UNIVERSITÄT HAMBURG

Testing hypotheses under a generalized Koziol-Green model with partially informative censoring

M. Brendel

Preprint–No. 2003-08 November 2003

FACHBEREICH MATHEMATIK

SCHWERPUNKT MATHEMATISCHE STATISTIK
UND STOCHASTISCHE PROZESSE

BUNDESSTR. 55, D – 20146 HAMBURG

Testing hypotheses under a generalized Koziol-Green model with partially informative censoring

Michael Brendel*

Universität Hamburg
Fachbereich Mathematik

Abstract. We consider the two-sample testing problem with randomly right censored data under the partial Koziol-Green model. It is distinguished between informative and non-informative censoring in this model. Rank and permutation tests that are distribution free under the hypothesis of randomness are derived and their optimality is discussed. Since the optimal scores of rank and permutation tests in the case of censored data usually depend on the distribution of the observations under the hypothesis of randomness we estimate the cumulative distribution function of the life times and the cumulative distribution function of the observations. With these estimators we can define rank and permutation tests whose optimality merely depends on the direction of the alternatives. Examples and simulations are discussed.

MSC 2000 Subject Classification: 62G10, 62G20
Key words: random censorship, informative and non-informative censoring, proportional hazard, asymptotically optimal tests, distribution free tests, LAN

1 Introduction

In this paper we consider the two-sample testing problem under the partial Koziol-Green (PKG) model introduced by Gather and Pawlitschko (1998) [3]. The PKG model extends the general random censorship model with right censoring which is quite common in survival analysis. Its key feature is the distinction between informative and non-informative censoring. We call a censoring informative, if it contains additional information about the distribution of the life times that we are interested in. A rigorous definition of informative censoring can be found in Andersen et al. (1993) [1, pp. 150]. Under the PKG model we assume that the survival functions of the informative censoring times are powers of the survival functions of the life times, that is to say the cumulative hazard functions are proportional.

In the next paragraphs we present the model, our notation and main assumptions. Let $T_{n,1}, \ldots, T_{n,n}$, $n \geq 2$, be non-negative, real-valued and stochastically independent (s.i.) random variables (r.v.s) denoting the unobservable life times of the pooled sample. We assume the first sample (control group) $T_{n,1}, \ldots, T_{n,m_n}$, $1 \leq m_n < n$, to be i.i.d. with continuous cumulative distribution function (c.d.f.) $F^{(c)}$ and the second sample (test group) $T_{n,m_n+1}, \ldots, T_{n,n}$ to be i.i.d. with continuous c.d.f. $F^{(t)}$. These life times are right censored by the informative censoring times $C_{1,n,1}, \ldots, C_{1,n,n}$ and the non-informative censoring

*Universität Hamburg, Fachbereich Mathematik, Schwerpunkt Mathematische Statistik und Stochastische Prozesse, Bundesstr. 55, 20146 Hamburg, Germany, Email: michael.brendel@math.uni-hamburg.de

times $C_{2,n,1},\ldots,C_{2,n,n}$. All censoring times are mutually s.i. and also s.i. of the $T_{n,i}$'s. Because of the Koziol-Green assumption of the model $C_{1,n,1},\ldots,C_{1,n,m_n}$ are i.i.d. with c.d.f. $1-(1-F^{(c)})^\beta$ and $C_{1,n,m_n},\ldots,C_{1,n,n}$ are i.i.d. with c.d.f. $1-(1-F^{(t)})^\beta$, whereas $\beta>0$ is an unknown but fixed model parameter. The non-informative censoring times are real-valued, i.i.d. r.v.s with continuous c.d.f. G. Under our model we can merely observe the censored life times $X_{n,i}=T_{n,i}\wedge C_{1,n,i}\wedge C_{2,n,i}$ and the censoring indicator $\Delta_{n,i}=\mathbb{1}_{[0,\infty)}\big((C_{1,n,i}\wedge C_{2,n,i})-T_{n,i}\big)-\mathbb{1}_{(0,\infty)}\big((C_{1,n,i}\wedge T_{n,i})-C_{2n,i}\big)$, $i=1,\ldots,n$. $a\wedge b$ stands for the minimum of $a,b\in\mathbb{R}$ and $\mathbb{1}_A$ denotes the indicator function of the set A. Hence, $\Delta_{n,i}=1$ implies that the i-th observation is uncensored, whereas $\Delta_{n,i}=0$ ($\Delta_{n,i}=-1$) means that it was informatively (non-informatively) censored.

We aim to develop tests for the null hypothesis of randomness $\mathscr{H}_0:\ F^{(c)}=F^{(t)}$ versus the non-parametric alternative $\mathscr{K}:\ F^{(c)}\geq F^{(t)},\ F^{(c)}\neq F^{(t)}$ (the underlying distribution of second sample $F^{(t)}$ is stochastically larger than the underlying distribution of the first sample $F^{(c)}$) and $\mathscr{H}:\ F^{(c)}\leq F^{(t)}$ versus $\mathscr{K}$. In order to achieve this aim we use asymptotic decision theory deriving rank and permutation tests that are distribution free under $\mathscr{H}_0$ and any continuous c.d.f. G.

Let $\mathfrak{F}=\{P_t\mid t\in(-\varepsilon_0,\varepsilon_0)\}$, $\varepsilon_0>0$, be a distribution family such that P_{t_2} is stochastically larger than P_{t_1}, that is to say $F_{t_1}\geq F_{t_2}$ and $F_{t_2}\neq F_{t_1}$, if and only if $t_1<t_2$, whereas F_t is the c.d.f. of P_t. In the sequel we call distribution families with this property as ordered. Furthermore, we assume that

$$\mathfrak{L}(T_{n,i})=P_{t^*\cdot c_{n,i}}\in\mathfrak{F}\tag{1}$$

for some $t^*\in\mathbb{R}$ holds true, whereas

$$c_{n,i}=\sqrt{\frac{m_n\cdot(n-m_n)}{n}}\begin{cases}-\frac{1}{m_n}, & 1\leq i\leq m_n,\\ \frac{1}{n-m_n}, & m_n+1\leq i\leq n.\end{cases}$$

Obviously, the $c_{n,i}$'s satisfy

$$\sum_{i=1}^n c_{n,i}=0\quad\text{and}\quad\sum_{i=1}^n c_{n,i}^2=1.\tag{2}$$

Suppose F_t denotes the continuous c.d.f. of $P_t\in\mathfrak{F}$. For $\beta>0$ we define the distribution $\widetilde{P}_{t,\beta}$ having the continuous c.d.f. $\widetilde{F}_{t,\beta}=1-(1-F_t)^\beta$ and the distribution family $\mathfrak{F}^\beta=\{\widetilde{P}_{t,\beta}\mid t\in(-\varepsilon_0,\varepsilon_0)\}$. As already mentioned above under the PKG model it is assumed that $\mathfrak{L}(C_{1,n,i})=\widetilde{P}_{t^*\cdot c_{n,i},\beta}$ for $i=1,\ldots,n$. The distribution of non-informative life times is denoted by $Q=\mathfrak{L}(C_{2,n,i})$ and its continuous c.d.f. by G. Under the model assumption (1) the testing problems $\mathscr{H}_0':\ t^*=0$ versus $\mathscr{K}':\ t^*>0$ and $\mathscr{H}':\ t^*\leq 0$ versus $\mathscr{K}':\ t^*>0$ are clearly subproblems of the general testing problems.

Under these assumptions and defining $Z:(\mathbb{R}^3,\mathbb{B}^3)\to(\mathbb{R}\times\{-1,0,1\},\mathbb{B}\otimes\mathcal{P}\{-1,0,1\})$, $(t,c_1,c_2)\longmapsto(t\wedge c_1\wedge c_2,\ \mathbb{1}_{[0,\infty)}\big((c_1\wedge c_2)-t\big)-\mathbb{1}_{(0,\infty)}\big((c_1\wedge t)-c_2\big)$, we can describe the distribution of a single observation by the distribution family $\mathfrak{R}^\beta=\{\mathscr{R}_t^\beta=(P_t\otimes\widetilde{P}_{t,\beta}\otimes Q)^Z\mid t\in(-\varepsilon_0,\varepsilon_0)\}$, that is to say $\mathfrak{L}(X_{n,i},\Delta_{n,i})\in\mathfrak{R}^\beta$, for $i=1,\ldots,n$. Consequently, the distribution of the pooled sample is given by

$$\mathscr{R}^\beta_{n,t^*\{c_{n,i}\}}:=\bigotimes_{i=1}^n R^\beta_{t^*\cdot c_{n,i}}.$$

The paper is organized as follows. In section 2 asymptotically optimal, parametric tests are developed by using LAN-theory. Therefore we state criteria ensuring that the sequence of statistical experiments

$$\{\Omega_n,\mathscr{A}_n,\mathfrak{R}_n^\beta\}=\left\{\bigtimes_{i=1}^n\big(\mathbb{R}\times\{-1,0,1\}\big),\bigotimes_{i=1}^n\big(\mathbb{B}\otimes\mathcal{P}\{-1,0,1\}\big),\big\{\mathscr{R}^\beta_{t\{c_{n,i}\}}\ \big|\ t\in(-\varepsilon_n,\varepsilon_n)\big\}\right\},$$

$\varepsilon_n:=(\max\{|c_{n,1}|,\ldots,|c_{n,n}|\})^{-1}\cdot\varepsilon_0$, is asymptotically normal. In section 3 rank and permutation tests for the testing problem in question are suggested. Especially, the situations in which sequences of the proposed rank and permutation tests are asymptotically optimal are discussed. Since optimal scores of rank tests in random censorship models depend on the c.d.f. of the censored life time $X_{n,i}$ and the c.d.f. of life time $T_{n,i}$ we use the modified Kaplan-Meyer estimator proposed by Gather and Pawlitschko

(1998) [3] for estimating the survival time of the $T_{n,i}$'s under the PKG model to obtain statistics whose asymptotic optimality properties merely depend on the direction of the alternatives. In section 4 examples and simulations are presented. The auxiliary results needed for the proofs of our main results are given in section 5.

Rank and permutation tests for models with randomly right censored data under local alternatives were already studied by Neuhaus (1988) [7], who considered translation families and positive scale families, and Janssen (1989) [6], who extended these results by assuming that both the distribution of the life times and the censoring times can be described by $\mathbb{L}_2$-differentiable distribution families. In both papers it turns out that for deriving asymptotically optimal rank and permutation tests, i.e. choosing optimal scores, one has to know the direction of the alternatives as well as the distribution of $X_{n,i}$ under the null hypothesis. The proofs concerning the asymptotic properties of the rank and permutation tests and the variance estimator in section 3 are extensions of the proofs given in Neuhaus (1988) [7] to our case.

2 Asymptotically optimal tests for parametric distribution families

In order to derive asymptotically optimal tests we apply asymptotic decision theory to the PKG model. This gives us parametric, asymptotically optimal tests that are asymptotically equivalent to the rank and permutation tests suggested in section 3. In a first step it is shown that the sequence of statistical experiments $\{\Omega_n, \mathscr{A}_n, \mathfrak{R}_n^\beta\}$, $n \geq 2$, is asymptotically normal by assuming that the distribution family $\mathfrak{R}^\beta$ is $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiable. Since $\mathfrak{R}^\beta$ is derived from $\mathfrak{F}$ and $\mathfrak{F}^\beta$ one is interested in guaranteeing this property by the $\mathbb{L}_2(P_0)$- and $\mathbb{L}_2(\widetilde{P}_{0,\beta})$-differentiability of the distribution families $\mathfrak{F}$ and $\mathfrak{F}^\beta$ respectively. The theorems 2.4-2.6 investigate this matter.

Note, a distribution family $\mathfrak{P} = \{P_t \mid t \in (-\varepsilon, \varepsilon)\}$ on the measurable space $(\mathfrak{X}, \mathscr{C})$ is called $\mathbb{L}_2(P_0)$-differentiable with $\mathbb{L}_2$-derivative $\dot{L}$, if $\dot{L}$ is a $\mathbb{L}_2(P_0)$-integrable mapping satisfying the conditions

$$\int_{\mathfrak{X}} \left(2\big(L_{0,t}^{1/2} - 1\big) - t \cdot \dot{L}\right)^2 \mathrm{d}P_0 = o(|t|) \tag{3}$$

and

$$P_t\big(N_{0,t}\big) = o(t^2), \tag{4}$$

whereas $(L_{0,t}, N_{0,t})$ is a Lebesgue-decomposition of P_t with respect to P_0. A definition of a Lebesgue-decomposition can be found in Strasser (1985) [8, p. 3]. In the sequel $\mathscr{R}^\beta_{n,0\{c_{n,i}\}}$ is abbreviated to $\mathscr{R}^\beta_{n,0}$ and we assume that $\frac{m_n}{n} \to \nu \in (0,1)$ as $n \to \infty$. The latter implies

$$\lim_{n\to\infty} \max\{|c_{n,1}|, \ldots, |c_{n,n}|\} = 0. \tag{5}$$

Together with (2) this means that the regression coefficients $c_{n,i}$, $i = 1, \ldots, n$, $n \geq 2$, satisfy the strict Noether-conditions. As an immediate consequence of Le Cam's second lemma we obtain

2.1 Theorem. *Let $\mathfrak{R}^\beta$ be a $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiable distribution family with $\mathbb{L}_2$-derivative $\dot{h}_\beta$. The sequence of statistical experiments $\{\Omega_n, \mathscr{A}_n, \mathfrak{R}_n^\beta\}$, $n \geq 2$, is asymptotically normal with central sequence $Z_n\big((x_i, \delta_i), i = 1, \ldots, n\big) := \sum_{i=1}^n c_{n,i} \dot{h}_\beta(x_i, \delta_i)$. More precisely, it holds $\mathfrak{L}_{\mathscr{R}^\beta_{n,0}}\big(Z_n\big) \xrightarrow{\mathfrak{L}} \mathcal{N}(0, \sigma^2)$, $\sigma^2 := \mathrm{Var}_{\mathscr{R}_0^\beta}(\dot{h}_\beta)$, and*

$$L^\beta_{0,t\{c_{n,i}\}} = \exp\left(t \cdot Z_n, t - \frac{1}{2}\sigma^2 \cdot t^2 + \zeta_{n,t}\right) \left[\mathscr{R}^\beta_{n,0}\right], \qquad \zeta_{n,t} \xrightarrow{\mathscr{R}^\beta_{n,0}} 0, \qquad \textit{for all } t \in \mathbb{R},$$

whereas $L^\beta_{0,t\{c_{n,i}\}}$ denotes a likelihood ratio of $\mathscr{R}^\beta_{n,t\{c_{n,i}\}}$ with respect to $\mathscr{R}^\beta_{n,0}$.

Proof. Cf. Witting and Müller-Funk (1995) [10, Satz 6.130, p. 317]. ∎

2.2 Corollary. *Let the assumptions of theorem 2.1 hold. Then* $\varphi_n^*\big((X_{n,i},\Delta_{n,i}), i=1,\dots,n\big)$, $n \geq 2$, *defined by*

$$\varphi_n^*\big((X_{n,i},\Delta_{n,i}), i=1,\dots,n\big) = \begin{cases} 1, \\ 0, \end{cases} \sum_{i=1}^{n} c_{n,i} \cdot \dot{h}_\beta(X_{n,i},\Delta_{n,i}) \begin{array}{c} \geq \\ < \end{array} u_\alpha \sqrt{\mathrm{Var}_{\mathscr{R}_0^\beta}(\dot{h}_\beta)},$$

is an asymptotically optimal sequence of tests of the level $\alpha \in (0,1)$ *for the testing problem* $\mathscr{H}'$ *versus* $\mathscr{K}'$, *i.e.*

$$\lim_{n\to\infty} \int \varphi_n^* \, \mathrm{d}\mathscr{R}^\beta_{n,t\{c_{n,i}\}} = \Phi\left(-u_\alpha + t \cdot \sqrt{\mathrm{Var}_{\mathscr{R}_0^\beta}(\dot{h}_\beta)}\right), \quad t \in \mathbb{R}.$$

Φ *denotes the c.d.f. of the normal distribution with mean* 0 *and variance* 1, $u_\alpha := \Phi^{-1}(\alpha)$.

Proof. Cf. Strasser (1985) [8, Definition 82.7 and Theorem 82.8, p. 430] □

2.3 Corollary. *Under the assumptions of theorem 2.1 the sequences of distributions* $\{\mathscr{R}^\beta_{n,0} \mid n \geq 2\}$ *and* $\{\mathscr{R}^\beta_{n,t\{c_{n,i}\}} \mid n \geq 2\}$ *are mutually contiguous.*

Proof. Le Cam's first and second lemmata give the result, cf. Witting and Müller-Funk (1995) [10, Satz 6.124 and Satz 6.130, p. 311, 317]. □

Criteria ensuring the $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiability of $\mathfrak{R}^\beta$ are given in the following theorems. In the proofs of theorem 2.5 and theorem 2.6 the $\mathbb{L}_2$-differentiability of the distribution families in question is directly verified. The most simple case is given, if both $\mathfrak{F}$ and $\mathfrak{F}^\beta$ are $\mathbb{L}_2$-differentiable.

2.4 Theorem. *Suppose the distribution family* $\mathfrak{F}$ *is* $\mathbb{L}_2(P_0)$*-differentiable with* $\mathbb{L}_2$*-derivative* $\dot{L}$. *If* $\mathfrak{F}^\beta$ *is also* $\mathbb{L}_2(\widetilde{P}_{0,\beta})$*-differentiable then* $\mathfrak{R}^\beta$ *is* $\mathbb{L}_2(R_0^\beta)$*-differentiable and its* $\mathbb{L}_2$*-derivative is given by*

$$\dot{h}_\beta(x,\delta) = \mathbb{1}_{\{0,1\}}(\delta)\dot{L}(x) + \mathbb{1}_{\{-1\}}(\delta)\frac{\int_{(x,\infty)} \dot{L}\,\mathrm{d}P_0}{1-F_0(x)} + \beta\frac{\int_{(x,\infty)} \dot{L}\,\mathrm{d}P_0}{1-F_0(x)} \quad [\mathscr{R}_0^\beta]. \tag{6}$$

Proof. Set $Q_t = Q$ for all $t \in (-\varepsilon_0,\varepsilon_0)$. Obviously, $\{Q_t \mid t \in (-\varepsilon_0,\varepsilon_0)\}$ is $\mathbb{L}_2(Q_0)$-differentiable. Moreover, one easily sees that $\big\{P_t \otimes \widetilde{P}_{t,\beta} \otimes Q_t, \mid t \in (-\varepsilon_0,\varepsilon_0)\big\}$ is $\mathbb{L}_2(P_0 \otimes \widetilde{P}_{0,\beta} \otimes Q_0)$-differentiable, cf. Witting (1985) [9, Satz 1.191, p. 177]. Because of $(P_t \otimes \widetilde{P}_{t,\beta} \otimes Q_t)^Z = \mathscr{R}_t^\beta$, $t \in (-\varepsilon_0,\varepsilon_0)$, the $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiability of $\mathfrak{R}^\beta$ is an immediate consequence of Witting (1985) [9, Satz 1.193, p. 178]. To complete the proof we just have to identify the $\mathbb{L}_2$-derivative. For this purpose let $\widetilde{L}_{0,t}$, $t \in (-\varepsilon_0,\varepsilon_0)$, be the likelihood ratio of $\mathscr{R}_t^\beta$ with respect to $\mathscr{R}_0^\beta$ given by Lemma 5.2. We show $2/t(\widetilde{L}_{0,t}^{1/2} - 1) \to \dot{h}_\beta$ in $\mathscr{R}_0^\beta$-probability as $t \to 0$. It holds $\mathscr{R}_0^\beta$-a.e. the identity

$$\begin{aligned} \frac{2}{t}(\widetilde{L}_{0,t}^{1/2} - 1) = {} & \underbrace{\mathbb{1}_{\{0,1\}}(\delta)\frac{2}{t}\left(L_{0,t}^{1/2} - 1\right)}_{=I_1(t)} + \underbrace{\mathbb{1}_{\{0,1\}}(\delta)\frac{2}{t}\left(L_{0,t}^{1/2} - 1\right)\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta}{2}} - 1\right)}_{=I_2(t)} \\ & + \underbrace{\mathbb{1}_{\{0,1\}}(\delta)\frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta}{2}} - 1\right)}_{=I_3(t)} + \underbrace{\mathbb{1}_{\{-1\}}(\delta)\frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta+1}{2}} - 1\right)}_{=I_4(t)}. \end{aligned} \tag{7}$$

Let us investigate the four terms separately. Using the $\mathbb{L}_2(P_0)$-differentiability of $\mathfrak{F}$ and Vitali's theorem we have for all $\varepsilon > 0$

$$\mathbb{1}_{\left\{\left|\frac{2}{t}\left(L_{0,t}^{1/2}-1\right)-\dot{L}\right|>\varepsilon\right\}} \xrightarrow{P_0} 0.$$

Applying lemma 5.1 and Lebesgue's theorem give

$$\mathscr{R}_0^\beta\left(\left|\frac{2}{t}(L_{0,t}^{1/2} - 1) - \dot{L}\right|\mathbb{1}_{\{0,1\}}(\delta) \geq \varepsilon\right) = \int \mathbb{1}_{\left\{\left|\frac{2}{t}\left(L_{0,t}^{1/2}-1\right)-\dot{L}\right|>\varepsilon\right\}} (\beta+1)(1-G)(1-F_0)^\beta \mathrm{d}P_0 \to 0,$$

i.e. $I_1(t) \to \mathbb{1}_{\{0,1\}}(\delta)\dot{L}$ in $\mathscr{R}_0^\beta$-probability as $t \to 0$. A Taylor-expansion at $t = 0$ using lemma 5.3 shows that $((1-F_t)/(1-F_0)^{\beta/2} - 1) \to 0$ $[\mathscr{R}_0^\beta]$ as $t \to 0$ implying $((1-F_t)/(1-F_0)^{\beta/2} - 1) \to 0$ in $\mathscr{R}_0^\beta$-probability. Having the previous computation in mind we have $I_2(t) \to 0$ in $\mathscr{R}_0^\beta$-probability as $t \to 0$. This proceeding can also be applied to $I_3(t)$ and $I_4(t)$ giving

$$I_3(t) \xrightarrow{\mathscr{R}_0^\beta} \mathbb{1}_{\{0,1\}}(\delta)\beta\frac{\int_{(\cdot,\infty)} \dot{L}\,\mathrm{d}P_0}{1-F_0} \quad \text{and} \quad I_4(t) \xrightarrow{\mathscr{R}_0^\beta} \mathbb{1}_{\{-1\}}(\delta)\,(\beta+1)\,\frac{\int_{(\cdot,\infty)} \dot{L}\,\mathrm{d}P_0}{1-F_0}.$$

Since $X_n \xrightarrow{\mathscr{R}_0^\beta} X^{(j)}$, $j = 1, 2$, implies $X^{(1)} = X^{(2)}$ $[\mathscr{R}_0^\beta]$ the proof is complete. □

Given the continuous c.d.f.s $F^{(1)}$ and $F^{(2)}$ such that $F^{(1)} \geq F^{(2)}$ and $F^{(1)} \neq F^{(2)}$ holds, one can construct an ordered family of distributions $\mathfrak{P} = \{P_t \mid t \in [\nu-1, \nu]\}$ such that $F_{\nu-1} = F^{(1)}$ and $F_\nu = F^{(2)}$, whereas F_t denotes the c.d.f. of P_t. For this purpose we define the c.d.f. $F_0 := \nu \cdot F^{(1)} + (1-\nu)\cdot F^{(2)}$ and the bounded function $b := (\mathrm{d}F^{(2)}/\mathrm{d}F_0 - \mathrm{d}F^{(1)}/\mathrm{d}F_0) \circ F_0^{-1}$, more precisely it holds $-1/\nu \leq b \leq 1/(1-\nu)$. By assuming that P_0 has the c.d.f. F_0 and that P_t has the P_0 densities $f_t = 1 + tb \circ F_0$ we obtain the distribution family $\mathfrak{P}$. Note that $\mathfrak{P}$ satisfies the conditions we imposed on $\mathfrak{F}$. In rank test theory in the uncensored case such distribution families are used to show that the optimality of rank tests merely depends on the direction of alternatives, i.e. on the choice of b, and is independent of the actual distribution of the r.v.s under the null hypothesis, s. Behnen and Neuhaus (1989) [2, pp. 18]. The same considerations can also be applied to our rank and permutation tests and this motivates the next result.

2.5 Theorem. *Suppose that all $P_t \in \mathfrak{F}$ have P_0-densities f_t given by $f_t = 1 + t\cdot b\circ F_0$ $[P_0]$, whereas F_0 denotes the continuous c.d.f. of P_0 and $b : ([0,1], \mathbb{B}\cap[0,1]) \to (\mathbb{R}, \mathbb{B})$ is a bounded function such that $\int_{[0,1]} b\,\mathrm{d}\lambda = 0$ holds. Then the following assertions hold true:*

(i) The distribution family $\mathfrak{F}$ is $\mathbb{L}_2(P_0)$-differentiable with $\mathbb{L}_2$-derivative $\dot{L} = b \circ F_0$ $[P_0]$.

(ii) The distribution family $\mathfrak{F}^\beta$, $\beta > 0$, is $\mathbb{L}_2(\widetilde{P}_{0,\beta})$-differentiable.

Proof. The first assertion is a well-known result, cf. e.g. Witting [9, Beispiel 1.200, p. 183]. It remains to prove that $\mathfrak{F}^\beta$ is $\mathbb{L}_2(\widetilde{P}_{0,\beta})$-differentiable. Let $(\widetilde{L}_{0,t}, \widetilde{N}_{0,t})$ be a Lebesgue-decomposition of $\widetilde{P}_t$ with respect to $\widetilde{P}_0$. It holds $P_0 \equiv P_t$ for all t sufficiently close to 0. Since $\widetilde{P}_{t,\beta} \ll P_t$ it results $\widetilde{P}_{t,\beta} \ll \widetilde{P}_{0,\beta}$. Thus (4) is trivially satisfied. Define

$$I(t) := \frac{2}{t}\left(\widetilde{L}_{0,t}^{1/2} - 1\right) - b\circ F_0 - (\beta-1)\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\cdot \mathbb{1}_{(0,1)}(F_0),$$

to complete the proof we show $\lim_{t\to 0} I(t) = 0$ $[\widetilde{P}_{0,\beta}]$ and $|I(t)| \leq C \in (0,\infty)$, so that Lebesgue's theorem yields (3). It holds $\widetilde{L}_{0,t} = \big((1-F_t)/(1-F_0)\big)^{\beta-1}\sqrt{1+t\cdot b\circ F_0}\cdot \mathbb{1}_{(0,1)}(F_0)$ $[\widetilde{P}_{0,\beta}]$. By using the chain of identities

$$\begin{aligned}
\frac{2}{t}\left(\widetilde{L}_{0,t}^{1/2} - 1\right) &\overset{[\widetilde{P}_0]}{=} \frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta-1}{2}}\cdot \mathbb{1}_{(0,1)}(F_0)\cdot L_{0,t}^{1/2} - 1\right)\\
&\overset{[\widetilde{P}_0]}{=} \frac{2}{t}\left(L_{0,t}^{1/2} - 1\right) + \frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta-1}{2}}\cdot \mathbb{1}_{(0,1)}(F_0) - 1\right)\cdot L_{0,t}^{1/2}\\
&\overset{[\widetilde{P}_0]}{=} \frac{2}{t}\left(L_{0,t}^{1/2} - 1\right) + \frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta-1}{2}}\cdot \mathbb{1}_{(0,1)}(F_0) - 1\right)\\
&\quad + \frac{2}{t}\left(L_{0,t}^{1/2} - 1\right)\cdot\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta-1}{2}}\cdot \mathbb{1}_{(0,1)}(F_0) - 1\right).
\end{aligned}$$

it results

$$\begin{aligned}
I(t) &\overset{[\widetilde{P}_0]}{=} \left(\frac{2}{t}\big(\sqrt{1+tb\circ F_0}-1\big)-b\circ F_0\right)\cdot \mathbb{1}_{(0,1)}(F_0)\\
&\quad+\left(\frac{2}{t}\left(\Big(\frac{1-F_t}{1-F_0}\Big)^{\frac{\beta-1}{2}}-1\right)-(\beta-1)\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\right)\cdot \mathbb{1}_{(0,1)}(F_0)\\
&\quad+\frac{2}{t}\big(\sqrt{1+tb\circ F_0}-1\big)\cdot\left(\Big(\frac{1-F_t}{1-F_0}\Big)^{\frac{\beta-1}{2}}-1\right)\cdot \mathbb{1}_{(0,1)}(F_0)\\
&= I_1+I_2+I_3.
\end{aligned}$$

Choose $\widetilde{\varepsilon}_0<\varepsilon_0$. Clearly, $\widetilde{\varepsilon}_0<|b|_\infty^{-1}$, $|b|_\infty:=\sup\{b(u)\mid u\in(0,1)\}$. A Taylor-expansion at $t=0$ gives

$$\frac{2}{t}\big(\sqrt{1+tb\circ F_0}-1\big)\cdot \mathbb{1}_{(0,1)}(F_0)=\frac{2}{t}\left(1+t\frac{b\circ F_0}{2}-\frac{t^2}{8}\frac{(b\circ F_0)^2}{\big(1+t(\cdot)\cdot b\circ F_0\big)^{3/2}}-1\right)\cdot \mathbb{1}_{(0,1)}(F_0),$$

$0<|t(\cdot)|<|t|$. Consequently, for all $0<|t|\le\widetilde{\varepsilon}_0$ we have the estimate

$$|I_1|=\frac{|t|}{4}\left|\frac{(b\circ F_0)^2}{\big(1+t(\cdot)\cdot b\circ F_0\big)^{3/2}}\right|\cdot \mathbb{1}_{(0,1)}(F_0)\le\frac{|t|}{4}\left|\frac{|b|_\infty^2}{\big(1-\widetilde{\varepsilon}_0|b|_\infty\big)^{3/2}}\right|\le\frac{\widetilde{\varepsilon}_0}{4}\left|\frac{|b|_\infty^2}{\big(1-\widetilde{\varepsilon}_0|b|_\infty\big)^{3/2}}\right|=:C_1,\quad (8)$$

whereas C_1 is some real number. Again, a Taylor-expansion $t=0$ gives

$$\begin{aligned}
\frac{2}{t}\left(\Big(\frac{1-F_t}{1-F_0}\Big)^{\frac{\beta-1}{2}}-1\right)\cdot \mathbb{1}_{(0,1)}(F_0)=&\frac{2}{t}\Bigg(1+\frac{t}{2}(\beta-1)\cdot\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\\
&+\frac{t^2}{8}(\beta-1)(\beta-3)\Big(\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\Big)^2\left(1+t(\cdot)\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\right)^{\frac{\beta-5}{2}}-1\Bigg)\cdot \mathbb{1}_{(0,1)}(F_0),
\end{aligned}$$

$0<|t(\cdot)|<|t|$. Because of the estimate

$$\left|\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\cdot \mathbb{1}_{(0,1)}(F_0)\right|\le\frac{\int_{(\cdot,\infty)}|b\circ F_0|\,\mathrm{d}P_0}{1-F_0}\cdot \mathbb{1}_{(0,1)}(F_0)\le\frac{\int_{(\cdot,\infty)}|b|_\infty\,\mathrm{d}P_0}{1-F_0}\cdot \mathbb{1}_{(0,1)}(F_0)\le|b|_\infty$$

we have for $0<|t|<\widetilde{\varepsilon}_0$

$$\begin{aligned}
|I_2|&\le\frac{|t|}{4}|\beta-1||\beta-3||b|_\infty^2\left(1+t(\cdot)\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\right)^{\frac{\beta-5}{2}}\cdot \mathbb{1}_{(0,1)}(F_0)\\
&\le\frac{|t|}{4}|\beta-1||\beta-3||b|_\infty^2\Big(\big(1+\widetilde{\varepsilon}_0|b|_\infty\big)^{\frac{\beta-5}{2}}+\big(1-\widetilde{\varepsilon}_0|b|_\infty\big)^{\frac{\beta-5}{2}}\Big)\\
&\le\frac{\widetilde{\varepsilon}_0}{4}|\beta-1||\beta-3||b|_\infty^2\Big(\big(1+\widetilde{\varepsilon}_0|b|_\infty\big)^{\frac{\beta-5}{2}}+\big(1-\widetilde{\varepsilon}_0|b|_\infty\big)^{\frac{\beta-5}{2}}\Big):\;=C_2,
\end{aligned}\quad (9)$$

whereas C_2 is also some real number. The estimates for I_1 and I_2 yields an upper bound for I_3:

$$\begin{aligned}
|I_3|&\le\frac{|t|}{2}\,|I_1+b\circ F_0|\cdot\left|I_2+(\beta-1)\frac{\int_{(\cdot,\infty)} b\circ F_0\,\mathrm{d}P_0}{1-F_0}\right|\cdot \mathbb{1}_{(0,1)}(F_0)\\
&\le\frac{|t|}{2}\left(C_1+|b|_\infty\right)\left(C_2+|\beta-1||b|_\infty\right)\le\frac{\widetilde{\varepsilon}_0}{2}\left(C_1+|b|_\infty\right)\left(C_2+|\beta-1||b|_\infty\right):=C_3.
\end{aligned}\quad (10)$$

Clearly, it holds $|I_3|\to0$ as $t\to0$. All in all we proved $|I_1+I_2+I_3|\le C_1+C_2+C_3$ for $0<|t|<\widetilde{\varepsilon}_0$. Moreover, (8), (9) and (10) also show, that $I_1+I_2+I_3\to0$ $[\widetilde{P}_0]$. Hence, (3) holds. □

The next result gives another condition ensuring that the $\mathbb{L}_2(P_0)$-differentiability of $\mathfrak{F}$ is sufficient for the $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiability of $\mathfrak{R}^\beta$. This result can be applied to distribution families with unbounded $\mathbb{L}_2$-derivatives. This result is also used to obtain optimal tests for the simulations in section 4.

2.6 Theorem. *We define $\tau_F := \sup\{x \mid F(x) < 1\}$ as the right end point of some c.d.f. F. If the distribution family $\mathfrak{F}$ is $\mathbb{L}_2(P_0)$-differentiable and $\tau_G < \tau_{F_0}$ then the distribution family $\mathfrak{R}^\beta$ is $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiable with the $\mathbb{L}_2$-derivative $\dot{h}_\beta$ given by formula (6).*

Proof. The following facts are used to prove the $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiability of $\mathfrak{R}^\beta$.

(i) Lemma 5.3 implies $\lim_{t\to 0} F_t(\tau_G) = F_0(\tau_G)$. $\tau_G < \tau_F$ gives $F_0(\tau_G) < 1$. Hence, it exists an $\varepsilon \in (0, \varepsilon_0)$, such that $1 - F_t(x) \geq K > 0$ for all $x \leq \tau_G$ and all $t \in (-\varepsilon, \varepsilon)$.

(ii) Given $\rho > 0$, the real function $x \mapsto x^\rho$ is Lipschitz-continuous on $[K, 1]$, i.e. it exists $M(\rho) \in \mathbb{R}$ such that $\left|a^\rho - b^\rho\right| \leq M(\rho)\left|a - b\right|$ for all $a, b \in [K, 1]$.

Let $(\widetilde{L}_{0,t}, \widetilde{N}_{0,t})$, $|t| \leq \varepsilon$, denote Lebesgue-decompositions according to lemma 5.2. Note that $\int_{\{F_0=1\}}(1-F_t)^{\beta+1}\,\mathrm{d}Q = 0$ follows from $\tau_G < \tau_{F_t}$ for all $t \in (-\varepsilon, \varepsilon)$. Thus, the $\mathbb{L}_2(P_0)$-differentiability of $\mathfrak{F}$ implies $0 \leq t^{-2}\mathscr{R}_t^\beta(\widetilde{N}_{0,t}) \leq t^{-2}(\beta+1)P_t(N_0,t) \to 0$ as $t \to 0$, i.e. (4). It remains to verify (3). Using the inequality $(a+b+c)^2 \leq 4\cdot(a^2+b^2+c^2)$ and (7) we have for $0 < |t| < \varepsilon$ the estimate

$$
\begin{aligned}
\int\left(\frac{2}{t}(\widetilde{L}_{0,t}^{1/2} - 1) - \dot{h}_\beta\right)^2 \mathrm{d}\mathscr{R}_0^\beta &= \sum_{\delta=-1}^{1}\int\left(\frac{2}{t}(L_{0,t}^{1/2}(x,\delta) - 1) - \dot{h}_\beta(x,\delta)\right)^2 h_0(x,\delta)\mathrm{d}\mu(x)\\
&\leq 4\cdot(\beta+1)\Bigg(\int\left(\frac{2}{t}(L_{0,t}^{1/2} - 1) - \dot{L}\right)^2(1-F_0)^\beta(1-G)\,\mathrm{d}P_0\\
&\quad + \int\left(\frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\beta/2} - 1\right) - \beta\frac{\int_{(\cdot,\infty)}\dot{L}\,\mathrm{P}_0}{1-F_0}\right)^2(1-F_0)^\beta(1-G)\,\mathrm{d}P_0\\
&\quad + \int\left(\frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\beta/2} - 1\right)(L_{0,t}^{1/2} - 1)\right)^2(1-F_0)^\beta(1-G)\,\mathrm{d}P_0\Bigg)\\
&\quad + \int\left(\frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\frac{\beta+1}{2}} - 1\right) - (\beta+1)\frac{\int_{(\cdot,\infty)}\dot{L}\,\mathrm{P}_0}{1-F_0}\right)^2(1-F_0)^{\beta+1}\,\mathrm{d}Q\\
&= 4\cdot(\beta+1)\big(I_1(t) + I_2(t) + I_3(t)\big) + I_4(t),
\end{aligned}
$$

so it suffices to show that $\lim_{t\to 0} I_j(t) = 0$, $i = 1, 2, 3, 4$. The $\mathbb{L}_2$-differentiability of the distribution family $\mathfrak{F}$ gives

$$
0 \leq I_1(t) \leq \int\left(\frac{2}{t}(L_{0,t}^{1/2} - 1) - \dot{L}\right)^2 \mathrm{d}P_0 \to 0 \qquad \text{as } t \to 0.
$$

Let us consider $I_2(t)$. For arbitrary chosen $\xi > 0$ lemma 5.3 implies the existence of $\widetilde{\varepsilon}$ with $0 < \widetilde{\varepsilon} \leq \varepsilon$, such that

$$
\left|\frac{F_0(x) - F_t(x)}{t} - \int_{(x,\infty)}\dot{L}\,\mathrm{d}P_0\right| \leq \xi
$$

holds for all $0 < |t| \leq \widetilde{\varepsilon}$ and all $x \in \mathbb{R}$. This fact is used to construct a function dominating the integrand, so that Lebesgue's theorem gives $\lim_{t\to 0} I_2(t) = 0$. The estimate $(a+b)^2 \leq 2(a^2+b^2)$ yields

$$
\left|\left(\frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\beta/2} - 1\right) - \beta\frac{\int_{(\cdot,\infty)}\dot{L}\,\mathrm{P}_0}{1-F_0}\right)^2(1-F_0)^\beta(1-G)\right| \leq
$$

$$
\underbrace{2(1-G)\left(\frac{2}{t}\left((1-F_t)^{\frac{\beta}{2}} - (1-F_0)^{\frac{\beta}{2}}\right)\right)^2}_{m_1(t)} + \underbrace{2\beta^2(1-G)\left(\int_{(\cdot,\infty)}\dot{L}\,\mathrm{d}P_0\right)^2(1-F_0)^{\beta-2}}_{m_2(t)}.
$$

For $m_1(t)$ we get using (i), (ii) and the Cauchy-Schwarz-inequality

$$\begin{aligned} m_1(t) &\le 8M^2(\beta/2)(1-G)\left(\frac{F_0-F_t}{t}\right)^2 \\ &\le 8M^2(\beta/2)(1-G)\left(\frac{F_0-F_t}{t} - \int_{(\cdot,\infty)} \dot{L}\,\mathrm{d}P_0 + \int_{(\cdot,\infty)} \dot{L}\,\mathrm{d}P_0\right)^2 \\ &\le 16M^2(\beta/2)(1-G)\left[\left(\frac{F_0-F_t}{t} - \int_{(\cdot,\infty)} \dot{L}\,\mathrm{d}P_0\right)^2 + \left(\int_{(\cdot,\infty)} \dot{L}\,\mathrm{d}P_0\right)^2\right] \\ &\le 16M^2(\beta/2)(1-G)\left(\xi^2 + (1-F_0)\cdot \mathrm{E}_{P_0}(\dot{L}^2)\right). \end{aligned}$$

This function is obviously integrable with respect to P_0. Applying the Cauchy-Schwarz-inequality to $m_2(t)$ yields another function that is P_0-integrable.

$$m_2(t) \le 2\beta^2(1-G)(1-F_0)^{\beta-1}\cdot \mathrm{E}_{P_0}(\dot{L}^2) \le \beta^2(1-G)(1+K^{\beta-1})\cdot \mathrm{E}_{P_0}(\dot{L}^2).$$

Finally, a Taylor-expansion at $t=0$ using lemma 5.3 shows

$$\lim_{t\to 0}\left(\frac{2}{t}\left(\left(\frac{1-F_t}{1-F_0}\right)^{\beta/2} - 1\right) - \beta\frac{\int_{(\cdot,\infty)} \dot{L}\,\mathrm{P}_0}{1-F_0}\right) = 0 \quad [P_0],$$

by Lebesgue's theorem we conclude $\lim_{t\to 0} I_2(t) = 0$. Concerning $I_3(t)$, for arbitrary $\xi > 0$ it exists $\widetilde{\varepsilon}$ with $0 < \widetilde{\varepsilon} \le \varepsilon$, such that $|F_t(x) - F_0(x)| \le \xi$ for all $x \in \mathbb{R}$ and $0 < t \le \widetilde{\varepsilon}$, s. lemma 5.3. By applying (i), (ii) and Vitali's theorem it results

$$\begin{aligned} 0 &\le \limsup_{t\to 0} I_3(t) \\ &= \limsup_{t\to 0} \int \left((1-F_t)^{\frac{\beta}{2}} - (1-F_0)^{\frac{\beta}{2}}\right)^2 \left(\frac{2}{t}(L_{0,t}^{1/2}-1)\right)^2 (1-G)\,\mathrm{d}P_0 \\ &\le \limsup_{t\to 0} M(\beta/2)^2 \int \left(F_0-F_t\right)^2 \left(\frac{2}{t}(L_{0,t}^{1/2}-1)\right)^2 (1-G)\,\mathrm{d}P_0 \\ &\le \lim_{t\to 0} M(\beta/2)^2\xi^2 \int \left(\frac{2}{t}(L_{0,t}^{1/2}-1)\right)^2 \mathrm{d}P_0 \\ &\le M(\beta/2)^2\xi^2 \mathrm{E}_{P_0}(\dot{L}^2). \end{aligned}$$

Since ξ was arbitrary, $\lim_{t\to 0} I_3(t) = 0$ follow from $\xi \downarrow 0$. $I_4(t)$ can be dealt with analogue to $I_2(t)$. □

The condition $\tau_G < \tau_{F_0}$ has a practical interpretation. If one models the time between the entry of a participant in a case study till the end of the study by the non-informative censoring time then the condition means that a participant can live longer than the duration of the study with positive probability. A condition which seems to be not too unrealistic.

3 Asymptotically optimal rank and permutation tests

This section contains the main results of this paper. Rank and permutation tests are developed and their optimality is discussed. However, before doing so, we need some more notation. $R_n := (R_{n,1},\ldots,R_{n,n})$ denotes the rank vector of the observations $X_{n,1},\ldots,X_{n,n}$, whereas the rank of the i-th observation is given by $R_{n,i} := \sum_{j=1}^n \mathbb{1}_{[0,\infty)}(X_{n,i}-X_{n,j})$. The inverse rank vector $D_n := (D_{n,1},\ldots,D_{n,n})$ is defined by the identities $R_{n,D_{n,i}} = D_{n,R_{n,i}} = i$, $i = 1,\ldots,n$. The order statistic of the observations and the order statistic of the censoring indicators are defined by $X_{n,\uparrow} := \left(X_{n:1},\ldots,X_{n:n}\right) := \left(X_{n,D_{n,1}},\ldots,X_{n,D_{n,n}}\right)$ and

$\Delta_{n,\uparrow} := (\Delta_{n:1}, \ldots, \Delta_{n:n}) := (\Delta_{n,D_{n,1}}, \ldots, \Delta_{n,D_{n,n}})$. One easily proves that the r.v.s R_n and $(X_{n\uparrow}, \Delta_{n\uparrow})$ are s.i. under $\mathscr{R}^{\beta}_{n,0}$. More precisely we have

$$\mathscr{R}^{\beta}_{n,0}(R_n = r) = \frac{1}{n!},\ r \in \mathfrak{S}_n, \quad \text{and} \quad \mathscr{R}^{\beta}_{n,0}(R_{n,i} = k) = \frac{1}{n},\ k \in \{1, \ldots, n\}, \tag{11}$$

whereas $\mathfrak{S}_n$ denotes the set of permutations of the numbers $1, \ldots, n$.

In the last section we saw that the asymptotically optimal sequences of tests contained the model parameter $\beta > 0$. Thus, for deriving tests it is essential to estimate this parameter. An important condition to prove consistency of the following estimator is that an observation remains uncensored with positive probability .

3.1 Lemma. *Define $p_k = \mathscr{R}^{\beta}_{n,0}(\{\Delta_{n,i} = k\})$, $k \in \{-1, 0, 1\}$, and suppose $p_1 > 0$, then it holds $p_0 = \beta p_1$, cf. Gather and Pawlitschko (1998) [3, Lemma 3.1]. With some $\widetilde{\beta} \geq 0$ and with*

$$\widehat{p}_{n,k} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{k\}}(\Delta_{n:i}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{k\}}(\Delta_{n,i}), \quad k \in \{-1, 0, 1\},$$

an estimator of β is given by

$$\widehat{\beta}_n = \begin{cases} \widehat{p}_{n,0}/\widehat{p}_{n,1}, & \widehat{p}_{n,1} > 0, \\ \widetilde{\beta}, & \widehat{p}_{n,1} = 0, \end{cases}$$

and it holds that $\widehat{\beta}_n \longrightarrow \beta$ in $\mathscr{R}^{\beta}_{n,t\{c_{n,i}\}}$-probability for all $t \in \mathbb{R}$.

Proof. Corollary 2.3 implies that the consistency needs to be proved under $\mathscr{H}_0'$, which immediately results from the weak law of large numbers. □

In the sequel we assume that $b : ([0,1], \mathbb{B} \cap [0,1]) \to (\mathbb{R}, \mathbb{B})$ is a bounded function and that the set $D(b) = \{u \in (0,1) \mid b \text{ is discontinuous in } u\}$ is countable. Furthermore, we define $B : ([0,1], \mathbb{B} \cap [0,1]) \to (\mathbb{R}, \mathbb{B})$ by $u \mapsto \int_{(u,1)} b \, \mathrm{d}\lambda \cdot (1-u)^{-1} \cdot \mathbb{1}_{[0,1)}(u)$. Obviously, B is bounded for all $u \in [0,1]$ and continuous for all $u \in (0,1)$. For $n \geq 2$ let $b_{n,i} = b_{n,i}(X_{n,\uparrow}, \Delta_{n,\uparrow})$ and $B_{n,i} = B_{n,i}(X_{n,\uparrow}, \Delta_{n,\uparrow})$, $i = 1, \ldots, n$, be scores that depend at most on the order statistics $X_{n,\uparrow}$ and $\Delta_{n,\uparrow}$ such that score-functions $b_n(t) = b_{n,\lceil nt \rceil}$, $t \in [0,1]$ and $B_n(t) = B_{n,\lceil nt \rceil}$, $t \in [0,1]$, $\lceil u \rceil := \inf\{z \in \mathbb{Z} \mid z \geq u\}$, satisfy the condition

$$\lim_{n\to\infty} \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}} \left(\int_{(0,1)} (b_n - b \circ F_0 \circ H^{-1})^2 \, \mathrm{d}\lambda \right) = 0 \text{ and } \lim_{n\to\infty} \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}} \left(\int_{(0,1)} (B_n - B \circ F_0 \circ H^{-1})^2 \, \mathrm{d}\lambda \right) = 0, \tag{12}$$

whereas $H(x) = \mathscr{R}^{\beta}_{n,0}(X_{n,i} \leq x) = 1 - (1 - F_0(x))^{\beta+1}(1 - G(x))$ is the c.d.f. of $X_{n,i}$ under $\mathscr{H}_0$ and H^{-1} is the pseudo-inverse of H. The next lemma yields score-functions satisfying the condition (12).

3.2 Lemma. *Define*

$$\widehat{b}_n(u) = b \circ \widehat{F}_n \circ \widehat{H}_n^{-1}(u) \quad \text{and} \quad \widehat{B}_n(u) = \frac{\int_{(\widehat{F}_n \circ \widehat{H}_n^{-1}(u),1)} b \, \mathrm{d}\lambda}{1 - \widehat{F}_n \circ \widehat{H}_n^{-1}(u)} \cdot \mathbb{1}_{[0,1)}(\widehat{F}_n \circ \widehat{H}_n^{-1}(u))$$

for all $u \in [0,1]$, whereas $\widehat{F}_n$ is an estimator of the c.d.f. F_0 such that $\|\widehat{F}_n - F_0\|_{[0,\tau_H]} := \sup\{|\widehat{F}_n(x) - F_0(x)| : x \in [0, \tau_H]\} \to 0$ in $\mathscr{R}^{\beta}_{n,0}$-probability and $\widehat{F}_n(x) = 0$ $[\mathscr{R}^{\beta}_{n,0}]$ for all $x < \inf\{x \mid F_0(x) > 0\}$. $\widehat{H}_n^{-1}$ is the pseudo-inverse of the empirical distribution function of $(X_{n,1}, \ldots, X_{n,n})$. The functions $\widehat{b}_n$ and $\widehat{B}_n$ are indeed score functions with the scores $\widehat{b}_{n,i} = \widehat{b}_n(i/(n+1))$ and $\widehat{B}_{n,i} = \widehat{B}_n(i/(n+1))$, $i = 1, \ldots, n$. If the condition $D(b) \cap K(F_0) = \emptyset$, $K(F_0) = \{u \in (0,1) \mid \exists x_1 < x_2 \text{ such that } F_0(x_1) = F_0(x_2) = u\}$, holds true then $\widehat{b}_n$ and $\widehat{B}_n$ satisfy the condition (12).

Proof. $\widehat{H}_n^{-1}(u) = X_{n:i}$, for all $u \in (\frac{i-1}{n}, \frac{i}{n}]$, implies that $\widehat{b}_n$ and $\widehat{B}_n$ are score functions, i.e. $\widehat{b}_n(u) = \widehat{b}_{n,\lceil nu \rceil}$ and $\widehat{B}_n(u) = \widehat{B}_{n,\lceil nu \rceil}$, $u \in (0,1)$. In a first step it is proved that $\int_{(0,1)} (\widehat{b}_n - b \circ F_0 \circ H^{-1})^2 \, \mathrm{d}\lambda \to 0$ in

$\mathscr{R}^{\beta}_{n,0}$-probability. Since this integral is bounded, Lebesgue's theorem gives

$$\lim_{n\to\infty} \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}} \left(\int_{(0,1)} (\widehat{b}_n - b \circ F_0 \circ H^{-1})^2 \,\mathrm{d}\lambda \right) = 0.$$

Without loss of generality it can be assumed that the r.v.s $(X_{n,i}, \Delta_{n,i})$, $i = 1, \ldots, n$, $n \geq 2$ are defined on the probability space $\{\bigtimes_{n=2}^{\infty} \Omega_n, \bigotimes_{n=2}^{\infty} \mathscr{A}_n, P := \bigotimes_{n=2}^{\infty} \mathscr{R}^{\beta}_{n,0}\}$. Therefore it is sufficient to show that there is a subsequence $\{n_k\}$ in every subsequence of the natural numbers such that $\int_{(0,1)} (\widehat{b}_{n_k} - b \circ F_0 \circ H^{-1})^2 \,\mathrm{d}\lambda \longrightarrow 0\ [P]$.

In every subsequence of the natural numbers we can find a subsequence $\{n_k\}$ such that

$$\|\widehat{F}_{n_k} - F_0\|_{[0,\tau_H]} := \sup\{|\widehat{F}_{n_k}(x) - F_0(x)| : x \in [0, \tau_H]\} \to 0\ [P]$$

and $\widehat{H}^{-1}_{n_k}(u) \to H^{-1}(u)\ [P]$ for all $u \in (0,1)\backslash D(H^{-1})$, whereas $D(H^{-1})$ denotes the set of discontinuity points of H^{-1}. This gives

$$\begin{aligned} |\widehat{F}_{n_k} \circ \widehat{H}^{-1}_{n_k}(u) - F_0 \circ H^{-1}(u)| &\leq |\widehat{F}_{n_k} \circ \widehat{H}^{-1}_{n_k}(u) - F_0 \circ \widehat{H}^{-1}_{n_k}(u)| + |F_0 \circ \widehat{H}^{-1}_{n_k}(u) - F_0 \circ H^{-1}(u)| \\ &\leq \underbrace{\|\widehat{F}_{n_k} - F_0\|_{[0,\tau_H]}}_{\longrightarrow 0\ [P]} + \underbrace{|F_0 \circ \widehat{H}^{-1}_{n_k}(u) - F_0 \circ H^{-1}(u)|}_{\longrightarrow 0\ [P]} \longrightarrow 0\ [P], \quad (13) \end{aligned}$$

$u \in (0,1)\backslash D(H^{-1})$, note that F_0 is continuous.

Suppose it exists $u \in (0,1)$ such that $F_0 \circ H^{-1}(u) = 0$, set $u_0 = H(x_0)$, $x_0 = \inf\{x \mid F_0(x) > 0\}$. Note $H^{-1}(u_0) = x_0$. It holds $F_0 \circ H^{-1}(u) = 0$ for all $u \in (0, u_0)$ and $0 < F_0 \circ H^{-1}(u) < 1$ for all $u \in (u_0, 1)$. We have

$$\widehat{H}^{-1}_{n_k}(u) \to H^{-1}(u) < H^{-1}(u_0) = x_0, \quad \text{for all } u \in (0, u_0)\backslash D(H^{-1}).$$

Since $\widehat{F}_n(x) = 0\ [P]$ for all $x < x_0$ it exists $k_0(u)$ such that $\widehat{F}_{n_k} \circ \widehat{H}^{-1}_{n_k}(u) = 0\ [P]$ for all $k \geq k_0(u)$. Consequently,

$$b \circ \widehat{F}_{n_k} \circ \widehat{H}^{-1}_{n_k}(u) \to b(0) = b \circ F_0 \circ H^{-1}(u)\ [P].$$

b being continuous for all $u \in (0,1)\backslash D(b)$ and (13) imply $\widehat{b}_{n_k}(u) \to b \circ F_0 \circ H^{-1}(u)\ [P]$ for all $u \in (u_0, 1)\backslash (D(H^{-1}) \cup M)$, $M = \{u \in (0,1) \mid F_0 \circ H^{-1}(u) \in D(b)\}$. In case it does not exist a $u \in (0,1)$ such that $F_0 \circ H^{-1}(u) = 0$ the above convergence holds for all $u \in (0,1)\backslash (D(H^{-1}) \cup M)$.

$D(b) \cap K(F_0) = \emptyset$ gives that the set M is countable. Suppose for $d \in D(b)$ it exists $u_1, u_2 \in (0,1)$ such that $u_1 \neq u_2$ and $F_0 \circ H^{-1}(u_i) = d$, $i = 1, 2$. Then either $H^{-1}(u_1) = H^{-1}(u_2)$ or $H^{-1}(u_1) \neq H^{-1}(u_2)$. The first contradicts the fact that H is continuous and the latter implies $d \in K(F_0)$. Thus, for every $d \in D(b)$ it exists exactly one $u \in (0,1)$ satisfying $F_0 \circ H^{-1}(u) = d$.

All in all it was proved $\widehat{b}_{n_k}(u) \to b \circ F_0 \circ H^{-1}(u)\ [P]$ for a.e. $u \in (0,1)$. Note that $(\widehat{b}_{n_k} - b \circ F_0 \circ H^{-1})^2$ is bounded, so Lebesgue's theorem yields $\int_{(0,1)} (\widehat{b}_{n_k} - b \circ F_0 \circ H^{-1})^2 \,\mathrm{d}\lambda \longrightarrow 0\ [P]$.

Since B is continuous on $(0,1)$ and bounded on $[0,1]$ one shows with the same arguments as above

$$\lim_{n\to\infty} \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}} \left(\int_{(0,1)} (\widehat{B}_n - B \circ F_0 \circ H^{-1})^2 \,\mathrm{d}\lambda \right) = 0.$$

□

3.3 Remark. The condition $D(b) \cap K(F_0) = \emptyset$ is satisfied, if either F_0 is a strictly monotone function on $[0, \tau_H]$ or b is continuous on $(0,1)$. A suitable estimator of the c.d.f. of F_0 is given by the modified Kaplan-Meyer estimator introduced by Gather and Pawlitschko (1998) [3],

$$\widehat{F}_n(t) = 1 - \left(\prod_{i=1}^{n} \left(1 - \frac{\mathbb{1}_{\{0,1\}}(\Delta_{n:i})}{n+1-i} \right)^{\mathbb{1}_{(-\infty,t]}(X_{n:i})} \right)^{\widehat{p}_n}, \qquad \widehat{p}_n = \frac{\widehat{p}_{n,1}}{\widehat{p}_{n,1} + \widehat{p}_{n,0}}.$$

This estimator satisfies the conditions $\|\widehat{F}_n - F_0\|_{[0,\tau_H]} \to 0$ in $\mathscr{R}^{\beta}_{n,0}$-probability, cf. Gather and Pawlitschko (1998) [3, Theorem 4.1], and $\widehat{F}_n(x) = 0\ [\mathscr{R}^{\beta}_{n,0}]$ for all $x < \inf\{x \mid F_0(x) > 0\}$ of lemma 3.2.

Having proposed appropriate scores we can concentrate on a rank statistic. A rank statistic for testing problems is given by

$$\widehat{S}_n(b_n, B_n) = \sum_{i=1}^{n} c_{n,i} \left(\mathbb{1}_{\{0,1\}}(\Delta_{n,i})\big(b_{n,R_{n,i}} + \widehat{\beta}_n B_{n,R_{n,i}}\big) + \mathbb{1}_{\{-1\}}(\Delta_{n,i})(1+\widehat{\beta}_n)B_{n,R_{n,i}} \right).$$

Of course, the structure of this statistic is closely related to the $\mathbb{L}_2$-derivative given in (6). Since we want to compute the asymptotic power function of tests depending on the statistic $\widehat{S}_n(b_n, B_n)$ it is necessary to compute the asymptotic distribution of the statistic $\widehat{S}_n(b_n, B_n)$. For doing so we need a bit more notation. Let $q_i : ((0,1),(0,1)\cap\mathbb{B}) \to (\mathbb{R},\mathbb{B})$, $i = 1,2,3$, be functions given by $q_i(u) = \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(\mathbb{1}_{\{2-i\}}(\Delta_{n,1}) \mid U_{n,1} = u\big)$, $i = 1,2,3$. With $\dot{L}$ of theorem 2.4 we define the optimal score generating functions $\widetilde{b}, \widetilde{B} : ([0,1],[0,1]\cap\mathbb{B}) \to (\mathbb{R},\mathbb{B})$,

$$\widetilde{b}(u) = \dot{L}\circ F_0^{-1}(u)\cdot \mathbb{1}_{(0,1)}(u) \quad \text{and} \quad \widetilde{B}(u) = \frac{\int_{(u,1)} \dot{L}\circ F_0^{-1}\,\mathrm{d}\lambda}{1-u}\cdot \mathbb{1}_{[0,1)}(u).$$

And for convenience let us introduce the abbreviations

$$\widetilde{k}_1 = \widetilde{b}\circ F_0\circ H^{-1} + \beta\widetilde{B}\circ F_0\circ H^{-1} \quad \text{and} \quad \widetilde{k}_2 = (1+\beta)\cdot\widetilde{B}\circ F_0\circ H^{-1}.$$

Note the close connection of these functions and the $\mathbb{L}_2$-derivative given in theorem 2.4 as well as the functions stated in lemma 3.2. If one considers the situation in theorem 2.5 one sees that $\widetilde{b}$ and $\widetilde{B}$ do not necessarily depend on F_0, i.e. the foot-point of the underlying parametric distribution family. For the next proof , we employ the following auxiliary statistic $S_n^*(b,B) = \sum_{i=1}^n c_{n,i}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n,i})\big(b\circ F_0\circ H^{-1}(U_{n,i}) + \beta B\circ F_0\circ H^{-1}(U_{n,i})\big) + \mathbb{1}_{\{-1\}}(\Delta_{n,i})(1+\beta)B\circ F_0\circ H^{-1}(U_{n,i})\big)$, whereas $U_{n,i} = H(X_{n,i})$. Note that the $U_{n,i}$, $i = 1,\dots,n$, are s.i. r.v.s that are uniformly distributed on $(0,1)$ under $\mathscr{H}_0'$. Furthermore, the abbreviations $k_1 = b\circ F_0\circ H^{-1} + \beta B\circ F_0\circ H^{-1}$ and $k_2 = (1+\beta)\cdot B\circ F_0\circ H^{-1}$ are used.

3.4 Theorem. *For the test statistic $\widehat{S}_n(b_n, B_n)$ it holds*

$$\mathfrak{L}_{\mathscr{R}^\beta_{n,t\{c_{n,i}\}}}\big(\widehat{S}_n(b_n,B_n)\big) \xrightarrow{\mathfrak{L}} \mathscr{N}(t\cdot\mu_b, \sigma_b^2) \quad \text{as } n\to\infty,$$

whereas

$$\mu_b := \int_{(0,1)} k_1\widetilde{k}_1 q_1 + k_1\widetilde{k}_1 q_2 + k_2\widetilde{k}_2 q_3\,\mathrm{d}\lambda$$

and

$$\sigma_b^2 := \int_{(0,1)} k_1^2 q_1 + k_1^2 q_2 + k_2^2 q_3\,\mathrm{d}\lambda - \left(\int_{(0,1)} k_1 q_1 + k_1 q_2 + k_2 q_3\,\mathrm{d}\lambda\right)^2.$$

Proof. Because of (2) and (5) we can apply Satz 5.112 in Witting and Müller-Funk (1995) [10, p. 112] obtaining

$$\mathfrak{L}_{\mathscr{R}^\beta_{n,0}}\big(S_n^*(b,B)\big) \xrightarrow{\mathfrak{L}} \mathscr{N}(0,\sigma_b^2) \quad \text{as } n\to\infty \tag{14}$$

with

$$\begin{aligned}
\sigma_b^2 &= \mathrm{Var}_{\mathscr{R}^\beta_{n,0}}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n,1})\cdot k_1(U_{n,1}) + \mathbb{1}_{\{-1\}}(\Delta_{n,1})\cdot k_2(U_{n,1})\big)\\
&= \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(\mathbb{1}_{\{1\}}(\Delta_{n,1})\cdot k_1^2 + \mathbb{1}_{\{0\}}(\Delta_{n,1})\cdot k_1^2 + \mathbb{1}_{\{-1\}}(\Delta_{n,1})\cdot k_2^2\big)^2\\
&\quad - \Big(\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(\mathbb{1}_{\{1\}}(\Delta_{n,1})\cdot k_1 + \mathbb{1}_{\{0\}}(\Delta_{n,1})\cdot k_1 + \mathbb{1}_{\{-1\}}(\Delta_{n,1})\cdot k_2\big)\Big)^2\\
&= \int_{(0,1)} k_1^2 q_1 + k_1^2 q_2 + k_2^2 q_3\,\mathrm{d}\lambda - \left(\int_{(0,1)} k_1 q_1 + k_1 q_2 + k_2 q_3\,\mathrm{d}\lambda\right)^2.
\end{aligned}$$

By Slutsky's lemma and (14) it results

$$\mathfrak{L}_{\mathscr{R}^{\beta}_{n,0}}\big(t\cdot S_n^*(\widetilde{b})\big) \xrightarrow{\mathfrak{L}} \mathscr{N}(0, t^2\sigma^2_{\widetilde{b}}).$$

Since the statistic in question is linear it holds

$$S_n^*(xb + yt\widetilde{b}, xB + yt\widetilde{B}) = x\cdot S_n^*(b,B) + y\cdot tS_n^*(\widetilde{b},\widetilde{B})$$

for all $x, y \in \mathbb{R}$. In particular we have

$$\mathfrak{L}_{\mathscr{R}^{\beta}_{n,0}}\big(x\cdot S_n^*(b,B) + y\cdot t\cdot S_n^*(\widetilde{b},\widetilde{B})\big) \xrightarrow{\mathfrak{L}} \mathscr{N}(0, \sigma^2_{xb+yt\widetilde{b}}),$$

because of (14). Applying the Cramér-Wold-device yields

$$\mathfrak{L}\big(S_n^*(b,B), t\cdot S_n^*(\widetilde{b},\widetilde{B})\big)^T \xrightarrow{\mathfrak{L}} \mathscr{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_b^2 & t\cdot\mu_b\\ t\cdot\mu_b & t^2\sigma^2_{\widetilde{b}}\end{pmatrix}\right) \quad \text{as } n\to\infty,$$

whereas

$$\begin{aligned}
\mu_b &= \mathrm{Cov}_{\mathscr{R}^{\beta}_{n,0}}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n,1})k_1(U_{n,1}) + \mathbb{1}_{\{-1\}}(\Delta_{n,1})k_2(U_{n,1}), \mathbb{1}_{\{0,1\}}(\Delta_{n,1})\widetilde{k}_1(U_{n,1}) + \mathbb{1}_{\{-1\}}(\Delta_{n,1})\widetilde{k}_2(U_{n,1})\big)\\
&= \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n,1})k_1(U_{n,1})\widetilde{k}_1(U_{n,1}) + \mathbb{1}_{\{-1\}}(\Delta_{n,1})k_2(U_{n,1})\widetilde{k}_2(U_{n,1})\big)\\
&\quad - \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n,1})k_1(U_{n,1}) + \mathbb{1}_{\{-1\}}(\Delta_{n,1})k_2(U_{n,1})\big)\cdot\\
&\qquad \underbrace{\mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n,1})\widetilde{k}_1(U_{n,1}) + \mathbb{1}_{\{-1\}}(\Delta_{n,1})\widetilde{k}_2(U_{n,1})\big)}_{\overset{(*)}{=}0}\\
&= \int_{(0,1)} k_1\widetilde{k}_1 q_1 + k_1\widetilde{k}_1 q_2 + k_2\widetilde{k}_2 q_3 \,\mathrm{d}\lambda.
\end{aligned}$$

The equality $(*)$ is a consequence of

$$S_n^*(\widetilde{b},\widetilde{B}) = \sum_{i=1}^n c_{n,i}\dot{h}_\beta(X_{n,i},\Delta_{n,i})$$

and

$$\mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n,1})\cdot\widetilde{k}_1(U_{n,1}) + \mathbb{1}_{\{-1\}}(\Delta_{n,1})\cdot\widetilde{k}_2(U_{n,1})\big) = \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}(\dot{h}_\beta(X_{n,1},\Delta_{n,1})) = 0,$$

cf. Witting (1985) [9, Hilfssatz 1.178, p. 164].

$S_n^*(\widetilde{b},\widetilde{B})$, $n\geq 2$, is a central sequence of a sequence of asymptotically normal experiments, s. theorem 2.1, thus it holds

$$\log\left(\frac{\mathrm{d}\mathscr{R}^{\beta}_{n,t\{c_{n,i}\}}}{\mathrm{d}\mathscr{R}^{\beta}_{n,0}}\right) - \left(t\cdot S_n^*(\widetilde{b},\widetilde{B}) - \frac{1}{2}t^2\sigma^2_{\widetilde{b}}\right) \xrightarrow{\mathscr{R}^{\beta}_{n,0}} 0 \quad \text{as } n\to\infty.$$

Using lemma 5.4 and Slutsky's lemma one obtains

$$\mathfrak{L}\left(\widehat{S}_n(b,B), \log\left(\frac{\mathrm{d}\mathscr{R}^{\beta}_{n,t\{C_{n,i}\}}}{\mathrm{d}\mathscr{R}^{\beta}_{n,0}}\right)\right)^T \xrightarrow{\mathfrak{L}} \mathscr{N}\left(\begin{pmatrix}0\\ -\frac{1}{2}\sigma^2_{\widetilde{b}}\end{pmatrix}, \begin{pmatrix}\sigma_b^2 & t\cdot\mu_b\\ t\cdot\mu_b & t^2\sigma^2_{\widetilde{b}}\end{pmatrix}\right) \quad \text{as } n\to\infty.$$

Le Cam's third lemma, cf .Witting and Müller-Funk (1995) [10, Korollar 6.139, p. 329], yields the assertion

$$\mathfrak{L}_{\mathscr{R}^{\beta}_{n,t\{c_{n,1}\}}}\big(\widehat{S}_n(b,B)\big) \xrightarrow{\mathfrak{L}} \mathscr{N}(t\mu_b, \sigma_b^2).$$

□

3.5 Corollary. *The sequence of tests* $(\varphi_n(b_n, B_n))_{n\in\mathbb{N}}$ *defined by*

$$\varphi_n(b_n, B_n) = \begin{cases} 1, \\ 0, \end{cases} \widehat{S}_n(b_n, B_n) \begin{matrix} \geq \\ < \end{matrix} u_\alpha \cdot \sqrt{\sigma_b^2}$$

is asymptotically of the level $\alpha \in (0,1)$ *and has the asymptotic power function*

$$\lim_{n\to\infty} \mathrm{E}_{\mathscr{R}^\beta_{n,t\{c_{n,i}\}}}(\varphi_n) = \Phi\left(-u_\alpha + t\frac{\mu_b}{\sqrt{\sigma_b^2}}\right). \tag{15}$$

Suppose it exists some $\eta > 0$ *such that* $\eta b = \widetilde{b}$ *then the sequence of tests is asymptotically optimal for the testing problem* $\mathscr{H}'$ *vs.* $\mathscr{K}'$.

Proof. The first assertion is an immediate consequence of theorem 3.4. The existence of some $\eta > 0$ such that $\eta b = \widetilde{b}$ implies $\eta B = \widetilde{B}$. Since the rank statistic in question is linear we have the following identities $\sigma_b^2 = \sigma_{\widetilde{b}}^2 \cdot \eta^2$ and $\mu_b = \mu_{\widetilde{b}} \cdot \eta$. Moreover, using the definitions of $\widetilde{b}$ and $\widetilde{B}$ it holds $\mu_{\widetilde{b}} = \sigma_{\widetilde{b}}^2 = \mathrm{Var}_{\mathscr{R}_0^\beta}(\dot{h}_\beta)$. Hence, the test sequence $(\varphi_n(b_n, B_n))_{n\in\mathbb{N}}$ has the asymptotic power function

$$\Phi\left(-u_\alpha + t\frac{\mu_b}{\sqrt{\sigma_b^2}}\right) = \Phi\left(-u_\alpha + t\frac{\mu_{\widetilde{b}}}{\sqrt{\sigma_{\widetilde{b}}^2}}\right) = \Phi\left(-u_\alpha + t\sqrt{\mathrm{Var}_{\mathscr{R}_0^\beta}(\dot{h}_\beta)}\right).$$

That is to say the asymptotic power function of an asymptotically optimal sequence of tests, s. corollary 2.2. □

However, this sequence of tests still depends on the unknown variance σ_b. The next lemma yields a consistent estimator of this variance which can be employed to derive under $\mathscr{H}_0$ asymptotically distribution free tests.

3.6 Lemma. *Using the abbreviation*

$$\widehat{V}_n(i) = \mathbb{1}_{\{0,1\}}(\Delta_{n,i})\big(b_{n,R_{n,i}} + \widehat{\beta}_n B_{n,R_{n,i}}\big) + \mathbb{1}_{\{-1\}}(\Delta_{n,i})(1+\widehat{\beta}_n)\cdot B_{n,R_{n,i}}, \tag{16}$$

a consistent estimator of the variance σ_b^2 *is given by*

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n \widehat{V}_n^2(i) - \left(\frac{1}{n}\sum_{i=1}^n \widehat{V}_n(i)\right)^2,$$

i.e. $\widehat{\sigma}_n^2 \to \sigma_b^2$ *in* $\mathscr{R}^\beta_{n,t\{c_{n,i}\}}$*-probability for all* $t \in \mathbb{R}$.

Proof. Because of corollary 2.3 it suffices to show the assertion under $\mathscr{H}'$. Let $V_{n,\beta}(i) = \mathbb{1}_{\{0,1\}}(\Delta_{n,i})\big(b_{n,R_{n,i}} + \beta B_{n,R_{n,i}}\big) + \mathbb{1}_{\{-1\}}(\Delta_{n,i})(1+\beta)\cdot B_{n,R_{n,i}}$ and $W_{n,\beta}(i) = \mathbb{1}_{\{0,1\}}(\Delta_{n,i})\big(b\circ F_0\circ H^{-1}(U_{n,i}) + \beta B\circ F_0\circ H^{-1}(U_{n,i})\big) + \mathbb{1}_{\{-1\}}(\Delta_{n,i})(1+\beta)\cdot B\circ F_0\circ H^{-1}(U_{n,i})$, $i = 1,\ldots,n$, be auxiliary r.v.s. The weak law of large numbers yields

$$\frac{1}{n}\sum_{i=1}^n W_{n,\beta}(i)^2 - \left(\frac{1}{n}\sum_{i=1}^n W_{n,\beta}(i)\right)^2 \xrightarrow{\mathscr{R}^\beta_{n,0}} \mathrm{Var}_{\mathscr{R}^\beta_{n,0}}\big(W_{n,\beta}(1)\big) = \sigma_b^2,$$

so that by showing

$$\frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i)^2 - \frac{1}{n}\sum_{i=1}^n W_{n,\beta}(i)^2 \xrightarrow{\mathscr{R}^\beta_{n,0}} 0, \qquad \left(\frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i)\right)^2 - \left(\frac{1}{n}\sum_{i=1}^n W_{n,\beta}(i)\right)^2 \xrightarrow{\mathscr{R}^\beta_{n,0}} 0 \tag{17}$$

and

$$\frac{1}{n}\sum_{i=1}^n \widehat{V}_n(i)^2 - \frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i)^2 \xrightarrow{\mathscr{R}^\beta_{n,0}} 0, \qquad \left(\frac{1}{n}\sum_{i=1}^n \widehat{V}_n(i)\right)^2 - \left(\frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i)\right)^2 \xrightarrow{\mathscr{R}^\beta_{n,0}} 0 \tag{18}$$

we obtain the assertion. Using the abbreviations $k_{n,1} = b_{n,R_{n,1}} - b \circ F_0 \circ H^{-1}(U_{n,i})$ and $K_{n,1} = B_{n,R_{n,1}} - B \circ F_0 \circ H^{-1}(U_{n,i})$, (28) and (29) imply

$$\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(V_{n,\beta}(1) - W_{n,\beta}(1)\big)^2 \leq 4\Big(\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(k^2_{n,1}\big) + \beta^2 \cdot \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(K_{n,1}\big)^2 + (\beta+1)^2 \cdot \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(K_{n,1}\big)^2\Big) \to 0 \quad \text{as } n \to \infty. \quad (19)$$

Obviously, it holds $\sup_{n\in\mathbb{N}}\Big(\mathrm{E}_{\mathscr{R}^\beta_{n,0}}(V_{n,\beta}(1) + W_{n,\beta}(1))^2\Big) < \infty$, so that applying the Cauchy-Schwarz-inequality and (19) give

$$\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(|V^2_{n,\beta}(1) - W^2_{n,\beta}(1)|\big) = E_{\mathscr{R}^\beta_{n,0}}\big(|V^2_{n,\beta}(1) + W^2_{n,\beta}(1)| \cdot |V^2_{n,\beta}(1) - W^2_{n,\beta}(1)|\big) \leq \sqrt{\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(V_{n,\beta}(1) + W_{n,\beta}(1)\big)^2} \cdot \sqrt{\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(V_{n,\beta}(1) - W_{n,\beta}(1)\big)^2} \to 0.$$

Note that under $\mathscr{R}^\beta_{n,0}$ the r.v.s $(V_{n,\beta}(i), W_{n,\beta}(i))$, $i = 1, \ldots, n$, are identically distributed , s. (24) and (25). It holds

$$\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\left|\frac{1}{n}\sum_{i=1}^n V^2_{n,\beta}(i) - \frac{1}{n}\sum_{i=1}^n W^2_{n,\beta}(i)\right| = \frac{1}{n}\sum_{i=1}^n \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\left|V^2_{n,\beta}(i) - W^2_{n,\beta}(i)\right| = \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\left|V^2_{n,\beta}(1) - W^2_{n,\beta}(1)\right| \to 0$$

implying the first part of (17). Applying the same arguments one derives the inequality

$$\begin{aligned}
&\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\left|\left(\frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i)\right)^2 - \left(\frac{1}{n}\sum_{i=1}^n W_{n,\beta}(i)\right)^2\right| \\
&= \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\left|\left(\frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i) + \frac{1}{n}\sum_{i=1}^n W_{n,\beta}(i)\right)\left(\frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i) - \frac{1}{n}\sum_{i=1}^n W_{n,\beta}(i)\right)\right| \\
&\leq \frac{1}{n^2}\sum_{i,j=1}^n \mathrm{E}_{\mathscr{R}^\beta_{n,0}}\left|\big(V_{n,\beta}(i) + W_{n,\beta}(i)\big)\big(V_{n,\beta}(j) - W_{n,\beta}(j)\big)\right| \\
&\leq \frac{1}{n^2}\sum_{i,j=1}^n \sqrt{\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(V_{n,\beta}(i) + W_{n,\beta}(i)\big)^2}\sqrt{\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(V_{n,\beta}(j) - W_{n,\beta}(j)\big)^2} \\
&= \sqrt{\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(V_{n,\beta}(1) + W_{n,\beta}(1)\big)^2}\sqrt{\mathrm{E}_{\mathscr{R}^\beta_{n,0}}\big(V_{n,\beta}(1) - W_{n,\beta}(1)\big)^2} \longrightarrow 0,
\end{aligned}$$

i.e. the second part of (17). Since the sums $\frac{1}{n}\sum_{i=1}^n |b_{n,i} \cdot B_{n,i}|$, $\frac{1}{n}\sum_{i=1}^n B^2_{n,i}$ and $\frac{1}{n}\sum_{i=1}^n b^2_{n,i}$ are bounded, applying lemma 3.1 to

$$\frac{1}{n}\sum_{i=1}^n \widehat{V}_n(i)^2 - \frac{1}{n}\sum_{i=1}^n V_{n,\beta}(i)^2 = \frac{1}{n}\sum_{i=1}^n\Big(\mathbb{1}_{\{0,1\}}(\Delta_{n,i})(\beta^2 - \widehat{\beta}^2_n)B^2_{n,R_{n,i}} + 2(\beta - \widehat{\beta}_n)b_{n,R_{n,i}}B_{n,R_{n,i}} + \mathbb{1}_{\{-1\}}(\Delta_{n,i})\big((1+\beta)^2 - (1+\widehat{\beta}_n)^2\big)B^2_{n,R_{n,i}}\Big)$$

gives the first assertion of (18). By the same arguments one shows the second part of (18). □

Obviously, corollary 3.5 also holds for the sequence of tests $\phi_n = \mathbb{1}(\widehat{S}_n(b_n, B_n) \geq u_\alpha\sqrt{\widehat{\sigma}^2_n})$, $n \geq 2$. In the next step we are going to derive tests that are distribution free under $\mathscr{H}_0$ with finite sample sizes. Consider the test statistic

$$\widehat{T}_n\big(D_n, X_{n,\uparrow}, \Delta_{n,\uparrow}\big) = \sum_{i=1}^n c_{n,D_{n,i}}\big(\mathbb{1}_{\{0,1\}}(\Delta_{n:i}) \cdot (b_{n,i} + \widehat{\beta}_n B_{n,i}) + \mathbb{1}_{\{-1\}}(\Delta_{n:i}) \cdot (1+\widehat{\beta}_n) \cdot B_{n,i}\big).$$

Clearly, it holds $\widehat{T}_n(D_n, X_{n,\uparrow}, \Delta_{n,\uparrow}) = \widehat{S}_n(b_n, B_n)$. Note that the estimators $\widehat{\beta}_n$, $b_{n,i}$ and $B_{n,i}$, $i = 1, \ldots, n$, merely depend on the order statistics $X_{n,\uparrow}$ and $\Delta_{n,\uparrow}$. Given the order statistics, the distribution of $\widehat{T}_n(D_n, X_{n,\uparrow}, \Delta_{n,\uparrow})$ is just dependent on the inverse rank vector which is uniformly distributed on the set of Permutations $\mathfrak{S}_n$, i.e. the distribution of the statistic is principally known and can easily be approximated by simulations. This leads to the introduction of conditional permutation tests.

Let $F_{n,x_{n,\uparrow}\delta_{n,\uparrow}}$ denote the c.d.f. of the distribution $\mathfrak{L}_{\mathscr{R}^{\beta}_{n,0}}\big(\widehat{T}_n(D_n, x_{n,\uparrow}, \delta_{n,\uparrow})\big)$, $\delta_{n,\uparrow} \in \{-1, 0, 1\}^n$ and $x_{n,\uparrow} \in \mathbb{R}^n_<$. For every $\alpha \in (0,1)$ there are real numbers $\gamma^n_\alpha(x_{n,\uparrow}, \delta_{n,\uparrow})$ and $c^n_\alpha(x_{n,\uparrow}, \delta_{n,\uparrow})$ satisfying the condition

$$\int \mathbb{1}_{\{c^n_\alpha(x_{n,\uparrow},\delta_{n,\uparrow})\}} + \gamma^n_\alpha(x_{n,\uparrow}, \delta_{n,\uparrow}) \cdot \mathbb{1}_{(c^n_\alpha(x_{n,\uparrow},\delta_{n,\uparrow}),\infty)} \,\mathrm{d}F_{n,\delta_{n,\uparrow}} = \alpha.$$

The sequence of tests defined by

$$\psi_n(b_n, B_n, X_{n,\uparrow}, D_n, \Delta_{n,\uparrow}) = \begin{cases} 1, & \qquad > \\ \gamma^n_\alpha(X_{n,\uparrow}, \Delta_{n,\uparrow}), & \widehat{T}_n(D_n, X_{n,\uparrow}, \Delta_{n,\uparrow}) \;=\; c^n_\alpha(X_{n,\uparrow}\Delta_{n,\uparrow}) \\ 0, & \qquad < \end{cases}$$

is called conditional permutation test of the level α. Obviously, this test holds the level, i.e. $\mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(\psi_n(b_n, B_n, D_n, X_{n,\uparrow}, \Delta_{n,\uparrow})\big) = \alpha$, for all $n \geq 2$. The next result enables us to characterize the asymptotic properties of these tests.

3.7 Theorem. *If $\sigma^2_b > 0$ it holds for all $t \in \mathbb{R}$ that*

$$\sup_{y\in\mathbb{R}} \left| F_{n,X_{n,\uparrow}\Delta_{n,\uparrow}}(y) - \Phi\left(\frac{y}{\sqrt{\sigma^2_b}}\right) \right| \longrightarrow 0 \quad \text{in } \mathscr{R}^{\beta}_{n,t\{c_{n,i}\}}\text{-probability.}$$

Proof. The proof given here relies on the same ideas as the proof of theorem 5.2 in Neuhaus (1988) [7]. Because of corollary 2.3 it suffices to verify the assertion under $\mathscr{H}'_0$. Without loss of generality, we can assume the r.v.s $(X_{n,i}, \Delta_{n,i})$, $i = 1, \ldots, n$ are defined on the probability space $(\Omega, \mathscr{A}, P) = \{\bigtimes_{n=2}^\infty \Omega_n, \bigotimes_{n=2}^\infty \mathscr{A}_n, \bigotimes_{n=2}^\infty \mathscr{R}^{\beta}_{n,0}\}$

For the proof another probability space $(\Omega', \mathscr{A}', P')$ is needed on which the r.v.s $U_{n,i}$, $i \leq 1 \leq n$, $n \geq 2$, are defined. It is assumed that $U_{n,1}, \ldots, U_{n,n}$ are i.i.d. and uniformly distributed on $(0,1)$. By $R'_n = (R'_{n,1}, \ldots, R'_{n,n})$ and $D'_n = (D'_{n,1}, \ldots, D'_{n,n})$ the rank vector and the inverse rank vector of $(U_{n,1}, \ldots, U_{n,n})$ are denoted. Analogue to equation (16) we define for every $\omega \in \Omega$, $i \in \{1, \ldots, n\}$ and $n \geq 2$ the quantities $v_n(i,\omega) = \mathbb{1}_{\{0,1\}}(\Delta_{n:i}(\omega)) \cdot (b_{n,i}(\omega) + \widehat{\beta}_n(\omega) B_{n,i}(\omega)) + \mathbb{1}_{\{-1\}}(\Delta_{n:i}(\omega)) \cdot (1 + \widehat{\beta}_n(\omega)) \cdot B_{n,i}(\omega)$ and $\overline{v}_n(\omega) = \frac{1}{n}\sum_{i=1}^n v_n(i,\omega)$. It holds $\frac{1}{n}\sum_{i=1}^n (v_n(i,\cdot) - \overline{v}_n(\cdot))^2 = \widehat{\sigma}^2_n$, whereas $\widehat{\sigma}^2_n$ is the variance estimator of lemma 3.6.

Because of lemma 3.1, lemma 3.6 and (12) it exists a subsequence $\{n_k\}$ in every subsequence of the natural numbers and a set $N \in \mathscr{A}$, $P(N) = 1$, such that following assertions hold true:

$$\lim_{k\to\infty} \int_{(0,1)} (b_{n_k} - b \circ F_0 \circ H^{-1}) \,\mathrm{d}\lambda = 0, \qquad \lim_{k\to\infty} \int_{(0,1)} (B_{n_k} - B \circ F_0 \circ H^{-1}) \,\mathrm{d}\lambda = 0,$$

$$\lim_{k\to\infty} \widehat{\beta}_{n_k}(\omega) = \beta \text{ and } \lim_{k\to\infty} \frac{1}{n_k}\sum_{i=1}^{n_k} (v_{n_k}(i,\omega) - \overline{v}_{n_k}(\omega))^2 = \sigma^2_b$$

for all $\omega \in N$.

For fixed $\omega \in N$ we define the linear statistic $T_n(\omega) = \sum_{i=1}^n c_{n,i} v_n(\lfloor n \cdot U_{n,i} \rfloor + 1, \omega)$, $n \geq 2$ and show by the central limit theorem of Lindeberg-Feller, cf. Witting and Müller-Funk (1995) [10, Korollar 5.102, p. 103], that it converges in distribution to some normal distribution. Because of (2), (5) and $U_{n,i}$, $1 \leq i \leq n$, being i.i.d. it holds for all $n \geq 2$ $\mathrm{E}_{P'}\big(T_n(\omega)\big) = 0$ and $s^2_n := \mathrm{Var}_{P'}\big(T_n(\omega)\big) = \sum_{i=1}^n c^2_{n,i} \mathrm{Var}_{P'}\big(v_n(\lfloor n \cdot U_{n,i} \rfloor + 1, \omega)\big) = \mathrm{Var}_{P'}\big(v_n(R'_{n,1}, \omega)\big) = \frac{1}{n}\sum_{i=1}^n (v_n(i,\omega) - \overline{v}_n(\omega))^2$, whereas the fact $\lfloor nU_{n,i} + 1 \rfloor \sim R'_{n,i}$ is used. In particular it holds

$$\lim_{k\to\infty} s^2_{n_k} = \sigma^2_b. \tag{20}$$

In the next step the Lindeberg condition is verified. For $\varepsilon > 0$ we have the estimate $|v_n(\lfloor nu+1 \rfloor, \omega)| \leq |b_{n,\lfloor nu+1 \rfloor}| + (1+\widehat{\beta}_n(\omega)) \cdot |B_{n,\lfloor nu+1 \rfloor}| = |b_n(u,\omega)| + (1+\widehat{\beta}_n(\omega))| \cdot B_n(u,\omega)| =: w_n(u,\omega)$ for a.e. $u \in (0,1)$. Since $U_{n,1}, \dots, U_{n,n}$ are i.i.d. for $n \geq 2$ it holds

$$\begin{aligned}\frac{1}{s_n^2}\sum_{i=1}^n \int_{\{|c_{n,i}v_n(\lfloor nU_{n,i}+1 \rfloor,\omega)| \geq \sqrt{s_n^2}\}} & c_{n,i}^2 v_n^2(\lfloor nU_{n,i}+1 \rfloor, \omega)\, \mathrm{d}P' \\ &= \frac{1}{s_n^2}\sum_{i=1}^n c_{n,i}^2 \int_{\{\max_{1\leq i\leq n}\{|c_{n,i}|\}\cdot w_n(U_{n,1}) \geq \sqrt{s_n^2}\}} w_n^2(U_{n,i},\omega)\, \mathrm{d}P' \\ &= \underbrace{\frac{1}{s_n^2} \int_{\{\max_{1\leq i\leq n}\{|c_{n,i}|\}\cdot w_n(U_{n,1}) \geq \sqrt{s_n^2}\}} w_n^2(U_{n,1},\omega)\, \mathrm{d}P'}_{I_n}.\end{aligned}$$

It holds $\lim_{k\to\infty} I_{n_k} = 0$, since firstly the functions b_n and B_n converge to the functions b and B in quadratic mean, secondly $\lim_{k\to} s_{n_k}^2 = \sigma_b^2 > 0$ and thirdly (5) holds. Thus, the Lindeberg condition is valid and the central limit theorem gives

$$\mathfrak{L}_{P'}\big(T_{n_k}(\omega)\big) \xrightarrow{\mathfrak{L}} \mathscr{N}(0,\sigma_b^2) \quad \text{as } k \to \infty. \tag{21}$$

For the next step in the proof we need the following assertion

$$\frac{1}{\sqrt{n_k}} \max_{1\leq i\leq n_k} |v_{n_k}(i,\omega) - \overline{v}_{n_k}(\omega)| \to 0. \tag{22}$$

It holds

$$\begin{aligned}\frac{1}{\sqrt{n_k}} \max_{1\leq i\leq n_k} |v_{n_k}(i,\omega) - \overline{v}_{n_k}(\omega)| \leq \frac{1}{\sqrt{n_k}} \max_{1\leq i\leq n_k} |v_{n_k}(i,\omega)| + \frac{1}{\sqrt{n_k}} \max_{1\leq i\leq n_k} |\overline{v}_{n_k}(\omega)| \leq \\ 2\frac{1}{\sqrt{n_k}} \max_{1\leq i\leq n_k} |v_{n_k}(i,\omega)| \leq (1+\widehat{\beta}_{n_k}(\omega)) \cdot \frac{1}{\sqrt{n_k}} \max_{1\leq i\leq n_k} |B_{n_k,i}(\omega)| + \frac{1}{\sqrt{n_k}} \max_{1\leq i\leq n_k} |b_{n_k,i}(\omega)| =: z_{n_k}(\omega).\end{aligned}$$

Since ${n_k}^{-1/2}\max\{|b_{n_k,i}(\omega)| : i=1,\dots,n_k\}$ and ${n_k}^{-1/2}\max\{|B_{n_k,i}(\omega)| : i=1,\dots,n_k\}$ converge to 0 as $k \to \infty$, cf. Neuhaus (1988) [7, Proof of theorem 5.2], we have $\lim_{k\to\infty} z_{n_k}(\omega) = 0$.

Because of (20) and (22) theorem 3.1 in Hájek (1961) [4] can be applied and one obtains

$$\mathrm{E}_{P'}\left(\sum_{i=1}^{n_k} c_{n_k,i} v_{n_k}(R'_{n_k,i},\omega) - T_{n_k}(\omega)\right)^2 \longrightarrow 0 \quad \text{as } k \to \infty.$$

This result and (21) lead to

$$\mathfrak{L}_{P'}\big(\sum_{i=1}^{n_k} c_{n_k,i} v_{n_k}(R'_{n_k,i},\omega)\big) \xrightarrow{\mathfrak{L}} \mathscr{N}(0,\sigma_b^2) \quad \text{as } k \to \infty. \tag{23}$$

With $x_{n,\uparrow} = X_{n,\uparrow}(\omega) \in \mathbb{R}^n_<$ and $\delta_{n,\uparrow} = \Delta_{n,\uparrow}(\omega) \in \{-1,0,1\}^n$ it holds the following identities

$$\begin{aligned}\mathfrak{L}_{P'}\big(\sum_{i=1}^n c_{n,i} v_n(R'_{n,i},\omega)\big) = \mathfrak{L}_{P'}\big(\sum_{i=1}^n c_n v_n(D'_{n,i},\omega)\big) = \\ \mathfrak{L}_{P'}\big(\widehat{T}_n(D'_n, X_{n,\uparrow}(\omega)\Delta_{n,\uparrow}(\omega)\big) = \mathfrak{L}_P\big(\widehat{T}_n(D_n, x_{n,\uparrow}, \delta_{n,\uparrow})\big).\end{aligned}$$

The statistic $\widehat{T}_{n_k}(D_{n_k}, x_{n_k,\uparrow}, \delta_{n_k,\uparrow})$ has c.d.f. $F_{n_k, X_{n_k,\uparrow}(\omega)\Delta_{n_k,\uparrow}(\omega)}$, therefore (23) implies

$$\sup_{x\in\mathbb{R}} \left| F_{n_k, X_{n_k,\uparrow}(\omega)\Delta_{n_k,\uparrow}(\omega)}(x) - \Phi\left(\frac{x}{\sqrt{\sigma_b^2}}\right)\right| \to 0 \quad \text{as } k \to \infty,$$

by subsequence principle it results the assertion. □

3.8 Corollary. *It holds* $c_\alpha^n(X_{n,\uparrow}, \Delta_{n,\uparrow}) \longrightarrow u_\alpha \cdot \sqrt{\sigma_b^2}$ *in* $\mathscr{R}^\beta_{n,t\{c_{n,i}\}}$*-probability for all* $t \in \mathbb{R}$.

Proof. Since Φ is a strictly monotone function, Φ^{-1} is a continuous function. Theorem 3.7 and Satz 5.76 in Witting and Müller-Funk (1995) [10, p. 71] give the result. ☐

One easily sees that the sequences of tests φ_n and ψ_n, $n \geq 2$, are asymptotically equivalent, therefore the assertions of corollary 3.5 also hold for the sequence of tests ψ_n, $n \geq 2$.

4 Examples and Simulations

In this section we present some examples and the results of our simulations. Typical families of survival time distributions are given by families of Weibull distributions.

4.1 Example. Let $\mathscr{W}(\alpha,\gamma)$ be a distribution on $(\mathbb{R},\mathbb{B})$ with Lebesgue density

$$f(x) = \alpha\gamma x^{\gamma-1}\exp\left(-\alpha x^\gamma\right) \cdot \mathbb{1}_{(0,\infty)}(x) \quad [\lambda], \qquad \alpha,\gamma > 0.$$

$\mathscr{W}(\alpha,\gamma)$ is obviously a Weibull distribution. Satz 1.194 in Witting (1985) [9, p. 179] gives that the distribution family $\{\mathscr{W}(\alpha - t,\gamma) \mid t \in (-\alpha,\alpha)\}$ is $\mathbb{L}_2(\mathscr{W}(\alpha,\gamma))$-differentiable with $\mathbb{L}_2$-derivative $\dot{L}(x) = x^\gamma - 1/\alpha$ $[\mathscr{W}(\alpha,\gamma)]$. Assume that the non-informative censoring time distribution is bounded, i.e. $\tau_G \leq c \in (0,\infty)$. Theorem 2.6 gives that the distribution family $\mathfrak{R}^\beta$ is $\mathbb{L}_2(\mathscr{R}_0^\beta)$-differentiable with $\mathbb{L}_2$-derivative

$$\dot{h}_\beta(x,\delta) = \mathbb{1}_{\{0,1\}}(\delta)\left(x^\gamma - \frac{1}{\alpha}\right) + \mathbb{1}_{\{-1\}}(\delta)\, x^\gamma + \beta x^\gamma \quad [\mathscr{R}_0^\beta].$$

Set $\widetilde{c} = F_0(c)$, $F_0(x) := (1 - \exp(-\alpha \cdot x^\gamma)) \cdot \mathbb{1}_{(0,\infty)}(x)$. It holds the identity

$$\dot{h}_\beta(x,\delta) = \mathbb{1}_{\{0,1\}}(\delta)\widetilde{b} \circ F_0(x) + \mathbb{1}_{\{-1\}}(\delta)\widetilde{B} \circ F_0 + \beta\widetilde{B} \circ F_0 \quad [\mathscr{R}_0^\beta],$$

whereas $\widetilde{b}, \widetilde{B} : ((0,1), \mathbb{B} \cap (0,1)) \to (\mathbb{R},\mathbb{B})$

$$\widetilde{b}(u) := -\frac{1}{\alpha} \cdot \begin{cases} 1 + \log(1-u), & u \leq \widetilde{c} \\ \log(1-\widetilde{c}), & u > \widetilde{c} \end{cases} \qquad \widetilde{B}(u) := -\frac{1}{\alpha} \cdot \begin{cases} \log(1-u), & u \leq \widetilde{c} \\ \log(1-\widetilde{c}), & u > \widetilde{c} \end{cases}.$$

Because of corollary 3.5 we can substitute the functions b and B for $\widetilde{b}$ and $\widetilde{B}$

$$b(u) := -\begin{cases} 1 + \log(1-u), & u \leq \widetilde{c} \\ \log(1-\widetilde{c}), & u > \widetilde{c} \end{cases} \qquad B(u) := -\begin{cases} \log(1-u), & u \leq \widetilde{c} \\ \log(1-\widetilde{c}), & u > \widetilde{c} \end{cases}$$

Applying the results of the previous section we can derive asymptotically optimal rank and permutation tests. In the case that the distributions of the survival times belong to a family of Weibull distributions $\{\mathscr{W}(\alpha - t,\gamma) \mid t \in (-\alpha,\alpha)\}$ that satisfy the condition $\mathscr{W}(\alpha,\gamma)\{(0,\tau_G]\} \leq \widetilde{c}$ the sequences of test ϕ_n and ψ_n, $n \geq 2$, are asymptotically optimal. Note, the functions b and B do not depend on the parameters of the Weibull distribution family.

In the simulations observations $X = ((X_{n,i}, \Delta_{n,i}),\ = 1,\dots,n)$ satisfying $\mathfrak{L}(X) = \mathscr{R}^\beta_{t\{c_{n,i}\}}$, $t \in \{-0.9 + j \cdot 0.1 \mid j = 0,\dots,60\}$, were generated. The power functions of the different tests are estimated with these observations. For given t the test was evaluated 20 000 times. The tests employed are of the level 5 %. Note that $n = n_1 + n_2$ is the size of the pooled sample.

4.2 Simulation. In this simulation the distribution of the survival times belongs to the distribution family $\{\mathscr{W}(1 - t, 1) \mid t \in (-1,1)\}$, i.e. the survival times are distributed according to an exponential law. The non-informative censoring time is uniformly distributed on the interval $(0,4)$, $\beta = 1.3$ and $n_1 = n_2 = 50$. In this situation the optimal parametric test is used, s. corollary 2.3.

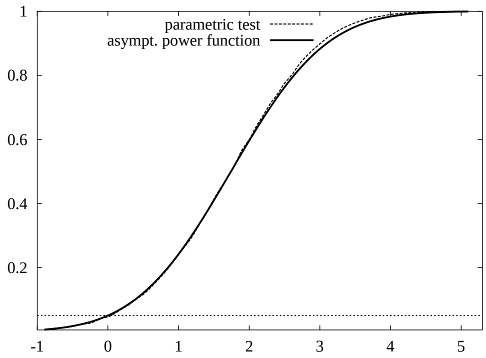

Fig. 1: $n_1 = n_2 = 50$

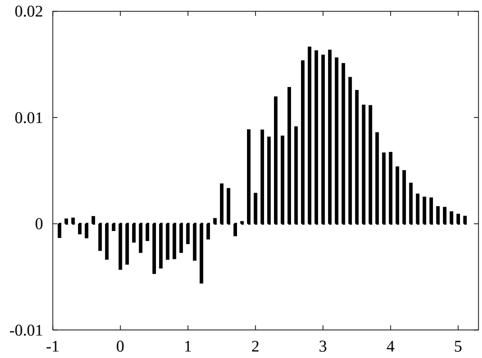

Fig. 2: $n_1 = n_2 = 50$

Figure 1 shows the asymptotic power function (solid line) and the simulated power function of the optimal parametric test (dashed line). In figure 2 the difference between the asymptotic and the simulated power function is displayed, i.e. $\Phi(-u_{0.05} + t_j \cdot \sigma_b) - \widehat{\eta}(t_j)$, $t_j = -0.9 + j \cdot 0.1$, $j = 0, \ldots, 60$, whereas $\widehat{\eta}$ denotes the simulated power function. One sees that the simulated power function fits well the asymptotic power function.

4.3 Simulation. The same setting as in simulation 4.2 is used. Instead of the parametric test, the optimal rank test ϕ_n, with estimated scores, estimated variance and the estimated model parameter β is utilized.

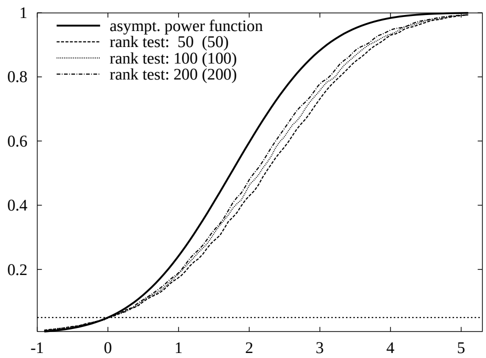

Fig. 3: power functions

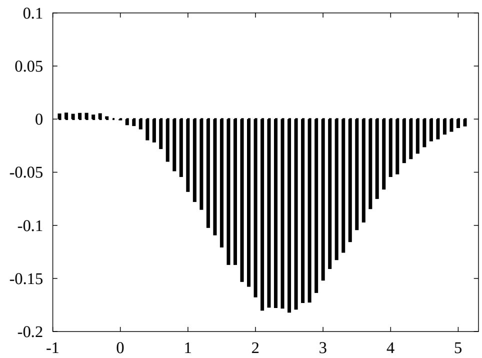

Fig. 4: $n_1 = n_2 = 50$

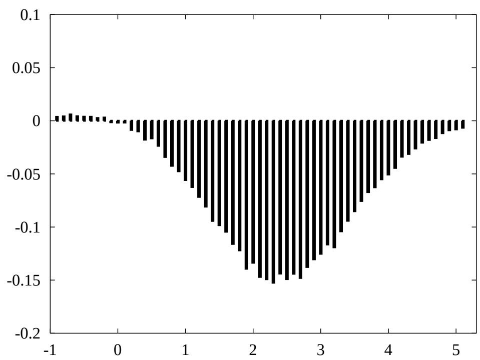

Fig. 5: $n_1 = n_2 = 100$

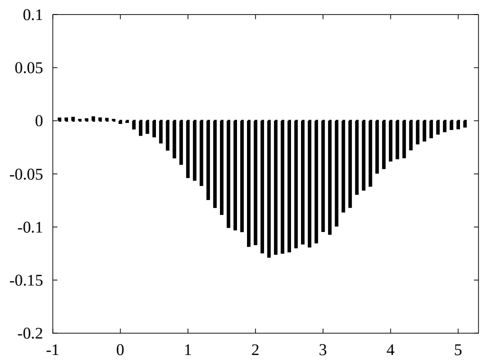

Fig. 6: $n_1 = n_2 = 200$

Figure 3 shows the asymptotic power function and the simulated power functions of the optimal rank test for different sample sizes. In the figures 4-6 the differences between the asymptotic power function and the simulated power functions are displayed. With increasing sample sizes one observes that this difference decreases. Note that under $\mathscr{H}_0'$ about 11 % of the observations are non-informatively censored.

4.4 Simulation. In this simulation the same setting as in simulation 4.3 is used. But this time the distribution of the maximum of two s.i., uniformly on the interval $(0, 4)$ distributed r.v.s is used as law of the non-informative censoring time. In this situation only about 3 % of the observations are non-informatively censored. This leads to significant gain of power of the rank test in comparison to simulation 4.3.

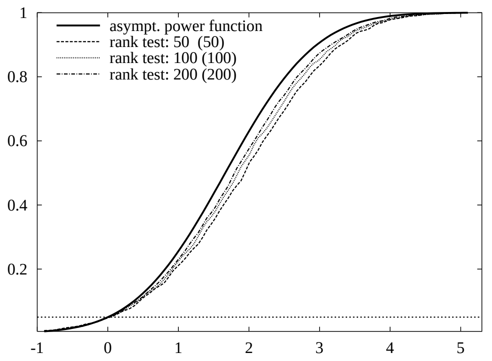

Fig. 7: power functions

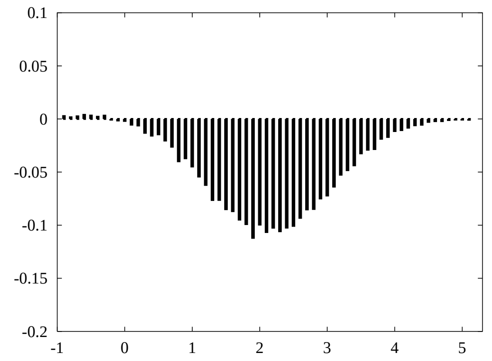

Fig. 8: $n_1 = n_2 = 50$

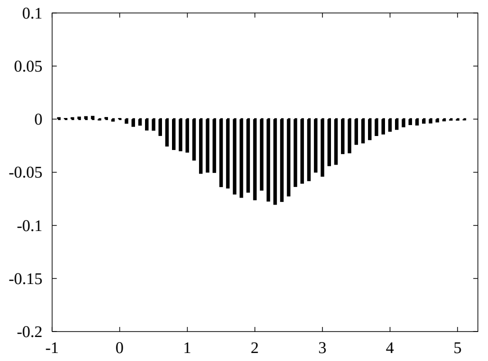

Fig. 9: $n_1 = n_2 = 100$

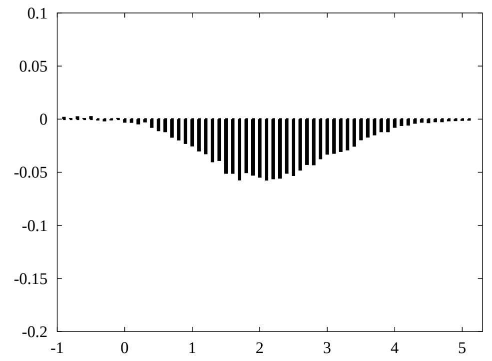

Fig. 10: $n_1 = n_2 = 200$

Again, the difference between the simulated power function and the asymptotic power functions decreases with increasing sample sizes, s. figures 8 - 10.

4.5 Simulation. In the setting of simulation 4.3 the properties of the optimal rank test are investigated, if the control and the test group have different sizes.

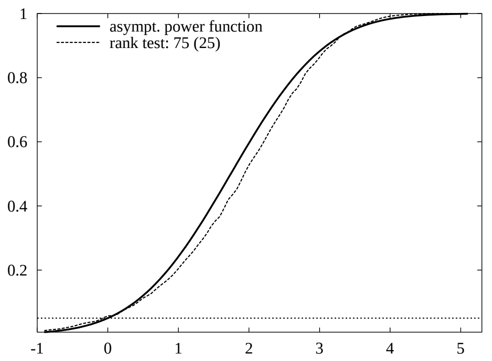

Fig. 11: power functions

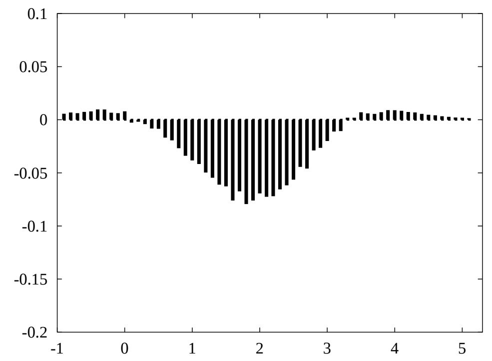

Fig. 12: $n_1 = 75$ $n_2 = 25$

In the case that there are more observations in the control group, the power of test the is better in comparison to simulation 4.3, s. figure 11 and 12. If the control group is smaller than the test group we see a decrease in the power of the optimal rank test. However, the difference between the simulated power function and the asymptotic power function decreases with increasing number of observations.

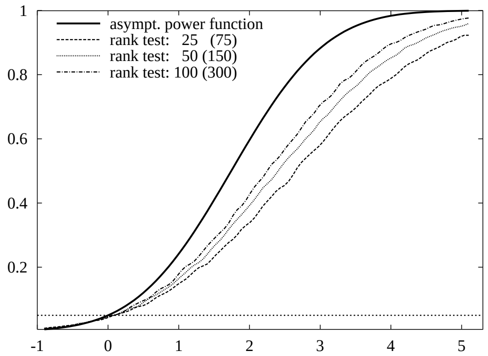

Fig. 13: power functions

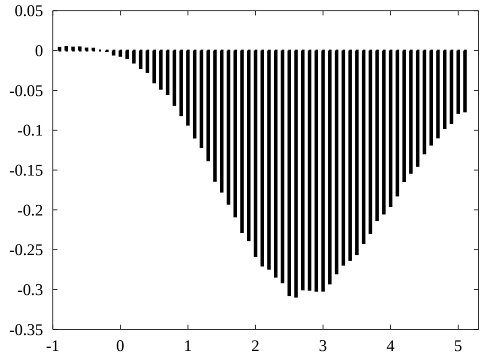

Fig. 14: $n_1 = 25$ $n_2 = 75$

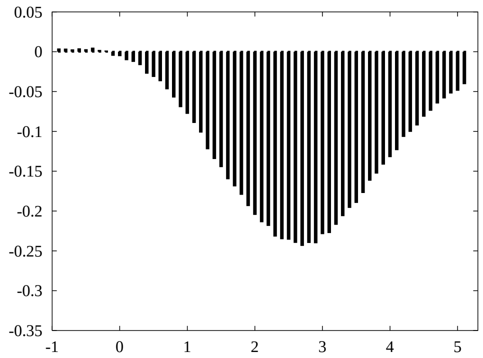

Fig. 15: $n_1 = 50$ $n_2 = 150$

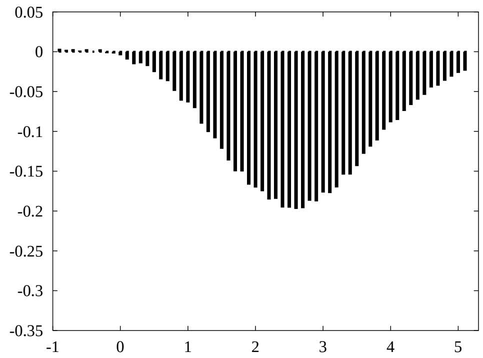

Fig. 16: $n_1 = 100$ $n_2 = 300$

4.6 Simulation. In the situation of simulation 4.5 the optimal permutation test instead of the optimal rank test is used. For estimating the 95 % quantile of the permutation test 6000 random permuations were generated. In figure 17 we see the simulated power functions of the rank test (solid line) and the permutation in the case $n_1 = 75$ and $n_2 = 25$. In figure 18 the difference between the simulated power functions of the permutation test and the rank test is displayed.

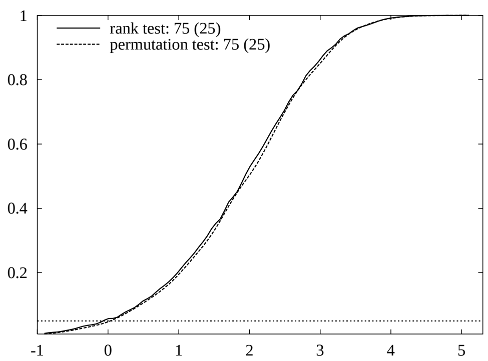

Fig. 17: power functions

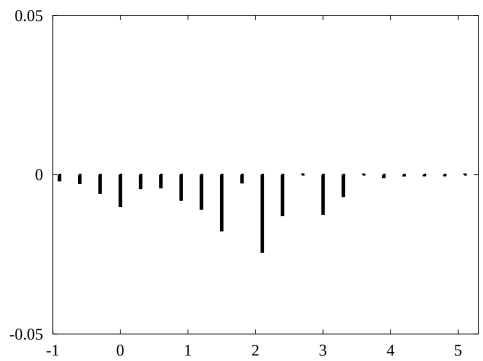

Fig. 18: $n_1 = 75$ $n_2 = 25$

The results for sample sizes ($n_1 = 25$ and $n_2 = 75$) are shown in figure 19 and 20.

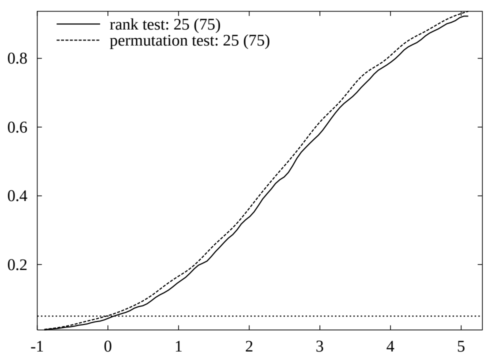

Fig. 19: power functions

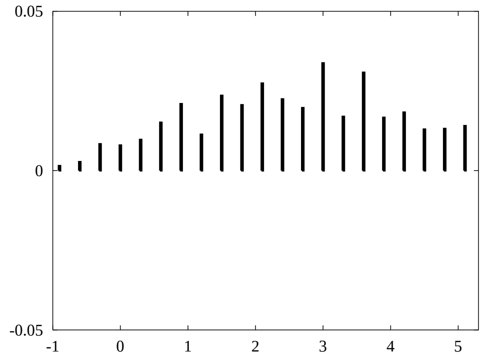

Fig. 20: $n_1 = 25$ $n_2 = 75$

In this case there are no considerable differences between the performance of the rank and the permutation test.

Another typical example for survival time distributions are log-logistic distributions.

4.7 Example. Let $\mathscr{L}(\alpha, \gamma)$ denote a distribution on $(\mathbb{R}, \mathbb{B})$ with Lebesgue density

$$f(x) = \frac{\gamma}{x} \cdot \left(\frac{x}{\alpha}\right)^{\gamma} \cdot \left(1 + \left(\frac{x}{\alpha}\right)^{\gamma}\right)^{-2} \cdot \mathbb{1}_{(0,\infty)}(x) \quad [\lambda], \qquad \alpha, \gamma > 0.$$

$\mathscr{L}(\alpha, \gamma)$ is called a log-logistic distribution. Satz 1.194 in Witting (1985) [9, p. 179] gives that the distribution family $\mathfrak{F} = \{\mathscr{L}(\alpha + t, \gamma) \mid t \in (-\alpha, \alpha)\}$ is $\mathbb{L}_2(\mathscr{L}(\alpha, \gamma))$ with $\mathbb{L}_2$-derivative

$$\dot{L} = -\frac{\gamma}{\alpha} \left(1 - 2\left(1 + \left(\frac{x}{\alpha}\right)^{\gamma}\right)^{-1} \left(\frac{x}{\alpha}\right)^{\gamma}\right) \cdot \mathbb{1}_{(0,\infty)}(x) \quad [\mathscr{L}(\alpha, \gamma)].$$

Using the same theorem one verifies that the distribution family $\mathfrak{F}^{\beta}$, $\beta > 0$, is $\mathbb{L}_2$-differentiable, hence theorem 2.4 yields that the distribution family $\mathfrak{R}^{\beta}$ is $\mathbb{L}_2(\mathscr{R}_0^{\beta})$ differentiable with $\mathbb{L}_2$-derivative

$$\dot{h}_{\beta}(x, \delta) = \mathbb{1}_{\{0,1\}}(\delta) \cdot \widetilde{b} \circ F_0(x) + \mathbb{1}_{\{-1\}}(\delta) \cdot \widetilde{B} \circ F_0(x) + \beta \cdot \widetilde{B} \circ F_0 \quad [\mathscr{R}_0^{\beta}],$$

whereas

$$F_0(x) = \left(1 + \left(\frac{x}{\alpha}\right)^{\gamma}\right)^{-1} \cdot \left(\frac{x}{\alpha}\right)^{\gamma} \cdot \mathbb{1}_{(0,\infty)}(x), \quad \widetilde{b}(u) = \frac{\gamma}{\alpha} \cdot (2u - 1) \quad \text{and} \quad \widetilde{B}(u) = \frac{\gamma}{\alpha} \cdot u.$$

Note, F_0 is the c.d.f. of $\mathscr{L}(\alpha, \gamma)$. Because of corollary 3.5 we can substitute the functions $b(u) = 2u - 1$ and $B(u) = u$ for $\widetilde{b}$ and $\widetilde{B}$. Applying the results of the previous section we obtain asymptotically optimal rank and permutation tests provided that the distribution of the survival times belongs to a family of log-logistic distributions.

4.8 Simulation. In this simulation the distribution of the survival times belongs to the distribution family $\{\mathscr{L}(1.8 + t, 2.2) \mid t \in (-1.8, 1.8)\}$, i.e. the survival times are distributed according to a log-logistic law. The non-informative censoring time is uniformly distributed on the interval $(0, 4)$, $\beta = 1.3$ and $n_1 = n_2 = 50$. In this setting the optimal parametric test, s. corollary 2.3, and the optimal rank test ϕ_n with estimated variance and estimated model parameter β are used. The parameters of distributions are chosen in a way that a lot of the observations are non-informatively censored. Under $\mathscr{H}_0'$, about 37 % of the observations are non-informatively censored. In figure 21 we see the simulated power functions of the parametric and the rank test as well as the asymptotic power function. The rank test performs considerably worse than the parametric test. The number of non-informatively censored observations has apparently a significant impact on the power of the rank test, s. also simulation 4.3 and 4.4.

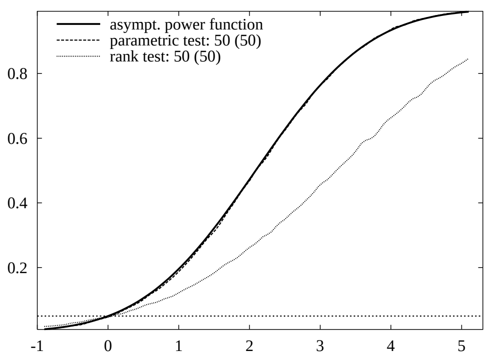

Fig. 21: power functions

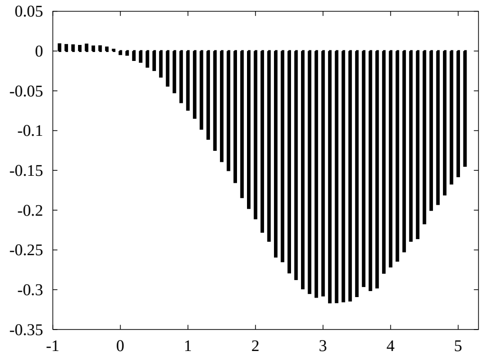

Fig. 22: $n_1 = n_2 = 50$

4.9 Simulation. In this simulation the same setting as in the previous simulation is used, but instead of a uniformly distributed non-informative censoring time the censoring time of simulation 4.4 is employed. As a consequence there are less non-informatively censored observations in this situation. Under $\mathscr{H}_0'$ merely about 14 % of the observations are non-informatively censored.

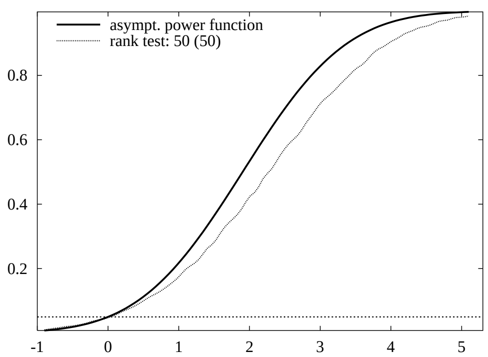

Fig. 23: power functions

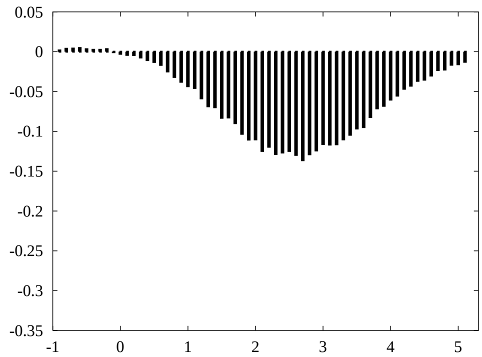

Fig. 24:$n_1 = n_2 = 50$

In figure 23 the simulated and the asymptotic power function are displayed. In figure 24 the difference between them is shown. The power of the rank test is better compared to simulation 4.8, s. also simulation 4.3 and 4.4

4.10 Simulation. Analogue to simulation 4.5 the properties of the rank test are investigated in case that the sample sizes of the control and the test group are different. We chose the cases $(n_1, n_2) = (75, 25)$ and $(n_1, n_2) = (25, 75)$, again. The results of the simulation in the setting of simulation 4.8 is displayed in figure 25. In figure 26 the result of the simulations in the setting of simulation 4.9 is shown.

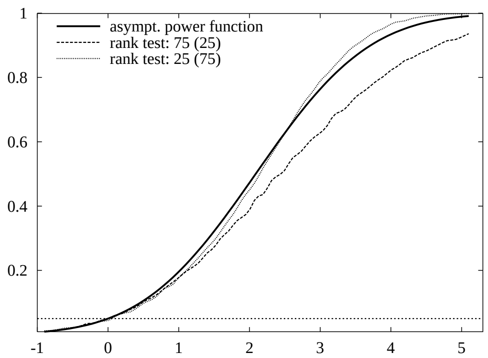

Fig. 25: power functions

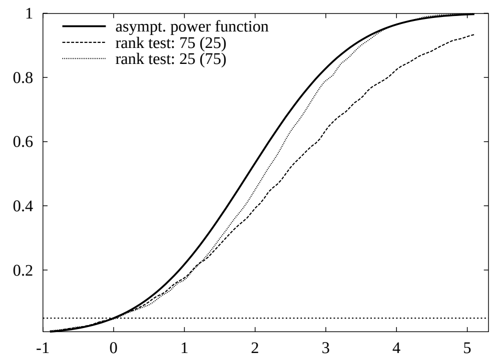

Fig. 26: power functions

In simulation 4.5 it turned out that the power of the rank test was better, if there were more observations in the control group than is the tests group, i.e. $n_1 > n_2$. In this simulation the contrary is can be observed.

4.11 Discussion. The simulations show that the simulated power functions seem to converge to the asymptotic power function, if the size of the pooled sample size is increased, s. simulation 4.3, 4.4 and 4.5. This underlines the results of section 2 and 3. Moreover, the power of the optimal rank test is considerably worse than the power of the optimal parametric test. However, this is the price to pay for the fact that rank tests are asymptotically distribution free under the null hypothesis. The power of the optimal rank test seems to correlate with the number of non-informatively censored observations. The more observations are non-informatively censored the less powerful is the corresponding rank test, s. simulation 4.3, 4.4, 4.8 and 4.9. In case that the number of observations in the control and test group was different, we obtained inconsistent results, s. simulation 4.5 and 4.10. In such situations one can hardly predict the properties of the rank test, if it is applied to small sample sizes. Certainly, one can easily overestimate the asymptotics.

5 Auxiliary Results

In this section the auxiliary results used in previous proofs are given.

5.1 Lemma. *For $B_1 \in \mathbb{B}$ and $B_2 \in \mathcal{P}\{-1,0,1\}$ it holds $\mathscr{R}_t^\beta\big(B_1 \times B_2\big) = \mathbb{1}_{B_2}(-1)\int_{B_1}(1-F_t)^{\beta+1}\,\mathrm{d}Q + \mathbb{1}_{B_2}(0)\int_{B_1}\beta(1-F_t)^\beta(1-G)\,\mathrm{d}P_t + \mathbb{1}_{B_2}(1)\int_{B_1}(1-F_t)^\beta(1-G)\,\mathrm{d}P_t$.*

Proof. Straightforward. □

5.2 Lemma. *Suppose f_t and f_0 are μ-densities of P_t and P_0 respectively, e.g. $\mu = P_t + P_0$. Without loss of generality it can be assumed that $f_t = 0$ on the set $\{F_t \in \{0,1\}\}$. By $L_{0,t} := f_t/f_0\mathbb{1}_{(0,\infty)}(f_0)$ and $N_{0,t} := \{x \in \mathbb{R} \mid f_0(x) = 0\}$ we define*

$$\widetilde{L}_{0,t}(x,\delta) := \mathbb{1}_{\{0,1\}}(\delta)\bigg(\frac{1-F_t(x)}{1-F_0(x)}\bigg)^\beta \cdot L_{0,t}(x) + \mathbb{1}_{\{-1\}}(\delta)\bigg(\frac{1-F_t(x)}{1-F_0(x)}\bigg)^{\beta+1} \cdot \mathbb{1}_{[0,1)}(F_0(x))$$

and $\widetilde{N}_{0,t} := N_{0,t} \times \{0,1\} \cup \{F_0 = 1\} \times \{-1\}$. The tuple $(\widetilde{L}_{0,t}, \widetilde{N}_{0,t})$ is a Lebesgue-decomposition of $\mathscr{R}_t^\beta$ with respect to $\mathscr{R}_0^\beta$.

Proof. For $B_1 \in \mathbb{B}$ and $B_2 \in \mathcal{P}\{-1,0,1\}$ we have $\mathscr{R}_t^\beta\big(B_1 \times B_2\big) = \mathscr{R}_t^\beta\big(B_1 \cap N_{0,t}^\complement \times B_2 \cap \{0,1\}\big) + \mathscr{R}_t^\beta\big(B_1 \cap \{F_0 < 1\} \times B_2 \cap \{-1\}\big) + \mathscr{R}_t^\beta\big(B_1 \cap N_{0,t} \times B_2 \cap \{0,1\}\big) + \mathscr{R}_t^\beta\big(B_1 \cap \{F_0 = 1\} \times B_2 \cap \{-1\}\big)$. Applying lemma 5.1 to the first two terms gives

$$\begin{aligned}
&\mathscr{R}_t^\beta\big(B_1 \cap N_{0,t}^\complement \times B_2 \cap \{0,1\}\big) + \mathscr{R}_t^\beta\big(B_1 \cap \{F_0 < 1\} \times B_2 \cap \{-1\}\big) \\
&\quad = \mathbb{1}_{B_2}(0)\int_{B_1 \cap N_{0,t}^\complement} \beta(1-F_t)^\beta(1-G)f_t\mathrm{d}\mu + \mathbb{1}_{B_2}(1)\int_{B_1 \cap N_{0,t}^\complement}(1-F_t)^\beta(1-G)f_t\mathrm{d}\mu \\
&\qquad + \mathbb{1}_{B_2}(-1)\int_{B_1 \cap \{F_0<1\}}(1-F_t)^{\beta+1}\mathrm{d}Q \\
&\quad = \mathbb{1}_{B_2}(0)\int_{B_1 \cap N_{0,t}^\complement}\bigg(\frac{1-F_t}{1-F_0}\bigg)^\beta \frac{f_t}{f_0}\cdot\mathbb{1}_{(0,\infty)}(f_0)\beta(1-F_0)^\beta(1-G)f_0\mathrm{d}\mu \\
&\qquad + \mathbb{1}_{B_2}(1)\int_{B_1 \cap N_{0,t}^\complement}\bigg(\frac{1-F_t}{1-F_0}\bigg)^\beta \frac{f_t}{f_0}\cdot\mathbb{1}_{(0,\infty)}(f_0)\beta(1-F_0)^\beta(1-G)f_0\mathrm{d}\mu \\
&\qquad + \mathbb{1}_{B_2}(-1)\int_{B_1 \cap \{F_0<1\}}\bigg(\frac{1-F_t}{1-F_0}\bigg)^{\beta+1}\cdot\mathbb{1}_{[0,1)}(F_0)\cdot(1-F_0)^{\beta+1}\mathrm{d}Q \\
&\quad = \int_{B_1 \times B_2}\widetilde{L}_{0,t}\,\mathrm{d}\mathscr{R}_0^\beta.
\end{aligned}$$

Thus, $\mathscr{R}_t^\beta(B_1 \times B_2) = \int_{B_1\times B_2} \widetilde{L}_{0,t}\,\mathrm{d}\mathscr{R}_0^\beta + \mathscr{R}_t^\beta\big((B_1 \times B_2) \cap \widetilde{N}_{0,t}\big)$. Since $\mathscr{R}_0^\beta(\widetilde{N}_{0,t}) = 0$ the proof is complete. ☐

5.3 Lemma. *Let $\mathfrak{F} = \{P_t \mid t \in (-\varepsilon_0, \varepsilon_0)\}$, $\varepsilon_0 > 0$, be a $\mathbb{L}_2(P_0)$-differentiable distribution family on $(\mathbb{R}, \mathbb{B})$ having the $\mathbb{L}_2$-derivative $\dot{L}$. It holds*

$$\nabla F_0 := \frac{\partial F_t}{\partial t}\Big|_{t=0} = \lim_{t\to 0} \frac{F_t(\cdot) - F_0(\cdot)}{t} = \int_{(-\infty,\cdot]} \dot{L}\,\mathrm{d}P_0 = -\int_{(\cdot,\infty)} \dot{L}\,\mathrm{d}P_0,$$

whereas the difference quotient converges uniformly.

Proof. Let $(L_{0,t}, N_{0,t})$ be a Lebesgue-decomposition of P_t with respect to P_0, $t \in (-\varepsilon_0, \varepsilon_0)$. Since $\mathbb{L}_2(P_0)$-differentiability implies $\mathbb{L}_1(P_0)$-differentiability we have

$$\begin{aligned}\lim_{t\to 0}\left|\int_{(-\infty,\cdot]} \frac{1}{t}\Big(L_{0,t} - 1\Big) - \dot{L}\,\mathrm{d}P_0\right| &\le \lim_{t\to 0}\int_{(-\infty,\cdot]} \left|\frac{1}{t}\Big(L_{0,t} - 1\Big) - \dot{L}\right| \mathrm{d}P_0\\ &\le \lim_{t\to 0}\int \left|\frac{1}{t}\Big(L_{0,t} - 1\Big) - \dot{L}\right| \mathrm{d}P_0 = 0\end{aligned}$$

and

$$\lim_{t\to 0}\left|\frac{1}{t}P_t\big(N_{0,t} \cap (-\infty,\cdot]\big)\right| \le \lim_{t\to 0}\frac{1}{|t|}P_t\big(N_{0,t}\big) = 0.$$

Given $\varepsilon > 0$ we can choose $t_0 > 0$ such that for all t, $0 < |t| \le t_0$, and for all $x \in \mathbb{R}$

$$\left|\int_{(-\infty,x]} \frac{1}{t}\Big(L_{0,t} - 1\Big) - \dot{L}\,\mathrm{d}P_0\right| \le \frac{\varepsilon}{2} \qquad \text{and} \qquad \left|\frac{1}{t}P_t\big(N_{0,t} \cap (-\infty,x]\big)\right| \le \frac{\varepsilon}{2}$$

hold. This means

$$\begin{aligned}\left|\frac{F_t(x) - F_0(x)}{t} - \int_{(-\infty,x]} \dot{L}\,\mathrm{d}P_0\right| &= \left|\int_{(-\infty,x]} \frac{1}{t}\big(L_{0,t} - 1\big) - \dot{L}\,\mathrm{d}P_0 + \frac{1}{t}P_t\big(N_{0,t} \cap (-\infty,x]\big)\right|\\ &\le \left|\int_{(-\infty,x]} \frac{1}{t}\big(L_{0,t} - 1\big) - \dot{L}\,\mathrm{d}P_0\right| + \left|\frac{1}{t}P_t\big(N_{0,t} \cap (-\infty,x]\big)\right| \le \varepsilon\end{aligned}$$

for all $x \in \mathbb{R}$ and for all t satisfying $0 < |t| \le t_0$. Since $\int \dot{L}\,\mathrm{d}P_0 = 0$, cf. e.g. Witting [9, Hilfssatz 1.178, p. 164], we have $\int_{(-\infty,x]} \dot{L}\,\mathrm{d}P_0 = -\int_{(x,\infty)} \dot{L}\,\mathrm{d}P_0$. ☐

The proof of theorem 3.4 depends crucially on the following result.

5.4 Lemma. *It holds $S_n^*(b,B) - \widehat{S}_n(b_n, B_n) \longrightarrow 0$ in $\mathscr{R}_{n,0}^\beta$-probability.*

Proof. The r.v.s $(X_{n,i}, \Delta_{n,i})$, $1 \le i \le n$, are i.i.d. under $\mathscr{R}_{n,0}^\beta$, so it holds

$$\mathfrak{L}_{\mathscr{R}_{n,0}^\beta}\big((R_{n,i}, U_{n,i}, \Delta_{n,i}), (X_{n,\uparrow}, \Delta_{n,\uparrow})\big) = \mathfrak{L}_{\mathscr{R}_{n,0}^\beta}\big((R_{n,j}, U_{n,j}, \Delta_{n,j}), (X_{n,\uparrow}, \Delta_{n,\uparrow})\big) \tag{24}$$

and

$$\begin{aligned}&\mathfrak{L}_{\mathscr{R}_{n,0}^\beta}\big((R_{n,i}, U_{n,i}, \Delta_{n,i}), (R_{n,j}, U_{n,j}, \Delta_{n,j}), (X_{n,\uparrow}, \Delta_{n,\uparrow})\big) =\\ &\qquad \mathfrak{L}_{\mathscr{R}_{n,0}^\beta}\big((R_{n,k}, U_{n,k}, \Delta_{n,k}), (R_{n,m}, U_{n,m}, \Delta_{n,m}), (X_{n,\uparrow}, \Delta_{n,\uparrow})\big)\end{aligned} \tag{25}$$

for $i \neq j$ and $k \neq m$. With the auxiliary statistic

$$S_n(b_n, B_n) = \sum_{i=1}^n c_{n,i}\left(\mathbb{1}_{\{0,1\}}(\Delta_{n,i})\big(b_{n,R_{n,i}} + \beta B_{n,R_{n,i}}\big) + \mathbb{1}_{\{-1\}}(\Delta_{n,i})(1+\beta)B_{n,R_{n,i}}\right)$$

we show $S_n(b_n, B_n) - S_n^*(b, B) \to 0$ in $\mathscr{R}_{n,0}^{\beta}$-probability in the first part of the proof. The Cauchy-Schwarz-inequality gives for two identically distributed r.v.s Y and Z

$$|\mathrm{E}(YZ)| \le \sqrt{\mathrm{E}(\mathrm{Y}^2) \cdot \mathrm{E}(\mathrm{Z}^2)} = \mathrm{E}(\mathrm{Y}^2). \tag{26}$$

Using the abbreviations $k_{n,i} = b_{n,R_{n,i}} - b \circ F_0 \circ H^{-1}(U_{n,i})$ and $K_{n,i} = B_{n,R_{n,i}} - B \circ F_0 \circ H^{-1}(U_{n,i})$ the inequality $(x+y)^2 \le 2(x^2+y^2)$ yields the estimate

$$\begin{aligned}
&\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(S_n(b_n, B_n) - S_n^*(b, B)\big)^2 \\
&\quad = \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(\sum_{i=1}^{n} c_{n,i}\mathbb{1}_{\{0,1\}}(\Delta_{n,i})(k_{n,i} + \beta K_{n,i}) + \sum_{i=1}^{n} c_{n,i}\mathbb{1}_{\{-1\}}(\Delta_{n,i})(\beta+1) \cdot K_{n,i}\right)^2 \\
&\quad \le 4\left(\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(\sum_{i=1}^{n} c_{n,i}k_{n,i}\right)^2 + \beta^2\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(\sum_{i=1}^{n} c_{n,i}K_{n,i}\right)^2 + (\beta+1)^2\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(\sum_{i=1}^{n} c_{n,i}K_{n,i}\right)^2\right).
\end{aligned}$$

Applying (2), (24), (25) and (26) one obtains

$$\begin{aligned}
\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(\sum_{i=1}^{n} c_{n,i}k_{n,i}\right)^2 &= \sum_{i=1}^{n} c_{n,i}^2 \cdot \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,i}\big)^2 + \sum_{i=1}^{n} c_{n,i} \cdot \sum_{\substack{j=1 \\ i\neq j}}^{n} c_{n,j} \cdot \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,i}k_{n,j}\big) \\
&\le \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,1}\big)^2 + \sum_{i=1}^{n} c_{n,i} \sum_{\substack{j=1 \\ i\neq j}}^{n} c_{n,j} \cdot \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,i}k_{n,j}\big) \\
&= \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,1}\big)^2 + \sum_{i=1}^{n} c_{n,i} \sum_{\substack{j=1 \\ i\neq j}}^{n} c_{n,j} \cdot \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,1}k_{n,2}\big) \\
&= \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,1}\big)^2 - \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,1}k_{n,2}\big) \\
&\le 2 \cdot \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,1}\big)^2.
\end{aligned}$$

By the same arguments one shows the estimate

$$\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(\sum_{i=1}^{n} c_{n,i}K_{n,i}\right)^2 \le 2 \cdot \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(K_{n,1}\big)^2. \tag{27}$$

Because of (5) $\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(S_n(b_n, B_n) - S_n^*(b, B)\right)^2 \to 0$ is implied by

$$\lim_{n\to\infty} \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(k_{n,1}\big)^2 = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\big(K_{n,1}\big)^2 = 0$$

which is shown in the next lines. Let $b_{n,i}^*$ be the exact scores, cf. Hájek and Šidák (1967) [5, p. 157 formula (12)], and $b_n^*(t) = b_{n,\lceil nt\rceil}^*$. Remember $b_{n,i} = b_{n,i}(X_{n,\uparrow}, \Delta_{n,\uparrow})$. Exploiting the fact that the rank vector and the order statistics are s.i. and (11) one obtains

$$\begin{aligned}
&E_{\mathscr{R}_{n,0}^{\beta}}(b_{n,R_{n,1}} - b_{n,R_{n,1}}^*)^2 \\
&\quad = \int \underbrace{\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left[\big(b_{n,R_{n,1}}(X_{n,\uparrow}, \Delta_{n,\uparrow}) - b_{n,R_{n,1}}^*\big)^2 \mid (X_{n,\uparrow}, \Delta_{n,\uparrow}) = (x,\delta)\right]}_{=\mathrm{E}_{\mathscr{R}_{n,0}^{\beta}}\left(b_{n,R_{n,1}}(x,\delta) - b_{n,R_{n,1}}^*\right)^2} d\mathscr{R}_{n,0}^{\beta\;(X_{n,\uparrow}, \Delta_{n,\uparrow})}(x,\delta) \\
&\quad = \mathrm{E}_{\mathscr{R}_{n,0}^{\beta}} \int_{(0,1)} (b_n - b_n^*)^2 \,\mathrm{d}\lambda.
\end{aligned}$$

Furthermore it holds

$$E_{\mathscr{R}^{\beta}_{n,0}}(b_{n,R_{n,1}} - b^*_{n,R_{n,1}})^2 \leq 2 \cdot \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}} \int_{(0,1)} (b_n - b \circ F_0 \circ H^{-1})^2 \,\mathrm{d}\lambda + 2 \cdot \int_{(0,1)} (b \circ F_0 \circ H^{-1} - b^*_n)^2 \,\mathrm{d}\lambda \to 0,$$

whereas the first term converges to 0 because of (12) and the second term converges because of Hájek and Šidak (1967), [5, Theorem V.1.5.b]. Using these results and theorem V.1.4.a in Hájek and Šidak (1967), [5] one obtains

$$\begin{aligned} \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(k_{n,1}\big)^2 &= \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(b_{n,R_{n,1}} - b \circ F_0 \circ H^{-1}(U_{n,1})\big)^2 \\ &\leq 2 \cdot \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(b_{n,R_{n,1}} - b^*_{n,R_{n,1}}\big)^2 + 2 \cdot \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(b^*_{n,R_{n,1}} - b \circ F_0 \circ H^{-1}(U_{n,1})\big)^2 \longrightarrow 0. \end{aligned} \tag{28}$$

By the same arguments it is proved that

$$\lim_{n\to\infty} \mathrm{E}_{\mathscr{R}^{\beta}_{n,0}}\big(K_{n,1}\big)^2 = 0. \tag{29}$$

Since $\mathbb{L}_2$-convergence implies convergence in probability we get $S_n(b_n, B_n) - S^*_n(b, B) \overset{\mathscr{R}^{\beta}_{n,0}}{\to} 0$. Because of $\widehat{S}_n(b_n, B_n) - S^*_n(b, B) = \widehat{S}_n(b_n, B_n) - S_n(b_n, B_n) + S_n(b_n, B_n) - S^*_n(b, B)$, it remains to show that $S_n(b_n, B_n) - \widehat{S}_n(b_n, B_n) \overset{\mathscr{R}^{\beta}_{n,0}}{\longrightarrow} 0$. For the r.v.s $T_n = \sum_{i=1}^n c_{n,i} \cdot B_{n,R_{n,i}}$ and $T^*_n = \sum_{i=1}^n c_{n,i} \cdot B(U_{n,i})$ the estimate (27) and the limit (29) imply $T_n - T^*_n \longrightarrow 0$ in $\mathscr{R}^{\beta}_{n,0}$-probability. Because of (2) and (5) the convergence theorem for linear statistics in Witting and Müller-Funk (1995) [10, Satz 5.112, p.112] can be applied. This gives $\mathfrak{L}_{\mathscr{R}^{\beta}_{n,0}}(T^*_n) \overset{\mathfrak{L}}{\longrightarrow} \mathscr{N}(0, \sigma^2)$, whereas $\sigma^2 = \mathrm{Var}_{\mathscr{R}^{\beta}_{n,0}}\big(B \circ F_0 \circ H^{-1}(U_{n,1})\big)$. Slutsky's lemma yields $\mathfrak{L}_{\mathscr{R}^{\beta}_{n,0}}(T_n) \overset{\mathfrak{L}}{\longrightarrow} \mathscr{N}(0, \sigma^2)$. By applying Slutsky's lemma and lemma 3.1 it results $S_n(b_n, B_n) - \widehat{S}_n(b_n, B_n) = (\beta - \widehat{\beta}_n) \sum_{i=1}^n c_{n,i} \cdot B_{n,R_{n,i}} \longrightarrow 0$ in $\mathscr{R}^{\beta}_{n,0}$-probability, that is to say the assertion. ◻

Acknowledgment

I want to thank Prof. Dr. Georg Neuhaus for his advice and encouragement.

References

[1] Per Kragh Andersen, Ørnulf Borgan, Richard D. Gill, and Niels Keiding. *Statistical models based on counting processes*. Springer, 1993.

[2] Konrad Behnen and Georg Neuhaus. *Rank tests with estimated scores and their applications*. B.G. Teubner, 1989.

[3] Ursula Gather and Jörg Pawlitschko. Estimating the survival function under a generalized Koziol-Green model with partially informative censoring. *Metrika*, 48:189–207, 1998.

[4] Jaroslav Hájek. Some extensions of the Wald-Wolfowitz-Noether theorem. *Annals of Mathematical Statistics*, 32:506–523, 1961.

[5] Jaroslav Hájek and Zbyněk Šidák. *Theory of Rank Tests*. Academic Press, Inc., 1967.

[6] Arnold Janssen. Local asymptotic normality for randomly censored data with applications to rank tests. *Statistica Neerlandica*, 43(2):109–125, 1989.

[7] Georg Neuhaus. Asymptotically optimal rank tests. *Communications in Statistics: Theory and Methods*, 17(6):2037–2058, 1988.

[8] Helmut Strasser. *Mathematical Theory of Statistics – Statistical Eperiments and Asymptotic Decision Theory*. Walter de Gruyter, 1985.

[9] Hermann Witting. *Mathematische Statistik I – Parametrische Verfahren bei festem Stichprobenumfang.* B.G. Teubner, 1985.

[10] Hermann Witting and Ulrich Müller-Funk. *Mathematische Statistik II – Asymptotische Statistik: Parametrische Modelle und nicht-parametrische Funktionale.* B.G. Teubner, 1995.