$ and $\hat{\phi}'$ the functional applied to $\hat{\Lambda}'_k$, $k = R, T$. Now we can use the Hadamard differentiability of ϕ and get, given $\hat{\Lambda}_k$, $k = R, T$,

$$\begin{aligned}\sqrt{\kappa_m}(\hat{\phi}^{*\prime} - \hat{\phi}') &= \sqrt{\kappa_m}(\hat{\phi}^{*\prime} - \phi) - \sqrt{\frac{\kappa_m}{\kappa_n}}\sqrt{\kappa_n}(\hat{\phi}' - \phi) \\ &\xrightarrow{\|\cdot\|_\infty} dT_{(\Lambda_T,\Lambda_R)}(\mathbb{Z}_T^*, \mathbb{Z}_R^*) \stackrel{\mathcal{D}}{=} \mathbb{W}_\phi \qquad a.s., \text{ as } n_k, m_k \to \infty,\ k = R, T.\end{aligned}$$

Then it follows that, given $\hat{\Lambda}_k$, $k = R, T$,

$$\begin{aligned}&\sqrt{\kappa_m}\left(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}^{*\prime}(t) - \sup_{t\in[\tau_0,\tau_1]} \hat{\phi}'(t)\right) \\ \stackrel{\mathcal{D}}{=}\ &\sqrt{\kappa_m}\left(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}^{*\prime}(t) - \sup_{t\in[\tau_0,\tau_1]} \phi(t)\right) - \sqrt{\kappa_m}\left(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}'(t) - \sup_{t\in[\tau_0,\tau_1]} \phi(t)\right) \qquad (9) \\ \xrightarrow{\mathcal{D}}\ &\sup_{t\in K_\phi^+} \mathbb{W}_\phi(t), \qquad a.s., \text{ as } n_k, m_k \to \infty,\ k = R, T.\end{aligned}$$

To this end, repeat the arguments in the proof of Lemma 2.4 and use the fact that the second summand in (9) equals

$$-\sqrt{\frac{\kappa_m}{\kappa_n}}\sqrt{\kappa_n}\left(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}'(t) - \sup_{t\in[\tau_0,\tau_1]} \phi(t)\right),$$

which tends to zero in probability. From this the weak consistency of the bootstrap approximation follows, i.e.

$$\rho_P\Big(\mathcal{L}^*\Big[\sqrt{\kappa_m}(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}^*(t) - \sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t))\Big], \mathcal{L}\Big[\sup_{t\in K_\phi^+} \mathbb{W}_\phi(t)\Big]\Big) \xrightarrow{P} 0$$

$a.s.,$ as $n_k, m_k \to \infty$, $k = R, T$. □

The results of Lemma 3.2 give rise to the use of the hybrid bootstrap method for addressing the testing problem (2). To this end the upper $(1-\alpha)$-confidence bound for $\sup_{t\in[\tau_0,\tau_1]} \phi(t)$ is constructed as follows. Denote the α-quantile of the distribution of $\sup_{t\in K_\phi^+} \mathbb{W}_\phi(t)$ by w_α. Then we have

$$P\left(\sqrt{\kappa_n}\Big(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t) - \sup_{t\in[\tau_0,\tau_1]} \phi(t)\Big) \geq w_\alpha\right) \geq 1 - \alpha. \qquad (10)$$

Now, w_α can be estimated by the α-quantile of B bootstrap replications

$$\sqrt{\kappa_m}\Big(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}_b^*(t) - \sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t)\Big), \quad b = 1, ..., B.$$

If $[x_b]_\alpha$ denotes the α-quantile of a sample $\{x_b,\ b = 1, ..., B\}$, then

$$\Big[\sqrt{\kappa_m}(\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}_b^*(t) - \sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t))\Big]_\alpha = \sqrt{\kappa_m}\Big[\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}_b^*(t)\Big]_\alpha - \sqrt{\kappa_m}\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t).$$

Plugging this into (10) yields

$$\sup_{t\in[\tau_0,\tau_1]} \phi(t) \leq \frac{\sqrt{\kappa_n} + \sqrt{\kappa_m}}{\sqrt{\kappa_n}} \sup_{t\in[\tau_0,\tau_1]} \hat{\phi}(t) - \frac{\sqrt{\kappa_m}}{\sqrt{\kappa_n}}\Big[\sup_{t\in[\tau_0,\tau_1]} \hat{\phi}_b^*(t)\Big]_\alpha \stackrel{\text{def}}{=} \phi^{(1-\alpha)},$$

which is the upper confidence bound for $\sup_{t\in[\tau_0,\tau_1]} \phi(t)$ as suggested in Freitag et al. [1].

References

[1] G. Freitag, S. Lange, and A. Munk. Nonparametric assessment of noninferiority with censored data. *(submitted)*, 2003.

[2] Non-Inferiority Trials: Advances in Concepts and Methodology. In R. B. D'Agostino et al., editors, *Stat. Med.*, volume 22 (2), 2003.

[3] CPMP. Biostatistical Methodology in Clinical Trials in Applications for Marketing Authorizations for Medicinal Products. CPMP Working Party on Efficacy of Medicinal Products Note for Guidance III/3630/92-EN. *Stat. Med.*, 14:1659–1682, 1995.

[4] CPMP. Points to Consider on Switching between Superiority and Non-inferiority. *CPMP / EWP / 482 / 99*, 2000.

[5] ICH. ICH Harmonised Tripartite Guideline. Statistical Principles for Clinical Trials (E9). *Stat. Med.*, 18:1905–1942, 1999.

[6] CPMP. (Draft) Points to Consider on the Choice of Non-Inferiority Margin. *CPMP / EWP / 2158 / 99*, 2004.

[7] ICH. ICH Harmonized Tripartite Guideline. Choice of Control Group and Related Issues in Clinical Trials (E10). *http: // www.fda.gov / cder / guidance / iche10.pdf. (Electronic Citation)*, 2001.

[8] G. Freitag. Methods of assessing noninferiority with censored data. *Biom. J. (to appear)*, 2004.

[9] P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding. *Statistical Methods Based on Counting Processes*. Springer Verlag, New York, 1993.

[10] G. R. Shorack and J. A. Wellner. *Empirical Processes with Applications to Statistics*. Series in Probability and Mathematical Statistics. Wiley, New York, 1986.

[11] R. D. Gill. Non- and semi-parametric maximum likelihood estimators and the von Mises method – I. *Scand. J. Stat.*, 16:97–128, 1989.

[12] A. W. van der Vaart and J. A. Wellner. *Weak Convergence and Empirical Processes*. Series in Statistics. Springer-Verlag, New York, 1996.

[13] M. Raghavachari. Limiting distributions of Kolmogorov-Smirnov-type statistics under the alternative. *Ann. Stat.*, 1:67–73, 1973.

[14] D. Pollard. *Convergence of Stochastic Processes*. Series in Statistics. Springer-Verlag, New York, 1984.

[15] B. Efron. Censored data and the bootstrap. *J. Amer. Stat. Assoc.*, 76:312–319, 1981.

[16] M. G. Akritas. Bootstrapping the Kaplan-Meier estimator. *J. Amer. Stat. Assoc.*, 81:1032–1038, 1986.

[17] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, New York, 1993.

[18] J. Shao and D. Tu. *The Jackknife and Bootstrap*. Series in Statistics. Springer-Verlag, New York, 1995.