

## ESTIMATING JOINTLY SYSTEM AND COMPONENT RELIABILITIES USING A MUTUAL CENSORSHIP APPROACH<sup>1</sup>

BY HANI DOSS, STEVEN FREITAG AND FRANK PROSCHAN

*Florida State University*

Let  $F$  denote the life distribution of a coherent structure of independent components. Suppose that we have a sample of independent systems, all having the same structure. Each system is continuously observed until it fails. For every component in each system, either a failure time or a censoring time is recorded. A failure time is recorded if the component fails before or at the time of system failure; otherwise a censoring time is recorded. We introduce a method for finding estimates for  $F(t)$ , quantiles and other functionals of  $F$ , based on the censorship of the component lives by system failure. We use the theory of counting processes and stochastic integrals to obtain limit theorems that enable the construction of confidence intervals for large samples. Our approach extends and gives a novel application of censoring methodology.

**1. Introduction and summary.** Consider a system of independent components labeled 1 through  $m$ . We assume that the system forms a coherent structure, which we denote by  $\phi$ . In particular the system and each component are in either a functioning state or a failed state, and the state of the system depends only on the states of the components; see Barlow and Proschan (1981), Chapters 1 and 2, for definitions and basic facts relating to coherent systems.

Suppose that we have a sample of  $n$  independent systems, each with the same structure  $\phi$ . Each system is continuously observed until it fails. For every component in each system, either a failure time or a censoring time is recorded. A failure time is recorded if the component fails before or at the time of system failure. A censoring time is recorded if the component is still functioning at the time of system failure. From these failure times and censoring times we wish to estimate  $F$ , the distribution of the system lifelength.

Let

$$T_i = \text{lifelength of system } i.$$

We note that  $F$  may be estimated by the empirical estimator

$$(1.1) \quad \hat{F}^{\text{emp}}(t) = \frac{1}{n} \sum_{i=1}^n I(T_i \leq t),$$

where  $I(A)$  is the indicator of the set  $A$ . However,  $\hat{F}^{\text{emp}}$  is unsatisfactory in that

---

Received February 1986; revised September 1988.

<sup>1</sup>Research supported by Air Force Office of Scientific Research Grant AFOSR 88-0040 and by Office of Naval Research Contract N00014-86-K-0156.

AMS 1980 subject classifications. Primary 62N05; secondary 62E20, 62G05.

*Key words and phrases.* Reliability, coherent structure, Kaplan–Meier estimator, martingale central limit theorem, stochastic integral, component life distribution, system life distribution, multivariate counting process.

it does not fully utilize the information contained in the sample. Specifically, it does not use the identity of the components still functioning at system failure time, nor the failure times of the components failing before system failure time.

The purpose of this article is two-fold. First, we present an estimator  $\hat{F}$  of  $F$  that does use all the information contained in the sample. Second, we apply (in Section 2) the statistical theory of counting processes initiated by Aalen (1978) to obtain the asymptotic distribution of  $\hat{F}$  in the observational scheme described above, and show (in Section 5) how this theory can be used to study several other problems involving related observational schemes.

To construct our estimator  $\hat{F}$ , we first obtain estimators of the distributions of the component lifelengths, and then combine these in a suitable way. To describe  $\hat{F}$  fully, we introduce some notation.

- $X_{ij}$  is the lifelength of component  $j$  in system  $i$ ;
- $F_j$  is the distribution of the lifelength of component  $j$  (thus,  $X_{1j}, X_{2j}, \dots, X_{nj}$  are i.i.d.  $\sim F_j$ );
- $Z_{ij} = \min(X_{ij}, T_i)$ ;
- $\delta_{ij} = I(X_{ij} \leq T_i)$ ;
- $H_j$  is the common distribution of the independent random variables  $Z_{1j}, \dots, Z_{nj}$ .

Here and throughout the article the letter  $i$  indexes systems and  $j$  indexes components;  $i$  ranges over  $1, \dots, n$ , and  $j$  over  $1, \dots, m$ . For each  $j$ , let  $Z_{(1)j} \leq Z_{(2)j} \leq \dots \leq Z_{(n)j}$  be the ordered values of  $Z_{1j}, Z_{2j}, \dots, Z_{nj}$ . Define

$$(1.2) \quad \delta_{(i)j} = \begin{cases} 1 & \text{if } Z_{(i)j} \text{ corresponds to an uncensored lifetime,} \\ 0 & \text{if } Z_{(i)j} \text{ corresponds to a censored lifetime.} \end{cases}$$

(When an uncensored and a censored observation are tied, the uncensored observation is considered to have occurred first.) Let  $\hat{F}_j$  denote the Kaplan–Meier estimator of  $F_j$ ,

$$(1.3) \quad \hat{F}_j(t) = 1 - \prod_{i: Z_{(i)j} \leq t} \left( \frac{n - i}{n - i + 1} \right)^{\delta_{(i)j}}.$$

The definition above differs from the usual definition of the Kaplan–Meier estimator in that  $\hat{F}_j(t)$  is not arbitrarily defined to be 1 for  $t \geq Z_{(n)j}$ .

For each coherent structure  $\phi$  of independent components, there corresponds a function  $h_\phi$ , called the *reliability function*, such that

$$(1.4) \quad \bar{F}(t) = h_\phi(\bar{F}_1(t), \dots, \bar{F}_m(t)).$$

[For a distribution function  $G$ ,  $\bar{G}(t)$  denotes  $1 - G(t)$ .] A more detailed description of reliability functions is provided in Chapter 2 of Barlow and Proschan (1981). The estimator  $\hat{F}$  is defined by

$$(1.5) \quad \hat{F}(t) = \begin{cases} 1 - h_\phi(\hat{F}_1(t), \dots, \hat{F}_m(t)) & \text{if } t < T_{(n)}, \\ 1 & \text{if } t \geq T_{(n)}. \end{cases}$$

Here,  $T_{(n)} = \max(T_1, T_2, \dots, T_n)$ . The estimator  $\hat{F}$  has obvious intuitive appeal.

The properties of the Kaplan–Meier estimator have been studied extensively by various authors. Under the assumption that the censoring variables and the lifelengths are independent, the Kaplan–Meier estimator is the maximum likelihood estimator [Kaplan and Meier (1958) and Johansen (1978)]. Regarded as a stochastic process, it is strongly uniformly consistent [Földes, Rejtő and Winter (1980)] and when normalized converges weakly to a Gaussian process [Breslow and Crowley (1974), Aalen (1976) and Gill (1983)].

The main results of this article can now be stated. Let  $D[0, T]$  be the space of all real-valued functions defined on  $[0, T]$  that are right-continuous and have left limits, with the Skorohod metric topology.  $D^m[0, T]$  denotes the product metric space.

**THEOREM 1.** *Suppose  $F_1, F_2, \dots, F_m$  are continuous, and let  $T$  be such that  $F_j(T) < 1$  for  $j = 1, 2, \dots, m$ . Then as  $n \rightarrow \infty$ ,*

$$n^{1/2}(\hat{F}_1 - F_1, \hat{F}_2 - F_2, \dots, \hat{F}_m - F_m) \rightarrow (W_1, W_2, \dots, W_m)$$

*weakly in  $D^m[0, T]$ , where  $W_1, \dots, W_m$  are independent mean 0 Gaussian processes. The covariance structure of  $W_j$  is given by*

$$(1.6) \quad \text{Cov}(W_j(t_1), W_j(t_2)) = \bar{F}_j(t_1)\bar{F}_j(t_2) \int_0^{t_1} \frac{dF_j(u)}{\bar{H}_j(u)\bar{F}_j(u)} \quad \text{for } 0 \leq t_1 \leq t_2 \leq T.$$

Since in general the dependence among the  $\hat{F}_j$ 's may be complex, Theorem 1 is not a trivial extension of the corresponding result for the individual Kaplan–Meier estimators  $\hat{F}_j$ .

Theorem 1 together with an application of the delta method yields weak convergence of the estimator  $\hat{F}$ . Let

$$(1.7) \quad I_j(t) = \left. \frac{\partial h_\phi}{\partial u_j}(u_1, \dots, u_m) \right|_{u_j = \bar{F}_j(t), j=1, \dots, m}$$

**THEOREM 2.** *Suppose  $F_1, F_2, \dots, F_m$  are continuous and  $T$  is such that  $F_j(T) < 1$  for  $j = 1, 2, \dots, m$ . Then as  $n \rightarrow \infty$ ,*

$$n^{1/2}(\hat{F} - F) \rightarrow W \quad \text{weakly in } D[0, T],$$

*where  $W$  is a mean 0 Gaussian process with covariance structure given by*

$$(1.8) \quad \text{Cov}(W(t_1), W(t_2)) = \sum_{j=1}^m I_j(t_1)I_j(t_2)\bar{F}_j(t_1)\bar{F}_j(t_2) \int_0^{t_1} \frac{dF_j(u)}{\bar{H}_j(u)\bar{F}_j(u)} \quad \text{for } 0 \leq t_1 \leq t_2 \leq T.$$

The commonly used estimate of the variance of the Kaplan–Meier estimate is given by Greenwood's formula [see Chapter 3 of Miller (1981)]. Since this estimate is known to be consistent [see Hall and Wellner (1980)] it follows that for fixed  $t$ , the variance of  $\hat{F}(t)$  given in Theorem 2 can be consistently

estimated. This enables the construction of confidence intervals for  $F(t)$  in large samples.

The competing risks model corresponds to a series system. Aalen (1976) showed that for this model, the vector of Kaplan–Meier estimates  $(\hat{F}_1, \dots, \hat{F}_m)$ , when normalized, converges to a multidimensional Gaussian process, whose components are independent. This result corresponds to our Theorem 1 for the case of a series structure.

We note that (assuming  $F_1, \dots, F_m$  to be continuous) the data from the competing risks model is the death time of the system and the identity of the component causing the death of the system. Meilijson (1981) expanded this to the case of a coherent system in an “autopsy model”: At system death time  $T$ , an autopsy is performed which reveals the set  $D$  of dead components (the identity of the component causing system death is not given by the autopsy). This model yields less information than does ours. Meilijson considered the identifiability question of when does the distribution of  $(T, D)$  determine  $F_1, \dots, F_m$ .

The present article is organized as follows. In Section 2 we prove Theorems 1 and 2. In Section 3 we give an application of our results to system design methods. In Section 4 we discuss the efficiency of our estimator vs. the empirical estimator (1.1). In Section 5 we discuss Meilijson’s model as well as other models where the data can yield more information than just  $T_1, \dots, T_n$ . The Appendix gives a proof of a result used in Section 2.

## 2. Weak convergence results.

**2.1. Random censoring and preliminaries.** Corresponding to a generic system, we define generic random variables  $X_j$ ,  $Z_j$ ,  $\delta_j$  and  $T$ , such that the random vector  $(X_j, Z_j, \delta_j, T)$  has the same distribution as  $(X_{ij}, Z_{ij}, \delta_{ij}, T_i)$  for  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, m$ .

In Section 1 it is noted that the strong consistency and weak convergence results for the Kaplan–Meier estimator are valid under the assumption that the lifelengths and the censoring random variables are independent. In our model, for each  $j$ ,  $X_j$  is censored by  $T$ , and for a coherent structure these two random variables are dependent. However, it is possible to redefine the censoring variables to circumvent this difficulty. This is best explained in terms of a simple example. Consider the structure shown diagrammatically in Figure 1.

In the example  $T = X_1 \wedge (X_2 \vee X_3)$ , where  $x \wedge y = \min(x, y)$  and  $x \vee y = \max(x, y)$ . Consider now component 1. Clearly,  $X_1$  is censored by  $Y_1 = X_2 \vee X_3$ ,

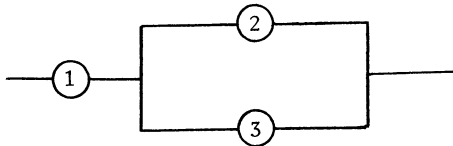


FIG. 1.

which is independent of  $X_1$ . Similarly,  $X_2$  is censored by  $Y_2 = X_1$ , and  $X_3$  by  $Y_3 = X_1$ .

A central result of this article, shown in the Appendix, is that this construction can always be made: In general, for each  $j = 1, \dots, m$ , there is a random variable  $Y_j$  such that

$$(2.1) \quad (Z_j, \delta_j) = (X_j \wedge Y_j, I(X_j \leq Y_j))$$

and

$$(2.2) \quad X_j \text{ and } Y_j \text{ are independent.}$$

Roughly speaking,  $Y_j$  is the lifelength of the system if  $X_j$  is replaced by  $\infty$ . Proposition A.1 states that to determine whether or not component  $j$ 's lifelength has been censored by time  $t$ , it is enough to know the history of the other components (or of  $Y_j$ ) up to time  $t$ . We will refer to  $Y_j$  as the censoring variable of  $X_j$ .

In order to describe the distribution of  $Y_j$  we introduce some notation. For  $y = (y_1, \dots, y_m) \in [0, 1]^m$ ,  $\alpha \in [0, 1]$  and  $j = 1, \dots, m$ , let

$$(2.3) \quad (\alpha_j, y) = (y_1, \dots, y_{j-1}, \alpha, y_{j+1}, \dots, y_m).$$

Let  $\bar{\mathbf{F}}(t) = (\bar{F}_1(t), \dots, \bar{F}_m(t))$  and recall that  $H_j$  is the distribution of  $Z_j$ . In the Appendix it is shown that

$$(2.4) \quad P(Y_j > t) = h_\phi(1_j, \bar{\mathbf{F}}(t)),$$

where  $h_\phi$  is the reliability function [see (1.4)]. Thus,

$$(2.5) \quad \bar{H}_j(t) = \bar{F}_j(t) h_\phi(1_j, \bar{\mathbf{F}}(t)).$$

We now review some terminology from reliability theory [see, e.g., Barlow and Proschan (1981)] to be used in the proofs of consistency and weak convergence of  $\hat{F}$ . For a coherent system of  $m$  components, the states of the components correspond to a vector  $U = (U_1, \dots, U_m)$ , where  $U_j = I$  (component  $j$  is in a functioning state). The *structure function* is defined by  $\phi(U) = I$  (system functions when  $U$  describes the states of the components) for  $U \in A_m$ , where  $A_m = \{0, 1\}^m$ . It is well-known (and easy to see) that for  $p = (p_1, \dots, p_m) \in [0, 1]^m$ ,

$$(2.6) \quad h_\phi(p) = \sum_{U \in A_m} \phi(U) \prod_{j=1}^m p_j^{U_j} (1 - p_j)^{1 - U_j},$$

where  $0^0 = 1$  by definition.

The Kaplan–Meier estimates  $\hat{F}_j$  given by (1.3) will be denoted  $\hat{F}_{j,n}$  when we want to emphasize the dependence on  $n$ ; similarly for the estimate  $\hat{F}$  of system life distribution. Also  $\hat{\mathbf{F}}^n$  will denote the vector  $(\hat{F}_1^n, \dots, \hat{F}_m^n)$ .

The following lemma is needed in the proof of strong uniform consistency and of weak convergence of  $\hat{F}$ .

**LEMMA 2.1.** *For any structure of  $m$  independent components, the corresponding reliability function  $h_\phi$  is twice continuously differentiable over  $[0, 1]^m$ , and the first and second partial derivatives are bounded in absolute value by 1 uniformly over  $[0, 1]^m$ .*

PROOF. For  $p = (p_1, \dots, p_m) \in [0, 1]^m$  and  $k = 1, 2, \dots, m$ , we have by (2.6),

$$(2.7) \quad \left. \frac{\partial h_\phi}{\partial p_k} \right|_p = h_\phi(1_k, p) - h_\phi(0_k, p).$$

From (2.7) we have

$$(2.8) \quad \left. \frac{\partial^2 h_\phi}{\partial p_k^2} \right|_p = 0$$

and for  $l \neq k$ ,

$$(2.9) \quad \left. \frac{\partial^2 h_\phi}{\partial p_k \partial p_l} \right|_p = \{h_\phi(1_k, 1_l, p) - h_\phi(1_k, 0_l, p)\} \\ - \{h_\phi(0_k, 1_l, p) - h_\phi(0_k, 0_l, p)\},$$

in an obvious extension of the notation (2.3). By (2.6)  $h_\phi$  is continuous over  $[0, 1]^m$ . This fact together with (2.7), (2.8) and (2.9) imply that the first and second partial derivatives are continuous on  $[0, 1]^m$ ; hence, by Theorem 6.18 of Apostol (1964),  $h_\phi$  is twice continuously differentiable on  $[0, 1]^m$ . Equation (2.7) implies that the first partials are bounded in absolute value by 1. Since each of the two quantities inside the braces on the right-hand side of (2.9) is between 0 and 1, it follows that the second partials are also bounded in absolute value by 1. The lemma follows since  $p$  is arbitrary.  $\square$

We now establish the strong uniform consistency of  $\hat{F}$  and give the rate of convergence.

PROPOSITION 2.1. Let  $T > 0$  be such that for  $j = 1, \dots, m$ ,  $\min(F_j(T), \bar{F}_j(T)) > 0$ . Then:

(a)  $n^{1/4}(\ln n)^{-1/2} \sup_{0 \leq t \leq T} |\hat{F}(t) - F(t)| \rightarrow 0 \text{ a.s.}$

(b) If  $F_1, \dots, F_m$  are continuous we get the stronger result  $n^{1/2}(\ln \ln n)^{-1/2} \sup_{0 \leq t \leq T} |\hat{F}(t) - F(t)|$  is bounded a.s.

PROOF. The convergence results for the individual  $\hat{F}_j$ 's given in Földes, Rejtő and Winter (1980) and Földes and Rejtő (1981), respectively, together with Lemma 2.1 yield parts (a) and (b), respectively.  $\square$

2.2. Proofs of Theorems 1 and 2. A review of the counting process and martingale theory used below is given in the survey paper of Andersen and Borgan (1985). Throughout, we adopt the convention that  $0/0 = 0$ . The index  $n$  defining a process is suppressed whenever possible.

To prove Theorem 1, we will show that for any  $T > 0$  satisfying  $\max_{1 \leq j \leq m} F_j(T) < 1$ ,

$$(2.10) \quad n^{1/2} \left( \frac{\hat{F}_1 - F_1}{\bar{F}_1}, \dots, \frac{\hat{F}_m - F_m}{\bar{F}_m} \right) \rightarrow_d (W_1^*, \dots, W_m^*),$$

where  $W_1^*, \dots, W_m^*$  are independent mean 0 Gaussian processes with covariance given by

$$(2.11) \quad \text{Cov}(W_j^*(t_j), W_j^*(t_2)) = \int_0^{t_1} \frac{dF_j(u)}{\bar{H}_j(u)\bar{F}_j(u)} \quad \text{for } 0 \leq t_1 \leq t_2 \leq T.$$

(Now and henceforth, the symbol  $d$  signifies weak convergence in  $D^m[0, T]$ .) Theorem 1 is an easy consequence of (2.10) and Theorem 5.1 of Billingsley (1968).

We prove (2.10) by a general method introduced by Aalen (1978) and later refined by Gill (1980). We define the stopped process  $F_j^*$  on  $[0, \infty)$ ,  $j = 1, 2, \dots, m$ , by

$$(2.12) \quad F_j^*(t) = F_j(t \wedge Z_{(n)j}) \quad \text{and} \quad \bar{F}_j^*(t) = 1 - F_j^*(t)$$

and show that

$$(2.13) \quad n^{1/2} \left( \frac{\hat{F}_1 - F_1^*}{\bar{F}_1^*}, \dots, \frac{\hat{F}_m - F_m^*}{\bar{F}_m^*} \right) \rightarrow_d (W_1^*, \dots, W_m^*).$$

This is enough to prove (2.10), as it is easy to see that the difference between the left-hand side of (2.10) and the left-hand side of (2.13) converges to 0 in probability in  $D^m[0, T]$ .

To show (2.13), we first establish that for each  $n$ ,

$$(2.14) \quad \left\{ n^{1/2} \left( \frac{\hat{F}_j(t) - F_j^*(t)}{\bar{F}_j^*(t)} \right); t \in [0, T] \right\}, \quad j = 1, \dots, m,$$

are orthogonal square integrable martingales with respect to an appropriate family of  $\sigma$ -fields. Weak convergence in (2.13) then follows from a multivariate version of a martingale central limit theorem due to Rebolledo (1980). The orthogonality allows a simple extension to the multidimensional case of the proof of weak convergence given by Gill (1983) for the one-dimensional case.

To show that the processes in (2.14) are orthogonal martingales, we will need to be careful about the families of  $\sigma$ -fields that we use.

We define the following processes on  $[0, \infty)$ ; these correspond to the situation where we imagine that no censoring occurs:

$$(2.15) \quad N_{ij}(t) = I(X_{ij} \leq t),$$

$$(2.16) \quad N_j(t) = \sum_{i=1}^n N_{ij}(t),$$

$$(2.17) \quad V_{ij}(t) = I(X_{ij} \geq t),$$

$$(2.18) \quad V_j(t) = \sum_{i=1}^n V_{ij}(t),$$

$$(2.19) \quad A_{ij}(t) = \int_0^t \left[ \frac{V_{ij}(s)}{\bar{F}_j(s)} \right] dF_j(s),$$

$$(2.20) \quad A_j(t) = \int_0^t \left[ \frac{V_j(s)}{\bar{F}_j(s)} \right] dF_j(s) \quad \left( = \sum_{i=1}^n A_{ij}(t) \right),$$

$$(2.21) \quad M_{ij}(t) = N_{ij}(t) - A_{ij}(t),$$

$$(2.22) \quad M_j(t) = N_j(t) - A_j(t) \quad \left( = \sum_{i=1}^n M_{ij}(t) \right).$$

We further define the filtration

$$(2.23) \quad \mathcal{F}_t = \text{completion of } \sigma(N_{ij}(s); 1 \leq i \leq n, 1 \leq j \leq m, s \leq t).$$

In Section 5 we will consider censoring schemes that are more complicated than the ones involved in Theorem 1. Therefore, we wish to give a proof of the fact that the processes in (2.14) are orthogonal martingales that can apply to more general censoring mechanisms. Thus, corresponding to an arbitrary censoring mechanism for component  $j$  of system  $i$ , let

$$(2.24) \quad C_{ij}(t) = I(X_{ij} \text{ is under observation at time } t).$$

In the present situation,

$$(2.25) \quad C_{ij}(t) = I(Y_{ij} \geq t),$$

where  $Y_{ij}$  is defined by (A.2).

We define the following entities; these correspond to (2.15)–(2.23) for the case of censoring:

$$(2.26) \quad N_{ij}^c(t) = \int_0^t C_{ij}(s) dN_{ij}(s),$$

$$(2.27) \quad N_j^c(t) = \sum_{i=1}^n N_{ij}^c(t),$$

$$(2.28) \quad V_{ij}^c(t) = C_{ij}(t) V_{ij}(t),$$

$$(2.29) \quad V_j^c(t) = \sum_{i=1}^n V_{ij}^c(t),$$

$$(2.30) \quad A_{ij}^c(t) = \int_0^t \frac{V_{ij}^c(s)}{\bar{F}_j(s)} dF_j(s),$$

$$(2.31) \quad A_j^c(t) = \int_0^t \frac{V_j^c(s)}{\bar{F}_j(s)} dF_j(s) \quad \left( = \sum_{i=1}^n A_{ij}^c(t) \right),$$

$$(2.32) \quad M_{ij}^c(t) = N_{ij}^c(t) - A_{ij}^c(t),$$

$$(2.33) \quad M_j^c(t) = N_j^c(t) - A_j^c(t) \quad \left( = \sum_{i=1}^n M_{ij}^c(t) \right),$$

$$(2.34) \quad \mathcal{F}_t^c = \text{completion of } \sigma(N_{ij}^c(s); 1 \leq i \leq n, 1 \leq j \leq m, s \leq t).$$



We further define

$$(2.35) \quad J_j^c(t) = I(V_j^c(t) > 0).$$

The following proposition [Theorem 3.1 of Aalen and Johansen (1978) and (3.2.12) of Gill (1980)] is fundamental in establishing that the processes in (2.14) are orthogonal martingales.

**PROPOSITION 2.2.** *Suppose  $F_1, \dots, F_m$  are continuous and  $t \geq 0$  is such that  $\max_{1 \leq j \leq m} F_j(t) < 1$ . Then for each  $j$  and for all  $n$ ,*

$$(2.36) \quad n^{1/2} \left( \frac{\hat{F}_j(t) - F_j^*(t)}{\bar{F}_j^*(t)} \right) = n^{1/2} \int_0^t \left[ \frac{J_j^c(s)}{V_j^c(s)} \frac{\hat{F}_j(s-)}{\bar{F}_j(s)} \right] dM_j^c(s).$$

Our plan is to show that  $M_j^c(t)$ ,  $j = 1, \dots, m$ , are orthogonal martingales with respect to either  $\{\mathcal{F}_t; t \in [0, T]\}$  or  $\{\mathcal{F}_t^c; t \in [0, T]\}$ . Then we will use Proposition 2.2 to obtain the same conclusion for the processes in (2.14).

The following general lemma, which is well-known [see, e.g., Andersen, Borgan, Gill and Keiding (1982)], is an essential part of our proof that the processes in (2.14) are orthogonal martingales.

**LEMMA 2.2.** *Let  $M_{ij}(t)$ ,  $C_{ij}(t)$  and  $M_{ij}^c(t)$  be defined by (2.21), (2.24) and (2.32), respectively. Let  $T$  satisfy  $F_j(T) < 1$  for each  $j$ , and let  $\{\mathcal{A}_t; t \in [0, T]\}$  be a filtration such that*

- (i)  $C_{ij}(\cdot)$  is  $\mathcal{A}_t$ -predictable;
- (ii)  $\{(M_{ij}(t), \mathcal{A}_t); t \in [0, T]\}$  is a square integrable martingale.

Then

$$\{(M_{ij}^c(t), \mathcal{A}_t); t \in [0, T]\} \text{ is a square integrable martingale.}$$

**PROOF.** From (2.32), (2.26), (2.30), (2.28), (2.21) and (2.19) it follows that

$$(2.37) \quad M_{ij}^c(t) = \int_0^t C_{ij}(s) dM_{ij}(s).$$

Since the integrand in (2.37) is  $\mathcal{A}_t$ -predictable and  $M_{ij}$  is square integrable on  $[0, T]$  (in fact uniformly bounded), the Stieltjes integral in (2.37) is a stochastic integral and is a square integrable martingale [see Aalen (1978), page 703 for a discussion].  $\square$

Let us see how Lemma 2.2 applies to the present situation, where  $C_{ij}(t) = I(Y_{ij} \geq t)$ , with  $Y_{ij}$  defined by (A.2). Part (ii) of Proposition A.1 (in the Appendix) asserts that this censoring process is  $\mathcal{F}_t$ -predictable, where  $\mathcal{F}_t$  is defined by (2.23). Now it is well-known that if

$$\mathcal{F}_t^{(i,j)} = \text{completion of } \sigma(N_{ij}(s); s \leq t),$$

then  $(M_{ij}(t), \mathcal{F}_t^{(i,j)})$  is a (bounded) martingale (on  $[0, T]$ ); see, for example,

Davis (1983), pages 136–137. Since  $X_{ij}$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ , are independent, this implies that  $(M_{ij}(t), \mathcal{F}_t)$  is a martingale on  $[0, T]$ . Lemma 2.2 therefore asserts that  $\{(M_{ij}^c(t), \mathcal{F}_t); t \in [0, T]\}$  is a martingale. It is easy to check that  $M_{ij}^c(t) \in \mathcal{F}_t^c$  with  $\mathcal{F}_t^c$  defined by (2.34). Using the fact that  $\mathcal{F}_t^c \subset \mathcal{F}_t$ , it is easy to see that  $\{(M_{ij}^c(t), \mathcal{F}_t^c); t \in [0, T]\}$  is also a martingale. Since the sum of martingales is a martingale, this gives that

$$(2.38) \quad \{(M_j^c(t), \mathcal{F}_t); t \in [0, T]\} \text{ and } \{(M_j^c(t), \mathcal{F}_t^c); t \in [0, T]\}$$

are square integrable martingales,

with

$$(2.39) \quad \langle M_{j_1}^c, M_{j_2}^c \rangle(t) = \begin{cases} A_{j_1}^c(t) & \text{if } j_1 = j_2, \\ 0 & \text{if } j_1 \neq j_2, \end{cases}$$

the orthogonality in (2.39) resulting from the fact that the counting processes  $N_1^c(t), \dots, N_m^c(t)$  have no common jump points with probability 1 [see, e.g., Theorem 2.3.1 of Gill (1980)].

Let us now apply Proposition 2.2. It is easy to see that for each  $j$ , the left-continuous processes  $J_j^c(s)$ ,  $\hat{F}_j(s-)$  and  $V_j^c(s)$  are adapted to either  $\{\mathcal{F}_t\}$  or  $\{\mathcal{F}_t^c\}$ . Therefore, the integrand on the right-hand side of (2.36) is predictable with respect to either  $\{\mathcal{F}_t\}$  or  $\{\mathcal{F}_t^c\}$ . Since this integrand is also bounded and since the square integrable martingale  $M_j^c$  is of bounded variation, the integral in (2.36) is a square integrable martingale, with covariation process

$$(2.40) \quad \left\langle n^{1/2} \left( \frac{\hat{F}_{j_1} - F_{j_1}^*}{\bar{F}_{j_1}^*} \right), n^{1/2} \left( \frac{\hat{F}_{j_2} - F_{j_2}^*}{\bar{F}_{j_2}^*} \right) \right\rangle(t) \\ = n \int_0^t \left[ \frac{J_{j_1}^c(s) \hat{F}_{j_1}(s-)}{V_{j_1}^c(s) \bar{F}_{j_1}(s)} \right] \left[ \frac{J_{j_2}^c(s) \hat{F}_{j_2}(s-)}{V_{j_2}^c(s) \bar{F}_{j_2}(s)} \right] d\langle M_{j_1}^c, M_{j_2}^c \rangle(s).$$

Equations (2.39) imply that this quantity is 0 if  $j_1 \neq j_2$ , and that for each  $j = 1, \dots, m$ ,  $t \in [0, T]$ ,

$$(2.41) \quad \left\langle n^{1/2} \left( \frac{\hat{F}_j - F_j^*}{\bar{F}_j^*} \right), n^{1/2} \left( \frac{\hat{F}_j - F_j^*}{\bar{F}_j^*} \right) \right\rangle(t) \\ = n \int_0^t \left[ \frac{J_j^c(s) \hat{F}_j(s-)}{V_j^c(s) \bar{F}_j(s)} \right]^2 dA_j^c(s) \\ = n \int_0^t \left[ \frac{J_j^c(s) (\hat{F}_j(s-))^2}{V_j^c(s) (\bar{F}_j(s))^3} \right] dF_j(s).$$

The proof of the weak convergence result for each individual  $n^{1/2}(\hat{F}_j - F_j^*)/\bar{F}_j^*$  is now the same as the proof of Theorem 1.1 in Gill (1983), which uses Theorem V.1 of Rebolledo (1980). The joint weak convergence in (2.13) follows from the orthogonality result in (2.39) and (2.40), and a simple multivariate extension of Rebolledo's (1980) Theorem V.1.

We now prove Theorem 2. The uniform bound for the first two partial derivatives of  $h_\phi$  given by Lemma 2.1 together with Taylor's theorem imply that for each  $t \in [0, T]$ ,

$$(2.42) \quad \begin{aligned} & n^{1/2} \left| \hat{F}(t) - F(t) - \sum_{j=1}^m I_j(t) (\hat{F}_j(t) - F_j(t)) \right| \\ & \leq \frac{n^{1/2}}{2} \sum_{j_1=1}^m \sum_{j_2=1}^m \left( \sup_{0 \leq t \leq T} |\hat{F}_{j_1}(t) - F_{j_1}(t)| \right) \left( \sup_{0 \leq t \leq T} |\hat{F}_{j_2}(t) - F_{j_2}(t)| \right). \end{aligned}$$

The right-hand side of (2.42) converges to 0 a.s. by the results of Földes and Rejtő (1981). We use the fact that convergence in sup norm implies convergence in the Skorohod topology to conclude that the process

$$n^{1/2} \left| \hat{F}(t) - F(t) - \sum_{j=1}^m I_j(t) (\hat{F}_j(t) - F_j(t)) \right| \rightarrow 0 \quad \text{a.s. in } D[0, T].$$

Thus, the proof follows by showing that

$$n^{1/2} \sum_{j=1}^m I_j(t) (\hat{F}_j(t) - F_j(t)) \rightarrow_d W,$$

which is a consequence of Theorem 5.1 of Billingsley (1968) and Theorem 1.

To construct confidence intervals for  $F(t)$ , we define the following functions and processes on  $[0, \infty)$ :

$$\begin{aligned} G_j(t) &= (\bar{F}_j(t))^2 \int_0^t \frac{dF_j(s)}{\bar{F}_j(s) \bar{H}_j(s)}, \\ \hat{G}_j^n(t) &= \frac{(\hat{F}_{jj}(t))^2}{n} \int_0^t \frac{dN_j^c(s)}{\hat{H}_j(s-) \hat{H}_j(s)} \\ &= n (\hat{F}_j(t))^2 \left( \sum_{i: Z_{(i)j} \leq t} \frac{\delta_{(i)j}}{(n-i+1)(n-i)} \right), \end{aligned}$$

where

$$\hat{H}_j(t) = \frac{1}{n} \sum_{i=1}^n I(Z_{ij} > t).$$

Also define

$$(2.43) \quad \hat{I}_j(t) = \frac{\partial h_\phi}{\partial u_j} \Big|_{u_j = \hat{F}_j(t), j=1, \dots, m}.$$

The quantity  $\hat{G}_j(t)/n$  is called Greenwood's estimator of the variance of  $\hat{F}_j(t)$ .

**LEMMA 2.3.** *Suppose  $F_1, \dots, F_m$  are continuous and  $T > 0$  is such that  $\max_{1 \leq j \leq m} F_j(T) < 1$ . Then  $\sum_{j=1}^m \hat{I}_j^2(t) \hat{G}_j(t)$  is a strongly consistent estimator of  $\sum_{j=1}^m I_j^2(t) G_j(t)$ .*

We note that in view of (1.8), Theorem 2 and Lemma 2.3 enable the formation of asymptotic confidence intervals for  $F(t)$ ,  $t \in [0, T]$ .

**PROOF.** Part (a) of the proposition in Section 2 of Hall and Wellner (1980), together with the convergence results of Földes and Rejtő (1981), imply that  $\hat{G}_j(t)$  is a strongly consistent estimator of  $G_j(t)$ . Lemma 2.1 implies that the partial derivatives of  $h_\phi$  are continuous. Thus, it follows from Proposition 2.2 that  $\hat{I}_j(t)$  is a strongly consistent estimator of  $I_j(t)$ . The proof follows.  $\square$

It will often be more appropriate to consider a parametric formulation of the model, that is,  $F_j(t) = F_j(t; \theta_j)$ , where  $\theta_j \in \Theta_j \subset R^{p_j}$ ,  $j = 1, \dots, m$ . Let  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_m)$  be the maximum likelihood estimate of  $\theta = (\theta_1, \dots, \theta_m)$ . It would be of interest to develop an asymptotic theory for  $\hat{\theta}$ . Some results along this direction are obtained in Doss, Freitag and Proschan (1987), following the approach laid out in Borgan (1984) for parametric counting process models.

**3. Estimation of the reliability importance of components.** The quantity  $I_j(t)$  defined by (1.7) is called the reliability importance of component  $j$  at time  $t$ . Its natural estimate is  $\hat{I}_j(t)$  given by (2.42). Let  $\varepsilon_1, \dots, \varepsilon_m$  be small numbers. Note that

$$h_\phi(\bar{F}_1(t) + \varepsilon_1, \dots, \bar{F}_m(t) + \varepsilon_m) - h_\phi(\bar{F}_1(t), \dots, \bar{F}_m(t)) \doteq \sum_{j=1}^m \varepsilon_j I_j(t).$$

Thus, the reliability importance of components may be used to evaluate the effect of an improvement in component reliability on system reliability, and can therefore be very useful in system analysis in determining those components on which additional research can be most profitably expended. For details, see pages 26–28 of Barlow and Proschan (1981), and the review by Natvig (1987).

**PROPOSITION 3.1.** *Suppose  $F_1, \dots, F_m$  are continuous and  $T > 0$  is such that  $F_j(T) < 1$ ,  $j = 1, 2, \dots, m$ . Then*

$$\sqrt{n}(\hat{I}_1 - I_1, \dots, \hat{I}_m - I_m) \rightarrow_d (Y_1, \dots, Y_m),$$

where  $(Y_1, \dots, Y_m)$  is a vector of mean 0 Gaussian processes whose covariance

structure is given by

$$(3.1) \quad \text{Cov}(Y_{j_1}(t_1), Y_{j_2}(t_2)) = \sum_{k=1}^m \left( \frac{\partial^2 h}{\partial u_{j_1} \partial u_k} \Big|_{u=\bar{F}(t_1)} \right) \left( \frac{\partial^2 h}{\partial u_{j_2} \partial u_k} \Big|_{u=\bar{F}(t_2)} \right) \\ \times \bar{F}_k(t_1) \bar{F}_k(t_2) \int_0^{t_1} \frac{dF_k(u)}{\bar{H}_k(u) \bar{F}_k(u)}$$

for  $0 \leq t_1 \leq t_2 \leq T$  and  $j_1, j_2 = 1, \dots, m$ .

As before, the covariance terms in (3.1) can be estimated consistently, enabling the construction of confidence intervals for  $I_j(t)$ .

The proof of Proposition 3.1 is similar to that of Theorem 2 and is omitted.

**4. Efficiency of  $\hat{F}$  vs. the naive estimator  $\hat{F}^{\text{emp}}$ .** Define the asymptotic relative efficiency of  $\hat{F}$  vs.  $\hat{F}^{\text{emp}}$  [see (1.1)] at time  $t$  to be the ratio of the asymptotic variance of  $\hat{F}^{\text{emp}}(t)$  to that of  $\hat{F}(t)$ ,

$$(4.1) \quad \text{ARE}(t) = F(t) \bar{F}(t) \Big/ \sum_{j=1}^m I_j^2(t) \bar{F}_j^2(t) \int_0^t \frac{dF_j(u)}{\bar{H}_j(u) \bar{F}_j(u)}.$$

We assume implicitly that  $F_1, \dots, F_m$  are continuous. In (4.1) we assume that  $t$  is such that the denominator is not 0. The condition  $F_j(t) \in (0, 1)$  for all  $j$  is sufficient (but not necessary) to ensure this. Let  $\tau_j = \sup\{t; F_j(t) < 1\}$ , and let  $\tau = \max_j \tau_j$ . Thus,  $\tau \in (0, \infty]$ . It is difficult to study  $\text{ARE}(\cdot)$  since this quantity depends on the system as well as on  $F_1, \dots, F_m$  and on  $t$ . We shall consider in some detail the cases of series and parallel systems, since these are often considered extreme cases in coherent structure theory. Indeed, in our situation series systems give rise to maximum possible censoring, while parallel systems give minimum possible censoring (i.e., no censoring at all). Thus, a study of these special cases will give insight on the behavior of  $\text{ARE}(\cdot)$ .

For series systems,

$$(4.2) \quad \hat{F} = \hat{F}^{\text{emp}}.$$

The validity of (4.2) for a series system of two components is well-known [cf. the argument leading to (2.6) of Efron (1981)] and the extension to general series systems presents no difficulty. Thus, there is no added advantage gained by considering the component failure times when estimating the life distribution of a series system.

We now consider parallel systems. The reliability function is

$$(4.3) \quad h(p_1, \dots, p_m) = 1 - \prod_{j=1}^m (1 - p_j), \quad 0 \leq p_j \leq 1,$$

and since there is no censoring,  $\bar{H}_j(u) = \bar{F}_j(u)$ . Thus, from (1.8) the asymptotic variance of  $\hat{F}(t)$  is

$$(4.4) \quad \text{A.Var } \hat{F}(t) = \sum_{j=1}^m \left( \prod_{k \neq j} F_k^2(t) \right) F_j(t) \bar{F}_j(t).$$

This formula can also be obtained directly since the absence of censoring makes Theorems 1 and 2 trivial. The restriction of  $t$  to the interval  $[0, T]$  where  $T$  satisfies  $F_j(T) < 1$  for all  $j$  is superfluous. From (4.4) we see that

$$(4.5) \quad \text{ARE}(t) = \frac{\bar{F}(t)/F(t)}{\sum_{j=1}^m \bar{F}_j(t)/F_j(t)}.$$

PROPOSITION 4.1. *For parallel systems,*

$$(4.6) \quad (i) \text{ ARE}(t) \geq 1$$

for all  $t$  such that  $F_j(t) > 0$  and  $t < \tau$ ; the inequality is strict if there are at least two values of  $j$  such that  $F_j(t) < 1$  (i.e., at time  $t$ , we do effectively have a system of at least 2 components).

$$(ii) \lim_{t \nearrow \tau} \text{ARE}(t) = 1.$$

(iii) Assume that  $F_j(t) > 0$  for all  $t > 0$  and all  $j$ . Then

$$\lim_{t \rightarrow 0} \text{ARE}(t) = \infty.$$

PROOF. Let  $f(x) = e^x - 1$ . Since  $f$  is convex and satisfies  $f(0) = 0$ , it is (strictly) superadditive on  $[0, \infty)$ . Hence

$$(4.7) \quad \begin{aligned} \sum_{j=1}^m \frac{\bar{F}_j(t)}{F_j(t)} &= \sum_{j=1}^m f(-\log F_j(t)) \\ &\leq f\left(-\log \prod_{j=1}^m F_j(t)\right) \\ &= \frac{1}{\prod_{j=1}^m F_j(t)} - 1. \end{aligned}$$

This shows that

$$(4.8) \quad \sum_{j=1}^m \frac{\bar{F}_j(t)}{F_j(t)} \leq \frac{1 - \prod_{j=1}^m F_j(t)}{\prod_{j=1}^m F_j(t)},$$

which proves (4.5) and hence (i).

Since  $f'(0) = 1 \neq 0$ ,

$$\lim_{x_1, \dots, x_m \rightarrow 0} \frac{\sum_{j=1}^m f(x_j)}{f(\sum_{j=1}^m x_j)} = \lim_{x_1, \dots, x_m \rightarrow 0} \frac{\sum_{j=1}^m (x_j f'(0))}{(\sum_{j=1}^m x_j) f'(0)} = 1,$$

which proves (ii).

Part (iii) is a special case of Proposition 4.2 below, which states that the efficiency of the naive estimator  $\hat{F}^{\text{emp}}(t)$  can be arbitrarily low for systems where all the components are *inessential*:  $\varphi(0_j, 1) = 1$  for all  $j$ ; that is, failure of any particular component does not necessarily cause system failure. Examples of

such systems include  $k$  out of  $m$  systems for  $k \leq m - 1$ . We will assume that  $F_j(t) > 0$  for all  $t$  and all  $j$ ; this is to avoid problems involving division by 0.

**PROPOSITION 4.2.** *If all components are inessential, then*

$$\lim_{t \rightarrow 0} \text{ARE}(t) = \infty.$$

**PROOF.** It is more convenient to work with  $(\text{ARE}(t))^{-1}$ . From (4.1) we have

$$(\text{ARE}(t))^{-1} = \frac{\sum_{j=1}^m I_j^2(t) \bar{F}_j^2(t) \int_0^t [dF_j(u) / \bar{H}_j(u) \bar{F}_j(u)]}{F(t) \bar{F}(t)}.$$

Let  $\varepsilon > 0$ , and let  $\delta > 0$  be such that if  $t \in (0, \delta)$ , then  $\bar{H}_j(t) > 1 - \varepsilon$  for all  $j$ . Then, we also have  $\bar{F}(t) > 1 - \varepsilon$  and  $\bar{F}_j(t) > 1 - \varepsilon$  for all  $j$ .

Consider the expression

$$A(t) = \frac{\sum_{j=1}^m I_j^2(t) F_j(t)}{F(t)}.$$

It is easy to see that for  $t \in (0, \delta)$ ,

$$A(t)(1 - \varepsilon)^2 \leq (\text{ARE}(t))^{-1} \leq A(t)(1 - \varepsilon)^{-3}.$$

Thus, we will work with the simpler expression  $A(t)$ . Let

$$A_j(t) = \frac{I_j^2(t) F_j(t)}{F(t)}.$$

We will show that for each  $j$ ,  $A_j(t) \rightarrow 0$  as  $t \rightarrow 0$ . We now consider  $A_j(t)$ , and to ease the notation, assume without loss of generality that  $j = 1$ . Using (2.7) and (2.6), we have

$$\begin{aligned} (4.9) \quad A_1(t) &= \left\{ \sum_{U \in A_{m-1}} (\varphi(1, U) - \varphi(0, U)) \psi(U; t) \right\}^2 F_1(t) \\ &\div \left( \left\{ \sum_{U \in A_{m-1}} (1 - \varphi(1, U)) \psi(U; t) \bar{F}_1(t) \right\} \right. \\ &\quad \left. + \left\{ \sum_{U \in A_{m-1}} (1 - \varphi(0, U)) \psi(U; t) F_1(t) \right\} \right), \end{aligned}$$

where

$$(4.10) \quad \psi(U; t) = \prod_{j=2}^m \bar{F}_j(t)^{U_j} F_j(t)^{1-U_j}.$$

In (4.9) and (4.10),  $U$  is in  $A_{m-1}$  and the vector  $(\alpha, u)$  is in  $A_m$ , for  $\alpha = 0$  or  $1$ . Let  $\mathcal{C} = \{U \in A_{m-1}; \varphi(1, U) = 1 \text{ and } \varphi(0, U) = 0\}$ . Then, clearly, the numerator of (4.9) ranges only over  $U \in \mathcal{C}$ . Since deleting the sum inside the first set of

braces in the denominator of (4.9) increases the ratio (4.9), we obtain that

$$(4.11) \quad A_1(t) \leq \frac{\{\sum_{U \in \mathcal{C}} \psi(U; t)\}^2}{\sum_{U \in A_{m-1}} (1 - \varphi(0, U)) \psi(U; t)}.$$

Note that if  $U \in \mathcal{C}$ , then  $\varphi(0, U) = 0$ ; therefore, from (4.11) we have

$$(4.12) \quad A_1(t) \leq \frac{\{\sum_{U \in \mathcal{C}} \psi(U; t)\}^2}{\sum_{U \in \mathcal{C}} \psi(U; t)}.$$

Let  $U \in \mathcal{C}$ , and consider now  $\psi(U; t)$ . Note that since component 1 is an inessential component, the condition  $\varphi(0, U) = 0$  implies that there exists a  $j \in \{2, \dots, m\}$  such that  $U_j = 0$ . Therefore,  $\psi(U; t) \leq \varepsilon$ . This fact, together with (4.12) implies that  $A_1(t)$  is bounded above by  $\varepsilon$  times the cardinality of  $\mathcal{C}$ . Since  $\varepsilon$  was arbitrary, this completes the proof.  $\square$

Let us consider the implications of Proposition 4.2. Suppose that the life testing experiment is carried out over limited time span. If the system is made highly reliable through a great deal of redundancy, then even though we will see many component failures, we will see very few actual system failures over the time span of the experiment. Thus, information from component failures becomes important. In this case the cost of designing and running an experiment which permits continuous monitoring of component failures may be far outweighed by the greater accuracy gained by using the estimator  $\hat{F}$ .

Since, as was mentioned earlier, series and parallel systems are considered extreme cases in reliability theory, we are led to conjecture that for arbitrary coherent systems:

(i) Let  $I \subset (0, \infty)$  be the set of all  $t$  such that the denominator of (4.1) is nonzero. Then  $\text{ARE}(t) \geq 1$  for all  $t \in I$ , with equality for all  $t \in I$  only for series systems.

(ii)  $\lim_{t \nearrow \sup I} \text{ARE}(t) = 1$ .

(iii)  $\text{ARE}(\cdot)$  is monotonically decreasing.

The estimator  $\hat{F}$  may be thought of as the nonparametric maximum likelihood estimator (NPMLE) of  $F$ . Informally, this is because of the well-known result that the Kaplan–Meier estimate is the NPMLE of a distribution function when the data are right-censored. Thus for each  $j$ ,  $\hat{F}_j$  is the NPMLE of  $F_j$ ; an extension of this is that  $\hat{F}_1 \times \dots \times \hat{F}_m$  is the NPMLE of  $F_1 \times \dots \times F_m$  (the  $\times$  denotes the product measure). The invariance principle for maximum likelihood estimates implies that  $h_\phi(\hat{F}_1, \dots, \hat{F}_m)$  is the NPMLE of  $h_\phi(\bar{F}_1, \dots, \bar{F}_m)$ . Of course, this by itself does not in any way imply any asymptotic optimality result.

**5. Related observational schemes.** In Theorem 1 we assumed that the component lifelengths are censored by system lifelength. The fact that the censoring process  $C_{i,j}(t) = I(Y_{i,j} \geq t)$  is predictable was key to the development of our results. It is therefore clear that our treatment can be used to approach other problems which involve predictable censoring of  $X_{i,j}$ .



Examples of predictable censoring include experiment termination after a prespecified time, censoring by independent outside causes of the lifetimes of certain components, experiment termination after a prespecified number of observed failures of component  $j$ , etc. An especially interesting class of examples is given by systems in which the failure of a subsystem prevents further monitoring of any of the components comprising the subsystem. For component  $j$  in system  $i$ , let  $T_i^{(j)}$  denote the lifetime of the subsystem whose failure would prevent further monitoring of component  $j$ . Then  $X_{ij}$  is censored by  $T_i \wedge T_i^{(j)}$ . The proof that for each  $j$  the censoring process is  $\mathcal{F}_t$ -predictable is only notationally different from the proof of part (ii) of Proposition A.1. The processes corresponding to (2.14) are still orthogonal martingales with respect to  $\{\mathcal{F}_t\}$ . The weak convergence result is the same except for the calculation of the  $\bar{H}_j$ 's. Formulas for  $\bar{H}_j$  can be readily obtained by using the concept of a modular decomposition of a coherent system [see pages 16–17 of Barlow and Proschan (1981)]. These formulas are, however, notationally rather messy since in general a component may be a member of several subsystems and one therefore has to specify for each  $j$  the subsystem whose failure would prevent further monitoring of component  $j$ . It should be kept in mind that precise formulas for the asymptotic distribution are not necessary for the construction of confidence intervals: these are obtained through Greenwood's formula (cf. Lemma 2.3).

Consider now Meilijson's (1981) model. Let  $D_i$  be the set of components that are found dead at the autopsy of system  $i$ . The data is then  $\{(T_i, D_i); i = 1, \dots, n\}$ . The censoring of  $X_{ij}$  by  $T_i$  is now quite complex: there is not only both right-censoring ( $j \notin D_i$ ) and left-censoring ( $j \in D_i$ ), but if  $j \in D_i$  we do not know if  $X_{ij} < T_i$  or if  $X_{ij} = T_i$  (component  $j$  caused system death). The distributions  $F_1, \dots, F_m$  are not necessarily identifiable from the distribution of  $(T_i, D_i)$ . For example, for parallel systems,  $D_1$  is always  $\{1, \dots, m\}$ ; it is clear that  $F_1, \dots, F_m$  cannot be identified from  $\prod_{j=1}^m F_j$ , the distribution of  $T_1$ . Meilijson gave conditions on the system and on  $F_1, \dots, F_m$  for identifiability. The problem of estimating  $F$  from the data is very interesting, but does not seem amenable to our approach.

## APPENDIX

**Random censorship.** In Section 2 we assumed the existence of a censoring random variable  $Y_j$  that satisfies (2.1), (2.2) and (2.4). Here we define  $Y_j$ , formally prove that it satisfies (2.1), (2.2) and (2.4), and prove that censoring is  $\mathcal{F}_t$ -predictable, where  $\mathcal{F}_t$  is defined by (2.23). Define the binary function  $\phi_j$  by

$$(A.1) \quad \phi_j(u_1, \dots, u_m) = \phi(1_j, u_1, \dots, u_m), \quad u_k = 0, 1, k = 1, 2, \dots, m,$$

where  $\phi$  is the structure function. [See the paragraph preceding (2.6).] The censoring random variable  $Y_j$  is defined as

$$(A.2) \quad Y_{ij} = \sup\{t; \phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\}.$$

**PROPOSITION A.1.** (i) For each  $j$ ,  $Y_{1j}, Y_{2j}, \dots$ , are i.i.d. random variables satisfying (2.1), (2.2) and (2.4).

(ii) The censoring process defined by  $C_{ij}(t) = I(Y_{ij} \geq t)$  is  $\mathcal{F}_t$ -predictable.

**PROOF OF (i).** It follows from (A.2) that  $Y_{ij}$  is a function of the vector  $(1_j, I(X_{i1} > t), \dots, I(X_{im} > t))$ . Thus it follows that  $Y_{1j}, Y_{2j}, \dots$ , are i.i.d. and that  $Y_{ij}$  satisfies (2.2).

We proceed to prove (2.4). The structure function  $\phi$  is increasing in its arguments [see Definition 2.1, page 6 of Barlow and Proschan (1981)] and hence a fortiori  $\phi_j$  is increasing in its arguments. Thus

$$(A.3) \quad \{Y_{ij} > t\} = \{\phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\}$$

and so

$$(A.4) \quad P(Y_{ij} > t) = P(\phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1).$$

It is easy to see that the right-hand side of (A.3) is equal to  $h_\varphi(1_j, \bar{F}(t))$  and so  $Y_{ij}$  satisfies (2.4). To prove that  $Y_{ij}$  satisfies (2.1), we consider two cases:  $\delta_{ij} = 1$  and  $\delta_{ij} = 0$ . We first prove (2.1) for the case  $\delta_{ij} = 1$ . Since  $\phi$  is increasing in its arguments,

$$(A.5) \quad \begin{aligned} & \sup\{t; \phi(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\} \\ & \leq \sup\{t; \phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\}. \end{aligned}$$

It is clear that the left-hand side of (A.5) equals  $T_i$  and the right-hand side of (A.5) equals  $Y_{ij}$ . Hence

$$(A.6) \quad T_i \leq Y_{ij}.$$

Since  $\delta_{ij} = 1$ ,

$$(A.7) \quad X_{ij} \leq T_i.$$

It is immediate from (A.6) and (A.7) that  $X_{ij} \leq Y_{ij}$ , which implies that (2.1) holds for this case. We now prove that (2.1) is satisfied if  $\delta_{ij} = 0$ . Since  $\delta_{ij} = 0$ , it follows that  $X_{ij} > T_i = Z_{ij}$ . Hence  $0 = \phi_j(I(X_{i1} > Z_{ij}), \dots, I(X_{im} > Z_{ij}))$ . Thus it follows from (A.2) that

$$(A.8) \quad Y_{ij} \leq Z_{ij}.$$

It is easy to see that (A.6) holds for this case. Thus  $Y_{ij} = Z_{ij}$ , which implies that (2.1) is satisfied for this case.  $\square$

**PROOF OF (ii).** From (A.3) we see that  $\{Y_{ij} > t\} \in \mathcal{F}_t$ , and this implies that  $\{Y_{ij} \geq t\} \in \mathcal{F}_t$ . The left-continuity of  $C_{ij}$  now gives the  $\mathcal{F}_t$ -predictability.  $\square$

**Acknowledgments.** We are very grateful to an Associate Editor whose suggestions led to a great simplification of our original proof of Theorem 1. We also thank Referee 1 for the present simple proof of parts (i) and (ii) of Proposition 4.1, and Referee 2 and Ian McKeague for helpful comments.

## REFERENCES

- AALLEN, O. (1976). Nonparametric inference in connection with multiple decrement models. *Scand. J. Statist.* **3** 15–27.
- AALLEN, O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6** 701–726.
- AALLEN, O. and JOHANSEN, S. (1978). An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scand. J. Statist.* **5** 141–150.
- ANDERSEN, P. K. and BORGAN, Ø. (1985). Counting process models for life history data: A review (with discussion). *Scand. J. Statist.* **12** 97–158.
- ANDERSEN, P. K., BORGAN, Ø., GILL, R. D. and KEIDING, N. (1982). Linear nonparametric tests for comparison of counting processes, with applications to censored survival data (with discussion). *Internat. Statist. Rev.* **50** 219–258. Correction **52** 225.
- APOSTOL, T. M. (1964). *Mathematical Analysis*. Addison-Wesley, Reading, Mass.
- BARLOW, R. E. and PROSCHAN, F. (1981). *Statistical Theory of Reliability and Life Testing*. To Begin With, Silver Spring, Md.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- BORGAN, Ø. (1984). Maximum likelihood estimation in parametric counting process models, with applications to censored failure time data. *Scand. J. Statist.* **11** 1–16. Correction **11** 275.
- BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimators under random censorship. *Ann. Statist.* **2** 437–453.
- DAVIS, M. H. A. (1983). The martingale theory of point processes and its application to the analysis of failure-time data. In *Electronic Systems Effectiveness and Life Cycle Costing*. NATO ASI Series (J. K. Skwirzinski, ed.) **F3** 135–155.
- DOSS, H., FREITAG, S. and PROSCHAN, F. (1987). Estimating jointly system and component life distributions using a mutual censorship approach. Technical Report M717R, Dept. Statistics, Florida State Univ.
- EFRON, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.* **76** 312–319.
- FÖLDES, A. and REJTÖ, L. (1981). A LIL type result for the product limit estimator. *Z. Wahrsch. verw. Gebiete* **56** 75–86.
- FÖLDES, A., REJTÖ, L. and WINTER, B. B. (1980). Strong consistency properties of nonparametric estimators for randomly censored data. I. *Period. Math. Hungar.* **11** 233–250.
- GILL, R. D. (1980). *Censoring and Stochastic Integrals*. Mathematical Centre Tracts **124**. Mathematisch Centrum, Amsterdam.
- GILL, R. D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.* **11** 49–58.
- HALL, W. J. and WELLNER, J. A. (1980). Confidence bands for a survival curve from censored data. *Biometrika* **67** 133–143.
- JOHANSEN, S. (1978). The product limit estimator as maximum likelihood estimator. *Scand. J. Statist.* **5** 195–199.
- KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.
- MEILIJSON, I. (1981). Estimation of the lifetime distribution of the parts from the autopsy statistics of the machine. *J. Appl. Probab.* **18** 829–838.
- MILLER, R. G., JR. (1981). *Survival Analysis*. Wiley, New York.
- NATVIG, B. (1987). Reliability importance of components. *Encyclopedia of Statistical Sciences* **8** 17–19. Wiley, New York.
- REBOLLEDO, R. (1980). Central limit theorem for local martingales. *Z. Wahrsch. verw. Gebiete* **51** 269–289.

DEPARTMENT OF STATISTICS  
 FLORIDA STATE UNIVERSITY  
 TALLAHASSEE, FLORIDA 32306-3033