

AD-A183 355

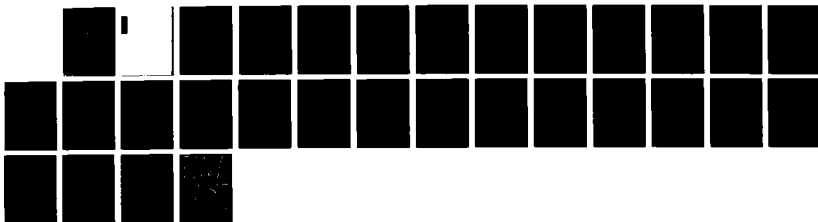
ESTIMATING JOINTLY SYSTEM AND COMPONENT RELIABILITIES  
USING A MUTUAL CENSORSHIP APPROACH(U) STANFORD UNIV CA  
DEPT OF STATISTICS H DOSS ET AL. 07 JUL 87 TR-393  
N00014-86-K-0156

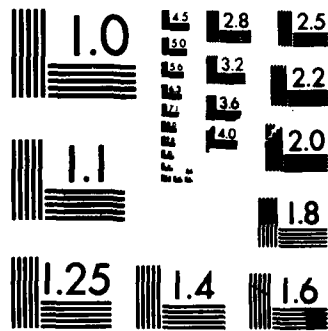
1/1

UNCLASSIFIED

F/G 12/3

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

ESTIMATING JOINTLY SYSTEM AND COMPONENT  
RELIABILITIES USING A MUTUAL CENSORSHIP APPROACH

BY

HANI DOSS, STEVEN FREITAG, and FRANK PROSCHAN

TECHNICAL REPORT NO. 393  
JULY 7, 1987

Prepared Under Contract  
N00014-86-K-0156 (NR-042-267)  
For the Office of Naval Research  
  
Herbert Solomon, Project Director

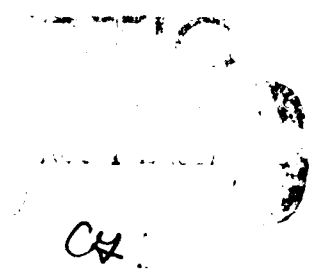
Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Reproduction in Whole or in Part is Permitted  
for any purpose of the United States Government



Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA



## 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

$T_i =$  lifelength of system  $i$ .

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

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

where  $I(A)$  is the indicator of the set  $A$ . However,  $\hat{F}^{\text{emp}}$  is unsatisfactory in that 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 paper 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 6) 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:

$X_{i,j}$  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 paper 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

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

(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  $F_j$ :

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

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

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

(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

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

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; 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 paper 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  $n \rightarrow \infty$

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

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

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

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

$$I_j(t) = \frac{\partial h_\phi(u_1, \dots, u_m)}{\partial u_j} \Big|_{u_j = F_j(t), j=1, \dots, m}. \quad (1.7)$$

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

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

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

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

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 multi-dimensional 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 paper 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 present results for a parametric formulation of our model. In Section 6 we discuss Meilijson's model as well as other models where the data can yield more information than just  $T_1, \dots, T_m$ . The appendix gives a proof of a result used in Sections 2 and 5.

## 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  $(X_{i_j}, Z_{i_j}, \delta_{i_j}, T_i)$  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 2.1.

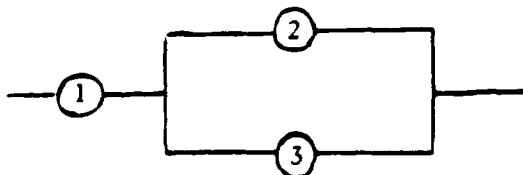


Figure 2.1.

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  $Y_1 = X_2 \vee X_3$ , 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 paper, 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

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

and

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

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

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

Let  $\underline{F}(t) = (F_1(t), \dots, F_m(t))$  and recall that  $H_j$  is the distribution of  $Z_j$ . In the appendix it is shown that

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

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

$$\hat{H}_j(t) = F_j(t) h_\phi(1_j, \underline{F}(t)). \quad (2.5)$$

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 stated of the components) for  $U \in A_m$ , where  $A_m = \{0, 1\}^m$ . It is well-known (and easy to see) that for  $p = (p_1, \dots, p_m) \in [0, 1]^m$ ,



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

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  $\hat{F}^n$  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),

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

From (2.7) we have

$$\frac{\partial^2 h_\phi}{\partial p_k^2} \Big|_p = 0, \quad (2.8)$$

and for  $l \neq k$ ,

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

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 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. ■

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

$$(a) \quad P\left(\sup_{0 \leq t \leq T} |\hat{F}(t) - F(t)| = o\left(\frac{(\ln n)^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right)\right) = 1.$$

(b) If  $F_1, \dots, F_m$  are continuous, the rate

$$o\left(\frac{(\ln n)^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right) \text{ may be replaced by } o\left(\sqrt{\frac{\ln \ln n}{n}}\right).$$

**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. ■

## 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  $\frac{0}{0} = 0$ . The index  $n$  defining a process is suppressed whenever possible.

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

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

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

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

(Now and henceforth, the symbol  $d$  signifies weak convergence in  $D^m[0, T]$ .) Theorem 1 is a 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

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

and show that

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

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

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

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

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.

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

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

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

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

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

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

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

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

We further define the following filtration:

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

In Section 6 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

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

In the present situation,

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

where  $Y_{i,j}$  is defined by (A.2).

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

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

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

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

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

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

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

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

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

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

We further define

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

The following proposition (Theorem 3.1 of Aalen and Johansen, 1978 and equation (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  $n$ ,

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

Our plan is to show that  $M_j^c(t), j = 1, \dots, m$ , are orthogonal martingales with respect to either  $\{\mathcal{F}_i; t \in [0, T]\}$  or  $\{\mathcal{F}_i^c; t \in [0, T]\}$ . 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 et al., 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,

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

**Proof:** From (2.32), (2.26), (2.30), (2.28), (2.21) and (2.19) it follows that

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

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, p. 703 for a discussion).

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 e.g. Davis (1983, pp. 136-7).

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 martingale. Since the sum of martingales is a martingale, this gives that

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

with

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

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 one (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 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

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

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

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

The proof of the weak convergence result for each individual  $n^{\frac{1}{2}}(\hat{F}_j - F_j^*)/\hat{F}_j^*$  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  $t \in [0, T]$ ,

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

The right 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 (see page 111 of Billingsley, 1968) to conclude that the process

$$n^{\frac{1}{2}} |\hat{F}(t) - F(t) - \sum_{j=1}^m I_j(t) (\hat{F}_j(t) - F_{j_1}(t))| \rightarrow 0 \text{ a.s. in } D[0, T].$$

Thus, the proof follows by showing that

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

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  $[0, \infty)$ .

$$G_j(t) = (\hat{F}_j(t))^2 \int_0^t \frac{dF_j(s)}{\hat{F}_j(s)\hat{H}_j(s)};$$

$$\hat{G}_j^n(t) = \frac{(\hat{F}_j(t))^2}{n} \int_0^t \frac{dN_j^c(s)}{\hat{H}_j(s-)\hat{H}_j(s)} \quad \left( = n(\hat{F}_j(t))^2 \sum_{i: Z_{(i)j} \leq t} \frac{\delta_{(i)j}}{(n-i+1)(n-i)} \right),$$

where

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

Also define

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

The quantity  $\frac{\hat{G}_j(t)}{n}$  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. ■

### 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  $\epsilon_1, \dots, \epsilon_m$  be small numbers. Note that

$$h_\phi(\hat{F}_1(t) + \epsilon_1, \dots, \hat{F}_m(t) + \epsilon_m) - h_\phi(\hat{F}_1(t), \dots, \hat{F}_m(t)) = \sum_{j=1}^m \epsilon_j I_j(t).$$

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 (1984).

**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

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

where  $(Y_1, \dots, Y_m)$  is a vector of mean zero Gaussian processes whose covariance structure is given by

$$\begin{aligned} \text{Cov}(Y_{j_1}(t_1), Y_{j_2}(t_2)) &= \sum_{k=1}^m \left( \frac{\partial^2 h}{\partial u_{j_1} \partial u_k} \Big|_{u=\mathbf{F}(t_1)} \right) \left( \frac{\partial^2 h}{\partial u_{j_2} \partial u_k} \Big|_{u=\mathbf{F}(t_2)} \right) \\ &\quad \bar{F}_k(t_1) \bar{F}_k(t_2) \int_0^{t_1} \frac{dF_k(u)}{\bar{H}_k(u) \bar{F}_k(u)}, \quad \text{for } 0 \leq t_1 \leq t_2 \leq T \text{ and } j_1, j_2 = 1, \dots, m. \end{aligned}$$

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 relation 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  $\hat{F}(t)$ :

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

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 insure 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  $ARE(\cdot)$  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  $ARE(\cdot)$ .



For series systems,

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

The validity of (4.2) for a series system of two components is well known (cf. the argument leading to equation (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

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

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

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

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

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

**PROPOSITION 4.1.** For parallel systems,

$$(i) \quad ARE(t) \geq 1 \quad (4.6)$$

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).

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

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

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

Proof of (i): By (4.5), we see that (4.6) is equivalent to

$$\sum_{j=1}^m \frac{F_j(t)}{\bar{F}_j(t)} \leq \frac{1}{\prod_{j=1}^m F_j(t)} - 1. \quad (4.7)$$

We prove (4.7) by an induction argument on  $m$ . For  $m = 1$ , (4.7) holds trivially. Assume (4.7) for  $m - 1$ , and consider the expression

$$B = \frac{1}{\prod_{j=1}^m F_j(t)} - 1 - \sum_{j=1}^m \frac{F_j(t)}{F_j(t)}. \quad (4.8)$$

We use the induction hypothesis on the second sum in (4.8) to obtain that

$$B \geq \frac{1}{\prod_{j=1}^m F_j(t)} - \frac{F_m(t)}{F_m(t)} - \frac{1}{\prod_{j=1}^{m-1} F_j(t)}. \quad (4.9)$$

The right side of (4.9) is equal to

$$\left( \frac{1}{\prod_{j=1}^{m-1} F_j(t)} - 1 \right) \frac{F_m(t)}{F_m(t)} \quad (4.10)$$

which is nonnegative. This establishes the induction step, and consequently (4.7). If there are two values of  $j$  such that  $F_j(t) < 1$ , then without loss of generality assume that one of these is  $j = m$ . Then, (4.10) is strictly positive, and hence  $ARE(t) > 1$ .

Proof (ii): From (4.5) we see that

$$ARE(t) = \frac{(1 - \prod_{j=1}^m F_j(t))/F(t)}{\sum_{j=1}^m F_j(t)/F_j(t)}. \quad (4.11)$$

Let  $\epsilon > 0$ , and let  $\delta > 0$  be such that if  $t \in (\tau - \delta, \tau)$ , then  $F_j(t) < \epsilon$ , for  $j = 1, \dots, m$ . Now let  $t \in (\tau - \delta, \tau)$ , and consider first the expression

$$\widetilde{ARE}(t) = \frac{1 - \prod_{j=1}^m F_j(t)}{\sum_{j=1}^m F_j(t)}. \quad (4.12)$$

Substituting  $1 - F_j(t)$  for  $F_j(t)$  in the product on the right side of (4.12) and expanding this product, we see that

$$\widetilde{ARE}(t) = 1 + S \quad (4.13)$$

where  $S$  is a sum of terms of the form

$$\frac{F_{j_1}(t)F_{j_2}(t)\dots F_{j_\ell}(t)}{\sum_{j=1}^m F_j(t)}, \quad \ell \geq 2. \quad (4.14)$$

Each of the terms in (4.14) is less than  $\epsilon$ . The result now follows from the fact that

$$\widetilde{ARE}(t)(1 - \epsilon) \leq ARE(t) \leq \widetilde{ARE}(t)(1 - \epsilon)^{-m}.$$

Part (iii) is a special case of Proposition 4.2 below, which states that the efficiency of the naive estimator  $\hat{F}^{naive}(t)$  can be arbitrarily low for systems where all the components are *inessential*:  $\varphi(0_j, 1) = 1$  for all  $j$ ; i.e. 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

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

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

$$(ARE(t))^{-1} = \frac{\sum_{j=1}^m I_j^2(t) F_j^2(t) \int_0^t \frac{dF_i(u)}{H_i(u) F_i(u)}}{F(t) \bar{F}(t)}. \quad (4.15)$$

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

Consider the expression

$$A(t) = \frac{\sum_{j=1}^m I_j^2(t) F_j(t)}{F(t)}. \quad (4.16)$$

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

$$A(t)(1 - \epsilon)^2 \leq (ARE(t))^{-1} \leq A(t)(1 - \epsilon)^{-2}. \quad (4.17)$$

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

$$A_j = \frac{I_j^2(t) F_j(t)}{F(t)}. \quad (4.18)$$

We will show that for each  $j$ ,  $A_j(t) \rightarrow 0$  as  $t \rightarrow 0$ . We now consider  $A_1(t)$ , and to ease the notation, assume without loss of generality that  $j = 1$ . Using (2.7) and (2.6) we have

$$A_1(t) = \frac{\{\sum_{U \in A_{m-1}} (\varphi(1, U) - \varphi(0, U)) \psi(U; t)\}^2 F_1(t)}{\{\sum_{U \in A_{m-1}} (1 - \varphi(1, U)) \psi(U; t) F_1(t)\} + \{\sum_{U \in A_{m-1}} (1 - \varphi(0, U)) \psi(U; t) F_1(t)\}}, \quad (4.19)$$

where

$$\psi(U; t) = \prod_{j=2}^m F_j(t)^{U_j} F_j(t)^{1-U_j}. \quad (4.20)$$

In (4.19) and (4.20),  $U$  is in  $A_{m-1}$  and the vector  $(\alpha, u)$  is in  $A_m$ , for  $\alpha = 0$  or  $1$ . Let  $C = \{U \in A_{m-1}; \varphi(1, U) = 1 \text{ and } \varphi(0, U) = 0\}$ . Then, clearly, the numerator of (4.19) ranges only over

$U \in C$ . Since deleting the sum inside the first set of braces in the denominator of (4.19) increases the ratio (4.19), we obtain that

$$A_1(t) \leq \frac{\{\sum_{U \in \mathcal{C}} \psi(U;t)\}^2}{\sum_{U \in \mathcal{A}_{m-1}} (1 - \varphi(0,U)) \psi(U;t)}. \quad (4.21)$$

Note that if  $U \in C$ , then  $\varphi(0,U) = 0$ ; therefore, from (4.21) we have

$$A_1(t) \leq \frac{\{\sum_{U \in C} \psi(U;t)\}^2}{\sum_{U \in C} \psi(U;t)}. \quad (4.22)$$

Let  $U \in C$ , 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,  $\psi(U;t) \leq \epsilon$ . This fact, together with (4.22) implies that  $A_1(t)$  is bounded above by  $\epsilon$  times the cardinality of  $C$ . Since  $\epsilon$  was arbitrary, this completes the proof. ■

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  $ARE(t) \geq 1$  for all  $t \in I$ , with equality for all  $t \in I$  only for series systems.
- (ii)  $\lim_{t \nearrow \sup I} ARE(t) = 1$ .
- (iii)  $ARE(\cdot)$  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 product measure). The invariance principle for maximum likelihood estimates implies that

$h_\phi(\hat{F}_1, \dots, \hat{F}_m)$  is the NPMLE of  $h_\phi(F_1, \dots, F_m)$ . Of course, this by itself does not in any way imply any asymptotic optimality result.

## 5. PARAMETRIC MODELS.

In this section we consider a parametric formulation of our problem. We assume that  $F_j(t) = F_j(t; \theta_j^0)$  is absolutely continuous, with density  $f_j(t; \theta_j^0)$ , where  $\theta_j^0 = (\theta_{j_1}^0, \dots, \theta_{j_{r_j}}^0)$  belongs to an open subset  $\Theta_j$  of  $\mathbb{R}^{p_j}$ , for  $j = 1, \dots, m$ . The distributions  $F_j(t; \theta_j)$  are not assumed to come from the same parametric family. Denote

$$(Z_j, \delta_j) = \{(Z_{1j}, \delta_{1j}), \dots, (Z_{nj}, \delta_{nj})\} \quad (5.1)$$

and let

$$L_j(\theta_j) = \prod_{i=1}^n f_j(Z_{ij}; \theta_j)^{\delta_{ij}} F_j(Z_{ij}; \theta_j)^{1-\delta_{ij}}. \quad (5.2)$$

Let  $\hat{\theta}_j$  be the value of  $\theta_j$  maximizing  $L_j(\theta_j)$ . The main results of this section are that under certain regularity conditions on the parametric families  $F_j(t; \theta_j)$ , the estimators  $\hat{\theta}_j$  are individually asymptotically normal and furthermore asymptotically independent. The results follow from the development of Section 2 and the results of Borgan (1984).

Frequently in reliability it is possible to specify, for certain indices  $j_1, \dots, j_h$ , a relationship among  $F_{j_1}, \dots, F_{j_h}$ , for example  $F_{j_1} = \dots = F_{j_h}$ . In such situations it may be possible to estimate  $\theta_{j_1}$  from the combined data  $(Z_{j_1}, \delta_{j_1}), \dots, (Z_{j_h}, \delta_{j_h})$ . This is the case for example for parallel systems of identically distributed components. We do not assume any known relationships among the  $F_j$ 's. It is also possible to specify parametric models for certain  $F_j$ 's but not for others. We do not consider this either. The techniques used below will, however, indicate how such problems may be addressed.

Let  $\mathcal{L}_j(\theta_1, \dots, \theta_m)$  be the (full) likelihood of  $(Z_j, \delta_j)$ . By the random censorship established by (2.1) and (2.2), we have

$$\mathcal{L}_j(\theta_1, \dots, \theta_m) = L_j(\theta_j) \tilde{L}_j(\theta_1, \dots, \theta_m) \quad (5.3)$$

where

$$\tilde{L}_j(\theta_1, \dots, \theta_m) = \prod_{i=1}^n \left\{ \left[ \frac{\partial}{\partial t} h_\phi(1, F_1(Z_{ij}; \theta_1), \dots, F_m(Z_{ij}; \theta_m)) \right]^{1-\delta_{ij}} \right. \\ \left. \left[ h_\phi(1, F_1(Z_{ij}; \theta_1), \dots, F_m(Z_{ij}; \theta_m)) \right]^{\delta_{ij}} \right\} \quad (5.4)$$

with  $h_\phi(1, \underline{F}(t))$  given by (2.4). Since we do not assume any known relationships among the  $F_j(t; \theta_j)$ 's, it follows that  $\tilde{L}_j$  does not depend on  $\theta_j$ . Thus, the only important part of the likelihood

is  $L_j(\theta_j)$ . (The condition that  $\bar{L}_j$  not depend on  $\theta_j$  is referred to in the literature as "noninformative censoring". If  $\bar{L}_j$  does depend on  $\theta_j$ , then  $L_j$  is viewed as a "partial likelihood"; see Chapter 5 of Kalbfleisch and Prentice, 1980).

Letting  $\alpha_j(t; \theta_j) = f_j(t; \theta_j) / \bar{F}_j(t; \theta_j)$  denote the hazard rate, (5.2) may be rewritten as

$$L_j(\theta_j) = \prod_{i=1}^n \{ \alpha_j(Z_{ij}; \theta_j)^{\delta_{ij}} \exp(-\int_0^{Z_{ij}} \alpha_j(u; \theta_j) du) \}. \quad (5.5)$$

By (2.27) and (2.29), we have

$$\log L_j(\theta_j) = \int_0^\infty \log \alpha_j(u; \theta_j) dN_j^c(u) - \int_0^\infty \alpha_j(u; \theta_j) V_j^c(u) du. \quad (5.6)$$

Let

$$U_{jh}(t; \theta_j) = \int_0^t \frac{\partial}{\partial \theta_{jh}} \alpha_j(u; \theta_j) / \alpha_j(u; \theta_j) dN_j^c(u) - \int_0^t \frac{\partial}{\partial \theta_{jh}} \alpha_j(u; \theta_j) V_j^c(u) du, \quad (5.7)$$

for  $j = 1, \dots, m, \quad k = 1, \dots, p_j$ .

Thus, differentiating under the integral sign in (5.6) (we will impose regularity conditions later) we obtain

$$\frac{\partial}{\partial \theta_{jh}} \log L_j(\theta_j) = U_{jh}(\infty; \theta_j). \quad (5.8)$$

Referring to equations (2.31) and (2.33), we have

$$U_{jh}(t; \theta_j) = \int_0^t \left( \frac{\partial}{\partial \theta_{jh}} \alpha_j(u; \theta_j) / \alpha_j(u; \theta_j) \right) dM_j^c(u), \quad j = 1, \dots, m, \quad k = 1, \dots, p_j. \quad (5.9)$$

Borgan (1984) shows that for fixed  $j$ , under certain regularity conditions (see below) if we assume only that

$$(M_j^c(t), \mathcal{F}_t) \text{ is a locally square integrable martingale with } \langle M_j^c(\cdot), M_j^c(\cdot) \rangle(t) = A_j^c(t), \quad (5.10)$$

then with probability tending to 1, the equations

$$U_{jh}(\infty; \theta_j) = 0, \quad k = 1, \dots, p_j, \quad (5.11)$$

have exactly one consistent solution (call it  $\hat{\theta}_j$ ), which is a local maximum of  $L_j(\theta_j)$ . Furthermore, this solution is asymptotically normal. In our setup  $L_j(\theta_j)$  is proportional to the likelihood of  $(Z_j, \delta_j)$ .

However, it is important to notice that the results concerning  $\hat{\theta}_j$  are valid whether or not this is the case.

Consider now Assumptions A-D of Borgan (1984) (with  $i = 1$ ). Assume that these hold for every  $j = 1, \dots, m$ , with  $Y_j(u)$  and  $y_j(u)$  in his paper given in our paper by  $V_j^c(u)$  and  $\tilde{H}_j(u)$ , respectively. Also make the notational change that  $t$  ranges over the positive axis instead of over  $[0, 1]$ .

**THEOREM 3** Under the above assumptions,

- (i) With probability tending to 1, for each  $j$  the likelihood equations (5.11) have exactly one consistent solution  $\hat{\theta}_j$ , and  $\hat{\theta}_j$  is a local maximum of  $L_j(\theta_j)$ .
- (ii)  $\sqrt{n}(\hat{\theta}_1 - \theta_1^0, \dots, \hat{\theta}_m - \theta_m^0)$  is asymptotically normal, with mean 0 and covariance matrix

$$\begin{pmatrix} \Sigma_1^{-1} & 0 & \dots & 0 \\ 0 & \Sigma_2^{-1} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \Sigma_m^{-1} \end{pmatrix}$$

where  $\Sigma_j^{-1}$  is the inverse of the  $p_j \times p_j$  matrix  $\Sigma_j$  whose  $k, \ell$ -th element is

$$\int_0^\infty \frac{\frac{\partial}{\partial \theta_{jk}} \alpha_j(u; \theta_j^0) \frac{\partial}{\partial \theta_{j\ell}} \alpha_j(u; \theta_j^0)}{\alpha_j(u; \theta_j^0)} \tilde{H}_j(u) du.$$

Furthermore,  $\Sigma_j$  may be consistently estimated by the matrix whose  $k, \ell$ -th element is

$$-\frac{1}{n} \frac{\partial^2}{\partial \theta_{jk} \partial \theta_{j\ell}} \log L_j(\theta_j) |_{\theta = \hat{\theta}_j}.$$

**Proof:** Our (somewhat sketchy) proof relies heavily on Theorems 1 and 2 of Borgan (1984). As was mentioned earlier, the starting point of Borgan's proof is the assumption that (5.10) above is valid. That this is the case follows by (2.38) and (2.39). Part (i) follows directly from Theorem 1 of Borgan (1984).

To obtain the distributional result of Part (ii) we proceed as in the proof of Theorem 2 of Borgan (1984), noting in addition that the orthogonality relationships given by (2.39) imply, via (5.9), that for  $j_1 \neq j_2$ ,  $k_1 = 1, \dots, p_{j_1}$ ,  $k_2 = 1, \dots, p_{j_2}$ , the locally square integrable martingales  $U_{j_1 k_1}(\cdot; \theta_{j_1})$  and  $U_{j_2 k_2}(\cdot; \theta_{j_2})$  are orthogonal. The distributional result follows by the argument in Billingsley (1961, Theorems 2.2 and 10.1). The consistency of the variance estimate follows directly from Theorem 2 of Borgan (1984).

## 6. RELATED OBSERVATIONAL SCHEMES.

In Theorem 1 (and also in Theorem 3) we assumed that the component lifelengths are censored by system lifelength. The fact that the censoring process  $C_{ij}(t) = I(Y_{ij} \geq t)$  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  $X_{ij}$ .

Examples of predictable censoring include experiment termination after a prespecified time, censoring of the lifelengths of certain components by independent outside causes, 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 lifelength 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  $\hat{H}_j$ 's. Formulas for  $\hat{H}_j$  can be readily obtained by using the concept of a modular decomposition of a coherent system (see pp. 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,  $D_i$  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

$$\phi_j(u_1, \dots, u_m) = \phi(1_j, u_1, \dots, u_m), \quad u_k = 0, 1, \quad k = 1, 2, \dots, m, \quad (A1)$$

where  $\phi$  is the structure function. (See the paragraph preceding equation (2.6).) The censoring random variable  $Y_{ij}$  is defined as follows:

$$Y_{ij} = \sup\{t; \phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\}. \quad (A2)$$

### 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

$$C_{ij}(t) = I(Y_{ij} \geq t) \text{ is } \mathcal{F}_t\text{-predictable.}$$

Proof of (i) : It follows from (A2) 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

$$\{Y_{ij} > t\} = \{\phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\} \quad (A3)$$

and so

$$P(Y_{ij} > t) = P(\phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1). \quad (A4)$$

It is easy to see that the right side of (A3) 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,

$$\begin{aligned} & \sup\{t; \phi(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\} \\ & \leq \sup\{t; \phi_j(I(X_{i1} > t), \dots, I(X_{im} > t)) = 1\}. \end{aligned} \quad (A5)$$

It is clear that the left side of (A5) equals  $T_i$  and the right side of (A5) equals  $Y_{ij}$ . Hence

$$T_i \leq Y_{ij}. \quad (A6)$$

Since  $\delta_{ij} = 1$ ,

$$X_{ij} \leq T_i. \quad (A7)$$

It is immediate from (A6) and (A7) 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 (A2) that

$$Y_{ij} \leq Z_{ij}. \quad (A8)$$

It is easy to see that (A6) hold for this case. Thus  $Y_{ij} = Z_{ij}$ , which implies that (2.1) is satisfied for this case.

Proof of (ii): From (A3) we see that

$$\{Y_{ij} > t\} \in \mathcal{F}_t, \text{ and this implies that } \{Y_{ij} \geq t\} \in \mathcal{F}_t.$$

The left continuity of  $C_{ij}$  now gives the  $\mathcal{F}_t$ -predictability.

## REFERENCES

- Aalen, O. (1976). Nonparametric inference in connection with multiple decrement models. *Scand. J. Statist.* **3**, 15-27.
- Aalen, O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6**, 701 - 726.
- Aalen, 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., Borgan, O., Gill, R. D. and Keiding, N. (1982). Linear non-parametric tests for comparison of counting processes, with applications to censored survival data (with discussion). *Int. Statist. Rev.* **50**, 219-258. Correction **52**, 225.
- Anderson, P. K. and Borgan, O. (1985). Counting process models for life history data: a review (with discussion). *Scand. J. Statist.* **12**, 97-158.
- Apostol, T.M. (1964). *Mathematical Analysis*. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts.
- Barlow, R. E. and Proschan, F. (1981). *Statistical Theory of Reliability and Life Testing. To Begin With*, Silver Springs, Maryland.
- Billingsley, P. (1961). *Statistical inference for Markov process*. The University of Chicago Press, Chicago.
- Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, Inc., 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, Vol F3, (J.K. Skwirziski, ed), 135-155.
- Efron, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.* **76**, 312-319.
- Földes, A. and Rejtő, L. (1981). A LIL type result for the product limit estimator. *Z. Wahrsch. verw. Gebiete* **56**, 75-86.
- Földes, A., Rejtő, L. and Winter, B.B. (1980). Strong consistency properties of nonparametric estimators for randomly censored data (Part I). *Periodica 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.
- Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. Wiley, New York.
- 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. Prob.* 18, 829-838.
- Miller, R. (1981). *Survival Analysis*. John Wiley & Sons, Inc., New York.
- Natvig, B. (1984). Reliability importance of components. Statistical Research Report No. 10, Institute of mathematics, University of Oslo. To appear in Vol. 6 of *Encyclopedia of Statistical Sciences*, S. Kotz and N. L. Johnson, eds. Wiley: New York.
- Rebolledo, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. verw. Gebiete* 51, 269-289.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 393	2. GOVT ACCESSION NO. <b>AD-A183355</b>	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Estimating Jointly System And Component Reliabilities Using A Mutual Censorship Approach	5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT	
7. AUTHOR(s) Hani Doss, Steven Freitag, and Frank Proshan.		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0156
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE July 7, 1987
		13. NUMBER OF PAGES 28
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Also issued as Statistics Department, Florida State University Technical Report, July 1987, under Contract No. AFOSR 85-C-0007		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) System life distribution, component life distribution, multivariate counting process, martingale central limit theorem		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) PLEASE SEE FOLLOWING PAGE.		

TECHNICAL REPORT NO. 393

## 20. • ABSTRACT

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.

END

9-87

Dtic