The Annals of Statistics
1993, Vol. 21, No. 4, 1760–1779

CONDITIONAL RANK TESTS FOR THE TWO-SAMPLE PROBLEM UNDER RANDOM CENSORSHIP

BY GEORG NEUHAUS

University of Hamburg

For the two-sample problem with randomly censored data, there exists a general asymptotic theory of rank statistics which are functionals of stochastic integrals with respect to certain empirical martingales. In the present paper a conditional counterpart of this theory is developed. The conditional martingales are versions of the original ones reduced to the unit interval having their jumps at fixed lattice points. The resulting conditional tests are strictly distribution free under the null hypothesis of randomness if the censoring distributions in both samples are equal and are asymptotically equivalent to their unconditional counterparts even if the censoring distributions are different. Simulations for linear rank statistics and Kolmogorov–Smirnov-type statistics show superiority of the conditional versions over their unconditional counterparts with respect to size and robustness under unequal censoring in both samples. At the same time the power of the conditional and unconditional tests is very similar in most cases.

1. Introduction. In the present paper we deal with the two-sample problem of testing the null hypothesis of randomness, that is, the equality of both sample distribution functions (d.f.'s) under random right censoring. This problem has received considerable interest in the past. A lot of different test procedures have been proposed and studied so far. Let us mention here only the log-rank test of Peto and Peto (1972), the generalized Wilcoxon test of Gehan (1965) and the tests of Harrington and Fleming (1982). These tests and many others belong to the class of generalized linear rank tests of Aalen (1978) and Gill (1980) which are functionals of certain empirical processes on $[0, \infty)$ being stochastic integrals with respect to basic martingales. By methods of continuous-time martingale theory, their asymptotic behavior has been derived by Gill (1980); see also Leurgans (1984) and the survey of Andersen, Borgan, Gill and Keiding (1982). Another class of functionals leads to Kolmogorov–Smirnov (KS) and Cramér–von Mises (CM) type tests which were reviewed and proposed by Schumacher (1984), who discussed, among others, the results of Koziol (1978), Koziol and Yuh (1982) and Gill (1980).

All the tests mentioned previously are usually applied using critical values of the asymptotic null distribution of standardized versions of the various test statistics. One major advantage of this approach is that these tests are

Received November 1990; revised December 1992.

AMS 1991 *subject classifications*. Primary 62G20; secondary 62G10.

Key words and phrases. Random censoring, finite sample validity under equal censoring, asymptotic validity under unequal censoring, permutation distribution, conditioning, discrete parameter martingales, linear rank tests, Kolmogorov–Smirnov tests, ties, simulation of size and power.

asymptotically valid even if the censoring distributions are different in both samples. On the other hand, the validity is an asymptotic one, and it has been reported in the literature that for finite sample numbers the nominal and actual level may differ substantially; see Latta (1981), Janssen and Brenner (1991) and Schumacher (1984).

In the case of equal censoring, one may construct strictly distribution-free tests under the null hypothesis by a conditioning device (permutation tests). In fact, the idea of permutation tests is an old one and has already been used in the framework of survival analysis under censoring by Gehan (1965) and Mantel (1967); see also Gill (1980), Section 3.3, and Andersen, Borgan, Gill and Keiding (1982), subsection 3.5. Gehan standardized his linear rank statistic by the exact permutation variance and proposed a normal approximation for large samples. Yet, under unequal censoring distributions this test is usually not distribution free, neither finite sample nor asymptotically. Neuhaus (1988) and Janssen (1991) made a systematic study of conditional rank statistics under random censoring in the framework of local asymptotic decision theory. Again, this theory is restricted to equal censoring and contiguous alternatives.

The aim of the present paper is to define and study conditional tests being strictly valid under equal censoring and being asymptotically valid under unequal censoring. The idea is simple enough: Take the unconditional test statistic T, say, standardized to make it asymptotically distribution free under the null hypothesis (even under unequal censoring). Then, keeping the observed *censoring pattern* fixed, perform a test based on the permutation distribution of T. It will turn out that the unconditional and the conditional tests will be asymptotically equivalent under very general circumstances including unequal censoring, while, by construction, the conditional tests are distribution free under the null hypothesis with equal censoring. Simulation results in Section 7 will show the pleasant behavior of the conditional tests with respect to size under equal *and* unequal censoring. At the same time the power behavior of the conditional and unconditional tests is very similar in most cases.

For the asymptotic treatment of the conditional tests, we will in a first step reduce the unconditional processes of Gill (1980) to the unit interval. Then, under the null hypothesis with equal censoring and conditionally under the observed censoring pattern, we develop an asymptotic theory paralleling the unconditional one of Gill (1980). In fact, since the reduced processes turn out to be discrete parameter martingales, the conditional theory becomes even simpler than the unconditional one. Andersen, Borgan, Gill and Keiding (1982) have already used the unreduced version of these martingales.

2. Unconditional rank tests. In this section we give a short review of unconditional tests which in Section 3 will be converted into conditional ones and state their asymptotic properties under the null hypothesis.

We assume the two-sample general censorship model. Observations are made on n_k individuals from population k, $k = 1, 2$, and all $n := n_1 + n_2$ observations are independent. ($a := b$ or $b =: a$ means that a equals b by definition.) The ith subject in sample k has nonnegative, independent latent survival and censoring times X_{ki}^0 and U_{ki}, respectively, whose distribution functions (d.f.'s) are F_k and G_k. For an easier presentation we will assume in this section that all d.f.'s are continuous, and give comments on the treatment of ties in Section 5. Let $S_k := (1 - F_k)$ and the cumulative hazard function $\Lambda_k(x) = \int_0^x (1 - F_k)^{-1}\, dF_k$. The observable random variables (r.v.'s) are $X_{ki} := \min(X_{ki}^0, U_{ki})$ and $\Delta_{ki} := 1\{X_{ki}^0 \le U_{ki}\}$, where $1\{E\}$ is 1 if the event E occurs and 0 otherwise. Let $Y_k(x)$ be the number of X_{ki}'s with $X_{ki} \ge x$ and $N_k(x)$ the number of uncensored X_{ki}'s with $X_{ki} \le x$. Put $Y := Y_1 + Y_2$ and $N := N_1 + N_2$. Always $0/0 := 0$.

Our hypothesis of interest is $\mathscr{H}_0$: $F_1 = F_2$ $(=: F)$ versus either the omnibus alternative $\mathscr{A}_0$: $F_1 \ne F_2$ or the one-sided alternative $\mathscr{A}_1$: $F_1 \ge F_2$, $F_1 \ne F_2$.

2.1. *Test statistics.* Many of the statistics which are usually applied for testing $\mathscr{H}_0$ depend on the process $\mathbb{W}_n$ given by

$$\mathbb{W}_n(x) = \int_0^x w_n \, d\mathbb{L}_n, \qquad x \ge 0, \tag{2.1}$$

with w_n a nonnegative stochastic weight function on $[0, \infty)$ and where the process $\mathbb{L}_n$, given by

$$\begin{aligned} \mathbb{L}_n(x) &= \left(\frac{n}{n_1 n_2}\right)^{1/2} \int_0^x \frac{Y_1 Y_2}{(Y_1 + Y_2)} \left\{\frac{dN_1}{Y_1} - \frac{dN_2}{Y_2}\right\} \\ &= \left(\frac{n}{n_1 n_2}\right)^{1/2} \left(N_1(x) - \int_0^x \frac{Y_1}{Y_1 + Y_2}\, dN\right), \qquad x \ge 0, \end{aligned} \tag{2.2}$$

will be called a *log-rank process* since $\mathbb{L}_n(\infty)$ is the well-known log-rank statistic. The preceding integrals are pathwise Stieltjes integrals. If the weight function w_n is left continuous and adapted to a natural filtration, then, under $\mathscr{H}_0$, $\mathbb{W}_n$ is a martingale and the powerful asymptotic methods of continuous-time martingale theory apply; see, for example, Aalen (1978), Gill (1980) and Andersen, Borgan, Gill and Keiding (1982).

An important class of weight functions is

$$w_n = \mathbb{S}_n^{\rho}((Y_1 + Y_2)/n)^{\kappa}(Y_1 Y_2/(n_1 n_2))^{-\lambda} 1\{Y_1 Y_2 > 0\}, \tag{2.3}$$

with $\rho, \kappa, \lambda \ge 0$ and $\mathbb{S}_n$ the left-continuous Kaplan–Meier estimate of the survivor function in the pooled sample. For $\kappa = \lambda = 0$ this is the class introduced by Harrington and Fleming (1982) containing the log-rank ($\rho = 0$) and Prentice's (1978) Wilcoxon ($\rho = 1$) cases. The case $\rho = \lambda = 0$ and $\kappa = 1$ corresponds to the Gehan–Wilcoxon test, while $\rho \ge 0.5$ and $\kappa = \lambda = 0.5$ lead to "approximately distribution free" statistics in Fleming, Harrington and O'Sullivan (1987). Finally, for $\rho = 0$, $\kappa = \lambda = 1$, $\mathbb{W}_n$ is the standardized Nelson estimator of $\Lambda_1 - \Lambda_2$ (stopped at $\hat{x} := \inf\{x: Y_1(x) \cdot Y_2(x) = 0\}$) and is called a *hazard process*. We estimate the variance of $\mathbb{W}_n(x)$ by

$$(2.4)\qquad \mathbb{V}_n(x) := \frac{n}{n_1 n_2}\int_0^x w_n^2 \frac{Y_1 Y_2}{(Y_1+Y_2)}\frac{d(N_1+N_2)}{Y_1+Y_2}, \qquad x \geq 0.$$

$\mathbb{V}_n$ is the estimator V_2 of Gill (1980), formula (3.3.12), specialized to the present continuous situation.

Many functionals of $(\mathbb{W}_n, \mathbb{V}_n)$ yield sensible test statistics for $\mathscr{H}_0$. For testing versus the one-sided alternative $\mathscr{A}_1$, the simplest functional is $\mathbb{W}_n(\infty)/\mathbb{V}_n^{1/2}(\infty)$, where $\mathbb{W}_n(\infty)$ is known as a *generalized linear rank statistic*. For testing versus the omnibus alternative $\mathscr{A}_0$, we will use as an example the Kolmogorov–Smirnov (KS) statistic

$$(2.5)\qquad \mathrm{KS}_n^{\vartheta} := \sup\{|(\mathbb{W}_n/(1+\mathbb{V}_n))(x)| : \mathbb{K}_n(x) \leq \vartheta\},$$

for some $\vartheta \in (0,1)$ and $\mathbb{K}_n := \mathbb{V}_n/(1+\mathbb{V}_n)$. For related test statistics, see Section 4.

2.2. *Unconditional tests.* The common way of using the preceding test statistics is to reject the null hypothesis if the observed value exceeds a critical value computed from the asymptotic null distribution which can be derived from the convergence properties of $\mathbb{W}_n$ and $\mathbb{V}_n$ in the Skorohod space $D[0,d]$ for suitable $d \in (0,\infty]$.

Though our main interest lies in conditional counterparts of the preceding unconditional tests, we need for the sake of comparison the asymptotic null distribution of the unconditional tests. Therefore, let us state some simple sufficient conditions ensuring convergence in distribution ($\to_{\mathscr{D}}$) of $\mathbb{W}_n$ and convergence in probability ($\to_P$) of $\mathbb{V}_n$. These follow immediately from the appendix of Fleming, Harrington and O'Sullivan (1987), which in turn specialize the results of Gill (1980). Let n tend to ∞ such that

$$(2.6)\qquad n_1/n \to \eta \quad \text{for some } \eta \in (0,1),$$

and $w_n \to_P w$ uniformly on each interval $[0,v]$ with $0 < v < \tau$, $\tau := \sup\{x: F(x) < 1, G_k(x) < 1, k = 1,2\}$, where w is some continuous function on $[0,\tau)$. Then, under $\mathscr{H}_0$, one has $\mathbb{W}_n \to_{\mathscr{D}} \mathbb{B}\circ V$ on $D[0,v]$ and $\mathbb{V}_n \to_P V$ uniformly on $[0,v]$, where $\mathbb{B}$ is Brownian motion and V the asymptotic variance function $V(x) = \int_0^x w^2(1-G_1)(1-G_2)/G\,dF$, with $G := \eta(1-G_1) + (1-\eta)(1-G_2)$. Consequently, if $V(v) > 0$, then

$$(2.7)\qquad \mathbb{W}_n(v)/\mathbb{V}_n^{1/2}(v) \to_{\mathscr{D}} \mathscr{N}(0,1),$$

where $\mathscr{N}(0,1)$ is the standard normal distribution and

$$(2.8)\qquad \mathrm{KS}_n^{\vartheta} \to_{\mathscr{D}} \sup_{0\leq t\leq \vartheta} |\mathbb{B}^0(t)|,$$

with $0 < \vartheta \leq K(v) := V(v)/(1+V(v))$ and $\mathbb{B}^0$ is the Brownian bridge. For (2.8) we have used the equality in distribution of the processes $\mathbb{B}$ and $((1+t)\mathbb{B}^0(t/(1+t)): t \geq 0)$ combined with a little extra argument for replacing the random set $\{x: \mathbb{K}_n(x) \leq \vartheta\}$ by the nonrandom set $\{x: K(x) \leq \vartheta\}$. If, additionally, the w_n's are uniformly bounded $\forall\ n$, then v may be replaced by $v = \infty$.

These results, which apply to the weights in (2.3), show that the unconditional tests based on the asymptotic distributions in (2.7), respectively (2.8), are asymptotically distribution free under $\mathscr{H}_0$ even if the censoring distributions G_1 and G_2 are different.

3. Conditional rank tests and their asymptotic behavior under the null hypothesis. In this section which contains the main results of the paper, we will convert the unconditional tests of the preceding section into conditional ones and derive their conditional asymptotics by considerations paralleling the unconditional case. In fact, the situation will become even simpler since the reduced processes given in this section can only jump at i/n, $1 \le i < n$. Let $X_{1:n} \le \cdots \le X_{n:n}$ be the order statistics of the pooled observations X_{kj}, put $X_{0:n} = 0$, and let $\mathbf{\Delta}_n = (\Delta_{1:n}, \dots, \Delta_{n:n})$ be the *censoring status vector* of the corresponding Δ's. Put $Z_i := 1$ (0) if the r.v. $X_{i:n}$ belongs to the first (second) sample and call $\mathbf{Z}_n = (Z_1, \dots, Z_n)$ the *sample status vector*.

Finally, put $p_i := (Y_1/Y)(X_{i:n})$ and $q_i := 1 - p_i$ for $0 \le i \le n$. We reduce all processes considered so far to the unit interval by transforming $X_{i:n}$ to i/n and let the reduced processes be constant in $[(i-1)/n, i/n)$. The reduced processes with paths in $D[0,1]$ are labeled by an overbar, that is, $\overline{\mathbb{L}}_n$, $\overline{\mathbb{W}}_n$, $\overline{w}_n$, $\overline{\mathbb{V}}_n$ and so on. We get

$$\overline{\mathbb{W}}_n(t) = \left(\frac{n}{n_1 n_2}\right)^{1/2} \sum_{i=1}^{[nt]} \overline{w}_n\left(\frac{i}{n}\right) \Delta_{i:n}(Z_i - p_i) \tag{3.1}$$

and

$$\overline{\mathbb{V}}_n(t) = \frac{n}{n_1 n_2} \sum_{i=1}^{[nt]} \overline{w}_n^2\left(\frac{i}{n}\right) \Delta_{i:n} p_i q_i, \qquad 0 \le t \le 1, \tag{3.2}$$

where $[x]$ denotes the integer part of x. Note that in (3.1) and trivially in (3.2) terms with $i/n > \overline{\tau}_n := \max\{j/n\colon\ p_j q_j > 0\} < 1$ vanish, so that these processes stay constant for $t \ge \overline{\tau}_n$. The predictability of w_n implies that $\overline{w}_n(i/n)$ depends solely on $\Delta_{1:n}, \dots, \Delta_{i-1:n}$ and $Z_1, \dots, Z_{i-1}$. Note that p_i and q_i depend likewise only on $Z_1, \dots, Z_{i-1}$. For example, the weight functions w_n from (2.3) reduce to

$$\overline{w}_n\left(\frac{i}{n}\right) = \overline{\mathbb{S}}_n^{\rho}\left(\frac{i}{n}\right)\left(\frac{n-i+1}{n}\right)^{\kappa - 2\lambda} \left(p_i q_i n^2/(n_1 n_2)\right)^{-\lambda} 1\{i \le \overline{\tau}_n\}, \tag{3.3}$$

with

$$\overline{\mathbb{S}}_n(i/n) = \prod_{1 \le j \le i-1} \left(1 - \Delta_{j:n}/(n-j+1)\right), \qquad \overline{\mathbb{S}}_n(1/n) := 1.$$

If $\mathbb{H}_n$ denotes the empirical d.f. of the $X_{i:n}$'s, then $\mathbb{L}_n = \overline{\mathbb{L}}_n \circ \mathbb{H}_n$, $\mathbb{W}_n = \overline{\mathbb{W}}_n \circ \mathbb{H}_n$

and so forth. We can rewrite the test statistics from (2.2) in reduced form as

$$\mathbb{W}_n(\infty)/\mathbb{V}_n^{1/2}(\infty) = \overline{\mathbb{W}}_n(1)/\overline{\mathbb{V}}_n^{1/2}(1) \tag{3.4}$$

and

$$\mathrm{KS}_n^{\vartheta} = \sup\left\{\left|\left(\overline{\mathbb{W}}_n/(1 + \overline{\mathbb{V}}_n)\right)\left(\frac{i}{n}\right)\right| : \overline{\mathbb{K}}_n\left(\frac{i}{n}\right) \le \vartheta\right\}. \tag{3.5}$$

In this way the previous formal dependence on the $X_{i:n}$'s has been removed. Hence these statistics depend solely on the sample status vector $\mathbf{Z}_n = (Z_1, \dots, Z_n)$ and the censoring status vector $\mathbf{\Delta}_n = (\Delta_{1:n}, \dots, \Delta_{n:n})$. The idea of conditioning relies on this fact and the following easy lemma.

LEMMA 3.1. *Under the hypothesis $\mathscr{H}_0$ with the additional assumption $G_1 = G_2$ (equal censoring) which will be called the restricted null hypothesis $\overline{\mathscr{H}}_0$, the vectors $\mathbf{\Delta}_n = (\Delta_{1:n}, \dots, \Delta_{n:n})$ and $\mathbf{Z}_n = (Z_1, \dots, Z_n)$ are independent and the r.v. $(Z_1, \dots, Z_n)$ is distributed as a random sample without replacement taken from a population consisting of n_1 members "1" and n_2 members "0."*

An arbitrary random vector $\mathbf{Z}_n \in \{0, 1\}^n$ is said to have *permutation distribution* if the latter property holds true. Let $T_n(\mathbf{Z}_n, \mathbf{\Delta}_n)$ be one of the test statistics (3.4) or (3.5) and $k(\alpha)$ the $(1 - \alpha)$-quantile of their asymptotic null distribution. The conditional counterpart of the corresponding tests is obtained by simply replacing $k(\alpha)$ by the $(1 - \alpha)$-quantile $k_n(\alpha, \delta_n)$ of the distribution of $T_n(\mathbf{Z}_n^*, \delta_n)$, where $\mathbf{\Delta}_n = \delta_n$ is the observed censoring status and $\mathbf{Z}_n^*$ is some r.v. having permutation distribution. According to Lemma 3.1, one has $\mathscr{L}(\mathbf{Z}_n) = \mathscr{L}(\mathbf{Z}_n^*)$ under $\overline{\mathscr{H}}_0$ (equal censoring). Thus the independence of $\mathbf{\Delta}_n$ and $\mathbf{Z}_n$ makes the conditional test finite sample distribution free under $\overline{\mathscr{H}}_0$.

However, the question arises whether the conditional version has any connection with the unconditional one or whether we have created just another test. The pleasing result is that the conditional and the unconditional versions are asymptotically equivalent even under the general null hypothesis $\mathscr{H}_0$ with possibly unequal censoring, $G_1 \neq G_2$, for which $\mathbf{\Delta}_n$ and $\mathbf{Z}_n$ are usually dependent and $\mathbf{Z}_n$ need not have permutation distribution. In fact, we will show $k_n(\alpha, \mathbf{\Delta}_n) \to_P k(\alpha)$ under $\mathscr{H}_0$ and under some mild extra conditions, entailing the asserted equivalence under $\mathscr{H}_0$ and under contiguous alternatives.

The key for proving these results is the observation that for fixed $\mathbf{\Delta}_n = \delta_n$ the process $\overline{\mathbb{W}}_n$ is a martingale if $\mathbf{Z}_n$ has permutation distribution. More exactly, let $\delta_n = (\delta_{1:n}, \dots, \delta_{n:n})$ be a fixed element of $\{0, 1\}^n$ and $\mathbf{Z}_n = (Z_1, \dots, Z_n)$ be any random vector on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, having permutation distribution. Then, putting $\mathbf{\Delta}_n = \delta_n$, the reduced process $\overline{\mathbb{W}}_n$ is a martingale on $[0, 1]$ with respect to the filtration $(\mathscr{F}_{[nt]}: 0 \le t \le 1)$, where $\mathscr{F}_i$ is the σ-algebra generated by $(Z_1, \dots, Z_i)$, $1 \le i \le n$, and $\mathscr{F}_0 := \{\varnothing, \Omega\}$. This follows immediately from the fact that, conditionally on $\mathscr{F}_{i-1}$, Z_i has binomial$(1, p_i)$ distribution with $p_i = (Z_i + \cdots + Z_n)/(n - i + 1)$, and that $\overline{w}_n(i/n)$ depends solely on the fixed δ_n and $Z_1, \dots, Z_{i-1}$, $1 \le i \le n$. Apparently, $\overline{\mathbb{W}}_n$ is the continuous-time version of the discrete-time martingale $(\overline{\mathbb{W}}_n(i/n), \mathscr{F}_i)$, $1 \le i \le n$.

Moreover, the conditional binomial nature of the Z_i's implies that for the reduced variance estimator $\overline{\mathbb{V}}_n$, see (3.2), one has

$$\overline{\mathbb{V}}_n(t) = \langle \overline{\mathbb{W}}_n, \overline{\mathbb{W}}_n \rangle(t) \quad \text{for } 0 \le t \le 1, \tag{3.6}$$

where $\langle \cdot, \cdot \rangle$ is the "predictable quadratic variation" of the underlying martingale; see, for example, Jacod and Shiryaev (1987), page 38. Equation (3.6) will ensure that $\overline{\mathbb{V}}_n$ is a consistent estimator of the limiting covariance function of $\overline{\mathbb{W}}_n$ under the permutation distribution for sequences δ_n, $n \ge 1$, specified in the following discussion, which is essential for showing that the unconditional and the conditional limiting null distribution of the statistics (3.4), respectively (3.5), coincide.

Now a well-known martingale CLT says that pointwise convergence of $\langle \overline{\mathbb{W}}_n, \overline{\mathbb{W}}_n \rangle$ with $[0, 1)$ to a continuous limit function $\overline{V}$, say, combined with a conditional Lindeberg condition, see (6.2), implies $\overline{\mathbb{W}}_n \to_{\mathscr{D}} \mathbb{B} \circ \overline{V}$ in the Skorohod space $D[0, 1)$; see Jacod and Shiryaev (1987), Theorem 8.3.33, or Gill (1980), Theorem 2.4.1. From this CLT we get a criterion for convergence in distribution of the conditional process $\overline{\mathbb{W}}_n$ in the space $D[0, 1)$, respectively $D[0, 1]$, which is in fact the main theoretical result of our paper. As a consequence we get in Theorem 3.3 asymptotic equivalence of our conditional and unconditional tests. All proofs will be given in Section 6.

THEOREM 3.2. *Assume $n_1/n \to \eta \in (0, 1)$ and let $\mathbf{Z}_n$, $n \ge 1$, have permutation distribution. Convergence in probability and in distribution in what follows refer to this assumption. Moreover, let $\delta_n = (\delta_{1:n}, \dots, \delta_{n:n}) \in \{0, 1\}^n$, $n \ge 1$, be a fixed sequence such that the functions $\overline{\mathbb{H}}_n^1(t) := \int_0^t \delta_{[ns]:n}\, ds$, $\delta_{0:n} := 0$, fulfill*

$$\overline{\mathbb{H}}_n^1(t) \to \overline{H}^1(t) \quad \textit{for } 0 \le t \le 1, \tag{3.7}$$

for some sub-d.f. $\overline{H}^1$ on $[0, 1]$. Finally, let $\overline{w}_{ni}$ be real numbers depending on δ_n and $Z_1, \dots, Z_{i-1}$, $1 \le i \le n$, such that the jump functions $\overline{w}_n(t) := \overline{w}_{n[nt]}$, $\overline{w}_{n0} := 0$, converge in quadratic mean in probability, that is,

$$\int_0^t (\overline{w}_n - \overline{w})^2(s)\, ds \to_P 0 \quad \textit{for } 0 < t < 1, \tag{3.8}$$

where $\overline{w}$ is a nonrandom function on $[0, 1]$ being square integrable on each subinterval $[0, t]$, $0 < t < 1$. Then

$$\overline{\mathbb{W}}_n \to_{\mathscr{D}} \mathbb{B} \circ \overline{V} \quad \textit{in } D[0, 1) \tag{3.9}$$

and

$$\overline{\mathbb{V}}_n(t) \to_P \overline{V}(t) \quad \textit{for } t \in [0, 1), \tag{3.10}$$

with

$$\overline{V}(t) = \int_{[0,t]} \overline{w}^2 \, d\overline{H}^1. \tag{3.11}$$

If, additionally, $\overline{V}(1) < \infty$ *and*

$$\lim_{t \uparrow 1} \limsup_{n \to \infty} \mathbb{P}\{\overline{V}_n(1) - \overline{V}_n(t) \geq \varepsilon\} = 0 \qquad \forall\, \varepsilon > 0, \tag{3.12}$$

then (3.9) *and* (3.10) *hold true in the Skorohod space* $D[0,1]$, *respectively on the interval* $[0,1]$.

Theorem 3.2 is easily applied to our situation: Using Glivenko–Cantelli theorems for the empirical d.f. of the $X_{i:n}$'s and the uncensored $X_{i:n}$'s, it will be shown in Section 6 that, under $\mathscr{H}_0$, there is a unique sub-d.f. $\overline{H}^1$, namely $\overline{H}^1 := H^1 \circ H^{-1}$ with $H^1 := \int_0^{\cdot}(1-G)\,dF$, $G := \eta G_1 + (1-\eta)G_2$, $H := 1 - (1-F)(1-G)$ and $H^{-1}(t) := \inf\{x\colon H(x) \geq t\}$, $0 \leq t \leq 1$, such that

$$\overline{\mathbb{H}}_n^1(t) = \int_0^t \Delta_{[ns]:n}\, ds \to_P \overline{H}^1(t) \quad \text{under } \mathscr{H}_0, \qquad 0 \leq t \leq 1, \tag{3.13}$$

$\Delta_{0:n} := 0$. By switching in (3.13) to subsequences of realizations fulfilling (3.7), Theorem 3.2 applies and yields the following theorem on convergence in probability of the conditional quantiles. Let us remark that under our assumptions in (3.13) one has even almost-sure convergence. Since in more general situations, for example, varying censoring d.f.'s in each sample, Glivenko–Cantelli theorems hold only in probability, we chose the preceding weaker formulation.

THEOREM 3.3. *Assume* $n_1/n \to \eta \in (0,1)$, (3.8), (3.12) *and* $0 < \overline{V}(1) < \infty$. *Let* $T_n(\mathbf{Z}_n, \boldsymbol{\Delta}_n)$ *be one of the test statistics* (3.4) *or* (3.5) *and* $k(\alpha)$ *the* $(1-\alpha)$-*quantile of their asymptotic null distribution* $\mathscr{N}(0,1)$, *respectively* $\mathscr{L}(\sup\{|\mathbb{B}^0(t)|\colon\ 0 \leq t \leq \vartheta\},\ 0 < \vartheta \leq \overline{K}(1) = \overline{V}(1)/(1+\overline{V}(1)))$. *Moreover, let* $k_n(\alpha, \delta_n)$ *be the* $(1-\alpha)$-*quantile of the distribution of* $T_n(\mathbf{Z}_n^*, \delta_n)$, *where* $\Delta_n = \delta_n$ *is the observed censoring status and* $\mathbf{Z}_n^*$ *is some random vector having permutation distribution. Then* $(0 < \alpha < 1)$

$$k_n(\alpha, \Delta_n) \to_P k(\alpha) \quad \textit{under } \mathscr{H}_0. \tag{3.14}$$

If (3.12) *does not hold or if* $\overline{V}(1) = \infty$, *then* (3.14) *remains true for the* KS *statistics* (3.5) *if* $\vartheta < \overline{K}(1)$.

Since the asymptotic null d.f.'s *are continuous and strictly increasing*, (3.14) *implies the asymptotic equivalence of the conditional test with their unconditional counterparts under* H_0.

Let us recall that asymptotic equivalence of the unconditional (unc) and the conditional (con) tests, that is, $E|\phi_{n,\text{unc}} - \phi_{n,\text{con}}| \to 0$ under $\mathscr{H}_0$, implies their asymptotic equivalence also under contiguous alternatives. Local shift and scale alternatives as described, for example, in Neuhaus (1988) and Janssen

(1991), are typical cases where contiguity holds true, even under unequal censoring.

Theorem 3.3 applies to the class of weights $\overline{w}_n$ from (3.3).

EXAMPLE 3.4. Under $n_1/n \to \eta \in (0,1)$ and the permutation distribution, condition (3.8) is shown in Section 6 to be fulfilled for weights $\overline{w}_n$ from (3.3) with arbitrary $\rho, \kappa, \lambda \geq 0$ and limiting function

$$(3.15) \qquad \overline{w}(s) = \overline{S}^{\rho}(s)(1-s)^{\kappa-2\lambda}, \qquad 0 \leq s < 1,$$

where $\overline{S} := S \circ H^{-1}$. The limiting variance function $\overline{V}(t)$ becomes, for $0 \leq t < 1$,

$$(3.16) \quad \overline{V}(t) = \int_0^t \overline{S}^{2\rho}(1-\mathrm{id})^{2\kappa-4\lambda}\, d\overline{H}^1 = \int_0^t \overline{S}^{2\rho-1}(1-\mathrm{id})^{2\kappa-4\lambda+1}\, d\overline{F},$$

where id := identity, $\overline{F} := 1 - \overline{S}$. The second equality in (3.16) follows from $(1-\mathrm{id}) = (1-\overline{F})(1-\overline{G})$ with $\overline{G} := G \circ H^{-1}$ and $d\overline{H}^1 = (1-\overline{G})\,d\overline{F}$. Since $\overline{S} \leq 1$, one has $\overline{V}(1) < \infty$ for all special choices of ρ, κ and λ mentioned after (2.3) except for the case $1 - \rho = \kappa = \lambda = 1$ leading to the hazard process. If $\lambda \leq 0.5$ and either $k - 2\lambda \geq 0$ or $2\rho \geq 1$ and $2\kappa - 4\lambda + 1 \geq 0$, then condition (3.12) is fulfilled. If the probability of no censoring, $\overline{H}^1(1) = (1 - \int G\,dF)$, is strictly positive, then $\overline{V}(1)$ is strictly positive, too.

Note that in the case $\lambda = \kappa = 0.5$, $\rho \geq 0.5$, and $\overline{F}(1) = 1$, one gets $\overline{V}(1) = 1/(2\rho)$, not depending on the censoring distributions. Related "approximately distribution free statistics" have been studied by Leurgans (1984) and Fleming, Harrington and O'Sullivan (1987).

4. Related omnibus tests. In contrast to our KS statistics which are based on Brownian bridge versions of the limiting processes, Gill (1980) as well as Fleming, Harrington and O'Sullivan (1987) chose Brownian motion versions as their starting point. This leads to a parallel class of so-called Rényi-type statistics with limiting distribution based on suprema of the Brownian motion instead of the Brownian bridge. All these statistics have conditional counterparts with analogous properties.

In the "omnibus" part of our simulations, we will concentrate on Brownian bridge versions based on the hazard process and the Kaplan–Meier process. An extensive comparative simulation study of unconditional tests and their conditional counterparts is beyond the scope of this paper.

4.1. *Statistics based on the Kaplan–Meier process.* So far, we have considered only statistics based on the log-rank process $\mathbb{L}_n$ which was motivated by the conditional martingale property of $\overline{\mathbb{L}}_n$. Likewise, one may introduce conditional KS tests based on the Kaplan–Meier process $\mathbb{X}_n = \overline{\mathbb{X}}_n \circ \mathbb{H}_n$, where $\overline{\mathbb{X}}_n$ is given by

$$(4.1) \quad \overline{\mathbb{X}}_n(t) = c\left(\prod_{i=1}^{[nt]}\left(1 - \Delta_{i:n}(1-Z_i)p_i^{-1}\right) - \prod_{i=1}^{[nt]}\left(1 - \Delta_{i:n}Z_i q_i^{-1}\right)\right),$$

with $c := (n_1 n_2/n)^{1/2}$. $\mathbb{X}_n$ is proportional to the difference of the Kaplan–Meier estimators of both samples; see, for example, Shorack and Wellner (1986), page 293. Similarly as in Section 2 one has convergence in distribution $\overline{\mathbb{X}}_n \to_{\mathscr{D}} S(1+V) \cdot \mathbb{W}^0 \circ K$ under $\mathscr{H}_0$ and, with $\overline{\mathbb{S}}_n$ from (3.3), one obtains in analogy to (2.8) that, under $\mathscr{H}_0$, for $0 < \vartheta < K(\tau)$,

$$\text{(4.2)}\quad \mathrm{KS}K_n^{\vartheta} := \sup_{i:\,\overline{\mathbb{K}}_n(i/n) \le \vartheta} \left| \overline{\mathbb{X}}_n / \left(\overline{\mathbb{S}}_n(1+\overline{\mathbb{V}}_n)\right)\left(\frac{i}{n}\right)\right| \to_{\mathscr{D}} \sup_{0 \le t \le \vartheta} |\mathbb{B}^0(t)|.$$

Now one may proceed as before and define unconditional and conditional tests being again asymptotically equivalent under $\mathscr{H}_0$. The proof relates $\overline{\mathbb{X}}_n$ to $\overline{\mathbb{S}}_n \cdot \overline{\mathbb{W}}_n$ by suitable Taylor expansions, $\mathbb{W}_n$ being here the hazard process. Details may be found in Neuhaus (1991a).

4.2. *Cramér–von Mises statistics.* Analogous considerations as for the KS statistics can be made for the CM statistics

$$\text{(4.3)}\qquad \mathrm{CM}K_n^{\vartheta} := \int \left(\overline{\mathbb{X}}_n^2 \overline{S}_n^{-2}\left(1+\overline{\mathbb{V}}_n\right)^{-2}\right) 1\left(\overline{\mathbb{K}}_n \le \vartheta\right) d\overline{\mathbb{K}}_n$$

and

$$\text{(4.4)}\qquad \mathrm{CM}_n^{\vartheta} := \int \left(\overline{\mathbb{W}}_n^2\left(1+\overline{\mathbb{V}}_n\right)^{-2}\right) 1\left(\overline{\mathbb{K}}_n \le \vartheta\right) d\overline{\mathbb{K}}_n.$$

The conditional and unconditional limiting distribution of these test statistics is that of $\int_0^{\vartheta} (\mathbb{B}^0(t))^2\, dt$.

5. Treatment of ties. In the sequel we describe briefly how the conditioning device has to be changed if the underlying d.f.'s are not continuous so that ties may occur. For details see Neuhaus (1991a, 1991b, 1992).

Put $(X_1, \dots, X_n) := (X_{11}, \dots, X_{1n_1}, X_{21}, \dots, X_{2n_2})$ and let r denote the number of different X_i's. In the continuous case $r = n$ but in general r is random. Define integers $T_0 := 0 < T_1 < \cdots < T_r = n$ and $\mathbf{T}_n = \{T_0, T_1, \dots, T_r\}$ by

$$X_{1:n} = \cdots = X_{T_1:n} < X_{T_1+1:n} = \cdots = X_{T_2:n} < \cdots < X_{T_{r-1}+1:n} = \cdots = X_{T_r:n}.$$

The observations X_i with $i \in I_j := \{k\colon X_k = X_{q:n}\}$, $q := T_j$, form the jth *tie group*. In the jth tie group the uncensored observations are ranked ahead of the censored ones. For $1 \le j \le r$ let D_j be the number of uncensored observations in the jth tie group, let A_j (B_j) be the number of uncensored (censored) observations from sample 1 in the jth tie group, and $p_j := (A_j + B_j + \cdots + A_r + B_r) \times (n - T_{j-1})^{-1}$, $q_j := 1 - p_j$. By transforming $X_{q:n}$,

$q := T_j$, to q/n and putting $\overline{w}_n(q/n) := w_n(X_{q:\,n})$, the processes $\overline{\mathbb{W}}_n$, respectively $\overline{\mathbb{V}}_n$, see (3.1), respectively (3.2), extend to

$$\overline{\mathbb{W}}_n(t) = \left(\frac{n}{n_1 n_2}\right)^{1/2} \sum_{j=1}^{l} \overline{w}_n\left(\frac{T_j}{n}\right)(A_j - D_j p_j) \tag{5.1}$$

and

$$\overline{\mathbb{V}}_n(t) = \frac{n}{n_1 n_2} \sum_{j=1}^{l} \overline{w}_n^2\left(\frac{T_j}{n}\right) p_j q_j D_j \frac{n - T_{j-1} - D_j}{n - T_{j-1} - 1}, \tag{5.2}$$

for $T_l \le n \cdot t < T_{l+1}$, $l = 0, \dots, r$, $T_{r+1} := \infty$. In fact, if ties occur, then a factor $(1 - (\Delta N - 1)/(Y - 1))$ with $\Delta N(x) := N(x) - N(x-)$ has to be added under the integral sign in the definition (2.4) of $\mathbb{V}_n$ in order to get Gill's estimator V_2 explaining the last factor in (5.2). The corresponding standardized linear rank statistic is $\overline{\mathbb{W}}_n(1)/\overline{\mathbb{V}}_n^{1/2}(1)$ and the KS statistic $\mathrm{KS}_n^{\vartheta}$ from (2.5) can be rewritten as

$$\mathrm{KS}_n^{\vartheta} = \sup\left\{\left|\left(\overline{\mathbb{W}}_n/(1 + \overline{\mathbb{V}}_n\right)(T_j/n)\right| : \overline{\mathbb{K}}_n(T_j/n) \le \vartheta\right\}. \tag{5.3}$$

Let us describe the generalized conditioning only for KS statistics. For other statistics analog considerations may be made. Apparently, $\mathrm{KS}_n^{\vartheta} = \mathrm{KS}_n^{\vartheta}(\mathbf{A}_n, \mathbf{B}_n, \mathbf{T}_n, \mathbf{D}_n)$ with $\mathbf{A}_n := (A_1, \dots, A_r)$ and similarly for $\mathbf{B}_n$, $\mathbf{T}_n$ and $\mathbf{D}_n$. It may be shown by simple symmetry considerations that under $\overline{\mathscr{H}}_0$ and given $(\mathbf{T}_n, \mathbf{D}_n) = (\mathbf{t}_n, \mathbf{d}_n)$, the vector $(\mathbf{A}_n, \mathbf{B}_n)$ has permutation distribution in the following sense. If t_j (d_j) are the components of $\mathbf{t}_n$ ($\mathbf{d}_n$) and $e_j := t_j - t_{j-1} - d_j$ $\forall j$, $t_0 := 0$, then $A_1, \dots, A_r, B_1, \dots, B_r$ are the number of red balls if $d_1, \dots, d_r, e_1, \dots, e_r$ balls are successively drawn at random from an urn with n_1 red and n_2 black balls. Now, having observed $(\mathbf{T}_n, \mathbf{D}_n) = (\mathbf{t}_n, \mathbf{d}_n)$, the conditional test rejects the null hypothesis if the observed $\mathrm{KS}_n^{\vartheta}$ exceeds the $(1 - \alpha)$-quantile, $k_n(\alpha, \mathbf{t}_n, \mathbf{d}_n)$ say, of the distribution of $\mathrm{KS}_n^{\vartheta}(\mathbf{A}_n^*, \mathbf{B}_n^*, \mathbf{t}_n, \mathbf{d}_n)$, where $(\mathbf{A}_n^*, \mathbf{B}_n^*)$ is some r.v. having permutation distribution in the preceding sense. By construction, the conditional test is distribution free under the restricted null hypothesis $\overline{\mathscr{H}}_0$. As in the continuous case, given $(\mathbf{T}_n, \mathbf{D}_n) = (\mathbf{t}_n, \mathbf{d}_n)$ and assuming that $(\mathbf{A}_n, \mathbf{B}_n)$ has permutation distribution, $\overline{\mathbb{W}}_n$ is a martingale with predictable quadratic variation $\overline{\mathbb{V}}_n$ for a suitable filtration. One may derive limiting results also in the present general case by reducing it to the continuous one. For the sake of brevity, we omit explicit formulations; see Neuhaus (1991b, 1992). Let us only remark that for discontinuous failure distributions F the limiting distribution of the unconditional as well as of the conditional KS statistic is of the form $\mathscr{L}(\sup\{|\mathbb{W}^0(t)|,\ 0 \le t \le \vartheta,\ t \in A\})$, where $A \subseteq [0, 1]$ may depend on F and the censoring d.f.'s G_1 and G_2. Consequently, the unconditional tests using the critical values of the distribution $\mathscr{L}(\sup|\mathbb{W}^0(t)|\colon 0 \le t \le \vartheta)$ is asymptotically not distribution free if F is discontinuous, not even under the restricted null hypothesis $\overline{\mathscr{H}}_0$, rather it is conservative; see also the related discussion in Fleming, Harrington and

O'Sullivan (1987). In contrast, the conditional tests *are* finite sample distribution free under $\mathscr{H}_0$.

REMARK 5.1. As mentioned previously, $\overline{\mathbb{V}}_n$ is the predictable quadratic variation of the conditional martingale $\overline{\mathbb{W}}_n$ under the permutation distribution. Following Gill (1980), (4.1.20), one may propose another estimator $\mathbb{V}_{1n} = \overline{\mathbb{V}}_{1n} \circ \mathbb{H}_n$ which is his estimator V_1. The reduced version $\overline{\mathbb{V}}_{1n}$ of $\mathbb{V}_{1n}$ is in the continuous case

$$\overline{\mathbb{V}}_{1n}(t) = \frac{n}{n_1 n_2} \sum_{i=1}^{[nt]} \overline{w}_n^2\left(\frac{i}{n}\right) \Delta_{i:n}(Z_i - p_i)^2. \tag{5.4}$$

If the failure-time distribution is continuous, $\overline{\mathbb{V}}_{1n}$ is the quadratic variation $[\overline{\mathbb{W}}_n, \overline{\mathbb{W}}_n]$ of the conditional martingale $\overline{\mathbb{W}}_n$ under the permutation distribution. This is not true in the general discontinuous case. Therefore, there is some theoretical support for $\overline{\mathbb{V}}_n$ in favor of $\overline{\mathbb{V}}_{1n}$. Beyond that, our simulations, performed for continuous distributions, showed superiority of $\overline{\mathbb{V}}_n$ over $\overline{\mathbb{V}}_{1n}$ in most situations; see the "omnibus" part of our simulations. If the conditional Lindeberg condition and (3.10) hold true, it follows also from Jacod and Shiryaev (1987), Theorem 8.3.33, that $\overline{\mathbb{V}}_{1n}(t) \to_P \overline{\mathbb{V}}(t)$.

6. Proofs. We use the previous notation. In particular, $\mathbf{Z}_n = (Z_1, \dots, Z_n)$ is assumed to have permutation distribution as defined after Lemma 3.1. The processes $Z_n(t) = (n_1 n_2/n)^{1/2}(M_{[nt]} - [nt]n_1/n)$, $0 \le t \le 1$, $M_i := Z_1 + \cdots + Z_i$, converge in distribution to the Brownian bridge $\mathbb{B}^0$ on $D[0,1]$; see Theorem 24.1 of Billingsley (1968). Consequently, if $n_1/n \to \eta \in (0,1)$,

$$\sup_{1 \le i \le nt} |p_i - \eta| \to_P 0, \qquad 0 < t < 1. \tag{6.1}$$

PROOF OF THE CONDITIONAL LINDEBERG CONDITION. The *conditional Lindeberg condition* mentioned before Theorem 3.2 is

$$\mathrm{CL}_n(\varepsilon, t) := \sum_{i=1}^{[nt]} E\big(U_i^2 1(|U_i| \ge \varepsilon) | \mathscr{F}_{i-1}\big) \to_P 0, \tag{6.2}$$

$$\forall \varepsilon > 0, 0 \le t < 1,$$

with $U_i := \overline{\mathbb{W}}_n(i/n) - \overline{\mathbb{W}}_n((i-1)/n)$. We show that (6.2) holds true for *all* sequences δ_n, $n \ge 1$. Since, conditionally on $\mathscr{F}_{i-1}$, Z_i has binomial$(1, p_i)$ distribution, one has with

$$d_i := (n/(n_1 n_2))\overline{w}_n^2\left(\frac{i}{n}\right)\delta_{i:n}$$

$$\le 2(\eta(1-\eta))^{-1} \max\left\{\overline{w}_n^2\left(\frac{i}{n}\right)/n : 1 \le i \le nt\right\} =: a_n,$$

the last inequality being true for large n, that

$$(6.3)\quad \begin{aligned} \mathrm{CL}_n(\varepsilon, t) &= \sum_{i=1}^{[nt]} d_i\big(1\{d_i q_i^2 \ge \varepsilon^2\}p_i + 1\{d_i p_i^2 \ge \varepsilon^2\}q_i\big) \\ &\le 1\{a_n \ge \varepsilon^2\} \sum_{i=1}^{[nt]} d_i \to_P 0 \quad \text{for } 0 < t < 1, \end{aligned}$$

where $\to_P 0$ follows since it is well known that quadratic mean convergence in probability in (3.8) implies $a_n \to_P 0$; see, for example, Neuhaus (1988), (5.15). Hence (6.2) is fulfilled for *all* sequences δ_n. □

PROOF THAT (3.7) AND (3.8) IMPLY (3.10). According to (3.2),

$$(6.4)\qquad \overline{\mathbb{V}}_n(t) = \int_0^t \overline{w}_n^2 f_n \, d\overline{\mathbb{H}}_n^1,$$

with $f_n(u) := (n^2/(n_1 n_2)) p_{[nu]} q_{[nu]} \to_P 1$ by (6.1) uniformly for $u \in [0, t]$. Combined with (3.8), one gets for $v_n := \overline{w}_n^2 f_n$ and $v := \overline{w}^2$ that $R_n := \int_0^t |v_n - v| \, d\lambda \to_P 0$, where λ is Lebesgue measure. For $\varepsilon > 0$ choose a continuous, bounded function g_ε with $b_\varepsilon := \int |g_\varepsilon - v| \, d\lambda < \varepsilon$. Hence

$$\left| \int_0^t v_n \, d\overline{\mathbb{H}}_n^1 - \int_0^t v \, d\overline{H}^1 \right| \le R_n + 2b_\varepsilon + \left| \int_0^t g_\varepsilon \, d\overline{\mathbb{H}}_n^1 - \int_0^t g_\varepsilon \, d\overline{H}^1 \right|,$$

where we have used $d\overline{\mathbb{H}}_n^1/d\lambda \le 1$ and $d\overline{H}^1/d\lambda \le 1$. The third term in the last sum tends to 0 because of (3.7). Since $\varepsilon > 0$ is arbitrary, (3.10) follows. □

According to the CLT cited before Theorem 3.2, the conditional Lindeberg condition and pointwise convergence of $\langle \overline{\mathbb{W}}_n, \overline{\mathbb{W}}_n \rangle$ to $\overline{V}$ imply weak convergence $\overline{\mathbb{W}}_n \to_{\mathscr{D}} \mathbb{B} \circ \overline{V}$ in $D[0,1)$. But $\langle \overline{\mathbb{W}}_n, \overline{\mathbb{W}}_n \rangle(t) = \overline{\mathbb{V}}_n(t) \to \overline{V}(t) \ \forall\, t$, according to (3.6) and (3.10).

EXTENSION OF (3.9) AND (3.10) TO $D[0,1]$, RESPECTIVELY $[0,1]$. Because of (3.12) and $\overline{V}(t) \to \overline{V}(1) < \infty$ as $t \uparrow 1$, Theorem 4.2 of Billingsley (1968) yields $\overline{\mathbb{V}}_n(1) \to_P \overline{V}(1)$. Following the reasoning of Gill (1980), one gets by Lenglart's inequality applied to $\overline{\mathbb{W}}_n^2$ and $\langle \overline{\mathbb{W}}_n, \overline{\mathbb{W}}_n \rangle = \overline{\mathbb{V}}_n$ that

$$\mathbb{P}\Big\{ \sup_{t \le s \le 1} |\overline{\mathbb{W}}_n(s) - \overline{\mathbb{W}}_n(t)| \ge \varepsilon \Big\} \le (\gamma/\varepsilon^2) + \mathbb{P}\{\overline{\mathbb{V}}_n(1) - \overline{\mathbb{V}}_n(t) \ge \gamma\} \qquad \forall\, \varepsilon, \gamma > 0.$$

Now, the same argument as for $\overline{\mathbb{V}}_n$ yields (3.9) on $D[0,1]$. □

PROOF OF (3.13). Let $\mathbb{H}_n$ be the empirical d.f. of the X_i's and let $\mathbb{H}_n^1$ be the empirical sub-d.f. of the uncensored X_i's. It follows from the Glivenko–Cantelli theorem (respectively, a simple extension of it) that, under $\mathscr{H}_0$, $\|\mathbb{H}_n - H\|_\infty \to_P 0$, respectively $\|\mathbb{H}_n^1 - H^1\|_\infty \to_P 0$, where $\|\cdot\|_\infty$ means supremum norm. By switching to subsequences, we may assume that $\mathbb{H}_n$ and $\mathbb{H}_n^1$, $n \geq 1$, are fixed sequences such that the preceding convergences hold true. Apparently, $\|\overline{\mathbb{H}}_n^1 - \mathbb{H}_n^1 \circ \mathbb{H}_n^{-1}\|_\infty \leq 1/n$. Since $\mathbb{H}_n^1 \circ \mathbb{H}_n^{-1}(t) \to H^1 \circ H^{-1}(t)$ for t from the dense set of continuity points of H^{-1}, (3.13) follows at once. □

PROOF OF (3.8) IN EXAMPLE 3.4. According to (6.1), the only nontrivial point is to show that $\overline{\mathbb{S}}_n(t) = \prod_{1 \leq j \leq i-1}(1 - \delta_{j:\,n}/(n-j+1))$, $i := [nt]$, tends to $\overline{S}(t)$, $0 < t < 1$. Put $f_n(u) = n/(n - [nu] + 1)$. Then

$$\begin{aligned}\left|-\ln \overline{\mathbb{S}}_n(t) - \int_0^{i/n} f_n \, d\overline{\mathbb{H}}_n^1\right| &= \sum_{j=1}^{i-1} \delta_{j:\,n}\left(\ln\left(1 - \frac{1}{n-j+1}\right) - \frac{1}{n-j+1}\right)\\ &\leq \sum_{1 \leq j < i} (n-j+1)^{-1}(n-j)^{-1}\\ &= (n-i+1)^{-1} - 1/n\\ &\leq (1-t)^{-1} n^{-1} \to 0,\end{aligned}$$

using the inequality $0 < -\ln(1 - 1/(x+1)) - 1/(x+1) < 1/x(x+1)$ for $0 < x < \infty$. As in (6.4),

$$\int_0^{i/n} f_n \, d\overline{\mathbb{H}}_n^1 \to \int_0^t (1 - id)^{-1} \, d\overline{H}^1 = -\ln \overline{S}(t).$$

For the last equality, see, for example, Shorack and Wellner (1986), page 295. Combining we get $\overline{\mathbb{S}}_n(t) \to \overline{S}(t)$. □

PROOF OF (3.12) IN EXAMPLE 3.4. If $\rho \geq 0$, $\kappa - 2\lambda \geq 0$ and $\lambda \leq 0.5$, then $\overline{w}_n^2(i/n) p_i q_i$, $1 \leq i \leq n$, $n \geq 1$, are uniformly bounded by some nonrandom constant, implying (3.12). If $\rho \geq 0.5$, $\kappa - 2\lambda + 1 \geq 0$ and $\lambda \leq 0.5$, one gets by using the equality $(n - i + 1)\Delta \overline{\mathbb{S}}_n(i/n) = -\overline{\mathbb{S}}_n(i/n)\delta_{i-1:\,n}$ that $\mathbb{V}_n(1) - \mathbb{V}_n(t) \leq \text{const}(\overline{\mathbb{S}}_n([nt]/n) - \overline{\mathbb{S}}_n(1))$. Since $\overline{\mathbb{S}}_n([nt]/n) \to \overline{\mathbb{S}}(t)$ for $0 < t < 1$, condition (3.12) is fulfilled if $\overline{\mathbb{S}}(1) = 0$. Otherwise, if $\overline{\mathbb{S}}(1) = 1 - F(H^{-1}(1)) > 0$, we use the fact that (unconditionally) $\prod_{1 \leq j \leq n}(1 - \Delta_{j:\,n}/(n - j + 1)) \to_P 1 - F(H^{-1}(1))$ under $\mathscr{H}_0$, which follows from a minor extension of Wang's (1987) results to the case with possibly $G_1 \neq G_2$. By switching to subsequences we may assume that $\liminf \overline{\mathbb{S}}_n(1) \geq \overline{S}(1)$. Then (3.12) follows from

$$\limsup_{n \to \infty}\left(\overline{\mathbb{S}}_n\left(\frac{[nt]}{n}\right) - \overline{\mathbb{S}}_n(1)\right) \leq \overline{S}(t) - \overline{S}(1) \to 0 \quad \text{as } t \to 0. \qquad \square$$

7. Simulations. In a separate simulation study the performance of various conditional and unconditional linear rank tests and omnibus tests has been worked out. Here we want to give a small excerpt and some conclusions.

7.1. *Linear rank tests.* In order to make comparisons with known results, we just repeated the simulation study of Janssen and Brenner (1991) with the following changes: First, due to our notation, we interchanged their sample numbering. Moreover, they used for the conditional test the nonstandardized statistic $\overline{\mathbb{W}}_n(1)$ and for the unconditional test the standardized version $\overline{\mathbb{W}}_n(1)/\overline{\mathbb{V}}_{1n}^{1/2}(1)$, whereas we use in both cases the standardized version $\overline{\mathbb{W}}_n(1)/\overline{\mathbb{V}}_n^{1/2}(1)$. At the outset our reason for preferring $\overline{\mathbb{V}}_n(1)$ to $\overline{\mathbb{V}}_{1n}(1)$ was that $\overline{\mathbb{V}}_n(1)$ is the predictable quadratic variation of $\overline{\mathbb{W}}_n$ in the continuous as well as in the discontinuous case, while $\overline{\mathbb{V}}_{1n}(1)$ is the quadratic variation only for continuous failure time distributions; see Remark 5.1. In fact, simulations not included here show the superiority of $\overline{\mathbb{V}}_n(1)$ in the conditional as well as in the unconditional case. The actual level was in practically all cases considered closer to the nominal one for $\overline{\mathbb{V}}_n(1)$; see also the following omnibus case. Furthermore, the estimated level of the conditional versions [with variance estimator $\overline{\mathbb{V}}_n(1)$] was usually more accurate than that of the unconditional version. In particular, for unbalanced cases ($n_1 = 50$, $n_2 = 10$) the unconditional versions were too conservative. Moreover, when the tests where performed for rounded variables (ties!), a case not treated by Janssen and Brenner (1991), the unconditional versions became highly anticonservative, while at the same time the conditional version's actual level was quite close to the nominal one; see Table 1.

TABLE 1

Actual levels of various unconditional (unc) and conditional (con) linear rank tests with predictable quadratic variation estimator (5.2). p_1 resp. p_2 is the probability of no censoring in sample 1 resp. sample 2

	p_1	p_2		$1/2 \mid 1/2$			$1/4 \mid 1/2$			$1/2 \mid 1$			$1/2 \mid 3/4$		
F	n_1	n_2		LR	GW	PW	LR	GW	PW	LR	GW	PW	LR	GW	PW
				Nominal level 5%											
E	10	10	unc	5.4	5.2	5.2	5.9	5.5	5.4	5.9	5.6	5.5	6.0	5.3	5.6
			con	5.2	5.0	4.9	5.2	5.6	5.6	5.4	5.5	5.2	4.8	4.8	5.4
E	50	10	unc	3.7	3.3	3.5	3.7	3.3	3.6	4.2	3.8	4.0	4.0	3.7	3.8
			con	5.4	4.9	5.3	5.4	5.8	6.8	5.3	5.3	6.0	5.5	5.4	5.7
W	10	10	unc	5.9	5.4	5.4	6.8	6.0	6.3	7.0	5.8	6.2	6.5	5.7	5.9
			con	4.7	5.3	5.0	5.5	5.8	6.4	6.0	5.9	6.9	5.4	5.1	6.5
W	50	10	unc	3.7	3.1	3.5	4.0	2.9	3.8	4.5	3.9	4.1	4.3	3.6	3.9
			con	5.1	4.8	4.8	6.2	6.5	7.5	5.8	5.4	6.9	5.8	5.2	5.5
L	10	10	unc	5.5	5.0	5.0	5.9	5.8	5.7	6.2	5.5	5.6	5.9	5.5	5.4
			con	5.3	4.7	4.6	5.0	5.5	6.5	5.6	5.4	5.4	5.6	4.9	5.4
L	50	10	unc	3.6	3.6	3.5	3.8	3.2	3.7	4.2	3.8	3.8	4.0	3.7	3.8
			con	5.8	4.4	5.3	5.6	6.0	6.6	5.7	5.5	5.9	5.8	5.4	6.5
				Typical cases with observations rounded up to 1 / 10											
E	10	10	unc	11.1	6.0	7.0	7.4	4.7	5.8	14.7	7.8	8.9	14.1	7.3	8.4
			con	5.0	4.7	5.2	4.5	5.0	4.9	3.5	4.6	4.9	3.8	5.1	5.0
W	50	10	unc	11.1	3.9	6.9	11.2	4.6	6.7	15.2	5.4	9.8	13.6	4.5	8.4
			con	5.5	4.2	4.8	6.4	5.7	6.3	5.8	6.3	6.2	5.7	5.9	4.6
L	50	10	unc	14.7	6.8	10.1	14.7	4.4	10.1	23.4	11.4	15.5	21.6	9.8	14.3
			con	5.2	5.1	5.0	5.9	6.3	6.7	6.3	5.9	6.1	5.7	6.0	5.8

The linear rank tests considered in the simulation study are the log-rank test (LR), the Gehan–Wilcoxon (GW) and the Prentice–Wilcoxon (PW) test, see (2.3), with variance estimator $\overline{\mathbb{V}}_n(1)$. The choice of distributions follows Latta (1981): Let F be either the exponential(β)-d.f., $F(x) = 1 - \exp(-\beta x)$, $x > 0$, or the Weibull(β^4, 4)-d.f., $F(x) = 1 - \exp(-(\beta x)^4)$, $x > 0$, or the log-normal(β)-d.f., $F(x) = \Phi(\log(x\beta))$, $x > 0$, with Φ the standard normal d.f. and parameter $\beta > 0$. The three classes of distributions will be abbreviated by $E = E(\beta)$, $W = W(\beta)$, $L = L(\beta)$, respectively. Under the null hypothesis we choose $F_1 = F_2 = F$ with $\beta = 1$ in the three cases and the censoring distribution G_k to be uniform on some interval $[0, T_k]$, where T_k is chosen such that the probabilities of no censoring, p_k, attain prescribed values $p_k = 1/4, 1/2, 3/4, 1$, for the distributions E, W and L, respectively, $k = 1, 2$.

Table 1 shows estimated levels under the null hypothesis $F_1 = F_2 = F$. The abbreviation "unc," respectively "con," means the unconditional, respectively conditional, version of the various tests. The nominal level is 5.0%. The second part of Table 1 repeats for some typical cases the computations with all observations rounded up to 1/10. Notice the inaccuracy of the unconditional tests when ties occur! The conditional tests for $p_1 = p_2 = 0.5$ have the exact level 5.0%. The variation in the corresponding part of Tables 1 and 3 is solely a consequence of simulation and may help judging the accuracy of the results for $p_1 \neq p_2$.

Table 2 shows the power of the tests in Table 1 at the 5.0% level for $n_1 = n_2 = 10$ under equal censoring $p_1 = p_2 = 0.5$ for various alternatives. It turns out (also in many other cases not reported here) that the conditional and unconditional tests have very similar power. Altogether the conditional versions appear to be preferable to their unconditional counterparts in most situations.

TABLE 2

Power of various unconditional (unc) and conditional (con) linear rank tests with predictable quadratic variation estimator (3.2) for sample sizes $n_1 = n_2 = 10$ at level 5% under various alternatives $(E(\beta), E(1))$, $(W(\beta), W(1))$ and $(L(\beta), L(1))$

	(E(β), E(1))				(W(β), W(1))				(L(β), L(1))			
β	LR	GW	PW	β	LR	GW	PW	β	LR	GW	PW	
0.5	0.6	0.6	0.5	0.9	1.8	1.7	1.6	0.5	0.3	0.3	0.3	unc
	0.6	0.8	0.6		1.6	1.8	1.6		0.3	0.2	0.6	con
1.0	5.1	4.8	5.1	1.0	6.1	5.3	5.5	1.0	5.3	5.0	5.2	unc
	4.9	5.1	4.9		4.4	4.9	5.0		5.8	5.0	5.0	con
1.5	16.6	14.8	15.5	1.1	14.9	13.0	13.4	1.5	18.5	18.6	18.9	unc
	14.4	14.8	13.9		12.9	12.5	12.5		17.7	18.9	13.0	con
2.0	31.7	27.2	29.2	1.4	64.6	55.5	58.2	2.0	36.3	36.7	36.6	unc
	30.5	27.0	25.2		60.8	54.6	54.4		34.4	36.7	25.7	con
3.0	59.6	52.1	54.6	1.6	87.1	79.1	81.3	3.0	65.6	67.9	67.7	unc
	57.3	50.8	48.1		84.2	78.5	78.8		65.1	67.0	45.2	con
4.0	77.9	69.6	72.3	1.8	95.9	91.2	92.7	4.0	82.7	84.9	84.8	unc
	75.7	68.4	68.4		94.4	90.9	91.6		81.7	83.8	61.1	con

TABLE 3

Estimated levels (in %) of various omnibus tests for sample sizes $n_1 = n_2 = 50$ and $n_1 = 10$, $n_2 = 50$. The nominal level is 5.0%. p_1 resp. p_2 is the probability of no censoring in sample 1 resp. sample 2. $d.f._1 = d.f._2$ is the common distribution function in both samples

p_1	p_2	1/2 | 1/2				1/2 | 1/4				1 | 1/2				3/4 | 1/2				
d.f.$_1$	d.f.$_2$	KSK	KS	CMK	CM	KSK	KS	CMK	CM	KSK	KS	CMK	CM	KSK	KS	CMK	CM	
		Unconditional version with pqv estimator																
Exp	Exp	2.7	2.7	5.5	5.2	2.8	2.5	4.6	4.4	3.7	3.4	5.9	5.4	3.2	3.2	5.3	5.0	$n_1 = 50$
W15	W15	3.2	3.0	5.2	4.7	3.0	2.9	4.6	4.3	2.8	2.8	5.1	4.8	2.8	2.6	4.5	4.1	
W05	W05	3.6	3.4	6.0	5.7	2.9	2.9	5.0	4.6	3.0	2.7	4.9	4.6	2.8	2.8	4.2	4.1	$n_2 = 50$
		Conditional version with pqv estimator																
Exp	Exp	5.2	5.2	5.3	5.3	5.3	5.5	4.6	4.5	5.1	5.2	5.3	5.3	4.8	4.9	4.9	4.9	$n_1 = 50$
W15	W15	4.8	4.9	4.9	4.9	5.3	5.4	4.9	4.9	5.3	5.2	5.0	5.1	4.2	4.2	4.4	4.4	
W05	W05	5.1	5.2	4.9	5.0	4.7	4.8	4.5	4.5	4.8	4.9	4.6	4.6	4.3	4.2	4.5	4.6	$n_2 = 50$
		Unconditional version with qv estimator																
Exp	Exp	19.9	14.6	12.5	10.3	13.3	9.6	9.9	8.4	16.4	12.4	11.4	9.4	17.2	12.6	10.8	8.9	$n_1 = 10$
W15	W15	20.0	14.2	12.2	9.5	15.8	11.4	12.3	10.5	16.6	11.5	11.4	9.3	17.5	12.4	11.5	9.5	
W05	W05	19.7	14.5	13.3	10.1	15.3	11.2	12.3	10.7	14.7	10.5	10.4	8.3	17.3	12.3	10.8	8.7	$n_2 = 50$
		Unconditional version with pqv estimator																
Exp	Exp	2.8	3.9	4.7	5.0	1.8	2.9	4.0	4.2	2.8	3.7	5.4	5.1	3.0	4.1	5.6	5.5	$n_1 = 10$
W15	W15	3.4	4.8	5.2	5.7	2.9	3.8	4.9	5.0	2.4	3.4	5.4	5.6	3.4	4.6	5.4	5.6	
W05	W05	4.4	5.5	6.0	6.4	2.7	3.4	5.2	5.2	3.4	4.1	6.1	6.3	2.5	3.6	5.3	5.2	$n_2 = 50$
		Conditional version with pqv estimator																
Exp	Exp	5.2	5.2	5.1	5.0	4.9	4.6	4.9	4.9	4.6	4.2	4.6	4.4	4.7	4.5	4.9	4.9	$n_1 = 10$
W15	W15	4.9	4.8	4.7	4.7	4.2	4.0	4.4	4.3	5.0	4.5	5.2	5.0	5.0	4.9	5.0	5.0	
W05	W05	4.8	5.1	4.7	4.6	4.2	4.3	4.8	4.8	4.7	4.2	5.0	4.8	4.9	4.7	5.2	5.2	$n_2 = 50$

7.2. *Omnibus tests.* Our simulations follow partially those of Schumacher (1984), who considered three types of alternatives: Proportional hazards with F_k = exponential(λ_k), $k = 1, 2$; large early difference with F_1 = Weibull(λ_1, 1.5), F_2 = Weibull(λ_2, 1); and crossing survival curves with F_1 = Weibull(λ_1, 0.5), F_2 = Weibull(λ_2, 1), abbreviated by (Exp, Exp), (W15, Exp) and (W05, Exp), respectively, in the tables. Under the null hypothesis $\mathscr{H}_0$: $\lambda_1 = \lambda_2 = 1$, in the three cases the censoring distributions are chosen in the same way as for linear rank tests. Under alternatives the censoring d.f. is the uniform distribution on the interval [3, 5] for both samples resulting in equal censoring, and λ_1, λ_2 chosen so that prescribed five-year survival probabilities q_1, q_2, say, result; see Schumacher (1984).

Several omnibus test statistics are considered: $\mathrm{KS}_n^{\vartheta}$ from (3.5) and $\mathrm{CM}_n^{\vartheta}$ from (4.4) ($\mathbb{W}_n$ the hazard process), $\mathrm{KS}K_n^{\vartheta}$ from (4.2) and $\mathrm{CM}K_n^{\vartheta}$ from (4.3), called KS, CM, KSK, CMK, respectively, in Tables 3 and 4, with the addition of the "pqv estimator" (predictable quadratic variation), respectively "qv estimator" (quadratic variation), if $\overline{\mathbb{V}}_n$, respectively $\overline{\mathbb{V}}_{1n}$, is used. In fact, in order to get results comparable to Schumacher's (1984) simulations, instead of $\overline{\mathbb{V}}_{1n}$ the slightly altered estimator $\overline{\mathbb{V}}_{1n}^*$ is used which is the reduced form of the estimator V_1; see Gill (1980), (3.3.11), with dN_k/Y_k replaced by $dN_k/(Y_k - 1)$, $k = 1, 2$. Moreover, the norming factor $(n_1 n_2/n)$ in all processes is replaced by "n" which influences the computation of $\overline{\mathbb{K}}_n = \overline{\mathbb{V}}_n/(1 + \overline{\mathbb{V}}_n)$ and $\overline{\mathbb{K}}_n^* := \overline{\mathbb{V}}_{1n}^*/(1 + \overline{\mathbb{V}}_{1n}^*)$. We follow the usual practice and stop the processes at the largest i/n with $Y_k(X_{i:n}) > 0$, $k = 1, 2$, when using $\overline{\mathbb{V}}_n$, respectively $Y_k(X_{i:n}) > 1$, $k = 1, 2$, when using $\overline{\mathbb{V}}_{1n}^*$.

Table 3 ($n_1 = n_2 = 50$) and ($n_1 = 10, n_2 = 50$) contains estimated levels, and Table 4 ($n_1 = n_2 = 50$) contains estimated powers of the various tests.

TABLE 4

Estimated power (in %) of various omnibus tests for sample sizes $n_1 = n_2 = 50$ and various distribution functions d.f.$_1$ and d.f.$_2$ with prescribed five-year survival probabilities q_1 and q_2 in sample 1 resp. sample 2. The censoring distribution is uniform (3, 5) in both samples

				Conditional tests				Conditional tests				Unconditional tests	
				KSK	KS	CMK	CM	KSK	KS	CMK	CM	CMK	CM
d.f.$_1$	d.f.$_2$	q_1	q_2	pqv estimator				qv estimator				qv estimator	
Exp	Exp	0.80	0.70	15.0	14.9	15.4	15.4	15.9	16.2	16.6	16.5	14.7	14.4
W15	Exp	0.80	0.70	28.7	28.7	30.6	30.6	28.5	28.4	30.4	30.6	43.9	42.6
W05	Exp	0.80	0.70	9.9	9.8	7.8	7.8	9.3	9.4	7.8	7.8	6.9	6.1
Exp	Exp	0.55	0.45	11.9	11.9	12.6	12.6	11.2	11.3	12.0	12.1	10.6	10.3
W15	Exp	0.55	0.45	42.4	42.4	46.5	46.5	42.1	42.0	46.1	46.0	49.0	47.6
W05	Exp	0.55	0.45	24.6	24.6	24.1	23.9	27.1	27.0	26.0	25.8	28.0	26.9
Exp	Exp	0.30	0.20	12.9	12.6	13.2	13.3	11.0	11.0	12.8	12.9	13.7	13.4
W15	Exp	0.30	0.20	69.0	68.9	75.1	75.2	68.1	68.1	75.3	75.3	75.2	74.2
W05	Exp	0.30	0.20	60.7	60.4	65.3	65.4	62.9	62.8	64.8	64.7	63.9	62.8

The nominal level is 5.0% in all tables. The unconditional versions are applied with asymptotical critical values of the usual KS and CM statistics.

Discussion. Several questions may be asked: What are the differences between the hazard processes and the Kaplan–Meier processes? Are the predictable quadratic variation (pqv) estimators preferable to the quadratic variation (qv) estimators? How sensitive are the different versions to unbalanced sample sizes? How robust are the different versions against unequal censoring?

It has turned out that the pqv estimators and the qv estimators behave quite similarly as long as the sample sizes are equal, whereas for unbalanced sample sizes the qv estimator may give disastrous estimated levels as shown in rows 7–9 of Table 3 with $n_1 = 10$, $n_2 = 50$ in contrast to rows 10–12. Therefore, the pqv estimator $\overline{\mathbb{V}}_n$ seems to be the better choice. Accepting $\overline{\mathbb{V}}_n$, it is seen in all of our simulations that the hazard process—and the Kaplan–Meier process—versions behave almost identically under the null hypothesis as well as under alternatives. Since the test statistics based on the hazard process $\overline{\mathbb{W}}_n$ are somewhat easier to compute, the latter versions seem to be the preferable ones.

As already found by Schumacher (1984) in his unconditional setting (with the qv estimator and equal sample sizes), the KS tests are too conservative compared to the corresponding CM tests. The same observation holds true for the *unconditional* version with the pqv estimator; see Table 3. In contrast, for the *conditional* versions the attainment of level is very good for the KS tests as well as for the CM tests under equal censoring *and* under unequal censoring, see the last three rows of Table 3 with $n_1 = n_2 = 50$ and $n_1 = 10$, $n_2 = 50$.

Let us look at the power results in Table 4. The unconditional tests CMK and CM with the qv estimator are Schumacher's tests $Q^0_{\mathrm{CM}\eta}$ and $Q^0_{\mathrm{CM}\xi}$, respectively. The powers of the unconditional CM tests of Schumacher are very similar to the powers of their conditional counterparts. Moreover, conditional KS tests (being nonapplicable in their unconditional version because of their too strong conservatism) have power properties similar to the corresponding CM tests. Among the tests considered here the conditional CM test with the pqv estimator seems to be the most attractive one. Each entry in our tables concerning unconditional tests is based on 2000 Monte Carlo repetitions. The conditional versions are based on 3000 Monte Carlo repetitions preceded by 2000 repetitions for estimating the conditional critical values, that is, $3000 \cdot 2000$ runs were made for each such entry. In order to make the computed powers of the different conditional tests comparable, we have used randomization to achieve exactly the nominal level.

Acknowledgments. I gratefully acknowledge the comments and remarks of the Associate Editor and two referees which led to a substantially improved presentation. I thank R. Moehring for his help during the simulations.

REFERENCES

AALEN, O. O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6** 77–85.

ANDERSEN, P. K., BORGAN, O., GILL, R. and KEIDING, N. (1982). Linear nonparametric tests for comparison of counting processes, with applications to censored survival data. *Internat. Statist. Rev.* **50** 219–258.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

FLEMING, T. R., HARRINGTON, D. P. and O'SULLIVAN, M. (1987). Supremum versions of the log-rank and generalized Wilcoxon statistics. *J. Amer. Statist. Assoc.* **82** 312–320.

GEHAN, E. (1965). A generalized Wilcoxon test for comparing arbitrarily single censored samples. *Biometrika* **52** 203–223.

GILL, R. D. (1980). *Censoring and Stochastic Integrals. Math. Centre Tracts* **124**. Math. Centrum, Amsterdam.

HARRINGTON, D. P. and FLEMING, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika* **69** 553–566.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin.

JANSSEN, A. (1991). Conditional rank tests for randomly censored data. *Ann. Statist.* **19** 1434–1456.

JANSSEN, A. and BRENNER, S. (1991). Monte Carlo results for conditional survival tests under randomly censored data. *J. Statist. Comput. Simulation* **39** 47–62.

KOZIOL, J. A. (1978). A two-sample Cramér–von Mises test for randomly censored data. *Biometrical J.* **20** 603–608.

KOZIOL, J. A. and YUH, Y. S. (1982). Omnibus two sample test procedures with randomly censored data. *Biometrical J.* **24** 743–750.

LATTA, R. B. (1981). A Monte Carlo study of some two-sample rank tests with censored data. *J. Amer. Statist. Assoc.* **76** 713–719.

LEURGANS, S. (1984). Asymptotic behavior of two-sample rank tests in the presence of random censoring. *Ann. Statist.* **12** 572–589.

MANTEL, N. (1967). Ranking procedures for arbitrarily restricted observation. *Biometrics* **23** 65–78.

NEUHAUS, G. (1988). Asymptotically optimal rank tests for the two-sample problem with randomly censored data. *Comm. Statist. Theory Methods* **17** 2037–2058.

NEUHAUS, G. (1991a). Conditional Kolmogorov–Smirnov tests for the two sample problem under random censorship. I. Preprint 91-1, Univ. Hamburg.

NEUHAUS, G. (1991b). Conditional Kolmogorov–Smirnov tests for the two sample problem under random censorship. II. Preprint 91-4, Univ. Hamburg.

NEUHAUS, G. (1992). Conditional rank tests for the two-sample problem under random censorship: treatment of ties. *Proceedings of the Fourth International Meeting of Statistics in the Basque Country*. To appear.

PETO, R. and PETO, J. (1972). Asymptotically efficient rank invariant test procedures. *J. Roy. Statist. Soc. Ser. A* **135** 185–206.

PRENTICE, R. L. (1978). Linear rank tests with right censored data. *Biometrika* **65** 165–179.

SCHUMACHER, M. (1984). Two sample tests of Cramér–von Mises and Kolmogorov–Smirnov-type for randomly censored data. *Internat. Statist. Rev.* **52** 263–281.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

WANG, J.-G. (1987). A note on the uniform consistency of the Kaplan–Meier estimator. *Ann. Statist.* **15** 1313–1316.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT HAMBURG
BUNDESSTRASSE 55, 20146 HAMBURG
GERMANY