stichting
mathematisch
centrum

MC

AFDELING STATISTIEK SW 72/81 JULI
(DEPARTMENT OF MATHEMATICAL STATISTICS)

PER KRAGH ANDERSEN, ØRNULF BORGAN,
RICHARD GILL & NIELS KEIDING

LINEAR NONPARAMETRIC TESTS FOR COMPARISON OF COUNTING PROCESSES, WITH APPLICATIONS TO CENSORED SURVIVAL DATA

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, 413 Kruislaan, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

1980 Mathematics subject classification: 62G10, 62M99, 62N05, 62P10

Linear nonparametric tests for comparison of counting processes, with applications to censored survival data

by

Per Kragh Andersen[1], Ørnulf Borgan[2,3], Richard Gill[1,4] & Niels Keiding[1].

ABSTRACT

This paper surveys linear nonparametric one- and k-sample tests for counting processes. The necessary probabilistic background is outlined and a master theorem proved, which may be specialized to most known asymptotic results for linear rank tests for censored data as well as to asymptotic results for one- and k-sample tests in more general situations, an important feature being that very general censoring patterns are allowed.

A survey is given of existing tests and their relation to the general theory, and we provide examples of applications to Markov processes. We also discuss the relation of the present approach to classical nonparametric hypothesis testing theory based on permutation distributions.

KEY WORDS & PHRASES: *asymptotic theory; censoring; Kruskal-Wallis test; k sample test; log-rank test; Markov process; martingale; one sample test; permutation distribution.*

---

1 Statistical Research Unit, Danish Medical and Social Science Research Councils, Copenhagen.

2 Universities of Copenhagen and Oslo.

3 Supported by The Norwegian Research Council for Science and the Humanities and by the Association of Norwegian Insurance Companies.

4 Mathematical Centre, Amsterdam

1. INTRODUCTION

Although actuarial computations of survivorship functions have been known for a long time, much of the recent biostatistical research in non-parametric statistical methods for censored data take the important review paper by KAPLAN & MEIER (1958) as starting point. Kaplan and Meier's empirical distribution (or survivorship) function provides a direct descriptive means of studying the simplest situation with censored observations. Out of n independent identically distributed lifetimes X_i, some are observed, but for the rest it is only known that they are larger than some time T_i: the observations are given as $(\min(X_i,T_i),\ I\{X_i \leq T_i\})$; $i = 1,\ldots,n$, where $I(\cdot)$ is the indicator function. The particular situation with the T_i's given as nonrandom quantities is called *fixed* censorship, but we shall not be restricted to that in this report.

Two important developments in nonparametric statistical analysis of censored data have emerged from the Kaplan-Meier estimator. The first sticks to the subject of independent lifetimes, but provides statistical methods for assessing the dependence of the survival distribution on individual characteristics. Examples are the modifications of standard nonparametric two- and k-sample tests by GEHAN (1965), BRESLOW (1970), and PETO & PETO (1972), and the semiparametric regression model by COX (1972). This development has been reviewed in the very important monograph by KALBFLEISCH & PRENTICE (1980), which contains a comprehensive set of references.

The second development generalizes the simple alive-dead situation to several states, modelling each individual's life history as a stochastic process; so far most often a Markov process. Again the models have roots far back in actuarial science and also demography (see e.g. HOEM, 1976); however a decisive breakthrough for nonparametric statistical theory for such models occurred only recently with Aalen's (1975) thesis, cf. AALEN (1978). The mathematical framework is multivariate counting processes and stochastic integrals and the asymptotic theory is based upon martingale central limit theory.

The present paper is devoted to an exposition of one- and k-sample tests for comparison of counting processes. We use the definitive martingale central limit theory by REBOLLEDO (1978), cf. also SHIRYAEV (1980), according to which one master theorem provides proofs of the asymptotic distributions of many estimators and test statistics known from the literature for the simple life expectancy problem, as well as their generalizations to the counting process framework. GILL (1980), following up the preliminary treatment by AALEN (1978, section 7), performed a detailed study of the two-sample problem (comparing two life distributions) using this probabilistic machinery, and added a study by similar methods of the one-sample estimation problem.

Along the way, we show that most, if not all, generalized linear rank tests for censored data may be treated in this general framework. (The word "linear" in the title of this paper serves to distinguish from Kolmogorov-Smirnov tests, such as discussed by FLEMING et al. (1980), and other non-linear nonparametric tests such as the recent k-sample median test by BROOKMEYER & CROWLEY (1980). Some of these can in fact also be treated in our framework, see GILL (1980).)

Our own original motivation (which at the same time provides a practical example) was an empirical study (AALEN et al., 1980) on the possible interaction between menopause and the onset of the chronical skin disease *pustulosis palmo-plantaris*. This application goes beyond the simple life-table set-up.

The plan of the paper is as follows. In Section 2 we review the basic results from the theory of counting processes, and we recapitulate Aalen's "multiplicative intensity" model for multivariate counting processes, including a discussion on how various types of censoring may be incorporated into the model. A brief summary of the necessary probabilistic background is deferred to Appendix I. Section 3A provides a general k-sample test statistic and its asymptotic distribution, while Section 3B is an exposition of simpler conservative approximations to the test statistic. In Section 3C we show how the classical nonparametric k-sample tests fit into our framework, deferring to Appendix III the detailed verification for the Kruskal-Wallis test of the conditions of our "master" Theorem 3.1. Section 3D surveys current literature on k-sample tests for censored data, including a discussion of the direct validity of the conservative approximations under asymptotically equal censorship. In Section 3E we discuss the connection between our approach and the conventional approach, only valid under equal censorship, via permutation distributions. Chapter 3 is concluded by a couple of examples from Markov processes, and thus outside of the censored survival data circles. Finally, Section 4A discusses a general one-sample test statistic with examples and discussion in Section 4B.

2. MULTIVARIATE COUNTING PROCESSES AND THE MULTIPLICATIVE INTENSITY MODEL

2A. Multivariate counting processes

The basic statistical model will be formulated in terms of *multivariate counting processes*; for a detailed review see BRÉMAUD & JACOD (1977). (The reader is referred to Appendix I for the necessary probabilistic definitions.) For an alternative approach to the theory attempting to minimize the dependence on general martingale theory we mention the recent comprehensive lecture notes by JACOBSEN (1981).

The discussion below is phrased in terms of processes defined on the time interval $[0,1]$. The theory thus trivially applies to any other compact interval $[0,t]$ but some care has to be observed when interpreting generalizations (via monotone transformation) to $[0,\infty]$, since the behaviour at ∞ may restrict the applicability of the results. We shall add some comments on this when necessary.

Briefly, we consider a complete probability space $(\Omega,\mathcal{F},P)$ and an increasing, right-continuous family $(\mathcal{F}_t,\ t \in [0,1])$ of sub-σ-algebras of $\mathcal{F}$. A stochastic process $\underset{\sim}{N} = ((N_1(t),\dots,N_k(t)),\ t \in [0,1])$ is called a k-dimensional *counting process* if each of the k component processes N_i has a sample function which is a right-continuous step function with a finite number of jumps, each of size +1, and if furthermore two different component processes cannot jump at the same time. We require that $\underset{\sim}{N}(0) = 0$, that $\underset{\sim}{N}$ is adapted to $(\mathcal{F}_t)$, and that $EN_i(1) < \infty$ for all i. (This last condition can actually be dropped if *local* martingale techniques are used; see MEYER (1976).) Since each component process N_i is increasing and integrable and hence a submartingale, we have the Doob-Meyer decomposition $N_i = A_i + M_i$, where A_i is a natural (or predictable) increasing process and M_i a martingale. We shall assume that A_i is absolutely continuous, more precisely that there exists a unique non-negative left-continuous stochastic process $\underset{\sim}{\Lambda} = ((\Lambda_1(t),\ldots,\Lambda_k(t)),\ t \in [0,1])$ adapted to $(\mathcal{F}_t)$ and with right-hand limits, such that $A_i(t) = \int_0^t \Lambda_i(s)ds$; $i = 1,\ldots,k$. Then (cf. AALEN, 1978, pp.704-705)

$$(2.1) \qquad M_i(t) = N_i(t) - \int_0^t \Lambda_i(s)ds, \qquad i = 1,\ldots,k,$$

are orthogonal square integrable martingales with variance processes

$$(2.2) \qquad \langle M_i,M_i\rangle(t) = \int_0^t \Lambda_i(s)ds, \qquad i = 1,\ldots,k.$$

The process $\underset{\sim}{\Lambda}$ is called the *intensity process* of $\underset{\sim}{N}$. This name is justified by the following property (valid whenever each of the components of $\underset{\sim}{\Lambda}$ are dominated by an integrable random variable):

$$\lim_{h\downarrow 0} \frac{1}{h} P(N_i(t+h) - N_i(t) = 1|\mathcal{F}_t) =$$

$$\lim_{h\downarrow 0} \frac{1}{h} E(N_i(t+h) - N_i(t)|\mathcal{F}_t) = \Lambda_i(t+), \qquad i = 1,\ldots,k.$$

2B. Asymptotic theory for square integrable martingales related to counting processes

The basic tool in deriving asymptotic results for the general k-sample statistic in Section 3 is the following theorem, which may be obtained from Rebolledo's (1978) Theorème I.3.4 by a fairly standard argument using the Cramér-Wold device (see also AALEN (1977, Lemma A.1)).

Consider a sequence $\underset{\sim}{N}^{(n)}$ of k-dimensional counting processes with intensity processes $\underset{\sim}{\Lambda}^{(n)}$ and define for each $n = 1,2,\ldots$ square martingales $\underset{\sim}{M}^{(n)} = (M_1^{(n)},\ldots,M_k^{(n)})$ by (2.1). Let furthermore $\underset{\sim}{H}^{(n)} = (H_{ij}^{(n)})$ be a sequence of (m×k)-matrices of predictable processes on [0,1] satisfying $E\int_0^1 [H_{ij}^{(n)}(s)]^2\,\Lambda_j^{(n)}(s)ds < \infty$ for all i,j, which for instance holds if $H_{ij}^{(n)}$ is bounded (cf. Appendix I). Then the following theorem holds.

THEOREM 2.1. *Suppose that the following two conditions are fulfilled: There exists an* $(m\times k)$*-matrix* $\underset{\sim}{g} = (g_{ij})$ *of non-negative Lebesgue-square integrable functions on* $[0,1]$ *such that*

$$(2.3)\qquad \sum_{j=1}^{k} \int_0^t H_{ij}^{(n)}(s)H_{\ell j}^{(n)}(s)\Lambda_j^{(n)}(s)ds \underset{n\to\infty}{\overset{P}{\to}} \sum_{j=1}^{k} \int_0^t g_{ij}(s)g_{\ell j}(s)ds$$

for all $i,\ell = 1,\ldots,m$, $t \in [0,1]$, *and*

$$(2.4)\qquad \int_0^1 \sum_{j=1}^{k} (H_{ij}^{(n)}(s))^2\ \Lambda_j^{(n)}(s)I(|H_{ij}^{(n)}(s)| > \varepsilon)ds \underset{n\to\infty}{\overset{P}{\to}} 0$$

for all $i = 1,\ldots,m$, $\varepsilon > 0$. *Then*

$$(2.5)\qquad \int \underset{\sim}{H}^{(n)}d\ \underset{\sim}{M}^{(n)} \underset{n\to\infty}{\overset{D}{\to}} \int \underset{\sim}{g}\ d\underset{\sim}{W},$$

where $\underset{\sim}{W} = (W_1,\ldots,W_k)$, *and* $W_1,\ldots,W_k$ *are independent standard Wiener processes.*

Also for all i, ℓ *and* t

$$\sum_{j=1}^{k} \int_0^t H_{ij}^{(n)}(s)H_{\ell j}^{(n)}(s)dN_j^{(n)}ds \underset{n\to\infty}{\overset{P}{\to}} \sum_{j=1}^{k} \int_0^t g_{ij}(s)g_{\ell j}(s)ds.$$

Here $\overset{D}{\to}$ denotes weak convergence in the space $D^m[0,1]$ of m-dimensional functions on $[0,1]$ with right-continuous real-valued components, equipped with the Skorohod product topology (cf. BILLINGSLEY, 1968, Chapter 3).

In practice the verification of the Lindeberg condition (2.4) is not always so direct, and it is useful to have alternative and more easily verifiable sets of conditions. Such conditions are given in the following corollaries.

COROLLARY 2.2. *Consider the situation of Theorem* 2.1 *and suppose that the following conditions are satisfied:*

$$(2.7)\qquad H_{ij}^{(n)}(s) \underset{n\to\infty}{\overset{P}{\to}} 0, \quad \textit{for all } i,j,s,$$

$$(2.8)\qquad \sum_{j=1}^{k} H_{ij}^{(n)}(s)H_{\ell j}^{(n)}(s)\Lambda_j^{(n)}(s) \underset{n\to\infty}{\overset{P}{\to}} \sum_{j=1}^{k} g_{ij}(s)g_{\ell j}(s) \quad \textit{for all } i,\ell,s,$$

and

$$(2.9) \qquad \sum_{j=1}^{k} H_{ij}^{(n)}(s)H_{\ell j}^{(n)}(s)\Lambda_j^{(n)}(s)$$

are integrable over $\{n,s,i,\ell\}$. *Then* (2.5) *and* (2.6) *hold.*

PROOF. In a similar way as AALEN & JOHANSEN (1978, Theorem 4.1) we may prove that (2.7) to (2.9) are sufficient for (2.3) and (2.4), and the corollary follows from Theorem 2.1. □

For some applications these conditions are a little too strong (they imply for instance that g_{ij} is bounded) but the following more complicated corollary can be applied:

COROLLARY 2.3. *Suppose that for each* $\tau < 1$ *the conditions of Corollary* 2.2 *hold with* s *restricted to the interval* $[0,\tau]$. *Suppose that the functions* g_{ij} *are Lebesgue square integrable on* $[0,1]$ *and that for all* i,j *and all* $\varepsilon > 0$

$$(2.10) \qquad \lim_{\tau\uparrow 1} \limsup_{n\to\infty} P\left(\int_\tau^1 [H_{ij}^{(n)}(s)]^2 \, \Lambda_j^{(n)}(s)ds > \varepsilon\right) = 0.$$

Then (2.5) *and* (2.6) *hold.*

PROOF. It is simple to check that (2.3) and (2.4) hold by applying Theorem 4.2 of BILLINGSLEY (1968). □

A sufficient condition for uniform integrability of a set of random variables which is often easier to verify is boundedness of the $(1+\varepsilon)$'th absolute moments of the variables for some $\varepsilon > 0$. Alternative sets of conditions in which the uniform integrability in (2.9) is replaced with uniform convergence in (2.7) and (2.8) were given by GILL (1980).

It should be noted that Corollary 2.2 can be applied mutatis mutandi to proving weak convergence on any finite time interval $[0,\tau]$, $\tau < \infty$. Theorem 2.1 and Corollary 2.3 are in fact also valid on the infinite time interval $[0,\infty]$.

2C. The multiplicative intensity model

We consider a statistical model $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$, that is, a family of probabilities on $(\Omega,\mathcal{F})$. Following AALEN (1978), it is assumed that for each $\theta \in \Theta$, the $((\mathcal{F}_t),P_\theta)$-intensity process $\underset{\sim}{\Lambda}^\theta$ exists and that there furthermore exists an $(\mathcal{F}_t)$-adapted stochastic process $\underset{\sim}{Y} = (Y_1,\ldots,Y_k)$ and functions $\underset{\sim}{\alpha}^\theta = (\alpha_1^\theta,\ldots,\alpha_k^\theta)$ such that

$$(2.11)\qquad \Lambda_i^\theta(t) = \alpha_i^\theta(t)Y_i(t), \quad i = 1,\ldots,k,\ t \in [0,1],$$

the point being that $\underset{\sim}{\alpha}^\theta$ is deterministic while $\underset{\sim}{Y}$ does not depend on θ. For each i it is required that the sample paths of Y_i are non-negative and left continuous with right-hand limits; α_i^θ is also non-negative.

Statistical inference in the multiplicative intensity model at time t is based on observing $(\underset{\sim}{N}(s),\underset{\sim}{Y}(s))$, $0 \le s \le t \le 1$, or more generally on observing the family $(\mathcal{F}_s, 0\le s\le t)$. We restrict here attention to inference concerning $\underset{\sim}{\alpha}^\theta$ (and drop the superscript θ) and refer to REBOLLEDO (1978, II.2) for a discussion of the more general problems of classifying statistical models with the structure (2.11), and of finding conditions for the α_i to parametrize the model.

Several examples of the multiplicative intensity model were given by AALEN (1975,1976,1978); we return below to some of these in connection with the discussion of the specific results. AALEN (1978) approached the estimation problem in the following way (we have slightly improved some of his arguments, in particular to make use of later developments in the theory of stochastic integrals).

Since for each $i = 1,\ldots,k$ we may rewrite (2.1) symbolically as "$dN_i(t) = \alpha_i(t)Y_i(t)dt$ + noise" a natural estimator of $\beta_i(t) = \int_0^t \alpha_i(s)ds$ would be $\int_0^t Y_i(s)^{-1}dN_i(s)$.

However, one may have $Y_i = 0$, and in order to deal systematically with this possibility the problem is rephrased by defining

$$(2.12)\qquad \beta_i^*(t) = \int_0^t \alpha_i(s)J_i(s)ds, \quad i = 1,\ldots,k, \quad t \in [0,1],$$

where

(2.13) $\quad J_i(s) = I(Y_i(s) > 0).$

Interpreting $J_i(t)/Y_i(t)$ as 0 whenever $Y_i(t) = 0$ and assuming that there exists a constant $c > 0$ such that $Y_i(t) < c \Rightarrow Y_i(t) = 0$ almost surely, estimators of β_i^* are defined by

$$(2.14) \qquad \hat{\beta}_i(t) = \int_0^t J_i(s)/Y_i(s)dN_i(s), \qquad i = 1,\dots,k,\ t \in [0,1].$$

It now holds that $\hat{\beta}_i - \beta_i^*$, $i = 1,\dots,k$, are orthogonal square integrable martingales with variance processes

$$(2.15) \qquad \langle\hat{\beta}_i-\beta_i^*,\hat{\beta}_i-\beta_i^*\rangle(t) = \int_0^t \alpha_i(s)J_i(s)/Y_i(s)ds, \qquad i = 1,\dots,k.$$

The asymptotic behaviour of the estimators and of tests for comparing α_i and α_j for $i \neq j$ may then be studied by means of the results of AALEN (1977).

2D. Censoring

One important advantage of the rather general formulation of the multiplicative intensity model is that it accomodates fairly general censoring patterns.

Thus assume that the time intervals where the process is observed (or at least, where its jumps are observed) are determined by an $(\mathcal{F}_t)$-adapted multivariate *indicator process* $\underset{\sim}{C} = (C_1,\dots,C_k)$ with values in $\{0,1\}^k$. (Formally, let $E_i \subseteq [0,1] \times \Omega$ be the set of (t,ω), where $N_i(t,\omega)$ and $Y_i(t,\omega)$ may be observed; then $C_i(t,\omega) = I_{E_i}(t,\omega)$.) If $\underset{\sim}{C}$ is predictable (for our purposes, left continuous), then the censored process

$$N_i^c(t) = \int_0^t C_i(s)dN_i(s)$$

has intensity process

$$\alpha_i(t)Y_i^c(t) = \alpha_i(t)C_i(t)Y_i(t).$$

This is seen by noting that

$$N_i^c(t) = \int_0^t C_i(s)\alpha_i(s)Y_i(s)ds + \int_0^t C_i(s)dM_i(s),$$

where the last term is a square integrable martingale, being a stochastic integral of a predictable process with respect to the square integrable martingale M_i given by (2.1) (cf. Appendix I).

A study of "censored observations from $\underset{\sim}{N}$" is thus equivalent to a study of the multivariate counting process $\underset{\sim}{N}^c$ with multiplicative intensity $(\alpha_i Y_i^c,\ i = 1,\ldots,k)$.

Two general remarks on censoring are important. The first is that we do *not* restrict attention to censoring processes adapted to the self-exciting family of σ-algebras (N_t) given by $N_t = \sigma\{N_i(u),\ u \le t,\ i = 1,\ldots,k\}$. This allows for the dependence of the censoring mechanism on outside random influences in addition to events in the counting process "before now". A particularly simple example is the so-called random censorship model, where the censoring process is stochastically independent of $\underset{\sim}{N}$.

The second remark concerns a fairly common situation where the index i of the multivariate counting process refers to a numbering $i = 1,\ldots,k$ of individuals whose intensity functions α_i are assumed equal (say = α), but where the censoring mechanism operates on each individual. That is, we have

$$dN_i(t) = \alpha(t)Y_i(t)dt + dM_i(t)$$

and, after censoring,

$$dN_i^c(t) = \alpha(t)Y_i(t)C_i(t)dt + dM_i^c(t).$$

In such situations it is quite common to restrict attention to the *aggregate* process $\bar{N} = \sum N_i$, that is, after censoring to the counting process $\sum N_i^c$. This, however, is seen to have intensity process given by

$$\alpha(t)\sum Y_i(t)C_i(t)$$

and is thus again covered by our framework. We return to specific examples of this and similar more complicated situations below. The sum $\sum Y_i(t)C_i(t)$ may often be interpreted as a *number at risk* at time t (for transition with the intensity α(t)).

3. THE k-SAMPLE TEST

3A. The general test statistic

Consider a k-dimensional ($k \geq 2$) counting process $\underset{\sim}{N} = ((N_1(t),\ldots,N_k(t)),$ $t \in [0,1])$ with intensity process $\underset{\sim}{\Lambda} = ((\Lambda_1(t),\ldots,\Lambda_k(t)),$ $t \in [0,1])$ satisfying the multiplicative intensity model (2.11), i.e. $\Lambda_i(t) = \alpha_i(t)Y_i(t)$, $i = 1,\ldots,k$, where the α_i's and Y_i's satisfy the regularity conditions given in Section 2C.

We want to test the hypothesis

$$H_0\colon \alpha_1(t) = \ldots = \alpha_k(t), \quad t \in [0,1].$$

The common value of $\alpha_1(t),\ldots,\alpha_k(t)$ will be denoted $\alpha(t)$. It is natural to construct a test statistic by comparing estimates of $\beta_i(t) = \int_0^t \alpha_i(s)ds$, $i = 1,\ldots,k$, with an estimate of the hypothesized common value $\int_0^t \alpha(s)ds$. Actually only β_i^* (see (2.12)) can be estimated so we have to modify this idea slightly, as follows.

Proceeding as in Section 2C we consider under the null hypothesis the aggregate processes

$$\bar{N} = \sum_{i=1}^{k} N_i, \quad \bar{Y} = \sum_{i=1}^{k} Y_i$$

and define $\bar{J}$ by $\bar{J}(t) = I\{\bar{Y}(t) > 0\}$. Instead of β we may estimate β^* defined by

$$\beta^*(t) = \int_0^t \alpha(s)\bar{J}(s)ds$$

by the estimator

$$\hat{\beta}(t) = \int_0^t \bar{J}(s)/\bar{Y}(s)d\bar{N}(s).$$

Let then $\hat{\beta}_i(t)$, $i = 1,\ldots,k$, be given by (2.14) and define

$$\bar{\beta}_i(t) = \int_0^t J_i(s)d\hat{\beta}(s) = \int_0^t J_i(s)/\bar{Y}(s)d\bar{N}(s), \quad i = 1,\ldots,k,$$

where $J_i(s) = I\{Y_i(s) > 0\}$ as defined by (2.13). When the α_i's are identical, $\bar{N}$ is a counting process with intensity process $\bar{\Lambda} = \alpha\bar{Y}$ (and $\bar{\beta}_i$ is a sensible estimator of β_i^*). Hence by (2.1) we have under H_0 that

$$\hat{\beta}_i(t)-\bar{\beta}_i(t) = \int_0^t J_i(s)/Y_i(s)dM_i(s) - \int_0^t J_i(s)/\bar{Y}(s)d\bar{M}(s), \quad i=1,\dots,k,$$

where $\bar{M} = \sum_{i=1}^k M_i$, are square integrable martingales.

Their *covariance processes* are given by

$$(3.1) \qquad \langle\hat{\beta}_i-\bar{\beta}_i,\hat{\beta}_j-\bar{\beta}_j\rangle(t) = \int_0^t J_i(s)J_j(s)\alpha(s)\left(\frac{\delta_{ij}}{Y_i(s)} - \frac{1}{\bar{Y}(s)}\right)ds,$$

where δ_{ij} is the Kronecker delta. To prove (3.1) we first note that under H_0

$$\hat{\beta}_i(t) - \bar{\beta}_i(t) = \sum_{j=1}^k \int_0^t J_i(s)\left(\frac{\delta_{ij}}{Y_i(s)} - \frac{1}{\bar{Y}(s)}\right)dM_j(s);$$

(3.1) follows by some straightforward calculations using (A.1) in Appendix I, (2.2) and the fact that M_i and M_j are *orthogonal* when $i \neq j$.

To construct a test statistic for H_0 we introduce the integrals

$$Z_i(t) = \int_0^t K_i(s)(d\hat{\beta}_i(s) - d\bar{\beta}_i(s)), \qquad i = 1,\dots,k,$$

where the K_i's are *almost surely bounded*, predictable processes. Under H_0 $Z_1,\dots,Z_k$ are square integrable martingales (cf. Appendix I).

Furthermore by (A.1) and (3.1) the covariance processes are given by

$$(3.2) \qquad \langle Z_i,Z_j\rangle(t) = \int_0^t K_i(s)K_j(s)J_i(s)J_j(s)\alpha(s)\left(\frac{\delta_{ij}}{Y_i(s)} - \frac{1}{\bar{Y}(s)}\right)ds.$$

In the following we derive the asymptotic distribution under H_0 of $\underset{\sim}{Z}^{(n)} = (Z_1^{(n)},\dots,Z_k^{(n)})$ assuming that we have a sequence $\underset{\sim}{N}^{(n)}$ of k-dimensional counting processes with intensity process $\underset{\sim}{\Lambda}^{(n)} = \alpha \cdot \underset{\sim}{Y}^{(n)}$, where α is the same for all n.

This is straightforward since $\underset{\sim}{Z}^{(n)}$ can be expressed as $Z^{(n)} = \int \underset{\sim}{H}^{(n)} d\underset{\sim}{M}^{(n)}$ with

$$(3.3)\qquad H_{ij}^{(n)} = K_i^{(n)} J_i^{(n)} \left(\frac{\delta_{ij}}{Y_i^{(n)}} - \frac{1}{\bar{Y}^{(n)}}\right), \qquad i,j=1,\dots,k$$

and

$$M_j^{(n)}(t) = N_j^{(n)}(t) - \int_0^t \alpha(s) Y_j^{(n)}(s)\,ds, \qquad j=1,\dots,k.$$

Thus Theorem 2.1 and Corollary 2.2 or 2.3 give sufficient conditions for the process $\underset{\sim}{Z}^{(n)}$ (properly normalized) to converge in distribution to a Gaussian process.

We will not, however, state these conditions explicitly except for a special case which seems to cover most examples of interest, namely the case where the $K_i^{(n)}$-processes have the form

$$(3.4)\qquad K_i^{(n)}(t) = Y_i^{(n)}(t) L^{(n)}(t), \qquad i = 1,\dots,k,$$

where $L^{(n)}$ is a predictable process that only depends on the process $(\bar{N}^{(n)}, \bar{Y}^{(n)})$. It is assumed that $L^{(n)}$ is zero where $\bar{Y}^{(n)}$ is zero, and we interpret $L^{(n)}/\bar{Y}^{(n)}$ as zero where $\bar{Y}^{(n)}$ is zero. (It should be noted that our $L^{(n)}(t)$ has a slightly different meaning from the L-function defined by AALEN (1978), Section 7.3).) When $K_i^{(n)}$ has the form (3.4) the matrix $\underset{\sim}{H}^{(n)}$ with elements given by (3.3) is singular and $\sum_{i=1}^k Z_i^{(n)} = 0$. Furthermore, since $Y_i^{(n)} J_i^{(n)} = Y_i^{(n)}$ the factor $J_i^{(n)}$ may be omitted from the expression (3.3), i.e.

$$Z_i^{(n)}(t) = \int_0^t L^{(n)}(s)\,dN_i^{(n)}(s) - \int_0^t L^{(n)}(s) \frac{Y_i^{(n)}(s)}{\bar{Y}^{(n)}(s)}\,d\bar{N}^{(n)}(s),$$

$i = 1,\dots,k.$

Theorem 3.1 below gives sufficient conditions based on Corollaries 2.2 and 2.3 for the weak convergence to take place and provides consistent estimators of the covariance function of the limiting Gaussian process.

<u>THEOREM 3.1</u>. *Assume that there exists a sequence of positive constants* (a_n) *and non-negative Lebesgue-square integrable functions* $h_1,\dots,h_k$ *on* $[0,1]$ *such that*

$$(3.5)\qquad a_n L^{(n)}(s) Y_i^{(n)}(s)/\bar{Y}^{(n)}(s) \underset{n\to\infty}{\overset{P}{\to}} 0 \quad \textit{for all } i,s,$$

$$(3.6)\qquad a_n^2 [L^{(n)}(s)]^2 Y_i^{(n)}(s) Y_\ell^{(n)}(s)\alpha(s)/\bar{Y}^{(n)}(s) \overset{P}{\to} h_i(s)h_\ell(s)$$

for all i,ℓ,s

and

$$(3.7)\qquad a_n^2 [L^{(n)}(s)]^2\ Y_i^{(n)}(s) Y_\ell^{(n)}(s)\alpha(s)/\bar{Y}^{(n)}(s)$$

are uniformly integrable over $\{n,s,i,\ell\}$.

Then under H_0

$$a_n(Z_1^{(n)},\ldots,Z_k^{(n)}) \underset{n\to\infty}{\overset{D}{\to}} (U_1,\ldots,U_k)$$

where

$$U_i(t) = \int_0^t (h_i(s)\bar{h}(s))^{\frac{1}{2}} dW_i(s) - \sum_{j=1}^{k} \int_0^t h_i(s)(h_j(s)/\bar{h}(s))^{\frac{1}{2}} dW_j(s),$$

$W_1,\ldots,W_k$ *are independent standard Wiener processes and* $\bar{h} = \sum_{i=1}^k h_i$.

If (3.5) *to* (3.7) *only hold for* $s \in [0,\tau]$ *for each* $\tau < 1$ *but also for all* i *and all* $\varepsilon > 0$ *we have*

$$(3.8)\qquad \lim_{\tau\uparrow 1} \limsup_{n\to\infty} P\left(\int_\tau^1 (a_n L^{(n)}(s) Y_i^{(n)}(s))^2 \alpha(s)/\bar{Y}^{(n)}(s)ds > \varepsilon\right) = 0$$

then the result still holds.

Under either set of conditions it also follows that

$$(3.9)\quad a_n^2 \int_0^t [L^{(n)}(s)]^2 Y_i^{(n)}(s) Y_\ell^{(n)}(s)/(\bar{Y}^{(n)}(s))^2 d\bar{N}^{(n)}(s) \underset{n\to\infty}{\overset{P}{\to}} \int_0^t h_i(s)h_\ell(s)ds$$

for all i,ℓ *and* t.

PROOF. By letting $g_{ij} = \delta_{ij}(h_i\bar{h})^{\frac{1}{2}} - h_i(h_j/\bar{h})^{\frac{1}{2}}$, (2.7) to (2.9) are easily verified from (3.5) to (3.7), while (3.8) implies (2.10).

Hence the first parts of the theorem follow from Corollary 2.2 and Corollary 2.3. The last part of the theorem now follows from (2.6) with $g_{ij} = h_i(h_j/\bar{h})^{\frac{1}{2}}$ and $\underset{\sim}{H}^{(n)}$ similarly modified. □

From Theorem 3.1 we find that under H_0, the statistic $\underset{\sim}{Z}^{(n)}(1) = (Z_1^{(n)}(1),\ldots,Z_k^{(n)}(1))$ is asymptotically normally distributed with mean $\underset{\sim}{0}$ and a (singular) covariance matrix $a_n^{-2}\underset{\sim}{\Sigma}$, where $\underset{\sim}{\Sigma}$ has elements

$$(3.10)\qquad \sigma_{ij} = \langle U_i,U_j\rangle(1) = \int_0^1 h_i(s)(\delta_{ij}\bar{h}(s) - h_j(s))ds.$$

A consistent estimator for $\underset{\sim}{\Sigma}$ is

$$a_n^2\underset{\sim}{V}^{(n)}(1) = a_n^2(V_{i\ell}^{(n)}(1))_{i,\ell=1}^k,$$

where

$$V_{i\ell}^{(n)}(t) = \int_0^t [L^{(n)}(s)]^2 \frac{Y_i^{(n)}(s)}{\bar{Y}^{(n)}(s)}(\delta_{i\ell} - \frac{Y_\ell^{(n)}(s)}{\bar{Y}^{(n)}(s)})d\bar{N}^{(n)}(s).$$

In Appendix II it is shown that under mild conditions on the h_i's $\underset{\sim}{\Sigma}$ has rank k-1, which is obviously its maximum rank. Likewise $\underset{\sim}{V}^{(n)}(1)$ has rank k-1 provided that for any i,ℓ there exists a time point where $\bar{N}^{(n)}$ jumps and such that $L^{(n)}$, $Y_i^{(n)}$ and $Y_\ell^{(n)}$ are positive. Since in any case $\underset{\sim}{V}^{(n)}(1)$ has rank $\le$ k-1 and is a consistent estimator of $\underset{\sim}{\Sigma}$, we have $P(\text{rank } \underset{\sim}{V}^{(n)}(1) = k-1) = \underset{n\to\infty}{\to} 1$. Hence under H_0 the statistic

$$(3.11)\qquad (X^{(n)})^2 = \underset{\sim}{Z}^{(n)}(1)^T\underset{\sim}{V}^{(n)}(1)^-\underset{\sim}{Z}^{(n)}(1),$$

where $\underset{\sim}{V}^{(n)-}$ is a generalized inverse, is asymptotically chi-squared distributed with k-1 degrees of freedom.

3B. Conservative approximations to the test statistics.

An important special case of the structure of the assumptions of

Theorem 3.1 occurs when

$$(3.12) \qquad h_i(s) = p_i\bar{h}(s), \qquad s \in [0,1]$$

for some positive constants p_i which sum to one and for some nonnegative Lebesgue-square integrable function $\bar{h}$.

It is easily seen that together with condition (3.6), (3.12) implies that

$$(3.13) \qquad Y_i^{(n)}(s)/\bar{Y}^{(n)}(s) \overset{P}{\to} p_i \quad \text{as } n \to \infty,$$

for each i and s such that $\bar{h}(s) > 0$. Conversely, if (3.6) holds and (3.13) holds for each i and s except possibly on a fixed (i.e. independent of n) set of s where $\bar{Y}^{(n)}(s) \equiv 0$, then (3.12) holds whatever the choice of L.

When (3.12) holds, some modifications of the test statistic (3.11) are possible. For in this situation the elements of the limiting covariance matrix reduce to

$$\sigma_{ij} = p_i(\delta_{ij} - p_j) \int_0^1 \bar{h}^2(s)ds;$$

that is, a constant times the covariance matrix of a multinomial random variable. By (3.9) $\int_0^1 \bar{h}^2(s)ds$ may be estimated consistently by $a_n^2 \int_0^1 [L^{(n)}(s)]^2 d\bar{N}^{(n)}(s)$. If we furthermore assume the existence of constants or random variables $\hat{p}_i^{(n)}$ such that $\hat{p}_i^{(n)} \overset{P}{\to} p_i$ as $n \to \infty$, then an alternative version of the statistic (3.11) is

$$(3.14) \qquad \left\{ \sum_{i=1}^{k} \frac{[Z_i^{(n)}(1)]^2}{\hat{p}_i^{(n)}} \right\}\left\{ \int_0^1 [L^{(n)}(s)]^2 d\bar{N}^{(n)}(s)\right\}^{-1}.$$

By (3.9) one may also estimate $p_i \int_0^1 \bar{h}^2(s)ds$ consistently by $a_n^2 \int_0^1 [L^{(n)}(s)]^2\, Y_i^{(n)}(s)/\bar{Y}^{(n)}(s)d\bar{N}(s)$. Hence, still another version of the test statistic is

$$(3.15) \qquad \sum_{i=1}^{k} \left\{[Z_i^{(n)}(1)]^2 \Big/ \int_0^1 [L^{(n)}(s)]^2\, Y_i^{(n)}(s)/\bar{Y}^{(n)}(s)d\bar{N}^{(n)}(s)\right\}.$$

This is actually (3.14) with the special choice of $\hat{p}_i(n)$

$$\hat{p}_i^{(n)} = \int_0^1 [L^{(n)}(s)]^2 Y_i^{(n)}(s)/\bar{Y}^{(n)}(s)d\bar{N}^{(n)}(s) \Big/ \int_0^1 [L^{(n)}(s)]^2 d\bar{N}^{(n)}(s).$$

By direct computations similar to those of CROWLEY & BRESLOW (1975) it may be verified that the difference between the covariance estimators used in (3.15) and (3.11) (and similarly between their asymptotic counterparts) is always positive semi-definite. Thus (3.15) always takes on smaller values than (3.11), while (3.11) always has the correct asymptotic distribution. A common terminology is to state that (3.15) is *conservative* except in the case where (3.13) holds (for all i and s except possibly on a fixed set of s where $\bar{Y}^{(n)}(s) \equiv 0$).

Nothing can be shown in general about the relationship between (3.11) and (3.14). One special case is that if $k = 2$ and $\hat{p}_i^{(n)} = \frac{1}{2}$ we have again a positive semi-definite difference. Thus with $k = 2$ and $\hat{p}_i^{(n)} = \frac{1}{2}$ (3.14) is also conservative except when (3.13) holds.

3C. Classical nonparametric k-sample tests.

As stated in the Introduction, the bulk of the literature on non-parametric tests for censored data concerns the classical k-sample situation where the object is to test the hypothesis $F_1 = \dots = F_k$ in the set-up with X_{ij}; $j=1,2,\dots,n_i$; $i=1,2,\dots,k$ (with $n=\Sigma\, n_i$) independent random variables with absolutely continuous distribution function F_i, density function f_i and hazard function α_i in group no. i. As an introduction to our more detailed exposition of tests for censored data in this situation, we briefly indicate how the theory for the uncensored case, including the asymptotic distribution results, fits into our framework.

Let $N_{ij}(t) = I\{X_{ij} \le t\}$. AALEN (1978), p.707) pointed out that

$$(3.16) \qquad N = (N_1,\dots,N_k) \text{ where } N_i(t) = \sum_{j=1}^{n_i} N_{ij}(t)$$

is a k-dimensional counting process with intensity process satisfying the multiplicative intensity model (2.11) with $Y_i(t) = n_i - N_i(t-)$. We want to test the hypothesis of identical α_i's.

Our theory only strictly applies if all observations with probability 1 fall into a finite time interval, say [0,1]. However, if $P[X_{ij} > t] > 0$ for all $t < \infty$, we can still apply our theory by first mapping $[0,\infty]$ continuously onto [0,1], e.g. by the transformation $x \to \frac{1}{1+x}$. In the same way, if the F_i's are continuous but not absolutely continuous we can still obtain this situation (under the null hypothesis at least) by, for example, the transformation $x \to F(x)$. All rank statistics are unaltered by such monotone time transformations. Alternatively, but we do not follow that course here, Aalen's theory can be extended to apply on $[0,\infty]$ with arbitrary F_i's, not even necessarily continuous (see Gill, 1980).

Returning to the situation at hand, let us first rewrite the expression for $Z_i(1)$ when K_i is given by (3.4), where we assume that $L(t) = L_0(\bar{Y}(t))$ for some fixed function L_0 (i.e. the same for all n). Let us also assume that F, the common distribution function of the X_{ij}'s, satisfies $F(1) = 1$. Since then $Y_i(t) = \int_{[t,1]} dN_i(u)$, we may write after a change of the order of integration

$$Z_i(1) = \int_0^1 L_0(\bar{Y}(s))dN_i(s) - \int_0^1 \int_{(0,u]} L_0(\bar{Y}(s))/\bar{Y}(s)d\bar{N}(s)dN_i(u).$$

If, furthermore, R_{ij} denotes the rank of X_{ij} in the combined sample, then $\bar{Y}(X_{ij}) = n + 1 - R_{ij}$, and it is seen that

$$Z_i(1) = \sum_{j=1}^{n_i} L_0'(n + 1 - R_{ij})$$

is a linear rank statistic with score function L_0' given by

$$L_0'(y) = L_0(y) - \sum_{v=y}^{n} L_0(v)/v.$$

Especially for $L_0(y) = y$ and $L_0(y) = 1$ one gets

$$Z_i(1) = n_i(n + 1) - 2 \sum_{j=1}^{n_i} R_{ij}$$

and

$$Z_i(1) = n_i - \sum_{j=1}^{n_i} \sum_{v=n+1-R_{ij}}^{n} 1/v,$$

respectively, corresponding to rank statistics of the Wilcoxon and Savage type (HÀJEK & ŠIDÁK, 1967, p.87 and p.97). For linear rank tests the distribution of the test statistic under the hypothesis is independent of the null distribution F, and hence the exact so-called permutation covariance matrix of $\underset{\sim}{Z}(1)$ may be found by a combinatorial argument. For our general statistic we have to be content with an *estimate* of the covariance matrix. General comments on the relation of the present approach to the permutation distribution are collected in Section 3E below. Let us as an example consider the situation $L_0(y) = y$ more specifically. One may show that the conditions in Theorem 3.1 in this case are fulfilled with $a_n = n^{-3/2}$ and $h_i(s) = p_i(1 - F(s))^{3/2}\alpha^{\frac{1}{2}}(s)$ when we assume that $\hat{p}_i = n_i/n \to p_i$ as $n \to \infty$ (see Appendix III). Hence, (3.12) holds and the version (3.14) of the test statistic applies. Since $\int_0^1 \bar{Y}^2(s)d\bar{N}(s) = \sum_{i=1}^{n} i^2 = n(n+1)(2n+1)/6$, it takes the form

$$\frac{12}{(n+\frac{1}{2})(n+1)} \sum_{i=1}^{k} \frac{1}{n_i} [n_i \frac{n+1}{2} - \sum_{j=1}^{n_i} R_{ij}]^2.$$

Except for a factor of $(n+\frac{1}{2})/n$ (which will be explained in Section 3E below) this equals the Kruskal-Wallis test, which is the k-sample generalization of the Wilcoxon test (HÀJEK & ŠIDÁK, 1967, p.104). The case $L_0(y) = 1$ may be analysed in a similar manner.

3D. Nonparametric k-sample tests for censored survival data. Asymptotically equal censorship.

Censoring introduces the modification of the situation discussed in the previous section that instead of observing the random variables X_{ij}; $j=1,\dots,n_i$, $i=1,\dots,k$, whose distributions are to be compared we observe only $(\min(X_{ij}, T_{ij}), I\{X_{ij} \le T_{ij}\})$ where several different assumptions may be made regarding how the T_{ij}'s are generated. We list below some such assumptions.

a. <u>Simple type I censorship</u>.

In each sample, observation is discontinued at a predetermined time t_i. That is, T_{ij} is degenerate at t_i, for all i and j.

b. <u>Progressive type I censorship (or: fixed censorship)</u>.

It is quite common in clinical trials for patients to enter consecutively, but to let observation stop at some fixed calendar time. The time variable of interest in this connection is usually time since entry, and the situation is then formalized by letting one deterministic censoring time, $T_{ij} = t_{ij}$, be attached to each patient. Usually entry times are random, but the above formalization is then justified by conditioning on the entry times and assuming independence of entry and survival distributions.

c. <u>Simple type II censorship</u>.

Observation in the i'th sample is discontinued at the r_i'th observed failure. That is $T_{ij} = X_{i(r_i)}$, the r_i'th smallest observation in sample i. For a particular industrial variant of the procedure, "Progressive type II Censorship", see e.g. GILL (1980, p.24).

The situation may be immediately generalized to T_{ij} being a *stopping time* on the process.

d. <u>Random censorship; competing risks</u>.

If all T_{ij} are random variables which are independent of one another and of the X_{ij}'s we are in the *random censorship* situation. One will usually assume T_{ij}, $j = 1,\ldots,n_i$ to be identically distributed (with distribution function G_i, say).

A similar situation will often arise when T_{ij} is the time to a competing cause of failure. It is then not always so clear that the assumption that the T_{ij} are independent of the X_{ij} is warranted. Recognition of this fact and of the ensuing difficulties of interpretation and identifiability have been made several times in the literature. We refer to PRENTICE et al. (1978) for a survey.

e. Testing with replacement.

GILL (1978; 1980, p.25) mentioned that the present theory *does not* apply for testing with replacement and gave alternative tools for handling this situation.

As described in general terms in Section 2D, the k-dimensional counting process (3.16) is modified by defining (left-continuous!) censoring processes

$$C_{ij}(t) = I\{T_{ij} \geq t\}$$

and the censored counting process in the i'th sample

$$N_i^{\,c}(t) = \sum_{j=1}^{n_i} \int_0^t C_{ij}(u)dN_{ij}(u) = \#\{X_{ij},\ j=1,\ldots,n_i,\ X_{ij} \leq T_{ij},\ X_{ij} \leq t\},$$

the number of uncensored observations ("observed failures") before t in group i. The k-dimensional counting process $N^c = (N_1^{\,c},\ldots,N_k^{\,c})$ has intensity process given by the i'th component $\alpha_i(t)Y_i^{\,c}(t)$ with

$$Y_i^{\,c}(t) = \sum_{j=1}^{n_i} C_{ij}(t)I\{X_{ij} \geq t\} = \#\{X_{ij},\ j=1,\ldots,n_i,\ X_{ij} \geq t,\ T_{ij} \geq t\}$$

the number still at risk at time t- in group i. Again α_i is the hazard function in group i. AALEN (1978) and GILL (1980) showed that the multiplicative intensity model holds under all the usual models for censoring, including those indicated above.

The various test statistics suggested in the literature may now be obtained by suitable choice of the process L and the asymptotic distribution theory then follows from Theorem 3.1. In the rest of this Section we omit the superscript c on $N_i^{\,c}$, $Y_i^{\,c}$, etc. for notational ease.

BRESLOW (1970, Section 3) obtained a generalization of the Kruskal-Wallis test ($L = \bar{Y}$) and appealed to routine asymptotic theory of U-statistics for proof of the asymptotic distribution of the test statistic under the model of random censorship (example d. above). PETO & PETO (1972, Section 9) briefly indicated how the log-rank test ($L = 1$) for two samples may be generalized to k samples. CROWLEY & THOMAS (1975, Section 4) provided a careful large-sample study of this k-sample log-rank rest, using classical Chernoff-Savage Theory (CHERNOFF & SAVAGE (1958)). TARONE & WARE (1977) suggested a family of test statistics generated by a certain function g, which is nothing but our L. These authors gave no formal asymptotic theory but indicated a comparison of the virtues of the different test statistics. For the two-sample case this was also done in an interesting case study by PRENTICE & MAREK (1979) and in a study of asymptotic properties by SCHOENFELD (1981).

If in the random censorship model (d. above) it is assumed that the same censoring distribution applies in each sample (i.e. $G_1 = \ldots = G_k$) it is easily seen that, under the null hypothesis, (3.13) holds with p_i estimated by n_i/n. Similarly, under the model of fixed censorship (a. and b. above) (3.13) holds if the discrete distributions of censoring times in each sample converge to the same distribution. Thus, under the null hypothesis (3.13) (for all i and s except possibly on a fixed set of s where $\bar{Y}(s) \equiv 0$) has the intuitive meaning that the patterns of censorship in the k samples are asymptotically the same. We take therefore (3.13) as a *definition* of *asymptotically equal censorship* in the survival analysis situation.

If the censoring patterns are asymptotically identical in all groups one may apply one of the modifications (3.14) and (3.15) of the test statistic. For $L = \bar{Y}$ and $\hat{p}_i = n_i/n$, (3.14) equals the version of the Kruskal-Wallis test considered by Breslow (1970, Section 4) for the case of identical censoring patterns. On the other hand for L = 1 (3.15) reduces to the well-known version $\Sigma_i(O_i - E_i)^2/E_i$ of the log-rank test, where we have introduced $O_i = \int_0^1 dN_i(s)$ and $E_i = \int_0^1 Y_i(s)/\bar{Y}(s)d\bar{N}(s)$. It has been shown (PETO & PIKE, 1973, CROWLEY & BRESLOW, 1975) that this version of the log-rank rest is conservative in the case of unequal censoring patterns (see also Section 3B). Further comments on simplified log-rank test statistics and inequalities between them were given by CROWLEY & THOMAS (1975, Section 4.2).

Let us finally mention that PRENTICE (1978), cf. KALBFLEISCH & PRENTICE (1980), developed a theory for linear rank statistics for tests on regression coefficients with censored data. Among other things he gave a generalization of the Kruskal-Wallis test which differs from BRESLOW'S (1970). Let us consider his test in some detail for the case where one wants to test the equality of k samples. Define then $D_i(t) = n_i - Y_i(t+)$, where n_i is the number of individuals in the i'th sample. If $B_i(t)$ counts the number of censored observations in group no. i in $[0,t]$, it is seen that $D_i(t) = N_i(t) + B_i(t)$. Furthermore, introduce $\tilde{S}(t) = \prod_{s\le t} (1 - \frac{d\bar{N}(s)}{\bar{Y}(s)+1})$ which is close to the Kaplan-Meier estimator. Then in our notation Prentice's Kruskal-Wallis statistic takes the form

$$\int_0^1 (2\tilde{S}(t) - 1)dN_i(t) + \int_0^1 (\tilde{S}(t) - 1)dB_i(t),$$

or

$$\int_0^1 \tilde{S}(t)dN_i(t) + \int_0^1 (\tilde{S}(t) - 1)dD_i(t), \qquad i=1,\ldots,k.$$

By partial integration the last integral equals $\int_0^1 Y_i(t)d\tilde{S}(t)$ or equivalently $-\int_0^1 \tilde{S}(t-)Y_i(t)/(\bar{Y}(t) + 1)d\bar{N}(t)$. Moreover, the first integral may be given as $\int_0^1 \tilde{S}(t-)\bar{Y}(t)/(\bar{Y}(t) + 1)dN_i(t)$. This follows since the only time points that give contributions to the integral are those where N_i jumps, and for such time points $\tilde{S}(t) = \tilde{S}(t-)\bar{Y}(t)/(\bar{Y}(t) + 1)$. Hence, Prentice's statistic is a special case of our general statistic with K_i given by (3.4) with $L(t) = \tilde{S}(t-)\bar{Y}(t)/(\bar{Y}(t) + 1)$ and its distributional properties follow from our Theorem 3.1. Note that for computational purposes, the statistic may be cast in the compact form

$$\int_0^1 \tilde{S}(t)dN_i(t) - \int_0^1 \tilde{S}(t)\frac{Y_i(t)}{\bar{Y}(t)} d\bar{N}(t), \qquad i=1,\ldots,k.$$

In a very recent report HARRINGTON & FLEMING (1981) introduced another class of test statistics G_ρ, having Prentice's Kruskal-Wallis generalization and the log-rank test as special cases.

Their test statistic G_ρ is obtained by taking $L(t) = (\hat{S}(t-))^\rho$ where $\hat{S}(t) = \prod_{s\le t}(1 - d\bar{N}(s)/\bar{Y}(s))$ is the usual Kaplan-Meier estimator of the survival distribution based on the combined samples. Thus G_0 is the log rank test while G_1 is very close to Prentice's Kruskal-Wallis generalization (and asymptotically equivalent to it). Harrington and Fleming only give asymptotic results in the two sample case, under random censorship, applying theorems of GILL (1980). However, they do include efficiency calculations (in this two sample case) showing that each G_ρ is appropriate for testing against a particular alternative in a continuum of alternatives ranging between the proportional hazards ($\rho = 0$) and logistic location alternatives ($\rho = 1$). How to specify ρ in practice is, however, not clear.

3E. "Exact" equal censoring and permutation distribution.

In several situations with censored survival data one has "exact" equal censorship allowing permutation tests (exact or approximate) to be used. This can arise for instance in Progressive Type I censorship (example b. above) when patients are assigned to treatment groups at random, or in random censorship (example d.) when it is known that under the null hypothesis $F_1 = \dots = F_k$, the censoring distributions $G_1,\dots,G_k$ are also all equal to one another.

In such cases one can conveniently test the null hypothesis by taking any of our test statistics $\underset{\sim}{Z}(1)$ and approximating its permutation distribution by a multivariate normal distribution with the same mean (which turns out to be zero) and the same covariance matrix. One then goes on to compute the analogue of (3.11), and refer this to the $\chi^2(k-1)$ distribution. In principle an exact test is also available, but will involve considerably more computations.

The permutation distribution referred to here is that obtained by holding the $n = \Sigma n_i$ observed values of $(\min(X_{ij}, T_{ij}), I\{X_{ij} \le T_{ij}\})$ fixed, but allocating them at random among the k groups of size $n_1,\dots,n_k$. Since we consider test statistics with K_i of the form (3.4) with L depending on $\bar{N}$ and $\bar{Y}$ only, the permutation expectation and covariance are simply calculated ($\bar{N}$ and $\bar{Y}$ and hence L, too, remain fixed under permutations). The result is a mean zero and a covariance matrix with (i,j)'th element

$$\frac{n}{n-1}\hat{p}_i(\delta_{ij} - \hat{p}_j)\int_0^1 L^2(s)\,\frac{\bar{Y}(s)-1}{\bar{Y}(s)}\,d\bar{N}(s)$$

where $\hat{p}_i = n_i/n$.

Thus we obtain, up to a factor $\frac{n}{n-1}\,\frac{\bar{Y}(s)-1}{\bar{Y}(s)}$, the version (3.14) of the test statistic. The permutation covariance matrix is slightly smaller than that used in (3.14). For instance, when there is no censoring and $L = \bar{Y}$ we obtain exactly the Kruskal Wallis test (cf. Section 3C.).

Taking the remarks in Section 3B into account, we see that this permutation test may be invalid (i.e. anticonservative) when censoring is not equal; we are guaranteed a conservative test when $k = 2$ and $\hat{p}_i = \frac{1}{2}$. PRENTICE & MAREK'S (1979) case study has strongly unequal censoring and strongly unequal sample sizes; the permutation variances are highly inflated and thus strongly conservative tests result.

Before a χ^2-type test based on the permutation distribution can be used one needs to know that $\underset{\sim}{Z}(1)$ is approximately normally distributed. It turns out that $\underset{\sim}{Z}(t)$ is still a (discrete time) martingale under this permutation distribution with an appropriate choice of $\mathcal{F}_t$ (specifying the allocation of observations to samples for those observations $\leq t$). So it seems likely that asymptotic normality can be invoked by the discrete time martingale central limit theorems of MCLEISH (1974).

3F. Applications to the analysis of discrete-state Markov processes.

As mentioned in the Introduction our original motivation for undertaking the investigation reported in the present paper was an empirical study of the possible interaction between menopause and the onset of the chronical skin disease pustulosis palmo-plantaris (AALEN et al. 1980). In that study the intensities of getting the disease when being in one of the three states M: natural menopause has occurred, I: induced menopause has occurred, 0: menopause has not occurred, were to be compared (the reader is referred to AALEN et al., 1980, for details).

In this Section we shall outline the Markov process application of the present methodology and in particular comment on the applicability of the conservative approximations to the test statistic discussed in Section 3B. We also present another empirical example of application of a Markov chain model.

Let Γ be a finite set of states and let α_{ij}, $i,j \in \Gamma$ be transition intensities satisfying the general regularity conditions from Section 2C. Let $(U(t),\ t \in [0,1])$ be a Markov process on Γ.

If $N_{ij}(t)$ is the number of direct transitions from i to j in [0,t] by U, $i \neq j$, $i,j \in \Gamma$, and if $Y_i(t) = I\{U(t-) = i\}$ specifies that U was in i just before t, then $N_{ij}(t)$, $i,j \in \Gamma$, $i \neq j$, is a p(p-1)-dimensional counting process (p being the number of points in Γ) with intensity $\alpha_{ij}(t)Y_i(t)$, $i,j \in \Gamma$, $i \neq j$.

In the simplest situation n independent identically distributed copies $U_1(t),\ldots,U_n(t)$ of such a Markov process are considered, though often it will be natural to condition on the initial states $U_1(0),\ldots,U_n(0)$. As described towards the end of Chapter 2, it will be quite common to have censoring and to restict attention to the aggregate processes, here given by (with obvious notation)

$$\bar{N}_{ij}(t) = \sum_{\nu=1}^{n} N_{ij}^{(\nu)}(t); \quad \bar{Y}_i(t) = \sum_{\nu=1}^{n} Y_i^{(\nu)}(t).$$

In the application to censored survival data, we saw in Section 3D above that it was quite often reasonable to assume (3.13); this situation was termed asymptotically equal censorship. In applications to Markov chain models, however, (3.13) will rarely be plausible. This is rather obvious when we, as in the application of AALEN et al. (1980), compare the intensities of *different* transitions in the *same* Markov chain. However, (3.13) will also often be violated under the hypothesis in the situation where we want to test the identity of the intensities of the same transition in k independent chains of the same structure. (An example of this is given below).

A sufficient condition for the validity of (3.13) in the latter situation is that all individuals are observed over exactly the same period of time, and the k Markov processes have identical initial distributions *and* the same set of transition probabilities. Without having numerical information, one may therefore expect that (3.15) will often give a strongly conservative test in applications to Markov chains.

EXAMPLE. We shall here comment on an application concerning admissions to psychiatric hospitals among women giving birth (ANDERSEN & RASMUSSEN, 1980). In that study it was investigated who among the about n = 70,000 Danish women giving birth to a child in 1975 had been admitted to a psychiatric hospital in the period ranging from 1 October 1973 to 31 December 1975 and the dates of admission and discharge respectively were registered. Moreover information on such demographic factors as age, marital status and parity (= number of children born before 1975) was available. Due to the fact that the exact date of birth was known only for the women who were actually admitted during the time span considered reliable information on admissions was only available in the time interval ranging from - 15 months = -456 days to 12 months = 366 days relative to the date of birth, and hence that interval is the relevant one to consider.

Let $Y_0^{(\nu)}(t) = 1$ if woman ν ($\nu=1,\dots,n$) is resident in a psychiatric hospital at time t- relative to the date of birth ($-\ 456$ days $\le t \le 366$ days) and let $Y_1^{(\nu)}(t) = 1$ otherwise; let $N_{10}^{(\nu)}(t)$ be the number of admissions for woman ν in the interval [- 456 days, t]. If we consider the two state Markov process model:

$$\boxed{1\text{: not admitted}} \quad \underset{\alpha_{01}(t)}{\overset{\alpha_{10}(t)}{\rightleftarrows}} \quad \boxed{0\text{: admitted}}$$

where $\alpha_{10}(t)$ and $\alpha_{01}(t)$ are the forces of transition, then $N_{10}^{(\nu)}(t)$ is a counting process with intensity process $\alpha_{10}(t)Y_1^{(\nu)}(t)$. Consider now a model where $\alpha_{10}(t)$ depends only on the parity $p = p(\nu)$ of the woman (ν) and define

$$\bar{N}_{10}^{p}(t) = \sum_{\substack{\nu=1\\ p(\nu)=p}}^{n} N_{10}^{(\nu)}(t), \quad \bar{Y}_1^{p}(t) = \sum_{\substack{\nu=1\\ p(\nu)=p}}^{n} Y_1^{(\nu)}(t)$$

where p takes the values 0, 1, 2 or 3+, (3+ denoting three or more). Then $(\bar{N}_{10}^{p}(t),\ p = 0, 1, 2, 3+)$ is a multivariate counting process, $\bar{N}_{10}^{p}(t)$ having intensity process $\alpha_{10}^{\ p}(t)\bar{Y}_1^{p}(t)$ and the hypothesis $\alpha_{10}^{\ 0} = \alpha_{10}^{\ 1} = \alpha_{10}^{\ 2} = \alpha_{10}^{\ 3+}$ that the intensity of being admitted does not depend on the parity of the woman may be tested by menas of (3.11), or rather for the ease of computation by means of the conservative approximation (3.15).

Choosing the weight function $L(t) = 1$ corresponding to the generalized log-rank test yields the highly significant χ^2-value 24.52 with 3 degrees of freedom. The same order of magnitude is obtained for the test statistics corresponding to $L(t) = \bar{Y}_1(t)$ or $L(t) = \bar{Y}_1(t))^{\frac{1}{2}}$ (cf. TARONE & WARE, 1977) namely 21.05 and 24.52 respectively. (Here $\bar{Y}_1 = \sum_p Y_1{}^p$.)

4. THE ONE-SAMPLE TEST

In this chapter we first consider the problem of testing whether the α-function of an intensity process is equal to a *known function*.

In Section 4B we show how this approach generalizes the one-sample test statistics suggested by BRESLOW (1975), HYDE (1977), HOLLANDER & PROSCHAN (1979), HARRINGTON & FLEMING (1981) and (in the earlier mentioned Markov process application) AALEN et al. (1980).

4A. A general one-sample test statistic.

Let $(N(t),\ t \in [0,1])$ be a (one-dimensional) counting process with intensity process $\alpha(t)Y(t)$. We want to test the hypothesis H_0: $\alpha = \alpha_0$, where α_0 is known. Let $J(t)$ and $\hat{\beta}(t)$ be given by (2.13) and (2.14) respectively, (where we have dropped the subscript i), and define $\hat{\beta}_0(t) = \int_0^t \alpha_0(s)J(s)ds$. Under H_0 we have that $\hat{\beta} - \hat{\beta}_0$ is a square integrable martingale with variance process

$$<\hat{\beta}-\hat{\beta}_0, \hat{\beta}-\hat{\beta}_0>(t) = \int_0^t J(s)\alpha_0(s)/Y(s)ds.$$

If K is an almost surely bounded predictable process, then under H_0

$$Z(t) = \int_0^t K(s)(d\hat{\beta}(s) - d\hat{\beta}_0(s))$$

is a square integrable martingale with

$$<Z,Z>(t) = \int_0^t K^2(s)J(s)\alpha_0(s)/Y(s)ds.$$

From Theorème 1.3.4 by REBOLLEDO (1978) we immediately derive the following Theorem 4.1, where we consider a sequence of counting processes $(N^{(n)}(t)$,

$t \in [0,1]$) as before.

THEOREM 4.1. *Assume that there exists a sequence of positive constants* (a_n) *and a Lebesgue-square integrable non-negative function* g *on* $[0,1]$ *such that*

$$(4.1) \qquad a_n^2 \int_0^t [K^{(n)}(s)]^2 J^{(n)}(s)\alpha_0(s)/Y^{(n)}(s)ds \underset{n\to\infty}{\overset{P}{\to}} \int_0^t g^2(s)ds, \quad t \in [0,1]$$

and

$$(4.2) \qquad a_n^2 \int_0^1 [K^{(n)}(s)]^2 J^{(n)}(s)\alpha_0(s)/Y^{(n)}(s) I\left(\left|\frac{a_n K^{(n)}(s)J^{(n)}(s)}{Y^{(n)}(s)}\right|>\varepsilon\right)ds \underset{n\to\infty}{\overset{P}{\to}} 0$$

Then $a_n Z^{(n)} \underset{n\to\infty}{\overset{D}{\to}} \int g dW$, *where* W *is a standard Wiener process.*

Thus we may use the asymptotically standard normally distributed statistic

$$(4.3) \qquad U^{(n)} = Z^{(n)}(1)(\langle Z^{(n)},Z^{(n)}\rangle(1))^{-\frac{1}{2}}$$

for testing H_0.

Note that in contrast to the k-sample situation, $\langle Z^{(n)},Z^{(n)}\rangle(1)$ is in this case directly observable. However, note that this variance estimator may be somewhat cumbersome to compute. If so, it could be replaced by the quantity

$$[Z^{(n)},Z^{(n)}](1) = \int_0^1 [K^{(n)}(s)]^2 J^{(n)}(s)/\{Y^{(n)}(s)\}^2 dN^{(n)}(s)$$

which is just a simple sum (the notation $[Z,Z]$ referring to general stochastic integral theory).

4B. Examples from survival studies.

In the present Section let $X_1,\dots,X_n$ be independent identically distributed life times with hazard function α. For the i'th individual we observe $X_i \wedge T_i = \min(X_i,T_i)$ and $I(X_i \le T_i)$ where $T_1,\dots,T_n$ are censoring times. For this situation BRESLOW (1975) via a maximum likelihood argument for a proportional hazard situation suggested the following statistic for testing the hypothesis H_0: $\alpha = \alpha_0$ where α_0 is a known hazard function:

$$(4.4) \qquad \frac{\left(\sum_{i=1}^{n} I(X_i \le T_i) + \sum_{i=1}^{n} \log(S_0(X_i \wedge T_i))\right)^2}{-\sum_{i=1}^{n} \log(S_0(X_i \wedge T_i))}$$

Here $S_0(t) = \exp(-\int_0^t \alpha_0(s)ds)$ is the survivorship function corresponding to the hazard function α_0. Breslow stated that (4.4), which is seen to have the form (observed-expected)^2/expected, has a limiting chi-squared distribution with one degree of freedom as $n \to \infty$.

We shall see how (4.4) is obtained as a special case of the general test statistic (4.3). Let then $K(t) = Y(t)$, where $Y(t)$ denotes the number at risk at time $t-$. When as in Section 3D $N(t)$ counts the observed failures in $[0,t]$ it is seen by partial integration that

$$-\sum_{i=1}^{n} \log(S_0(X_i \wedge T_i)) = \int_0^1 Y(s)\, \alpha_0(s)ds.$$

Since furthermore $\sum_{i=1}^{n} I(X_i \le T_i) = \int_0^1 dN(s)$ we find that (4.3) equals the square root of Breslow's test statistic (4.4). (As in Section 3A we have utilized that $YJ = Y$.) The asymptotic distribution of the statistic can be found by verifying the conditions (4.1) and (4.2) in Theorem 4.1. This is particularly easy in the model of random censorship in which $T_1,\ldots,T_n$ are independent identically distributed variables independent of the X_i's, if we also suppose that $P[T_i \le 1] = 1$ and α_0 is bounded on $[0,1]$. Then with the choice $a_n = n^{-\frac{1}{2}}$ and $g^2(s) = \alpha_0(s)P(X_1 \wedge T_1 > s)$ the conditions of AALEN (JOHANSEN (1978, Theorem 4.1), sufficient for (4.1) and (4.2), are easily verified.

HYDE (1977) considered a generalization of (4.4) to the case of *left truncation* where the i'th individual enters at risk at time $\nu_i \ge 0$. Using martingale arguments he proved that the statistic

$$(4.5) \qquad \left(\sum_{i=1}^{n} I(X_i \le T_i) - E_i\right) / \left(\sum_{i=1}^{n} E_i\right)^{\frac{1}{2}},$$

where $E_i = \log(S_0(\nu_i)) - \log(S_0(X_i \wedge T_i))$ has a limiting standard normal distribution when $n \to \infty$. Interpreting $Y(t)$ as the size of risk set at time t- (4.5) turns out to be a special case of (4.3).

The one sample limit $C = -\int_0^1 S_0(s)\, d\hat{S}(s)$ of EFRON (1975)'s two-sample test statistic considered by HOLLANDER & PROSCHAN (1979) may be written as $C = \int_0^1 S_0(s)\, \hat{S}(s-)\, d\hat{\beta}(s)$, where $\hat{S}$ denotes the Kaplan-Meier estimator obtained from the sample. Hence also this statistic turns out to be a special case of (4.3). For the HOLLANDER & PROSCHAN statistic, verifying the conditions of Theorem 4.1 is quite difficult except in the special case of random censorship with α_0 bounded on [0,1], $P[T_i < 1] < 1$ and $P[T_i \le 1] = 1$. The difficulties alluded to in a more general situation are however not specific to our approach; HOLLANDER & PROSCHAN'S proof of asymptotic normality is actually incomplete. For useful techniques for dealing with such situations, see Appendix III and GILL (1980). HARRINGTON & FLEMING (1981) recently suggested a family of one-sample test statistics, in our relation given by $K(t) = Y(t)(S_0(t))^\rho$. These authors indicated that the results of GILL (1980) could be used to derive the asymptotic properties of the resulting family of test statistics which have large power against proportional hazards alternative ($\rho = 0$) and logistic location alternatives ($\rho = 1$).

As we have seen in this section several recently suggested test statistics for the one-sample situation are special cases of (4.3) and hence their asymptotic distributions can all be found from our Theorem 4.1 (i.e. from Rebolledo's results). Of course the assumptions (4.1) and (4.2) have to be checked in each case, and this may be done as we have outlined for (4.4) above.

APPENDIX I. Square integrable martingales and stochastic integrals

Most of the results given in this Appendix are found in the short review by AALEN (1978, Section 2). For a more thorough discussion of the theory of square integrable martingales and stochastic integrals, see the references in Aalen's paper and MEYER (1976).

A stochastic process $M = (M(t),\ t \in [0,1])$, *adapted* to (F_t) i.e. $M(t)$ is F_t-measurable for each $t \in [0,1]$) satisfying $M(0) = 0$ and having right-continuous sample functions with left-hand limits is called a *square integrable martingale* if

$$E(M(t)\,|\,F_s) = M(s) \text{ for all } 0 \le s < t \le 1 \quad \text{and}$$

$$\sup_{t\in[0,1]} E[M(t)]^2 < \infty.$$

Let M be a square integrable martingale. Then M^2 is a right-continuous non-negative submartingale, and by the Doob-Meyer decomposition theorem there exists a unique natural (or predictable) increasing process $\langle M,M\rangle$, called the *variance process* of M, such that $M^2 - \langle M,M\rangle$ is a martingale. Furthermore, if M_1 and M_2 are square integrable martingales, then the *covariance process* $\langle M_1,M_2\rangle$ is defined by

$$\langle M_1,M_2\rangle = (\langle M_1+M_2,\ M_1+M_2\rangle - \langle M_1-M_2,\ M_1-M_2\rangle)/4,$$

and we say that M_1 and M_2 are *orthogonal* whenever $\langle M_1,M_2\rangle = 0$.

Let $H = (H(t),\ t \in [0,1])$ be a *predictable* process. We will not give the precise definition of this concept here. For our purpose it is enough to note that any left-continuous adapted process is predictable. Moreover, let M be a square integrable martingale. If the Stieltjes integral $\int H^2 d\langle M,M\rangle$ satisfies the condition

$$E\int_0^1 H^2(s)\,d\langle M,M\rangle(s) < \infty,$$

then the *stochastic integral*

$$\int HdM = ((\int HdM)(t), \qquad t \in [0,1])$$

is well defined, and it is itself a square integrable martingale. If furthermore the Stieltjes integral $\int_0^1 |H(s)||dM(s)|$ almost surely exists, then the stochastic integral $(\int HdM)$ (t) coincides with the corresponding Stieltjes integral $\int_0^t H(s)dM(s)$.

Finally we shall mention an important result valid for predictable processes H_i and K_j and square integrable martingales M_i, namely the formula

$$\text{(A.1)} \qquad <\sum_i \int H_i dM_i, \sum_j \int K_j dM_j> = \sum_i \sum_j \int H_i K_j d <M_i, M_j>.$$

From (A.1) it follows especially that $\int H_i dM_i$ and $\int K_j dM_j$ are orthogonal when M_i and M_j are orthogonal.

APPENDIX II. Rank of asymptotic covariance matrix.

The asymptotic covariance matrix $\underset{\sim}{\Sigma}$ of $a_n(Z_1^{(n)},\dots,Z_k^{(n)})$ has elements σ_{ij} given by (3.10).

To prove that $\underset{\sim}{\Sigma}$ has rank k-1 we use the technique of Breslow (1980, Appendix 5). It is enough to show that

$$\underset{\sim}{X}^T \underset{\sim}{\Sigma} \underset{\sim}{X} > 0$$

for any vector $\underset{\sim}{X}^T = (X_1,\dots,X_k)$ where not all the components X_i are equal. We will apply the elementary inequality

$$\text{(A.2)} \qquad (\Sigma r_i)(\Sigma X_i^2 r_i) > (\Sigma X_i r_i)^2$$

valid for non-negative r_i's provided that $X_i \neq X_j$ for some i,j such that both r_i and r_j are positive. By (3.10) we find

$$\underset{\sim}{X}^T \underset{\sim}{\Sigma} \underset{\sim}{X} = \int_0^1 [\bar{h}(s) \sum_i X_i^2 h_i(s) - (\sum_i X_i h_i(s))^2]ds.$$

Now $X_i \neq X_j$ for some i,j by assumption and if we assume that for any i,j there exists a set $A_{ij} \subseteq [0,1]$ with positive Lebesgue measure such that $h_i(t) > 0$ and $h_j(t) > 0$ for $t \in A_{ij}$, then (A.2) yields that $\underset{\sim}{X}^T \underset{\sim}{\Sigma} \underset{\sim}{X} > 0$. A similar argument shows that $V^{(n)}(1)$ has rank k-1 provided that for any i,j there exists a time point where $N^{(n)}$ jumps and such that $L^{(n)}$, $Y_i^{(n)}$ and $Y_j^{(n)}$ are positive.

APPENDIX III. Example of verification of conditions of Theorem 3.1.

The example is the ordinary Kruskal-Wallis test; i.e. the situation described at the beginning of Section 3C with no censoring and $L_0(y) = y$. The reason we give this argument in such detail is that it works in identical fashion for much more complicated situations.

Thus in Theorem 3.1 we set $a_n = n^{-3/2}$ and $L^{(n)}(s) = \bar{Y}^{(n)}(s)$. We have

$$Y_i^{(n)}(s) = \#\{j \le n_i : X_{ij} \ge s\} \qquad i=1,\ldots,k,$$

$$\bar{Y}^{(n)}(s) = \sum_{i=1}^{k} Y_i^{(n)}(s)$$

where X_{ij}, $j=1,\ldots,n_i$, $i=1,\ldots,k$ are i.i.d. with distribution function F. The sample sizes $n_i = n_i(n)$ are such that $n_i/n \to p_i \in [0,1]$ as $n \to \infty$. We suppose F is such that $F(1) = 1$, $F(t) < 1$ for $t < 1$, and F has a continuous density f on [0,1]. We define $\alpha(s) = f(s)/(1-F(s))$ for $s < 1$ and $\alpha(1) = 0$.

For condition (3.5) we have

$$a_n L^{(n)}(s) Y_i^{(n)}(s)/\bar{Y}^{(n)}(s) = n^{-\frac{1}{2}} Y_i^{(n)}(s)/\bar{Y}^{(n)}(s) \le n^{-\frac{1}{2}} \to 0 \text{ as } n \to \infty.$$

For condition (3.6) we have for $s < 1$

$$a_n^2 [L^{(n)}(s)]^2 Y_i^{(n)}(s) Y_\ell^{(n)}(s) \alpha(s)/\bar{Y}^{(n)}(s)$$

$$\text{(A.3)} \qquad = \frac{n_i}{n} \frac{n_\ell}{n} \frac{Y_i^{(n)}(s)}{n_i} \frac{Y_\ell^{(n)}(s)}{n_\ell} \frac{\bar{Y}^{(n)}(s)}{n} \alpha(s)$$

$$\underset{n\to\infty}{\overset{P}{\to}} p_i p_\ell (1-F(s))^2 f(s).$$

So we take $h_i(s) = p_i(1-F(s))f(s)^{\frac{1}{2}}$ for $s < 1$; $h_i(1) = 0$. Clearly h_i is square integrable on [0.1].

For (3.7) we note that by (A.3) for any $s \le \tau < 1$

$$a_n^2[L^{(n)}(s)]^2 Y_i^{(n)}(s) Y_\ell^{(n)}(s)\alpha(s)/\bar{Y}^{(n)}(s) \le \alpha(s) \le \frac{\sup_{t\in[0,1]} f(t)}{1-F(\tau)} \quad .$$

Thus we certainly have uniform integrability on $[0,\tau]$ for any $\tau > 1$. Uniformly integrability on [0,1] in fact only holds for certain choices of F; and it is more instructive to verify (3.8) instead.

First using (A.3) we bound the integral in (3.8) by

$$\int_\tau^1 \frac{\bar{Y}^{(n)}(s)}{n}\,\alpha(s)\,ds.$$

Next we apply a result on empirical distribution functions: for every $\beta \in [0,1]$,

$$\text{(A.4)} \qquad P\left(\frac{\bar{Y}^{(n)}(s)}{n} \le \frac{1}{\beta}\,(1-F(s)) \text{ for all } s\right) = 1 - \beta$$

(DANIELS (1945)).

Thus

$$P\left(\int_\tau^1 \frac{\bar{Y}^{(n)}(s)}{n}\,\alpha(s)ds > \frac{1}{\beta}\int_\tau^1 (1-F(s))\alpha(s)ds\right) < \beta.$$

Also $\frac{1}{\beta}\int_\tau^1 (1-F(s))\alpha(s)ds = \frac{1}{\beta}\int_\tau^1 f(s)ds = \frac{1}{\beta}(1-F(\tau))$, which for given β can be made arbitrarily small by taking τ sufficiently close to 1. Thus the left-hand side of (3.8) is smaller than β; but since β is arbitrary the required result holds.

When dealing with censored data it is often useful to known that analogues to (A.4) hold for an empirical distribution function based on independent but *not* identically distributed random variables (VAN ZUYLEN 1977, 1978) and for the product limit estimator (GILL, 1980).

REFERENCES

[1] AALEN, O.O. (1975), *Statistical inference for a family of counting processes*, Ph.D. Dissertation. Department of Statistics, University of California, Berkeley.

[2] AALEN, O.O. (1976), *Nonparametric inference in connection with multiple decrement models*, Scand. J. Statist. 3, 15-27.

[3] AALEN, O.O. (1977), *Weak convergence of stochastic integrals related to counting processes*, Z. Wahrscheinlichkeitsth. verw. Geb. 38, 261-277. Correction, Vol. 48 1979, 347.

[4] AALEN, O.O. (1978), *Nonparametric inference for a family of counting processes*, Ann. Statist. 6, 701-726.

[5] AALEN, O.O., BORGAN, Ø, KEIDING, N. & THORMANN, J. (1980), *Interaction between life history events. Nonparametric analysis for prospective and retrospective data in the presence of censoring*, Scand. J. Statist. 7, 161-181.

[6] AALEN, O.O. & JOHANSEN, S. (1978), *An empirical transition matrix for non-homogeneous Markov chains based on censored observations*, Scand. J. Statist. 5, 141-150.

[7] ANDERSEN, P.K. & RASMUSSEN, N.K. (1980), *Admissions to psychiatric hospitals among women giving birth and women having induced abortion*. Contributed paper, 13 Europ. Meet. Statist., Brighton.

[8] BILLINGSLEY, P. (1968), *Convergence of probability measures*, Wiley, New York.

[9] BRÉMAUD, P. & JACOD, J. (1977), *Processus ponctuels et martingales: Résultats récents sur la modélisation et filtrage*, Adv. Appl. Prob. 9, 362-416.

[10] BRESLOW, N.E. (1970), *A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship*, Biometrika 57, 579-594.

[11] BRESLOW, N.E. (1975), *Analysis of survival data under the proportional hazards model*, Int. Stat. Rev. 43, 45-58.

[12] BROOKMEYER, R. & CROWLEY, J. (1980), *A k-sample median test for censored data*, Tech. Rep. 11, Wisc. Clin. Cancer Center-

Biostatistics, University of Wisconsin - Madison.

[13] CHERNOFF, H. & SAVAGE, I.R. (1958), *Asymptotic normality and efficiency of certain nonparametric test statistics*, Ann. Math. Statist. 29, 972-994.

[14] COX, D.R. (1972), *Regression models and life-tables (with discussion)* J. Roy. Statist. Soc. B 34, 187-220.

[15] CROWLEY, J. & BRESLOW, N.E. (1975), *Remarks on the conservatism of* $\Sigma(O-E)^2/E$ *in survival data*. Biometrics 31, 957-961.

[16] CROWLEY, J. & THOMAS, D.R. (1975), *Large sample theory for the log rank test*. Tech. Rep. 415, Dept. of Statistics, University of Wisconsin-Madison.

[17] DANIELS, H.E. (1945), *The statistical theory of the strength of bundles of threads*, I. Proc. Roy. Soc. A 183, 405-435.

[18] EFRON, B. (1967), *The two sample problem with censored data*, Proc. Fifth Berk. Symp. Math. Statist. Prob. 4, 831-854.

[19] FLEMING, T.R., O'FALLON, J.R., O'BRIEN, P.C. & HARRINGTON, D.P. (1980), *Modified Kolmogorov-Smirnov test procedures with application to arbitrarily right censored data*, Biometrics 36, 607-625.

[20] GEHAN, E.A. (1965), *A generalized Wilcoxon test for comparing arbitrarily singly-censored samples*, Biometrika 52, 203-223.

[21] GILL, R.D. (1978), *Testing with replacement and the product limit estimator*, SW 57/78, Mathematisch Centrum, Amsterdam, Condensed version to appear in Ann. Statist., 1981.

[22] GILL, R.D. (1980), *Censoring and stochastic integrals*, Mathematical Centre Tracts 124, Mathematisch Centrum Amsterdam.

[23] HÁJEK, J. & ŠIDAK, Z. (1967), *Theory of rank tests*, Academic Press, New York.

[24] HARRINGTON, D.P. & FLEMING, T.R. (1981), *A class of rank test procedures for censored survival data*, Tech. Rep. Series no. 12, Sect. Med. Res. Statist., Mayo Clinic, Rochester, Minnesota.

[25] HOEM, J.M. (1976), *The statistical theory of demographic rates,* (with discussion), Scand. J. Statist. 3, 169-185.

[26] HOLLANDER, M. & PROSCHAN, F. (1979), *Testing to determine the underlying distribution using randomly censored data,* Biometrics 35, 393-401.

[27] HYDE, J. (1977), *Testing survival under right censoring and left truncation,* Biometrika 64, 225-230.

[28] JACOBSEN, M. (1981), *Lecture notes on counting processes,* Inst. Math. Statist., University of Copenhagen.

[29] KALBFLEISCH, J.D. & PRENTICE, R.L. (1980), *The statistical analysis of failure time data,* Wiley, New York.

[30] KAPLAN, E.L. & MEIER, P. (1958), *Nonparametric estimation from incomplete observations,* J. Amer. Statist. Assoc. 53, 457-481.

[31] MCLEISH, D.L. (1974), *Dependent central limit theorems and invariance principles,* Ann. Probab. 2, 620-628.

[32] MEYER, P.A. (1976), *Un cours sur let intégrales stochastiques,* Lecture Notes in Mathematics 511, 245-400, Springer Verlag, Berlin.

[33] PETO, R. & PETO, J. (1973), *Asymptotically efficient rank invariant test procedures* (with discussion). J. Roy. Statist. Soc. A 135, 185-206.

[34] PETO, R. & PIKE, M.C. (1973), *Conservatism of the approximation* $\Sigma(O-E)^2/E$ *in the log rank test for survival data or tumor incidence data,* Biometrics 29, 579-584.

[35] PRENTICE, R.L. (1978), *Linear rank tests with right censored data,* Biometrika 65, 167-179.

[36] PRENTICE, R.L., KALBFLEISCH, J.D., PETERSON, A.V., Jr., FOURNOY, N., FAREWELL, V.T. & BRESLOW, N.E. (1978), *The analysis of failure times in the presence of competing risks,* Biometrics 34, 541-554.

[37] PRENTICE, R.L. & MAREK, P. (1979), *A qualitative discrepancy between censored data rank tests,* Biometrics 35, 861-867. Correction note, Vol. 36, 1980, 751.

[38] REBOLLEDO, R. (1978), *Sur les applications de la théorie des martingales à l'étude statistique d'une famille de processus ponctuels*, Springer Verlag, Berlin.

[39] SCHOENFELD, D. (1981), *The asymptotic properties of nonparametric tests for comparing survival distributions*, Biometrika 68, 316-319.

[40] SHIRYAYEV, A.N. (1980), *Martingales: recent results, developments, applications*, Forum lectures, 13 Europ. Meet. Statist., Brighton. (to appear in Int. Statist. Rev.).

[41] TARONE, R.E. & WARE, J. (1977), *On distribution-free tests for equality of survival distributions*, Biometrika 64, 156-160.

[42] VAN ZUIJLEN, M.C.A. (1978), *Properties of the empirical distribution function for independent nonidentically distributed random variables*, Ann. Probability 6, 250-266.

[43] VAN ZUIJLEN, M.C.A. (1977), *Empirical distributions and rank statistics*, Mathematical Centre Tracts 79, Mathematisch Centrum, Amsterdam.