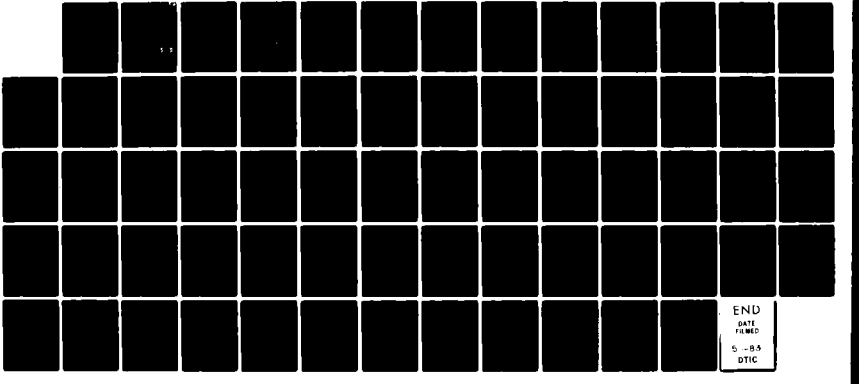
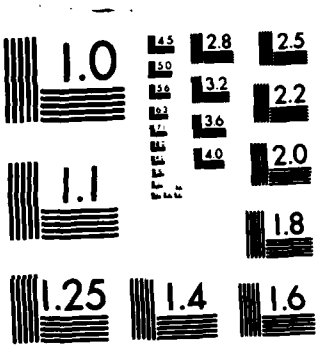
AD-A127 439 MULTIPLE STATE MODELS WITH COVARIATES: ANALYSIS AND SIMULATION(U) UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES DEPT OF ELECTRI.. V K MURTHY 01 APR 83 1/1

UNCLASSIFIED AFOSR-TR-83-0301 AFOSR-81-0162 F/G 12/1 NL





UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0301	2. GOVT ACCESSION NO. AD-A127439	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MULTIPLE STATE MODELS WITH COVARIATES: ANALYSIS AND SIMULATION		5. TYPE OF REPORT & PERIOD COVERED Final JUN 81-30 MAR 83
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) V.K. Murthy		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0162
9. PERFORMING ORGANIZATION NAME AND ADDRESS Electrical Engineering Department-Systems University of Southern California, University Park, Los Angeles CA 90089-0781		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Mathematical & Information Sciences Directorate Air Force Office of Scientific Research Bolling AFB DC 20332		12. REPORT DATE 1 APR 83
		13. NUMBER OF PAGES 62
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		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

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The research reported in this document deals with a multiple state, staged failure phenomenon with state specific covariates characteristic of the states. Different types of censoring was taken into consideration. The classical maximum likelihood estimation is compared with Cox's partial likelihood estimation for efficiency, since maximum likelihood estimation of covariate effects depends on the baseline failure rate, while partial likelihood estimation does not. Results of analysis as well as simulation are presented.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AFOSR-TR- 83-0301

Final Report

Department of the Air Force Grant AFOSR-81-0162

Multiple State Models With Covariates: Analysis and Simulation

Submitted by:

V. K. Murthy, Ph.D.
University of Southern California
Electrical Engineering Department-Systems
University Park
Los Angeles, California 90089-0781

April 1, 1983

Contents

1. Preface

2. Multiple State Cox Model. Analysis of Survival

3. Monte Carlo Comparison of Cox's Partial vs. Total Likelihood

Preface

The research reported in this document deals with a multiple state, staged failure phenomenon with state specific covariates characteristic of the states. Different types of censoring was taken into consideration. The classical maximum likelihood estimation is compared with Cox's partial likelihood estimation for efficiency, since maximum likelihood estimation of covariate effects depends on the baseline failure rate, while partial likelihood estimation does not. Results of analysis as well as simulation are presented.

Part of this work was supported at the University of Southern California by a grant from the Air Force Office of Scientific Research, Grant AFOSR-81-0162.

Multiple State Cox Model. Analysis of Survival

The vast majority of reliability analyses for components and/or systems assume that the equipment under consideration is in either one of two states: Functioning or Failed. In many situations one can measure and distinguish between various "levels of performance" for both the system and its components. Current literature is largely devoted to the probabilistic aspects of both static and dynamic multistate systems. Parallel investigations in clinical data analysis deal with reliability under the analogous concept of survival analysis.

In this section a model for the analysis of survival data of a staged failure phenomena with competing risks is considered. The model is a multiple state extension of the proportional hazard rate model of Cox (1972). It uses the theory of Markov processes as applied to staged failure phenomena described, for example, in Chiang (1968). Cox's method of partial likelihood was extended to obtain the estimates of covariate effects. Maximum likelihood estimation of the intensity matrix is developed. Using the results of Aalen (1975), the consistency of the estimator for the integrated intensity matrix is established. The asymptotic variance and the weak convergence to normality of the estimator is obtained. An illustration applying the model to a multiple transition situation is given.

1. Introduction

In the following, we consider the stochastic modeling of staged failure phenomena using the theory of Markov processes. Complicating the statistical picture is a set of observed covariates on each item sampled. The motivation for this work came from a desire to solve

certain problems that arise in the study of cancer data. Typically, for each patient in a cancer study, there exist a set of measured front end variables (e.g. initial white blood cell count, age, sex, etc.) as well as a set of periodically updated medical observations on each patient. The problem is to properly account for the effect of these observed covariates on the transition probabilities between various stages of the disease. This situation is most often modeled using just two states: disease and death. Cox (1972) considers this situation and uses the parametric model

$$\lambda_i(t) = \lambda_o(t) \exp\left[\beta' z_i(t)\right], \qquad (1.1)$$

where $\lambda_i(t)$ is the failure rate (hazard function) for the ith patient, $z_i(t)$ is the ith patient's observed set of covariates, and β' is the transpose of the vector representing the effect of the covariates. The function $\lambda_o(t)$ is a reference failure rate and may be considered as the failure rate of a patient with covariate vector $z_i(t) = 0$. If covariates are measured as deviations from a mean, then $\lambda_o(t)$ is an "average" failure rate. In a study where the determination of the effect of covariates is of paramount importance, then the method of partial likelihood (Cox 1975) is a very useful procedure for estimating the vector β in that it results in a formulation independent of the unknown $\lambda_o(t)$.

In this paper, we consider the multiple transition extension of a parametric covariate model in an analysis of life data (e.g. the Cox Model) using the theory of Markov processes applied to staged failure phenomena as described for example in Chiang (1968). The

failed states (e.g. death) in such a model will of course be absorbing states. Competing risk models can be handled by allowing for more than one death state where each death state corresponds to a different cause of death. For example, in the illustration discussed in Section 5, patients undergoing porta-caval shunt surgery to correct hepatic bleeding can at any particular time, either reside in one of two treatment states, (alive-free of hepatitis, or alive-having had hepatitis) or in one of three absorbing states (death from bleeding, death from other cause, or withdrawn-lost to follow-up).

2. STAGED FAILURE PHENOMENA: MARKOV PROCESSES

We consider the situation where each patient in a population must exist in one of several states. The states will be comprised of transient states (i.e. illness states such as various stages of a disease, various complications, etc.) and absorbing states (i.e. death states). Complicating the analysis of many clinical studies is the censored state, which here may be treated as another absorbing state. The Markov assumption is that the probability distribution over future states a patient might be in, is completely determined by his current state. Specifically it does not depend upon how the patient arrived at his current state or how long he has been there. (This statement can be modified if the state space is expanded to include these properties).

Let $A^{(o)}$ denote the set of mutually exclusive states i, $i \in A^{(o)}$. In our applications the cardinality of $A^{(o)}$ will be quite small (five for the liver patient example). The probability of being in

state j at time t given that one was in state i at time $\tau, \tau \leq t$ is denoted by the function $P_{ij}(\tau,t)$. The fact that a patient must occupy some state for each time t and that the states are mutually exclusive is characterized by the relations

$$\sum_j P_{ij}(\tau,t) = 1 \quad \text{for all} \quad \tau \leq t, \text{ and for all } i, \tag{2.1a}$$

and

$$P_{ij}(t,t) = \delta_{ij}, \tag{2.1b}$$

where δ_{ij} is the Kronecker delta.

The intensity matrix $\nu_{ij}(t)$ is given by

$$\nu_{ij}(\tau) = \frac{\partial}{\partial t} P_{ij}(\tau,t) \Big|_{t=\tau}. \tag{2.2}$$

From (2.1a), the intensity matrix satisfies

$$\sum_j \nu_{ij}(t) = 0 \quad \text{for all } i,\ t. \tag{2.3}$$

The intensity has the interpretation of being proportional to the probability of transition from state i to state j at time t. Specifically,

$$P_{ij}(t,t+\Delta) = \delta_{ij} + \nu_{ij}(t)\Delta + o(\Delta), \tag{2.4}$$

where $o(\Delta)$ is such that

$$\lim_{\Delta \to 0} o(\Delta)/\Delta = 0 \tag{2.5}$$

Letting $P(\tau,t)$ be the matrix $(P_{ij}(\tau,t))$ and $V(t)$ be the matrix $(\nu_{ij}(t))$, we have the following well known relationships between P and V.

I. Chapman-Kolmogrov forward differential equation:

$$\frac{\partial P(\tau,t)}{\partial t} = P(\tau,t)\ V(t);\ \tau \le t, \tag{2.6}$$

with initial condition $P(\tau,\tau) = I.$ (2.7)

II. Chapman-Kolmogrov backward differential equation:

$$\frac{\partial}{\partial \tau} P(\tau,t) = -V(\tau)\ P(\tau,t);\ \tau \le t \tag{2.8}$$

with "initial condition" $P(t,t) = I.$ (2.9)

The products PV in (2.6) and VP in (2.8) are the usual matrix inner products. The integral equation form of (2.7) to (2.9) is

$$P(\tau,t) = I + \int_{\tau}^{t} P(\tau,s)\ dB\ (\tau,s), \tag{2.10}$$

where $$B(\tau,t) = \int_{\tau}^{t} V(s)\ ds. \tag{2.11}$$

For a further discussion of the development of (2.6) through (2.9) see Chiang (1968), pp. 116-120.

In the usual application of the above, the intensity matrix $(\nu_{ij}(t))$ is assumed to be almost everywhere continuous so that equations (2.6) and (2.8) involve no irregularities. It will be seen later, however, that useful results are obtainable if we let $\nu_{ij}(t)$ have the form

$$\nu_{ij}(t) = \sum_{\ell=1}^{n_{ij}} a_{ij\ell}\ \delta(t-\xi_{ij\ell}), \tag{2.12}$$

where $\delta(\ )$ is the Dirac delta function and the $\xi_{ij\ell}$ are isolated points on the real line. (In particular, this is the natural "non parametric parameterization" of $\lambda_o(t)$ in (1.1). See Cox (1972)).

Equation (2.12) says that $i \to j$ transitions are only allowed to occur at the isolated points $\xi_{ij\ell}$, $\ell = 1, \ldots, n_{ij}$.

Computations are facilitated by ordering the $\xi_{ij\ell}$ for all $ij\ell$. Let the ordered $\xi_{ij\ell}$ be relabeled μ_k with $\mu_1 < \mu_2 < \ldots < \mu_m$. Let $P(\mu_k^-, \mu_k)$ be the probability transition matrix for the times $\mu_k = \xi_{ij\ell}$, $\mu_k^- = \mu_k - \varepsilon$ with $0 < \varepsilon < \mu_k - \mu_{k-1}$. It can be shown that $P(\mu_k^-, \mu_k)$ is the identity matrix except for the (i,j)th off diagonal element and the ith diagonal element which are given by

$$P_{ij}(\mu_k^-, \mu_k) = 1 - \exp(-a_{ij\ell}), \tag{2.12a}$$

and

$$P_{ii}(\mu_k^-, \mu_k) = \exp(-a_{ij\ell}), \tag{2.12b}$$

respectively. The Chapman-Kolmogrov equations imply for,

$$\mu_{k_1-1} \leq t_1 < \mu_{k_1} < \mu_{k_1+1} \ldots < \mu_{k_2} \leq t_2 < \mu_{k_2+1},$$

that

$$P(t_1, t_2) = \prod_{k=k_1}^{k_2} P(\mu_k^-, \mu_k). \tag{2.13}$$

The next section discusses the parametric estimation of the covariate effects. The following section discusses the nonparametric estimation of the transition probabilities themselves given knowledge of the covariate effects.

3. COVARIATE PARAMETERIZATION AND LIKELIHOOD ESTIMATION

From (2.7) or (2.8), it is seen that the transition probability matrix P is completely determined by the intensity matrix V. The parameterization of the covariate effect is thus facilitated through

a parameterization of V. Following (1.1), the Cox model parameterization for the pth patient for a transition from state i to state j is taken as

$$\nu_{ij}^{(p)}(t;\alpha_{ij}) = \nu_{ij}^{(o)}(t) \exp\left[\alpha_{ij}' z_p(t)\right], \quad i \neq j, \tag{3.1}$$

where α_{ij} is a vector of length M, the number of covariates in the vector $z_p(t)$. Note the constraint in (3.1) that i not equal j. For the case $i=j$, the parameterization of $\nu_{ij}^{(p)}(t)$ follows from (2.3). That is,

$$\nu_{ii}^{(p)}(t) = -\sum_{j \neq i} \nu_{ij}^{(p)}(t;\alpha_{ij}). \tag{3.2}$$

Likelihood estimation in both the case in which the $\nu_{ij}(t)$ are parameterized and the case in which they are left arbitrary is described in Fertig, Murthy, et al (1979). Here, we only consider the latter situation.

Let $W_p(t)$ be the right continuous state function for the pth patient. We observe $W_p(\tau)$ on $[0,T_p]$. Here we take T_p to be a nonrandom study time established apriori for the pth patient.

If the patient leaves the study because of death or is censored unexpectedly, he will be observed in that death or censored state until T_p. This approach explicitly takes into account a censored state that may be entered only randomly. Presumably, the transition probabilities of interest are those that would exist in the hypothetical situation of no censoring. Estimation of these probabilities is the thrust of the Kaplan-Meier estimator in the three state problem (life-death-censored). Fleming (1978) describes a multiple state extension to the *K-M* estimator that can be used in the current situation.

The state functions $W_p(\tau)$ are each seen to be step functions on $[0, T_p]$ with a finite number of jumps.

Cox's method of partial likelihood will be used to estimate the α_{ij} in (3.1). In this situation, one may specify as the conditioning variates $(S^{(j-1)}, X^{(j)})$, in Cox's notation, as respectively the risk set at each transition time (defined below) and the information that a specific transition occurred. With this approach we need to record: (1) Ordered event times: $t_1 < \ldots < t_m$. (2) Individuals undergoing transition: $i_1, \ldots, i_m$. (3) State being left: $x_1, \ldots, x_m$. (4) State being entered: $y_1, \ldots, y_m$. (5) "Risk" sets: $R_{x_1}(t_1), \ldots, R_{x_m}(t_m)$. The sets $R_{x_j}(t)$ for arbitrary t are defined by

$$R_{x_j}(t) = \left\{ p \,\middle|\, W_p(t^-) = x_j \right\}.$$

Each $R_{x_j}(t)$ contains as members all patients who are at "risk" of leaving state x_j.

Cox's partial likelihood is then

$$L = \prod_{j=1}^{M} \exp(\alpha' x_j y_j z_{i_j}) \Big/ \sum_{k \in R_{x_j}(t_j)} \exp(\alpha' x_j y_j z_k) \,. \quad (3.3)$$

The jth factor in the partial likelihood in a sense measures the likelihood that it is individual i_j who undergoes the transition $x_j \to y_j$ at time t_j given that some individual in $R_{x_j}(t_j)$ must undergo this transition. (This likelihood is written for the case in which there are no ties).

A useful feature of (3.3) is that it can be factored into independent expressions for each observed transition $x \to y$. Thus, the

estimation of α_{xy} for given (x,y) can be performed independently of the estimation of $\alpha_{x'y'}$ for $x \neq x'$ or $y = y'$. This factorization also holds for the information matrix based on (3.3). Specifically we have, for the qth component of α_{xy},

$$\frac{\partial \ln L}{\partial \alpha_{xyq}} = \sum_{t_j \in T_{xy}} \left[z_{i_j}^{(q)} - \sum_{k \in R_{x_j}(t_j)} z_k^{(q)} \exp(\alpha'_{xy} z_k) / \sum_{k \in R_{x_j}(t_j)} \exp(\alpha'_{xy} z_k) \right] \tag{3.4}$$

where $$T_{xy} = \left\{ t_j \mid x_j = x \text{ and } y_i = y \right\}.$$

For specified x,y the expression (3.4) for all q comprise the usual set of likelihood equations for estimating the q dimensional parameter vector of the Cox model when only two states x and y (besides the censoring state) are possible.

It can be seen from (3.4) that all off diagonal terms of the information matrix involving different transition states will be zero. Thus, the information matrix will be block diagonal with the (x,y) block given by

$$I_{xy}^{pq} = -\frac{\partial^2 \ln L}{\partial \alpha_{xy}(p)\, \partial \alpha_{xy}(q)} = \sum_{t_j \in T_{xy}} \left\{ \sum_{k \in R_{x_j}(t_j)} z_k^{(p)} z_k^{(q)} \exp(\alpha'_{ij} z_k) \right.$$

$$/ \sum_{K \in R_{x_j}(t_j)} \exp(\alpha'_{ij} z_k) - \left(\sum_{k \in R_{x_j}(t_j)} z_p^{(p)} \exp(\alpha'_{xy} z_k) \right)$$

$$\left. \left(\sum_{k \in R_{x_j}(t_j)} z_k^{(q)} \exp(\alpha'_{xy} z_k) / \left(\sum_{k \in R_{x_j}(t_j)} \exp(\alpha'_{xy} z_k) \right)^2 \right) \right\} \tag{3.5}$$

Of course, we only need consider those vectors α_{xy} and corresponding I_{xy}^{pq} for those transitions of interest. These ideas will be

made clearer in the next section.

Estimation of the transition probabilities require an estimation procedure for the $\nu_{ij}^{(o)}(t)$ for each $i \neq j$ of interest. Kaplan and Meier (1958) provide the nonparametric estimate in the three state problem with no covariates. Fleming (1978) gives the procedure for the multi-state case. Cox (1972) provides an adaptation of the Kaplan-Meier estimate when covariates are present. The estimator proposed here is the simple extension of the above procedures to the case where covariates exist in a multi-state environment. It is convenient to obtain the estimator as a maximum likelihood estimator in the space of all functions whose centers form a set with no cluster point. This procedure provides an estimator with certain intuitively nice properties. Specifically, the observed conditional frequencies of the sample will match those of the estimator's. Moreover, it reduces to the Kaplan-Meier estimate when there are only three states. It is recognized that the fact that the estimator is maximum likelihood may provide no advantage since we will be treating a case where the number of parameters and the number of patients on study grow at the same rate. Even though the estimation procedure seems to focus on the intensity matrix, the transition probabilities are of primary interest. This is fortunate since the presentation of a Dirac delta function as an estimate of intensity in its own right would be quite unsatisfactory in most situations. However, the transition probabilities essentially involve integrating over the estimated intensities and therefore, a great deal of smoothing is effected. This corresponds to the familiar case of deriving the empirical

distribution function as a maximum likelihood estimator. Its density is of course composed of delta functions.

With the above considerations in mind let us write the intensity for the pth patient as

$$\nu_{ij}(t;z_p) = \sum_{\ell \in A_{ij}} a_{ij\ell}\, \delta(t-\xi_{ij\ell}) \exp(\alpha_{ij}' z_p), \quad \text{for } i \neq j \tag{3.6}$$

where A_{ij} is the set of indices, ℓ, for which a transition is possible. The set $\overset{U}{i\neq j} A_{ij}$ which forms the set of centers of the Dirac delta functions, is assumed to have no cluster point. For a given clinical trial we can only consider the $\xi_{ij\ell}$ which are less than $\max_p T_p$. Thus we may take $\overset{U}{i\neq j} A_{ij}$ as a finite set.

The data that are observable with model (3.6) are given in the following table:

Table I: Observable Data with Model (3.6)

Event Times: $\xi_{ij1}, \xi_{ij2}, \ldots, \xi_{ijn}$

Risk Sets: $R_i(\xi_{ij1}), \ldots, R_i(\xi_{ijn})$

Transition Sets: $S_{ij}(\xi_{ij1}), \ldots, S_{ij}(\xi_{ijn})$

The above table is constructed for each i and j such that $i \neq j$. The risk set $R_i(t)$ is the set of all patients in state i at time t^-. The transition set $S_{ij}(t)$ is the set of all patients that undergo an $i \to j$ transition at time t. Either or both of these sets may be empty at any t. Thus if no patient undergoes a transition from $i \to j$ at $t = \xi_{ij\ell}$ say, then $\#[S_{ij}(\xi_{ij\ell})] = 0$. As in (2.12b) the probability that a patient in $R_i(\xi_{ij\ell})$ survives the $i \to j$ transition may be written

as $$P_{ii}(\xi_{ijl}, \xi_{ijl}; z_p) = \exp\left[-a_{ijl} \exp(\alpha_{ij}' z_p)\right]. \quad (3.7)$$

The probability of a transition is one minus this quality. With (3.7) it is simple to write the likelihood of the data in Table 1 as

$$L = \prod_{i,j \ni i \neq j} \prod_{l=1}^{n_{ij}} \left[\prod_{p \in S_{ij}(\xi_{ijl})} (1 - \exp a_{ijl}^{(p)}) \right] \left[\prod_{p \in R_i(\xi_{ijl}) - S_{ij}(\xi_{ijl})} \exp(-a_{ijl}^{(p)}) \right], \quad (3.8)$$

where $$a_{ijl}^{(p)} = a_{ijl} \exp(\alpha_{ij}' z_p). \quad (3.9)$$

We will proceed with estimation of a_{ijl} in the case that the α_{ij} vectors are known. The partial likelihood estimates of these vectors may be used since these estimates were derived independently of the a_{ijl}.

To estimate the a_{ijl}, we must maximize (3.8). If $R_i(\xi_{ijl})$ is empty, then the corresponding a_{ijl} cannot be estimated. By convention we take it to be zero. If $R_i(\xi_{ijl})$ is non-empty but $S_{ij}(\xi_{ijl})$ is, we see from (3.8) that the maximum occurs for $\hat{a}_{ijl} = 0$. Thus, we need consider only those times ξ_{ijl} for which events actually occurred. In order to emphasize this fact we relabel the ξ_{ijl} as t_{ijl} and consider them "event" times. Since we are estimating intensities but are primarily interested in transition probabilities, it is clear from the decoupled nature of the log of (3.8) that those transitions that are not going to be included in the final state

space (e.g. transitions into a censored state) may be ignored.

We are now reduced to the case where for each t_{ijl}, $\#(S_{ij}(t_{ijl}))>1$ and $R_i(t_{ijl})$ is non-empty. Setting the derivative of the log of (3.8) with respect to a_{ijl} to zero, we have

$$\sum_{p\in S_{ij}(t_{ijl})} \frac{\exp(\alpha_{ij}' z_p)\exp(-a_{ijl}\exp(\alpha_{ij}' z_p))}{1-\exp(-a_{ijl}\exp(\alpha_{ij}' z_p))}$$

$$= \sum_{p\in R_i(t_{ijl})-S_{ij}(t_{ijl})} \exp(\alpha_{ij}' z_p) \text{ for all } i\neq j, \text{ and } l. \quad (3.10)$$

In the case of "ties" at t_{ijl}, equation (3.10) may be solved iteratively for $\hat{a}_{ijl}$. In the important case of no ties, we have the following easily solved relation for $\hat{a}_{ijl}$:

$$\exp(-\hat{a}_{ijl}\exp(\alpha_{ij}' z_{p_{ijl}}))$$

$$= 1-\exp(\alpha_{ij}' z_{p_{ijl}}) \Big/ \sum_{p\in R_i(t_{ijl})} \exp(\alpha_{ij}' z_p), \quad (3.11)$$

where p_{ijl} is that patient in $S_{ij}(t_{ijl})$.

It can be seen from (3.11) that in the case of no covariate effects ($\alpha_{ij}=0$), we have

$$\exp(-a_{ijl}) = 1-\frac{1}{N_{ijl}} \quad (3.12)$$

where N_{ijl} is the number of patients in $R_i(t_{ijl})$. Expression (3.12) leads to the usual Kaplan-Meier estimate of the survival function in the three-state no covariate case. The estimator given in (3.11) is identical to that proposed by Kalbfleisch and Prentice (1973) in

the three state problem (life-death-censored). They develop their estimator by considering the grouping of continuous time to failure data.

From (3.11), we see that $\hat{a}_{ijl}$ is a function of the covariates of the patients in the risk set $R_i(t_{ijl})$. As mentioned earlier, the estimates $\hat{a}_{ijl}$ will in practice be found using the partial likelihood estimates $\hat{\alpha}_{ijl}$ of the α_{ij} vectors. Having estimated the α_{ij}'s and the ν_{ij}^{o}'s, we may write an estimate of the intensity matrix for a hypothetical person with covariate vector z as

$$\hat{\nu}_{ij}(t;z) = \nu_{ij}^{(o)}(t;\hat{a}) \exp\ (\hat{\alpha}_{ij}' z),\ i \neq j$$

$$\hat{\nu}_{ii}(t;z) = -\sum_{j \ni j \neq i} \hat{\nu}_{ij}(t;z)$$

Because each $\nu_{ij}^{(o)}$ is a sum of delta functions, we have

$$P(\tau, t;z) = \prod_{k=k_1}^{k_2} \hat{P}(\mu_k^-, \mu_k;z) \tag{3.13}$$

where as in section 2 the μ_k are the ordering of the t_{ijl} such that

$$\mu_{k_1-1} \leq \tau < \mu_{k_1} < \mu_{k_1+1} < \ldots < \mu_{k_2} \leq t < \mu_{k_2+1}$$

and for $t_{ijl} = \mu_k$, $\hat{P}(\mu_k^-, \mu_k;z)$ is the identity matrix except for the ith row which is given by

$$\hat{P}_{il}(\mu_k^-, \mu_k;z) = \begin{cases} \exp\ (-\hat{a}_{ijl} \exp\ (\hat{\alpha}_{ij}' z)), & l=i \\ 1-\exp\ (-\hat{a}_{ijl} \exp\ (\hat{\alpha}_{ij}' z)), & l=j \\ 0, & \text{otherwise} \end{cases} \tag{3.14}$$

These ideas are implemented in section 5 with an example. The

consistency, asymptotic variance and weak convergence of the integrated intensity matrix estimator

$$\hat{B}_{ij}^{F\text{-}M}(t,\Delta,z) = \int_{\Delta}^{t} \hat{\nu}_{ij}(t',z)\, dt'$$

are given in the next section.

4. CONSISTENCY OF THE INTEGRATED INTENSITY MATRIX ESTIMATOR

The notation used in this section is defined in Appendix A. The (i,j) element of the integrated intensity matrix is defined by

$$B_{ij}(t,\Delta,z) = \int_{\Delta}^{t} \nu_{ij}(t',z)\, dt' \tag{4.1}$$

From Section 2 the maximum likelihood estimator of (4.1) is given by

$$\begin{aligned}\hat{B}_{ij}^{F\text{-}M}(t,\Delta,z) &= \int_{\Delta}^{t} \hat{\nu}_{ij}^{(o)\ F\text{-}M}(t',z) h_{ij}\, dt' \\ &= \sum_{\Delta \le t_{jr} \le t} \ln\left(1 - \frac{h_{ijr}}{n_i t_{ijr} h_{ij} t_{ijr}}\right)^{-\frac{h_{ij}}{h_{ijr}}} \\ &\qquad \text{for } n_i t_{ijr} > 2\ , \\ &= 0 \quad \text{for } n_i t_{ijr} \le 2\end{aligned} \tag{4.2}$$

Aalen (1975) presents an estimator for B_{ij} of the form

$$\begin{aligned}\hat{B}_{ij}^{A}(t,\Delta,z) &= \int_{\Delta}^{t} \hat{\nu}_{i,j}^{(o)\ A}(t',z)\, h_{ij}\, dt' \\ &= \sum_{\Delta \le t_{ijr} \le t} \frac{h_{ij}}{n_i t_{ijr} h_{ij} t_{ijr}}\end{aligned} \tag{4.3}$$

Noting that for small x, $\ln(1-x) \approx x$, it can be intuitively seen that the two estimators given by (4.2) and (4.3) respectively are

nearly the same for large n_{it}. In Appendix B, this point is used to prove the following:

Theorem 4.1

The maximum likelihood estimate $\hat{B}^{F-M}(t,\Delta,z)$ is consistent for estimating $B(t,\Delta,z)$. i.e.,

$$\operatorname*{Plim}_{N\to\infty} \left| \hat{B}^{F-M}(t,\Delta,z) - B(t,\Delta,z) \right| = 0.$$

Thus, the asymptotic properties that hold for $\hat{B}^{A}$ also hold for $\hat{B}^{F-M}$. In particular we have

Theorem 4.2

$$\sqrt{N}\left(\hat{B}_{ij}^{F-M}(t,\Delta,z) - B_{ij}(t,\Delta,z)\right) \xrightarrow[N\to\infty]{D} \int g_{ij}(s)\, dW_{ij}(s)$$

where $W_{ij}(s)$ is the Weiner process and $\int g_{ij}(s)\, dW_{ij}(s)$ is a normal process with independent increments and

$$g_{ij}^2(s) = \frac{\nu_{ij}^o(s)}{\sum_{j'\neq i} \int_z g_{j'}(z)\, P_{j'i}(0,s,z) \prod_{j'} \exp(\alpha_{ij}' z)\, dz}$$

Theorem 4.3

$$\lim_{N\to\infty} \operatorname{Var}\left[\sqrt{N}\left(\hat{B}_{ij}^{F-M}(t,\Delta,z) - B_{ij}(t,\Delta,z)\right)\right] = \int_\Delta^t \frac{\nu_{ij}^o(s)\, ds}{\sum_{j'\neq i} \int_z g_{j'}(z)\, P_{j'i}(0,s,z) \prod_{j'} \exp(\alpha_{ij}' z)\, dz}$$

5. ILLUSTRATION OF MULTIPLE TRANSITION ESTIMATION

In order to illustrate the use of these techniques in the analysis

of data from an actual clinical trial, we have drawn data from a comparison of two protocols for treatment of hepatic hemorrhage.

The sample consists of 89 patients who were admitted to a local medical center with hepatic hemorrhaging which was in most cases a complication of alcholic liver disease. The patients were randomized into one of two treatment protocols. Patients in Group 1 received the standard medical treatment for this disorder, which consists of whole blood transfusion and antibiotics. Patients in Group 2 received, in addition to the standard treatment, a porta-caval shunt operation to route blood around the hemorrhaging portion of the liver. There was a lag time between admission to the study and randomization into treatment groups which ranged from 0 to 22 months. At the time of randomization, a complete blood chemistry workup as well as a disease history was taken. Demographic data for this clinical trial are contained in Fertig, Murthy, et al (1979).

During the follow-up period occurrence times of several events were noted. The relevant events for the purpose of this example are death from hepatic hemorrhage, death from other cause, withdrawal resulting in loss to follow-up, and a rise in blood bilirubin level over 3mg. which was taken as an indicator of hepatitis onset. These considerations define five mutually exclusive states:

State 1: Patient alive, no hepatitis since randomization.

State 2: Patient alive, hepatitis infection has occurred.

State 3: Dead from hepatic hemorrhage.

State 4: Dead from other causes.

State 5: Withdrawing lost to follow-up.

Since State 1 can never be reentered, and since states 3 through 5 are absorbing states, the only allowable transitions are 1-2, 1-3, 1-4, 1-5, 2-3, 2-4, and 2-5. All patients entered the study in state 1.

The above classification of states represents a competing risk model when there is the risk of dying from hepatic hemorrhaging and the risk of dying from other causes. The primary interest in this study is in the effect of treatment on the probability of dying from hepatic hemorrhaging. Since the treatment for each of the patients in this study involved whole blood transfusions, there was a large incidence of hepatitis, the presence of which could alter the effectiveness of the porta-caval shunt operation. Thus, it is important to be able to estimate the effect of treatment on the 1-3 transition intensity (no hepatitis → hepatic hemorrhage) separately from its effect on the 2-3 intensity (hepatitis → hepatic hemorrhage). The multiple transition model presented in this paper allows for just this separate estimation to be performed.

To this end, we define the first covariate, $z^{(1)}$, such that $z^{(1)}=0$ for those patients undergoing standard treatment, and $z^{(1)}=1$ for those patients undergoing a porta-caval shunt operation. For illustrative purposes, we consider a second covariate, $z^{(2)}$, defined as 0 for those patients whose time to randomization was less than the median time, and as 1 for those patients whose time to randomization was greater than the median.

The total number of transitions for all 89 patients was 142. A selected subset of these transitions, together with covariate

information is presented in Table 1 below.

TABLE 1

Selected Events for Porta-Caval Shunt Clinical Trial

Event No.	Patient I.D.	Event Time (Weeks)	From State i	To State j	Treatment Covariate Value $z^{(1)}$	Lag Time Covariate Value $z^{(2)}$	$R_i(t_{(\ell)})$
1	33	0.286	1	4	1	1	{1,2,...,89}
2	69	0.571	1	2	0	0	{33}
3	75	0.714	1	4	1	0	{33,69}
4	69	2.000	2	3	0	0	{69}
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
70	47	84.286	2	4	1	0	{7,13,18,20, 23,29,31,47, 54,56,62,64, 70,72,76,80, 81,85}
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
139	12	476.143	1	5	1	1	{13,9,12}
140	9	522.143	1	5	1	0	{1,3,9}
141	1	523.428	1	5	1	0	{1,3}
142	3	525.857	1	5	1	1	{3}

Transitions $i \to 5$, $i=1, 2$ represent loss to follow-up. These transitions supply no information pertinent to the study of treatment effect. State 5 is thus not treated as a competing risk. Rather, the time that a patient undergoes in $i \to 5$ transition is the time he is deleted from $R_i(t)$ (the risk set of individuals who may undergo a transition from state i). This is formally done by setting $a_{i5\ell}=0$ for all t_ℓ at which these transitions have occurred. In this case

the matrix whose elements are defined by equation (3.14) reduces to the identity, and the matrix of transition probabilities remains unchanged at time t_ℓ.

The raw estimates of the parameters α_{ij} were obtained by the partial likelihood methods discussed in Section 3 and are displayed in Table 2 along with their standard scores, obtained by dividing each estimate by its estimated asymptotic standard error.

TABLE 2

Covariate Parameter Estimates

From State	To State	Treatment Effect $\alpha_{ij}^{(1)}$ Estimate	Treatment Effect $\alpha_{ij}^{(1)}$ Est/S.E.	Lag Time Effect $\alpha_{ij}^{(2)}$ Estimate	Lag Time Effect $\alpha_{ij}^{(2)}$ Est/S.E.
1	2	-0.0139	-0.05	0.0486	0.17
1	3	$-\infty$	-	0.2432	0.33
1	4	0.0687	0.11	0.2837	0.46
2	3	-1.9088	-2.43	0.6218	1.07
2	4	0.8275	1.65	0.6937	1.55

The estimate of $\hat{\alpha}_{13}^{(1)} = -\infty$ resulted because all the, 1-3 transitions (eight of them) occurred among people in the "standard treatment" group. The permutation test of the hypothesis that the eight transitions are evenly distributed between the treatment and no treatment groups rejects this hypothesis at the 0.01 significance level.

Using the techniques of Section 3, the matrix of transition probabilities was estimated at each event time. For example, consider the first event in Table 1. A patient with covariate $z_{p_{141}} = (1,1)$ entered state 4 from state 1. From Table 2, we can compute $\hat{\alpha}_{14}' z_{p_{141}} = 0.3523$. At the time of this event, all the patients were in

state 1, so all were at risk. The value of $\sum \exp(\alpha'_{14} z_p)$ for this risk set is 107.283. The resulting estimate of a_{141} from equation (3.11) is $\hat{a}_{141} = -\ln(1-e^{0.3523/107.283})/e^{0.3523} = 0.009383$. The probability transition matrix, $P_{ij}(0,t;z)$ is the identity matrix for $t<t_{(1)}=20.286$. At $t=20.286$, it is updated according to equations (3.13) and (3.14). For the case $z=(0,0)$, we see from (3.14), that

$$P(0,t_1;z)=P(t_1^-,t_1;z=(0,0)) = \begin{bmatrix} 0.9907 & 0 & 0 & 0.00934 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where the P_{14} element is given by $1-[\exp -0.009383\, e^{\hat{\alpha}'_{14} z}]$ which reduces to 0.00934 for $z=(0,0)$.

As an illustration of the updated procedure, we consider the event at time $t_{(70)}=84.286$. Without specifying the details, we have $\hat{a}_{24t_{70}}=0.01987$. From this we can compute $P_{ij}(t_{70}^-, t_{70};z)$ from equation (3.13). The non identity portion of this matrix is

$$P_{24}(t_{70}^-, t_{70};z) = 1-\exp\left[-0.01987\, e^{\hat{\alpha}'_{24} z}\right] = 0.01967 \text{ for}$$

$$z=(0,0) \text{ and } P_{22}(t_{70}^-, t_{70};z) = 0.98033.$$

The updating of $P(0,t;z)$ to $t=t_{70}$ is done by noting

$$P(0,t_{70};z) = P(0,t_{69};z)\, P(t_{69},t_{70}^-;z)\, P(t_{70}^-, t_{70};z).$$

But $P(t_{69}, t_{70}^-;z) = I$. Thus, for $z=(0,0)$,

$$P(0,t_{70};z) = \begin{bmatrix} 0.403 & 0.223 & 0.264 & 0.111 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.9803 & 0 & 0.0197 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.403 & 0.218 & 0.264 & 0.115 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where $P(0,t_{69};z)$ is as indicated above.

We see from this, for example, that the probability a patient with covariate value $z=(0,0)$ remains in state 1 for $t_{70}=84.3$ weeks is estimated as 0.403. Note that the fifth state, lost-to-follow-up, has zero entry probability from states 1-4. This is a consequence of setting $a_{i5}=0$ for $i=1,\ldots,4$, and is in keeping with the Kaplan-Meier treatment of the censored state.

We present in Figures 1 and 2 plots of $P_{13}(0,t;z)$ and $P_{14}(0,t;z)$, respectively for all values of z. A literal interpretation of Figure 1 is that a person in the surgical treatment groups is much less likely to die from hepatic hemorrhage, than a person in the standard treatment group, which is the expected result. This large difference

is primarily due to the infinite estimate for $\alpha_{13}^{(1)}$. There is essentially no difference that can be attributed to lag time to randomization.

From Figure 2 we see a person in the surgical group is much more likely to eventually die from other causes than a person in the non-treatment group. This apparent negative consequence of treatment could easily be accounted for by the fact that those patients who were not treated had a higher incidence of death from hepatic hemorrhaging and thus were not available to die from "other causes". Inspection of Table 2 indicated, however, a very marginal positive $\alpha_{24}^{(1)}$. Though not statistically significant, an $\alpha_{24}^{(1)}$ of 0.8275 is an indication that treatment enhances the hazard of dying from other causes.

FIGURE 1

Probability of death due to hepatic hemorrhage in the presence of all other causes of death versus time on study.

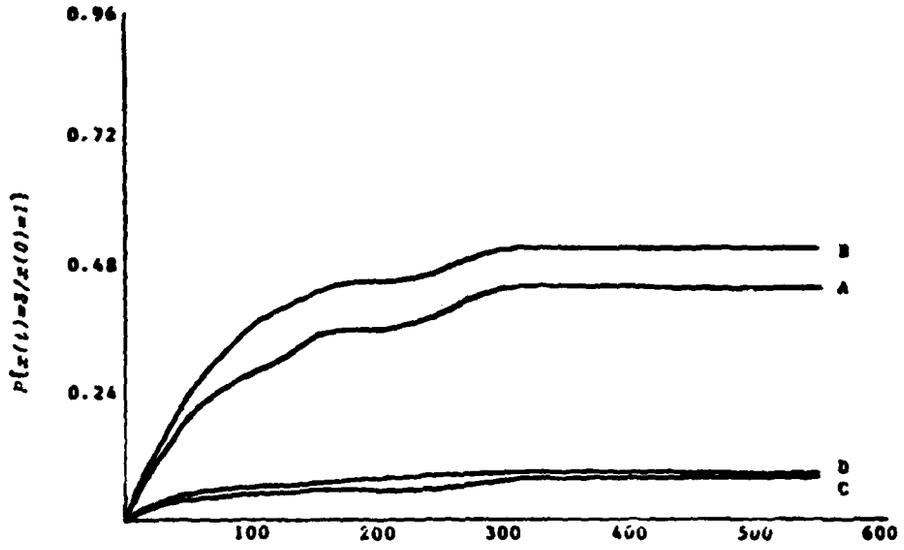
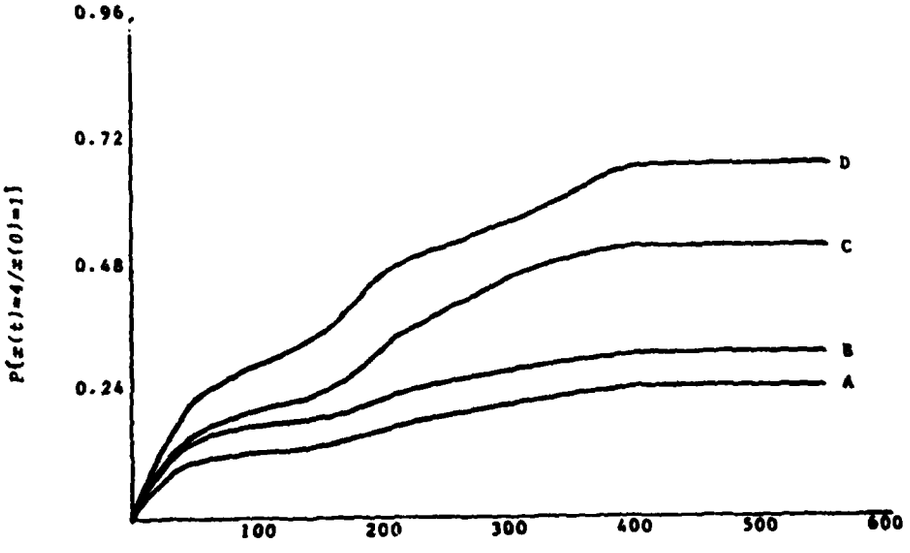


FIGURE 2

Probability of death due to other causes in the presence of all risks of death versus time on study.



LEGEND

		A	B	C	D
Treatment type	[0=Standard, 1=Portal Shunt]	0	0	1	1
Lag Time (ε)	[0=t<median, 1=t>median]	0	1	0	1

REFERENCES

Aalen, Odd Olai (1975). "Statistical Inference for a Family of Counting Processes". Ph.D. Dissertation, University of California, Berkeley 1975.

Chiang, C.L. (1968). "Introduction to Stochastic Processes in Biostatistics", John Wiley & Sons, New York.

Cox, D.R. (1972). "Regression Models and Life Tables". J.R. Stat. Soc., B., 34, pp. 187-220.

Cox, D.R. (1975). "Partial Likelihood". Biometrika, 61, pp. 31-37.

Fertig, K.W., Murthy, V.K., Sposto, R., Weiner, J. (1979). "Analysis of Survival: Multiple Transitions, Competing Risks, Covariates and Censoring". Technical Report No. 2, Dept. of Medicine, Cedars-Sinai Medical Center, Los Angeles, Ca.

Flemming, T.R. (1978). "Nonparametric Estimation for Non-homogeneous Markov Processes in the Problem of Competing Risks". Annals of Statistics, 6, 5, pp. 1057-70.

Kalbfleisch, J.D. and R.L. Prentice (1973). "Marginal Likelihoods based on Cox's Regression and Life Model". Biometrika, 60, 2.

Kaplan, E.L. & P. Meier (1958). "Nonparametric Estimation from Incomplete Observations". Journal American Statistical Assoc., 53, pp. 457-81.

APPENDIX A

Notation

$\hat{\nu}_{i,j}(t,z) = \nu_{ij}^{(o)}(t,\hat{a}) \exp(\alpha_{ij}' z)$, $i \neq j$

$\hat{\nu}_{ii}(t,z) = -\sum_{j \exists j \neq i} \hat{\nu}_{ij}(t,z)$

N : Total pool of patients.

α_{ij} : Covariate parameter vector for $i \to j$

$N_{ij}(t)$: # of $i \to j$ transitions up to t.

z : Covariate vector

h_{ij} : $\exp(\alpha_{ij}' z)$

t_{ijr} : rth i-j transition.

P_{ijr} : Patient labelled p_{ijr} undergoing i-j transition.

n_{it} : # of patients in $R_i(t)$.

$R_i(t)$: $\{p / U_p(t^-) = i\}$

h_{ijr} : $\exp(\alpha_{ij}' z_{p_{ijr}}(t_{ijr}))$

$h_{ijt_{ijr}}$: $\frac{1}{n_{it_{ijr}}} \sum_{p=1}^{N} I(U_p(t^-) = i)\, h_{ij}^{(p)}(t)$

$h_{ij}^{(p)}(t)$: $\exp(\alpha_{ij}' z_p(t))$

Z : Space of covariate vectors.

$g_i(z)$: Multivariate distribution of covariates z for patients entering state i.

$\overline{\Pi}_i$: Probability of entering state i.

APPENDIX B

Proof of Consistency and Weak Convergence

We will now prove the consistency of our estimator for the integrated intensity matrix

$$B_{ij}(t,\Delta,z) = \int_{\Delta}^{t} \nu_{ij}(t',z)\, dt' \qquad (1)$$

Assuming the α_{ij} vectors are known, our estimate for (1) denoted by $\hat{B}^{F\text{-}M}(t,\Delta,z)$ is given by

$$\hat{B}^{F\text{-}M}(t,\Delta,z) = \int_{\Delta}^{t} \hat{\nu}_{ij}^{(o)F\text{-}M}(t',z)\, h_{ij}\, dt'$$

$$= \sum_{\Delta \le t_{ijr} \le t} \ln\left(1 - \frac{h_{ijr}}{n_{it_{ijr}}\, \bar{h}_{ijt_{ijr}}}\right)^{\frac{-h_{ij}}{h_{ijr}}}$$

$$\text{for } n_{it_{ijr}} > 2$$

$$= 0 \qquad \text{for } n_{it_{ijr}} \le 2 \qquad (2)$$

We will prove the consistency of our estimator (2) for estimating (1) by proving that our estimator is equivalent in probability to Aalen's consistent estimator (1975) for estimating (1). Now Aalen's estimator for estimating (1) is given by

$$\hat{B}^{A}(t,\Delta,z) = \int_{\Delta}^{t} \hat{\nu}_{ij}^{oA}(t',z)\, h_{ij}\, dt'$$

$$= \sum_{\Delta \le t_{ijr} \le t} \frac{h_{ij}}{n_{it_{ijr}}\, \bar{h}_{ijt_{ijr}}} \qquad (3)$$

We will now demonstrate that the difference between (2) and (3) will converge to zero in probability. To simplify the proof we assume the following:

$$A_1 : \quad \emptyset = \inf_{i,j,z\in Z} \exp(\alpha_{ij}' z) > 0$$

$$A_2 : \quad \Theta = \sup_{i,j,z\in Z} \exp(\alpha_{ij}' z) < \infty$$

$$A_3 \quad : \quad \lim_{N\to\infty} P\{n_{it}>2,\ldots,n_{st}>2, \forall 0<\Delta\le t\le T\} = 1$$

$$A_4 \quad : \quad \sup_{i,j\in z\in Z} \nu_{ij}^{(o)}(t) \exp(\alpha_{ij}' z) = K < \infty .$$

The proposition to be established is then given by the following theorem and its proof.

Theorem:

$$\operatorname*{Plim}_{N\to\infty} \left| \hat{B}^{A}(t,\Delta,z) - \hat{B}^{F-M}(t,\Delta,z) \right| = 0$$

for all t such that $\Delta<t\le T$. (4)

Proof:

For given $\varepsilon>0$, $\delta>0$, in view of A_3 choose $N_o(T)$ such that for $N>N_o(t)$

$$P\{n_{it}>2,\ldots,n_{st}>2\ \forall_t\, 3 0<\Delta\le t\le T\} > 1-\delta/2 \tag{5}$$

Now, for fixed t with probability greater than $1-\delta/2$ we have

$$\left| \hat{B}^{A}(t,\Delta,z) - \hat{B}^{F-M}(t,\Delta,z) \right|$$

$$= \left| \sum_{\Delta\le t_{ijr}\le t} \frac{h_{ij}}{n_{it_{ijr}} \bar{h}_{ijt_{ijr}}} - \ln\left(1 - \frac{h_{ijr}}{n_{it_{ijr}} \bar{h}_{ijt_{ijr}}}\right)^{\frac{-h_{ij}}{h_{ijr}}} \right| \tag{6}$$

By $A_3, n_{it}>2$ which implies $\dfrac{h_{ijr}}{n_{it_{ijr}} \bar{h}_{ijt_{ijr}}} < 1$

Hence, the left hand side of (6) is dominated by

$$\left| \hat{B}^{A} - \hat{B}^{F-M} \right| \le \sum_{r\,3\Delta\le t_{ijr}\le t} \left| \sum_{q=2}^{\infty} \frac{1}{q} \frac{h_{ij}^{q-1} h_{ijr}}{\bar{h}_{ijt_{ijr}}^{q} n_{it_{ijr}}^{q}} \right|$$

$$\leq \sum_{r \exists \Delta \leq t_{ijr} \leq t} \left| \sum_{q=2}^{\infty} \frac{1}{q} \frac{h_{ij}}{h_{ijr}} \left(\frac{h_{ijr}}{h_{ijr} + (n_{it_{ijr}} - 1)\phi} \right)^q \right|$$

$$\leq h_{ij} \sum_{r \exists \Delta \leq t_{ijr} \leq t} \left| \sum_{q=2}^{\infty} \frac{1}{q\phi} \left(\frac{\theta}{\theta + (n_{it_{ijr}} - 1)\phi} \right)^q \right|$$

$$= \frac{h_{ij}}{\phi} \sum_{r \exists \Delta \leq t_{ijr} \leq t} \sum_{q=2}^{\infty} \frac{1}{q} \left(\frac{\theta}{\theta + (n_{it_{ijr}} - 1)\phi} \right)^q \qquad (7)$$

Now there exists a constant A such that

$$\frac{A_1}{n_{it}^2} \geq \sum_{q=2}^{\infty} \frac{1}{q} \left(\frac{\theta}{\theta + (n_{it} - 1)\phi} \right)^q$$

for $\forall n_{it} > 2$.

In particular choose

$$A_1 = \left(\frac{\theta}{\phi} \right)^2 \sum_{q=0}^{\infty} \frac{1}{q+2} \left(\frac{\theta}{\theta + \phi} \right)^q .$$

We then have

$$\left| \hat{B}^{A} - \hat{B}^{F-M} \right| \leq \frac{h_{ij}}{\phi} \sum_{r \exists \Delta \leq t_{ijr} \leq t} \frac{A_1}{n_{it_{ijr}}^2}$$

$$= \frac{h_{ij}}{\phi} A_1 \int_{\Delta}^{t} \frac{1}{n_{is}^2} dN_{ij}(s), \qquad (8)$$

where $N_{ij}(s)$ is a multivariate counting process with intensities

$$\nu_{ij}(s) = \sum_{\nu=1}^{N} \exp(\alpha_{ij}' z_\nu) I(U_\nu(s^-) = i) \nu_{ij}^{(o)}(s) J_\nu(s) \qquad (9)$$

and

$$M_{ij}(t) = N_{ij}(t) - \int_o^t \nu_{ij}(s)\, ds \qquad (10)$$

are orthogonal integrable martingales by Aalen's (1975) theorem 2.1 on page 14.

Let $$\lambda_{ij}^{*}(s) = \begin{cases} \nu_{ij}(s), n_{is} > 2 \\ 0 \ . \ n_{is} \leq 2 \end{cases} \tag{11}$$

Now, consider the random variable

$$\phi_1(\omega) = \int_{\Delta}^{t} \frac{1}{n_{is}^2(\omega)} \, dN_{ij}(s,\omega) \tag{12}$$

Since ϕ_1 is positive, we have

$$\underset{N\to\infty}{\text{Plim}} \ \phi_1(\omega) = 0 \quad \text{if we can show that}$$

$$\lim_{N\to\infty} \ E(\phi_1(\omega)) = 0$$

But

$$E(\phi_1(\omega)) = E \int_{\Delta}^{t} \frac{1}{n_{is}^2(\omega)} \, dN_{ij}(s,\omega)$$

$$= \int_{\Delta}^{t} E\left(\frac{\lambda_{ij}^{*}(s)}{n_{is}^2}\right) ds \tag{13}$$

because $M_{ij} = N_{ij}(t) - \int_{o}^{t} \lambda_{ij}^{*}(s) \, ds$ and $\int H(s,\omega) \, dM_{ij}(s,\omega)$ are square integrable martigales for any simple process $H(s,\omega)$ as defined by Aalen (1975) on page 8.

We have

$$M_{ij}(t) = N_{ij}(t) - \int_{\Delta}^{t} \sum_{\nu=1}^{N} h_{ij}^{(\nu)}(s) \ I(U_\nu(s^-)=i) \ J_\nu(s) \ \nu_{ij}^{o}(s) \, ds$$

$$\text{for} \quad n_{is} > 2$$

$$= 0 \quad \text{otherwise; and} \tag{14}$$

$$n_{is} = \sum_{\nu=1}^{N} I(U_\nu(s^-)=i) \, .$$

Now for $n_{is} > 2$

$$\lambda_{ij}^{*}(t) \leq \sum_{\nu=1}^{N} I(U_\nu(t^-)=i) \ h_{ij}^{(\nu)}(t) \ J_\nu(t) \ \nu_{ij}^{o}(t)$$

$$\leq \underset{i,j,t,\nu}{\text{Max}} \; h_{ij}^{(\nu)}(t)\; \nu_{ij}^{o}(t) \sum_{\nu=1}^{N} I(U_{\nu}(t^{-})=i)$$

$$= K\, n_{it} \tag{15}$$

where $K<\infty$ by A.4

Hence

$$\lim_{N\to\infty} E\left(\frac{\lambda_{it}^{*}}{n_{it}^{2}}\right) \leq \lim_{N\to\infty} E\left(\frac{K}{n_{it}}\right)$$

$$= \lim_{N\to\infty} E\left(\frac{K/N}{n_{it}/N}\right) \tag{16}$$

$$= \frac{\lim_{N\to\infty} K/N}{\underset{N\to\infty}{\text{Plim}}\left(\frac{n_{it}}{N}\right)} = \frac{0}{\Theta_{it}} = 0,$$

Since $\Theta_{it}>0$ by A.3. Combining equations (8), (13), and (16) we finally obtain that

$$\underset{N\to\infty}{\text{Plim}} \left| \hat{B}^{A}(t,\Delta,z) - \hat{B}^{F\text{-}M}(t,\Delta,z) \right| = 0 . \tag{17}$$

In other words our estimate $\hat{B}^{F\text{-}M}$ is asymptotically equivalent in probability to Aalen's (1975) estimate $\hat{B}^{A}(t,\Delta,z)$. But the consistency of Aalen's estimate is established in proposition 8.1 on page 69 of his thesis (1975) assuming that the risk sets grow large like N. Therefore, we obtain

Theorem

$\hat{B}^{F\text{-}M}(t,\Delta,z)$ defined by equation (2) is consistent for estimating $B(t,\Delta,z) = \int_{\Delta}^{t} \nu_{ij}(s,z)\, ds$.

Weak Convergence and Asymptotic Variance of $\hat{B}^{F\text{-}M}(t,\Delta,z)$

From the martingale convergence theorem given on page 29 and the theorem 8.2 on page 69 of Aalen (1975), we obtain that

$$\sqrt{N}\left(\hat{B}_{ij}^{F\text{-}M}(t,\Delta,z) - B_{ij}(t,\Delta,z)\right)$$

$$\xrightarrow[N\to\infty]{D} \int g_{ij}(s)\, dW_{ij}(s) \tag{18}$$

where $W_{ij}(s)$ is the Weiner process and $\int g_{ij}(s)\, dW_{ij}(s)$ is a normal process with independent increments where

$$g_{ij}(s) = \underset{N\to\infty}{\text{Plim}} \frac{\sum_{\nu=1}^{N} I(U_\nu(s^-)=i)\ h_{ij}^{(\nu)}(s)\ \nu_{ij}^{o}(s)}{\left(\sum_{\nu=1}^{N} I(U_\nu(s^-)=i\ h_{ij}^{(\nu)}(s)\right)^2}$$

$$\underset{N\to\infty}{\text{Plim}} \frac{\nu_{ij}^{o}(s)}{\sum_{\nu=1}^{N} I(U_\nu(s^-)=i)\ h_{ij}^{(\nu)}(s)\ /\ N} \qquad (19)$$

But

$$\underset{N\to\infty}{\text{Plim}} \sum_{\nu=1}^{N} I(U_\nu(s^-)=i)\ h_{ij}^{(\nu)}(s)\ /N$$

$$= E\left(I(U_\nu(s^-)=i)\ h_{ij}^{(\nu)}(s)\right)$$

$$= \sum_{j'\neq j'\neq i} \int_z g_{j'}(z)\ P_{j'i}(o,s,z) \prod_{j'} \exp(\alpha_{ij}^{'} z)\, dz \qquad (20)$$

Therefore,

$$g_{ij}^2(s) = \frac{\nu_{ij}^{o}(s)}{\sum_{j'\neq i} \int_z g_{j'}(z)\ P_{j'i}(0,s,z) \prod_{j'} \exp(\alpha_{ij}^{'} z)\, dz}$$

Hence the asymptotic variance of $\sqrt{N}\left(\hat{B}_{ij}^{F-M}(t,\Delta,z) - B_{ij}(t,\Delta,z)\right)$

is given by

$$\lim_{N\to\infty} \text{Var}\left[\sqrt{N}\left(\hat{B}_{ij}^{F-M}(t,\Delta,z) - B_{ij}(t,\Delta,z)\right)\right]$$

$$= \int_{\Delta}^{t} \frac{\nu_{ij}^{o}(s)\ ds}{\sum_{j'\neq i} \int_z g_{j'}(z)\ P_{j'i}(o,s,z,) \prod_{j'} \exp(\alpha_{ij}^{'} z)\, dz} \qquad (21)$$

Monte Carlo Comparison of Cox's Partial vs. Total Likelihood

The performance of the maximum total likelihood estimate and the Cox maximum partial likelihood estimate are compared using Monte Carlo computer simulation techniques. A wide variety of different experimental conditions, including those pertaining to an actual clinical trial, are used to evaluate the effects of changes in basic parameters on the size and power of three tests. The parameters considered include censoring (presence or absence), the baseline hazard rate, the number of covariates and their distributions, and the magnitude of the covariate effects. Distributional data for the estimates of the covariate effects, including confidence bounds for selected percentiles, are also given.

1. Introduction

In recent years much attention has been given to the problem of accounting for covariate information in survival analysis of clinical trials in which study participants may have been censored. Cox (1972) addresses this situation by proposing the proportional hazard rate model:

$$\lambda(t \mid z_i, \beta) = \lambda_0(t) \exp [\beta , z_i] ,$$

where $\lambda(t \mid z_i, \beta)$ is the failure rate (hazard function) for the i-th patient, z_i is the covariate vector for that patient, and β is the vector representing the effect of the covariates. $\lambda_0(t)$ is an arbitrary function representing the baseline failure rate -- the failure rate for a patient with covariate vector $z_i = \underline{0}$. The survival function, $S(t \mid z_i, \beta)$ and the failure density function $f(t \mid z_i, \beta)$ are then given by

$$S(t \mid z_i,\beta) = -\left\{\int_0^t \lambda_i(x|z_i,\beta)dx\right\} = \exp -\left\{\exp (\beta' z_i) \int_0^t \lambda_0(x)dx \right\}$$

and

$$f(t|z_i, \beta) = \lambda(t|z_i, \beta)\, S(t|z_i, \beta)$$

$$= \lambda_0(t) \exp\left\{\beta' z_i - \exp(\beta' z_i) \int_0^t \lambda_0(x)dx\right\},$$

respectively.

Much analysis has been done in the exponential case where $\lambda_0(t) = e^{\theta_0}$ (Glasser, 1967), Sprott and Kalbfleisch (1969), Breslov (1972) and Prentice (1973)). Simple departures from the constant hazard model are accomidated using the Weibull model where $\lambda_0(t) = e^{\theta_0+\theta_1 \ln t}$ (Prentice, 1973). When the concern is chiefly to estimate the covariate effects, Cox's partial likelihood (Cox, 1975) is attractive, for it ignores the formulation of $\lambda_0(t)$. Oakes (1977) and Efron (1977) establish conditions under which the partial likelihood is asymptotically fully efficient, and show that under many circumstances of interest, the partial likelihood suffers little loss of efficiency. In an important recent work in this area, Lindsay (1980) proposes the mixture model as a tool for evaluating the maximum partial likelihood estimates in the covariate case, and gives conditions under which it is fully efficient within that model, as well as providing a lower bound for the asymptotic variance in those cases. Peace and Flora (1978) did important work on the effectiveness of various tests in the small sample situation, using Monte Carlo methods to compare the size and power of tests based on Cox's maximum partial likelihood estimate (MPLE), $\hat{\beta}_p$, and the maximum total

likelihood estimate (MTLE), $\hat{\beta}_T$ of β, for the above models in an experimental situation. In this paper we expand and refine their work. We compare the performance of the partial and total likelihoods, using three test statistics based, respectively, upon the asymptotic normality (AN) of $\hat{\beta}$, the likelihood ratio (LR) criterion, and the asymptotic normality of the gradient of the likelihood function (DLOGL) evaluated at the null hypothesis. The baseline hazard $\lambda_0(t)$ is parametrized using zero, one and two of the nuisance parameters θ_j, respectively. The test statistics are compared based on the partial likelihood, the exponential total likelihood, and the Weibull total likelihood.

Our results are divided into parts. In the first part, we consider an experiment in which a single covariate takes the values zero and one. We explore the size and power characteristics of the LR, AN and DLOGL tests for the MPLE, the exponential MTLE, and the Weibull MTLE, and how they change due to censoring or variations in the null hypothesis, the sample size, or the covariate distribution. In the second part, we look at the distribution of $\hat{\beta}$ under these circumstances. In part three, the performance of the various tests is evaluated in a clinical situation, using the covariate distributions and censoring patterns found in a clinical data base.

2. Test Statistics

Let $\{t_i\}_{i=1}^{N}$ denote the termination times for the N test

participants and $\{x_i\}_{i=1}^{N}$ be an indicator variable where $x_i = 0$ if the participant was censored, and $x_i = 1$ if the termination was a failure. The log of the total likelihood is then given by:

$$\log L_T = \sum_{i=1}^{N} (x_i \ln \lambda_0(t_i) + \ln S(t_i|z_i, \beta))$$

and the log of the Cox partial likelihood is given by:

$$\log L_C = \sum_{i=1}^{N} x_i \left(\beta' z_i - \ln \sum_{j \in R(t_i)} e^{\beta' z_j} \right)$$

where $R(t_i)$ is the set of participants at risk at time t_i. Using a modified Newton-Raphson technique, the estimates of $\hat{\beta}$ and $\hat{\theta}$ were found, both under the null hypothesis and the alternative. Under certain regularity conditions we have, as $N \to \infty$,

$$(1)\quad 2[\log L(\hat{\rho}_0) - \log L(\hat{\rho})] \longrightarrow \chi^2_{(k)} \quad ,$$

$$(2)\quad -\sqrt{N}\,(\hat{\rho} - \hat{\rho}_0)' \left[\frac{\partial^2 \log L}{\partial \rho_i \, \partial \rho_j} \right]^{-1} \Bigg|_{\hat{\rho}_0} (\hat{\rho} - \hat{\rho}_0) \longrightarrow \chi^2_{(k)} \quad ,$$

$$(3)\quad -\frac{\partial \log L}{\partial \rho} \Bigg|'_{\hat{\rho}_0} \left[\frac{\partial^2 \log L}{\partial \rho_i \, \partial \rho_j} \right]^{-1} \Bigg|_{\hat{\rho}_0} \frac{\partial \log L}{\partial \rho} \Bigg|_{\hat{\rho}_0} \longrightarrow \chi^2_{(k)} \quad ,$$

where k is the number of parameters being tested under the null

hypothesis. Here, ρ is the vector of combined parameters

$$\rho = (\beta, \theta) = (\beta_1, \beta_2, \ldots, \beta_k, \ldots, \beta_n, \theta_1, \ldots, \theta_m)$$

where θ parameterizes $\lambda_0(t)$. The hat on ρ refers to the maximum likelihood estimate of ρ and the zero subscript refers to the maximum likelihood estimate under a specified null hypothesis. Thus, if the hypothesis is:

$$H_0 : \beta_i = \beta_{i0}, \quad 1 \leq i \leq k, \quad \text{for} \quad k \leq n$$

versus

$$K : \beta_i \neq \beta_{i0} \quad \text{for some } i, \ 1 \leq i \leq K$$

then

$$\hat{\rho}_0 = (\beta_{10}, \ldots, \beta_{k0}, \hat{\beta}_{(k+1)0}, \ldots, \hat{\beta}_{n0}, \hat{\theta}_{10}, \ldots, \hat{\theta}_{m0})$$

where $\hat{\beta}_{(k+1)0}, \ldots, \hat{\beta}_{n0}, \hat{\theta}_{10}, \ldots, \hat{\theta}_{m0}$ are the maximum likelihood estimates of these parameters under H_0.

3. The Simulation

Each experimental run consisted of the generation and analysis of 1000 samples. The survival times were generated randomly, according to the formula

$$t_i = S^{-1}(re^{-\tilde{\beta}z_i} \mid z = 0, \beta = 0)$$

where $S^{-1}(\cdot \mid z = 0, \beta = 0)$ is the inverse of the survival function, taken when $\beta = 0$ and $z = 0$. The quantity $\tilde{\beta}$ is the true value of β for the simulation and S^{-1} is generated tabularly with r taken from the uniform $(0, 1)$ random number generator;

$$u_i = \frac{n_i}{2^{31} - 1}$$

where

$$n_i = (n_{i-1} \cdot 630,360,016) \bmod (2^{31} - 1) \ .$$

As for the data from an active clinical trial, data was taken from the Western and Southeastern Cancer Study Groups' breast cancer studies. Three covariates were chosen:

Z_1 - treatment type (65.3% zeros, 34.7% ones) ,

Z_2 - years (y) free from cancer prior to entering study ,

= 0 if $0 < y \leq 1$ (45.2%)

= 1 if $1 < y \leq 5$ (45.2%)

= 2 if $5 < y$ (9.6%)

Z_3^* = 0 if liver metastases absent at entry (57.6%) ,

= 1 if liver metastases present at entry (42.4%) .

Two additional covariates were defined to measure interaction effects:

$$Z_3 = (1 - Z_3^*)(1 - Z_1) \quad \text{and} \quad Z_4 = (1 - Z_3^*)Z_1$$

which were present in the following proportions:

Z_1	Z_3	Z_4	Percent
0	0	0	18.1
0	1	0	10.5
1	0	0	24.3
1	0	1	47.1

The partial likelihood estimate of β over the breast cancer study is $\beta^* \equiv (0.47109, -0.34805, -0.26766, -0.44233)$ when (z_1, z_2, z_3, z_4) is the covariate vector. $\theta_E^* \equiv (-6.1771)$ is the nuisance parameter estimate under the exponential model, and $\theta_E^* \equiv (-6.9318, 0.13988)$ is the estimate under the Weibull model.

The method of censoring is also derived from the data base. The Kaplan-Meier (K-M), or product limit estimate of the censoring distribution for the data was computed and stored in tabular form. It is displayed in Figure 1. A random censoring time was evaluated for each study participant according to this distribution and, if it was less than the survival time for that participant, the survival time was replaced by the censoring time and the participant was considered to have been censored. This is a time dependent, and therefore, covariate dependent form of censoring. The effects of this type of censoring on different covariate combinations are discussed later in the paper.

3.1. The Results - Part 1

Tables 1 and 2 show the results of a set of ninety simulation runs for a model with one covariate, $z_1 \in \{0, 1\}$. In the exponential case $\tilde{\theta} = (0.0)$ for rows 1 - 6 of each table and $\theta = \theta_E^*$ for rows 7 - 9. In the Weibull case, $\tilde{\theta} = (1.0579, 1.0)$ in rows 1 - 6 and $\tilde{\theta} = \theta_W^*$ in rows 7 - 9. In the Appendix herein we show that the value of θ has, in the absence of censoring, no bearing on the value of $\hat{\beta}$ in either the exponential or the Weibull models. In lines 1 - 7, z_1 is chosen without replacement from a population of size N, containing half zeros and half ones. In line 8, z_1 is chosen with replacement from the same population. In line 9, z_1 is chosen with replacement from its empirical distribution in the clinical trial data base.

In the first three lines of each table, we see the effects on the test statistics (under $H_0 : \beta = 0$) of variations in the sample size in the absence of censoring. Power curves corresponding to the DLOGL test are chosen for visual comparison and are displayed in Figures 2 and 3. Upon inspection of either the tables or the graphs, it becomes immediately evident that the partial likelihood is much less powerful than either of the total likelihoods for N = 10. The AN test performs very poorly under the partial likelihood, the LR test does better, and the DLOGL test is best of all, though it still falls well short of the same test in the total likelihood analysis. On the other hand, the tests in the total likelihood analysis are very similar to each other, although the DLOGL test still appears

to be somewhat more powerful.

Comparing the sizes ot the tests, we observe those based on the Weibull model are very biased. Using the Kolmogorov-Smirnov (K-S) confidence intervals we reject the hypothesis with 95% confidence that any of them are truly $\chi^2_{(1)}$. A representative plot, using the DLOGL test is given in Figure 7. Under the partial likelihood, the DLOGL test is again not $\chi^2_{(1)}$, though the deviation is not so great, as is seen in Figures 4 and 6. The L.R. test appears to be $\chi^2_{(1)}$ and the AN test is borderline; failing the K-S test and passing it once. The tests based on the exponential model appear to be $\chi^2_{(1)}$.

When $N = 20$, the partial likelihood compares better against both of the total likelihoods. The various tests based on the partial likelihood are much closer to each other in power, and much closer to the corresponding test in either of the total likelihoods. The sizes of the various tests are stabilizing as well, although the tests based on the Weibull model are still rather far from $\chi^2_{(1)}$.

By the time $N = 40$, the irregularities are all smoothed out, almost down to random fluctuations. The tests are now almost close to $\chi^2_{(1)}$ and the powers differ only slightly between tests under a given model. The partial likelihood is now competing well against both the exponential and Weibull total likelihoods. The DLOGL test still appears to be the strongest test in all cases, and while the AN test appears to be slightly stronger than the L.R. test in the total likelihoods, the two tests are almost indistinguishable in the

partial likelihood, with perhaps a slight preference going to the L.R. test.

Lines 4 - 6 in Tables 1 and 2 give results for experiments identical to those above, except that they correspond to testing $H_0 : \beta_1 = 1$ against $H_1 : \beta_1 \neq 1$. An immediate observation is that for $N = 10$ the L.R. and AN tests have no power at all against $\beta_1 > 1$. This is due to the clustering of the covariates, so that all of the members with $z_1 = 1$ fail before any of the members with $z_1 = 0$. Under these circumstances, the MPLE of β is infinite under the alternative hypothesis and the analysis fails to converge. Our program had no means of dealing with these cases, except to throw them out, which, of course, seriously weakened these tests when it happened frequently; namely, when $\tilde{\beta}$ was large and N was small. The DLOGL test computes the MPLE only under the null hypothesis, therefore, it was unaffected by this problem. Even so, it too was asymetric in β_1 around 1, being more powerful against $\beta_1 < 1$ and less powerful against $\beta_1 > 1$. This bias is characteristic of the tests based on the partial likelihood, being evident for $N = 20$, 40 as well, with the L.R. test being least affected by it. On the other hand, the tests based on the exponential total likelihood are not visibly affected by the change in H_0, apart from random fluctuations, just as they did before. Tests based on the Weibull total likelihood showed a moderate asymetry, again with the L.R. test being relatively less affected than the others. The partial likelihood is nearly as strong as the Weibull total likelihood on the left,

though significantly less strong on the right, while the Weibull is weaker than the exponential at all points.

Line 7 in Tables 1 and 2 is identical to line 3, except that censoring is imposed using the K-M estimate of the censoring distribution of the clinical trial, as described in the previous section. When $z_1 = 0$, the censoring probability is .128 for all values of $\tilde{\beta}$; the censoring probability for $z_1 = 1$ varies with $\tilde{\beta}$ as shown in Table 3. The tests all lose power due to the censoring, more so where $\tilde{\beta} < 0$, due to the greater probability of a sample member being censored, than where $\tilde{\beta} > 0$.

Within each model the tests retain the same relative strengths. Tests based upon the Weibull MTLE suffered relatively less loss of power than the tests based on partial and exponential total likelihoods. This is, perhaps, due to the effect of time-varying hazard rate on the censoring distribution. Examining line 8 where the covariates are selected with replacement and line 3 where they are selected without replacement, we observe essentially no effect on either the power or size calculations. However, the additional change of the covariate distribution which was affected in line 9 has greater influence. The various tests are weakened, but not uniformly so. The L.R. test is weakened more on the right than on the left, while the AN and DLOGL tests are weakened more on the left than on the right. The bias of the L.R. test is observable on all of the models, but the bias of the AN and DLOGL tests is evident only on the MPLE and the Weibull MTLE.

3.2 Results - Part 2

Tables 4 and 5 give scaled confidence intervals for selected percentiles of $\hat{\beta}$ when $N = 40$, under conditions similar to those mentioned in Part 1. All of the values have been standardized by subtracting the median and dividing by a scaling factor equal to the interquartile range, divided by 1.349 (the interquartile range for the normal distribution with $\sigma^2 = 1$).

Deviations from normality occur sporadically throughout and are attributable, in many cases, to chance fluctuations. However, when $\tilde{\beta} = (1.0)$, we see a more pronounced trend away from normality among the MPLE and the Weibull MTLE. The right tails of the distributions are generally heavier and, in the Weibull MTLE, the median is biased high. However, these changes are not discernable in the exponential MTLE.

3.3 Results - Part 3

Table 6 gives the results of tests patterned after the data taken from the breast cancer data base mentioned previously. Samples of size 40, 60 and 100, were taken with replacement from the data base itself and analyzed both for size at $\tilde{\beta} = 0$ and power at $\tilde{\beta} = \beta^*$ (restricted to the number of covariates being analyzed) against the null hypothesis that $\tilde{\beta} = 0$. The covariate vectors used are (z_1), (z_1, z_2) and (z_1, z_2, z_3, z_4), with the distribution of the terms given in Section 3. In the exponential case $\tilde{\theta} = \theta_E^*$ and in the Weibull case $\tilde{\theta} = \theta_W^*$.

The performance of the partial likelihood in the one-covariate case is nearly indistinguishable from the performance of the total likelihoods, either with censoring or without. The DLOGL test is the most powerful test throughout, while retaining good size characteristics, and the L.R. test is the least powerful. Censoring has a minor impact on the power of the tests, but in this situation has no effect on their relative power. Even at $N = 100$, the tests are clearly not powerful enough to decide against $H_0 : \beta = 0$; larger samples are obviously required.

With two covariates, the tests are somewhat more powerful, but unchanged in relationship to each other. Censoring has less effect in this case because $\beta_2^* < 0$, which reduces the probability of censoring for many sample members.

The four-covariate case is not strictly comparable with the other cases due to the interdependence of the covariates. This can be seen in the power of the various tests, which is lower than in the two-covariate case because the combined effects of z_1 and z_4 nearly cancel each other, occurring 47.1% of the time.

The size of the DLOGL test, and, to a lesser extent the AN test, are severely affected, especially when $N = 40$. Whether this is attributable to the interdependence of the covariate values, or whether it is simply due to the large number of covariates and their distribution taken independently of each other, remains to be tested, although the latter appears likely.

4. Computational Note

The program which generated these results was written in Fortran IV and run on a PDP 11/34 with 96k of core. Run time for each simulation varied from about 45 minutes to over 20 hours of CPU time, depending on the complexity of the analysis. Well over a thousand CPU hours were required from program development and simulation run time. The authors wish to thank the Department of Obstetrics and Gynecology at Cedars-Sinai Medical Center, Los Angeles, California for the free use of their computer for this enormous usage of computer time. This work would have been impossible without their generous aid.

5. Appendix

Here, we show that when the covariates are time-independent and censoring is absent, then the MTLE of β is independent of θ in the exponential and Weibull models.

Consider first the Weibull model,

$$\lambda_0(t|\theta_0, \theta_1) = \exp(\theta_0 + \theta_1 \ln t) .$$

Reparametrize as follows:

$$a = \exp\left(\frac{\theta_1}{\theta_1 + 1}\right) \quad \text{and} \quad b = \theta_1 + 1$$

to get

$$\lambda_0(t|a, b) = abt^{b-1} .$$

Now, suppose t arises from the distribution:

$$F(t_1|a_1, b_1) = 1 - \exp-[a_1 t_1^{b_1} \exp(\beta' z)] .$$

If $(a_1, b_1) \rightarrow (a_2, b_2)$, then $t_1 \rightarrow t_2$ such that $a_1 t_1^{b_1} = a_2 t_2^{b_2}$. Now,

$$L(t_1|a_1, b_1) = \prod_{i=1}^{N} [a_1 b_1, t_1^{b_1-1} \exp[\beta' z_L - a_1 t_{1i}^{b_1} \exp(\beta' z_i)]] ,$$

where t_{1i} and z_i belong to the i-th member of a study of size N.

$$\frac{\partial \ln L(t_1|a_1, b_1)}{\partial \beta_j}$$

$$= \frac{\partial}{\partial \beta_j} \sum_{i=1}^{N} [\ln(a_1 b_1) - (b_1 - 1) \ln t_{1i}^{b_1} \exp(\beta' z_i)]$$

$$= \sum_{i=1}^{N} [z_{ij} - z_{ij} a_1 t_{1i}^{b_1} \exp(\beta' z_i) .$$

Since the likelihood is maximized when $a_1 = \hat{a}_1$, $b_1 = \hat{b}_1$, and $\beta = \hat{\beta}_1$,

$$\sum_{i=1}^{N} z_{ij} = \sum_{i=1}^{N} z_{ij}\, \hat{a}_1\, t_{1i}^{\hat{b}_1} \exp(\hat{\beta}_1 z_i) \quad .$$

Substituting in

$$t_{1i} = \left(\frac{a_2}{a_1} t_{2i}^{b_2}\right)^{1/b_1} \quad ,$$

gives

$$\sum_{i=1}^{N} z_{ij} = \sum_{i=1}^{N} z_{ij}\left[\hat{a}_1 \left(\frac{a_2}{a_1}\right)^{1/b_1}\right] t_{2i}^{(\hat{b}_1 b_2)/b_1} \exp(\hat{\beta}_1 z_i) \quad .$$

We now have the following identities:

$$\hat{a}_2 = \hat{a}_1 \left(\frac{a_2}{a_1}\right)^{1/b_1} \quad ,$$

$$\hat{b}_2 = \frac{\hat{b}_1 b_2}{b_1} \quad ,$$

$$\hat{\beta}_2 = \hat{\beta}_1 \quad .$$

The results for the exponential case may be seen as a special case of the above proof, where $b_1 = \hat{b}_1 = b_2 = \hat{b}_2 = 1$.

References

Breslow, Norman E.
(1972) "Covariance Analysis of Censored Survival Data," Biometrics, 3, 80-99.

Cox, David R.
(1972) "Regression Models and Life Tables," Journal of the Royal Statistical Society, Ser. B, 34, 187-220.

(1975) "Partial Likelihood," Biometrika, 62, 269-276.

Efron, Bradley
(1977) "The Efficiency of Cox's Likelihood Function for Censored Data," Journal of the American Statistical Association, 72, 557-565.

Glasser, Marion N.
(1967) "Exponential Survival with Covariance," Journal of the American Statistical Association, 62, 561-568.

Lindsay, B. G.
(1980) "Nuisance Parameters, Mixture Models, and the Efficiency of Partial Likelihood Estimators," Philosophical Transactions of the Royal Society of London, Ser. A, 296, 639-665.

Oakes, David
(1977) "The Asymptotic Information in Censored Survival Data," Biometrika, 64, 441-448.

Peace, Karl E. and Flora, Roger E.
(1978) "Size and Power Assessments on Tests of Hypotheses on Survival Parameters," Journal of the American Statistical Association, 73, 129-132.

Prentice, Ross L.
(1973) "Exponential Survivals with Censoring and Explanitary Variables," Biometrika, 60, 279-288.

Sprott, D. A. and Kalbfleisch, John D.
(1969) "Examples of Likelihoods and Comparison with Point Estimates and Large Sample Approximations," Journal of the American Statistical Association, 64, 468-489.

Legend for Tables

β_0 : value of β under null hypothesis

$\tilde{\beta}$: true value of β

Cov : method of selecting covariate values, i.e.,

F - without replacement from a group of size N containing half zeros and half ones

R - same as F except with replacement

E - with replacement from the values of z_1 taken from the data base

Cen : whether censoring is imposed - Y = yes and N = no

N : sample size

Test : L - likelihood ratio (LR) test

A - asymptotic normality (AN) of $\hat{\beta}$ test

D - asymptotic normality of gradient (DLOGL) test

	β_0	Cov	Cen	N	Test	Partial					Exponential				
						β_0-2	β_0-1	β_0	β_0+2	β_0+2	β_0-2	β_0-1	β_0	β_0+1	β_0+2
	0	F	N	10	L	62.0	21.5	5.5	21.5	56.5	87.2	31.0	5.0	33.5	85.0
1					A	45.0	12.5	3.5	14.5	42.5	88.0	32.5	6.0	36.5	86.5
					D	73.5	27.0	7.0	28.0	70.5	89.5	37.0	7.5	40.5	89.5
				20	L	96.0	52.5	6.5	51.5	95.5	99.0	59.0	5.5	60.0	99.0
2					A	95.5	51.0	6.5	50.5	95.0	99.0	60.0	6.0	61.0	99.0
					D	97.0	54.0	7.0	53.5	95.5	99.5	62.5	6.5	63.0	99.0
				40	L	100.0	83.5	4.5	85.0	100.0	100.0	88.5	4.0	87.0	100.0
3					A	100.0	83.5	4.5	85.0	100.0	100.0	89.0	4.5	87.0	100.0
					D	100.0	84.5	5.0	86.0	100.0	100.0	89.5	5.0	88.0	100.0
	1	F	N	10	L	83.0	30.0	3.0	0.0	0.0	88.0	32.0	5.0	34.0	85.0
4					A	84.0	30.5	3.0	0.5	0.0	90.0	33.5	5.0	37.0	86.5
					D	87.5	36.0	9.0	25.5	58.0	91.5	38.0	7.5	40.0	88.0
				20	L	99.0	54.5	5.0	36.0	79.0	99.5	59.5	4.5	62.0	99.5
5					A	99.0	56.0	4.5	24.5	64.0	99.5	61.0	4.5	63.5	99.5
					D	99.0	58.5	5.5	34.5	80.5	99.5	64.5	5.0	65.5	99.5
				40	L	100.0	85.0	6.5	69.5	99.0	100.0	89.0	6.0	88.5	100.0
6					A	100.0	85.5	6.0	66.0	99.0	100.0	89.5	6.0	89.0	100.0
					D	100.0	86.5	6.5	68.0	99.0	100.0	90.0	6.0	89.0	100.0
	0	F	Y	40	L	99.5	75.0	6.0	83.5	100.0	100.0	78.5	6.0	85.5	100.0
7					A	99.5	74.5	6.0	83.5	100.0	100.0	78.5	6.0	86.0	100.0
					D	99.5	75.0	6.5	84.5	100.0	100.0	79.5	6.5	86.5	100.0
	0	R	N	40	L	100.0	85.0	5.5	84.0	100.0	100.0	88.5	5.0	88.0	100.0
8					A	100.0	85.0	5.5	83.5	100.0	100.0	89.0	5.0	88.5	100.0
					D	100.0	85.5	6.0	84.5	100.0	100.0	89.5	6.0	88.5	100.0
	0	E	N	40	L	100.0	79.5	4.5	77.5	100.0	100.0	85.0	4.5	82.0	100.0
9					A	100.0	77.5	5.0	80.5	100.0	100.0	84.0	4.5	84.5	100.0
					D	100.0	78.5	5.5	81.5	100.0	100.0	84.5	5.5	85.0	100.0

Table 1

Size and Power Results for Tests Based on
the Partial and Exponential Total Likelihoods

						Partial					Weibull				
	β_0	Cov	Cen	N	Test	β_0-2	β_0-1	β_0	β_0+1	β_0+2	β_0-2	β_0-1	β_0	β_0+1	β_0+2
	0	F	N	10	L	57.5	21.0	5.5	22.0	55.5	84.5	38.0	9.5	36.5	84.5
1					A	41.0	12.5	2.5	14.0	40.0	84.5	38.0	9.5	36.5	84.5
					D	70.5	27.5	7.5	28.0	70.5	87.0	41.5	10.5	40.5	87.5
				20	L	96.5	51.5	6.5	55.0	97.0	98.5	58.5	7.0	60.0	99.0
2					A	96.0	50.0	6.0	54.0	96.5	98.5	59.0	7.5	60.5	99.0
					D	97.0	53.5	7.0	56.0	97.0	98.5	61.5	8.0	62.5	99.5
				40	L	100.0	83.0	5.5	83.0	100.0	100.0	86.5	5.5	86.0	100.0
3					A	100.0	83.0	5.5	83.0	100.0	100.0	86.5	5.5	86.0	100.0
					D	100.0	83.5	5.5	84.5	100.0	100.0	87.5	6.0	86.5	100.0
	1	F	N	10	L	84.0	31.0	3.5	0.5	0.0	89.5	36.0	11.0	35.5	73.5
4					A	84.0	31.0	3.5	0.5	0.0	91.0	38.0	10.0	31.0	69.5
					D	88.5	38.5	10.5	28.0	59.5	92.0	42.0	12.0	35.5	73.0
				20	L	98.5	53.0	6.0	34.5	77.5	99.0	56.0	8.0	53.5	96.0
5					A	99.0	54.0	5.0	22.0	63.0	99.0	57.5	8.0	52.0	95.0
					D	99.0	56.5	6.0	33.5	80.0	99.0	60.0	9.0	53.5	96.0
				40	L	100.0	83.5	5.0	69.5	99.5	100.0	84.5	5.5	81.0	100.0
6					A	100.0	83.5	5.0	67.0	99.5	100.0	85.0	5.0	81.0	100.0
					D	100.0	84.5	5.0	68.0	99.0	100.0	85.5	5.5	81.0	100.0
	0	F	Y	40	L	100.0	78.5	5.0	81.5	100.0	100.0	81.0	5.5	85.5	100.0
7					A	100.0	78.0	5.0	82.0	100.0	100.0	81.0	5.5	86.0	100.0
					D	100.0	79.5	6.0	82.5	100.0	100.0	82.0	6.0	86.0	100.0
	0	R	N	40	L	100.0	81.5	4.5	86.5	100.0	100.0	84.5	4.5	88.0	100.0
8					A	100.0	82.0	4.5	86.5	100.0	100.0	85.0	4.5	88.0	100.0
					D	100.0	82.5	5.0	87.5	100.0	100.0	85.0	5.0	88.5	100.0
	0	E	N	40	L	100.0	80.0	5.0	78.0	100.0	100.0	83.5	6.5	82.0	100.0
9					A	100.0	77.0	5.5	79.5	100.0	100.0	81.5	6.5	84.0	100.0
					D	100.0	78.5	5.5	81.0	100.0	100.0	82.5	7.0	85.0	100.0

Table 2

Size and Power Results for Tests Based On
the Partial and Weibull Total Likelihoods

	Probability of Censoring	
β_1	Exponential ($\theta = (-6.1771)$)	Weibull ($\theta = (-6.9318, 0.13988)$)
-2.0	0.715	0.688
-1.0	0.418	0.383
0.0	0.128	0.115
1.0	0.028	0.030
2.0	0.010	0.013

Table 3

Probability of Censoring for Various Values of $\tilde{\beta}_1$ When $z_1 = 1.0$

	β	Cov	Cen	Partial						Exponential					
				Median Scale	5%	25%	50%	75%	95%	Median Scale					
	0	F	N	-0.006	-1.569	-0.583	0.063	0.793	2.137	-0.003	-1.574	-0.572	0.061	0.800	1.785
1				-0.335	-1.664	-0.659	0.000	0.690	1.987	0.319	-1.668	-0.657	0.000	0.692	1.666
					-1.772	-0.746	-0.102	0.572	1.582		-1.776	-0.757	-0.070	0.572	1.582
	0	F	Y	-0.002	-1.545	-0.597	0.063	0.744	1.769	-0.006	-1.503	-0.626	0.072	0.721	1.698
2				0.364	-1.674	-0.695	0.000	0.654	1.612	0.363	-1.642	-0.707	0.000	0.643	1.517
					-1.917	-0.819	-0.052	0.565	1.515		-1.803	-0.758	-0.057	0.585	1.419
	0	R	N	-0.014	-1.489	-0.553	0.065	0.799	1.906	-0.020	-1.522	-0.542	0.063	0.852	1.939
3				0.346	-1.643	-0.627	0.000	0.722	1.755*	0.320	-1.641	-0.622	0.000	0.727	1.773*
					-1.785	-0.737	-0.085	0.602	1.666		-1.791	-0.741	-0.072	0.628	1.681
	0	E	N	0.018	-1.521	-0.575	0.096	0.763	1.953	0.016	-1.491	-0.561	0.107	0.799	1.964
4				0.355	-1.630	-0.667	0.000	0.682	1.728	0.338	-1.551	-0.657	0.000	0.692	1.660
					-1.776	-0.753	-0.073	0.599	1.587		-1.686	-0.788	-0.068	0.635	1.551
	1	F	N	1.026	-1.454	-0.579	0.075	0.781	2.115	1.007	-1.566	-0.589	0.076	0.785	1.784
5				0.365	-1.552	-0.660	0.000	0.689	1.915*	0.331	-1.676	-0.666	0.000	0.683	1.627
					-1.699	-0.734	-0.111	0.609	1.788		-1.830	-0.753	-0.097	0.572	1.487
	1	F	Y	1.024	-1.544	-0.573	0.081	0.856	2.255	1.022	-1.634	-0.631	0.084	0.734	2.075
6				0.351	-1.677	-0.637	0.000	0.712	2.021*	0.316	-1.791	-0.706	0.000	0.643	1.880*
					-1.836	-0.741	-0.092	0.623	1.889		-1.981	-0.807	-0.097	0.586	1.754
	1	R	N	1.017	-1.426	-0.595	0.083	0.774	2.116	0.989	-1.507	-0.574	0.073	0.774	1.817
7				0.374	-1.499*	-0.661	0.000	0.688	1.871*	0.324	-1.631	-0.649	0.000	0.700	1.630
					-1.630	-0.741	-0.069	0.592	1.681		-1.844	-0.720	-0.057	0.584	1.556
	1	E	N	1.037	-1.388	-0.560	0.056	0.746	2.055	1.013	-1.503	-0.593	0.080	0.751	1.842
8				0.391	-1.525	-0.672	0.000*	0.677	1.804	0.348	-1.556	-0.685	0.000	0.664	1.681
					-1.663	-0.747	-0.069	0.578	1.614		-1.695	-0.755	-0.078	0.598	1.556

*is non-normal

Table 4

Distributional Data for $\hat{\beta}$ for Partial and

Exponential Total Likelihoods

	$\tilde{\beta}$	Cov	Cen	Partial						Weibull					
				Median Scale	5%	25%	50%	75%	95%	Median Scale	5%	25%	50%	75%	95%
	0	F	N	0.022	-1.688	-0.596	0.086	0.764	1.823	0.019	-1.523	-0.568	0.090	0.730	1.852
1				0.322	-1.770*	-0.668	0.000	0.681	1.659	0.330	-1.723	-0.689	0.000	0.660	1.746
					-1.970	-0.817	-0.090	0.588	1.531		-1.882	-0.770	-0.066	0.600	1.566
	0	F	Y	0.028	-1.770	-0.615	0.080	0.754	1.873	0.029	-1.676	-0.633	0.060	0.704	1.723
2				0.329	-1.878*	-0.686	0.000	0.663	1.664	0.349	-1.723*	-0.726	0.000/	0.548	1.554
					-2.090	-0.848	-0.084	0.585	1.551		-1.930	-0.816	-0.745	0.548	1.451
	0	R	N	-0.027	-1.478	-0.551	0.074	0.803	1.855	-0.030	-1.470	-0.562	0.096	0.834	1.874
3				0.323	-1.551	-0.644	0.000/	0.705	1.739	0.340	-1.565	-0.644	0.000	0.731	1.730
					-1.678	-0.720	-0.076	0.632	1.638		-1.666	-0.724	-0.051	0.635	1.632
	0	E	N	-0.005	-1.506	-0.609	0.096	0.766	2.015	0.000	-1.443	-0.613	0.079	0.768	2.013
4				0.347	-1.666	-0.684	0.000	0.665	1.796	0.349	-1.652	-0.686	0.000	0.663	1.790*
					-1.809	-0.736	-0.099	0.590	1.644		-1.788	-0.772	-0.093	0.570	1.654
	1	F	N	0.991	-1.520	-0.602	0.073	0.771	2.084	1.025	-1.499	-0.569	0.073	0.736	2.065
5				0.367	-1.619	-0.655	0.000	0.694	1.847*	0.361	-1.585	-0.685	0.000	0.664	1.736
					-1.714	-0.735	-0.069	0.614	1.675		-1.709	-0.798	-0.088	0.594	1.577
	1	F	Y	1.020	-1.444	-0.581	0.091	0.799	2.065	1.041	-1.528	-0.610	0.074	0.764	1.995
6				0.385	-1.583	-0.648	0.000	0.701	1.737*	0.375	-1.624	-0.672	0.000/	0.677	1.838*
					-1.682	-0.743	-0.066	0.599	1.647		-1.834	-0.778	-0.067	0.068	1.680
	1	R	N	1.011	-1.472	-0.496	0.083	0.901	2.304	1.037	-1.444	-0.536	0.099	0.851	2.172
7				0.350	-1.562	-0.578*	0.000	0.771	1.989*	0.354	-1.529	-0.604	0.000/	0.745	1.878*
					-1.693	-0.668	-0.062	0.671	1.783		-1.652	-0.681	-0.078	0.626	1.715
	1	E	N	1.011	-1.396	-0.542	0.099	0.869	2.251	1.038	-1.386	-0.558	0.106	0.859	2.090
8				0.375	-1.448*	-0.606	0.000	0.743	1.916*	0.377	-1.478*	-0.620	0.000/	0.729	1.903*
					-1.618	-0.693	-0.058	0.625	1.794		-1.621	-0.720	-0.070	0.660	1.800

*is non-normal

Table 5

Distributional Data for Partial and

Weibull Total Likelihoods

	Cov	Cen	N	Test	Partial S	Partial P	Exponential S	Exponential P	Partial S	Partial P	Weibull S	Weibull P
	1	N	40	L	5.5	24.0	5.0	26.0	6.5	27.0	6.5	30.0
1				A	6.0	26.0	5.5	29.0	6.5	29.5	6.5	32.0
				D	6.0	27.0	6.0	30.0	6.5	31.0	7.0	32.5
			60	L	5.5	40.5	4.5	41.5	4.5	35.5	5.0	38.0
2				A	5.5	43.0	4.5	44.5	5.0	38.0	5.0	41.0
				D	5.5	44.0	4.5	45.5	5.0	38.5	5.0	42.0
			100	L	5.5	58.5	5.5	60.0	5.5	57.5	5.5	59.0
3				A	5.0	60.5	5.5	62.0	5.5	59.0	5.5	60.5
				D	5.0	61.0	5.5	62.5	5.5	60.0	6.0	61.0
	1	Y	40	L	5.0	26.5	5.5	27.0	5.0	23.5	5.0	26.0
4				A	4.5	28.5	5.5	29.5	5.5	24.5	5.5	28.5
				D	5.0	30.0	6.0	31.0	5.5	26.0	6.0	29.5
			60	L	4.5	37.5	4.5	40.0	5.5	33.0	5.5	35.0
5				A	4.5	39.5	4.5	41.5	5.5	34.5	5.5	37.0
				D	4.5	41.0	5.0	42.0	6.0	35.5	6.0	38.0
			100	L	5.5	56.0	4.5	56.5	6.0	53.5	6.0	55.0
6				A	5.0	57.0	5.0	59.0	6.0	56.0	6.0	56.5
				D	5.0	57.0	5.0	59.0	6.0	56.5	6.0	57.5
	2	N	40	L	5.5	39.0	4.5	42.5	7.5	37.0	7.0	44.0
7				A	5.5	39.5	4.5	44.5	7.5	38.0	7.0	44.5
				D	5.5	40.5	5.0	46.0	8.0	39.5	8.0	46.0
			60	L	6.0	56.5	6.0	60.0	5.5	54.5	6.0	56.0
8				A	6.0	56.0	6.0	60.0	5.5	54.5	6.0	58.0
				D	6.5	56.5	6.5	61.0	6.5	54.5	7.0	58.5
			100	L	6.0	75.0	5.0	78.0	5.5	78.5	5.5	80.0
9				A	6.0	75.5	5.0	78.5	6.0	79.0	5.5	81.0
				D	6.0	75.5	5.0	79.0	6.5	79.0	5.5	81.0
	2	Y	40	L	5.0	35.5	5.0	37.0	5.0	35.0	6.5	38.5
10				A	4.5	35.0	5.5	38.5	5.0	35.0	6.5	39.0
				D	5.0	36.0	6.5	41.0	5.5	37.0	7.0	40.0
			60	L	4.5	47.5	5.0	50.5	6.5	50.0	7.0	54.0
11				A	4.5	47.5	5.5	51.5	6.0	49.5	7.0	54.5
				D	5.0	48.0	5.5	52.5	6.5	51.5	7.5	56.0
			100	L	6.5	72.5	6.0	74.5	5.5	72.0	6.0	74.5
12				A	7.0	72.5	6.0	75.0	5.5	72.0	6.0	74.0
				D	7.0	73.0	6.0	75.5	6.0	72.5	6.5	74.5

Table 6

Size and Power Assessments for Tests Based on the Partial, Exponential Total, and Weibull Total Likelihoods, Under Experimental Conditions

	Cov	Cen	N	Test	Partial S	Partial P	Exponential S	Exponential P	Partial S	Partial P	Weibull S	Weibull P
	4	N	40	L	6.5	31.0	5.0	33.5	5.5	31.0	6.0	35.5
13				A	7.0	32.0	6.0	36.0	7.0	32.5	7.5	38.5
				D	10.0	37.0	8.5	41.0	9.0	38.0	11.0	43.5
			60	L	7.5	41.0	5.0	43.5	5.5	45.0	7.0	48.0
14				A	7.5	44.0	6.0	48.0	6.0	46.0	8.5	51.5
				D	9.0	47.5	7.5	51.0	7.6	50.5	9.5	54.0
			100	L	5.0	68.0	4.5	71.0	6.0	69.0	7.0	71.0
15				A	6.0	68.5	6.0	73.0	6.5	70.0	7.0	72.5
				D	6.5	71.0	6.0	74.5	7.0	71.0	8.5	74.0
	4	Y	40	L	7.0	25.5	5.5	28.5	7.0	26.5	8.5	30.5
16				A	7.5	27.0	7.5	31.0	7.0	27.0	10.5	32.0
				D	10.0	32.0	9.5	36.5	10.5	33.0	13.7	37.5
			60	L	7.5	36.5	7.5	37.0	7.0	35.5	8.0	40.0
17				A	8.0	38.0	8.0	40.0	7.0	37.0	8.0	42.5
				D	10.0	41.0	10.0	43.5	8.0	41.5	9.5	47.0
			100	L	5.0	58.5	4.5	60.0	5.5	60.0	6.0	63.0
18				A	5.0	59.5	5.0	62.0	6.5	60.0	6.5	63.5
				D	6.0	62.0	5.5	64.0	7.5	62.5	7.5	65.5

Table 6

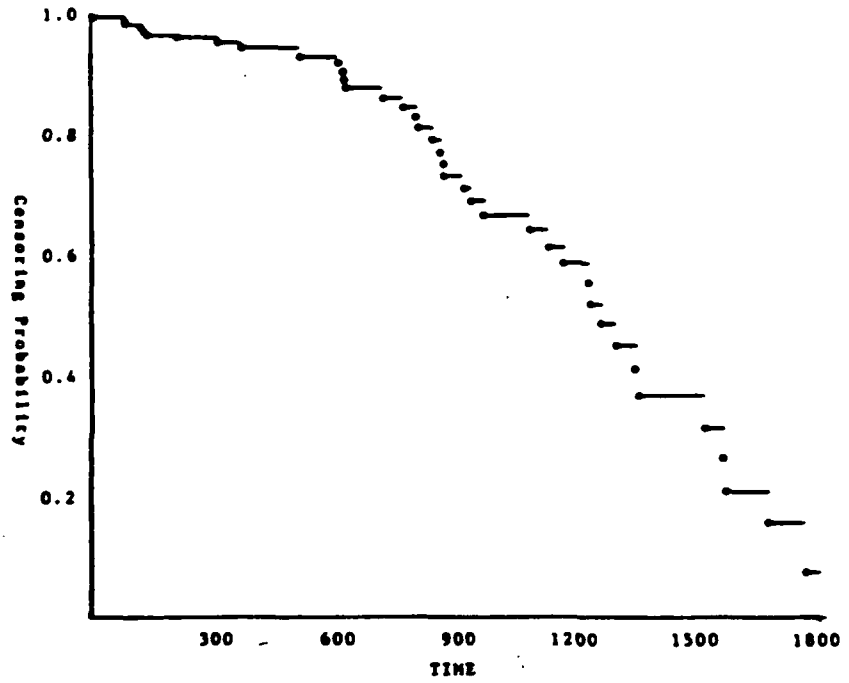
(Continued)

z_1	z_2	z_3	z_4	Percentage Occurrence	Censoring Probability	
					Exponential $\theta = (-6.1771)$	Weibull $\theta = (-6.9318, 0.13988)$
0	0	0	0	7.6	.128	.115
0	0	1	0	21.0	.189	.168
0	1	0	0	8.6	.211	.188
0	1	1	0	21.9	.291	.261
0	2	0	0	1.9	.317	.285
0	2	1	0	4.3	.406	.371
1	0	0	0	3.8	.060	.057
1	0	0	1	12.9	.122	.110
1	1	0	0	5.7	.105	.096
1	1	0	1	9.0	.203	.181
1	2	0	0	1.0	.178	.159
1	2	0	1	2.4	.307	.276

Table 7

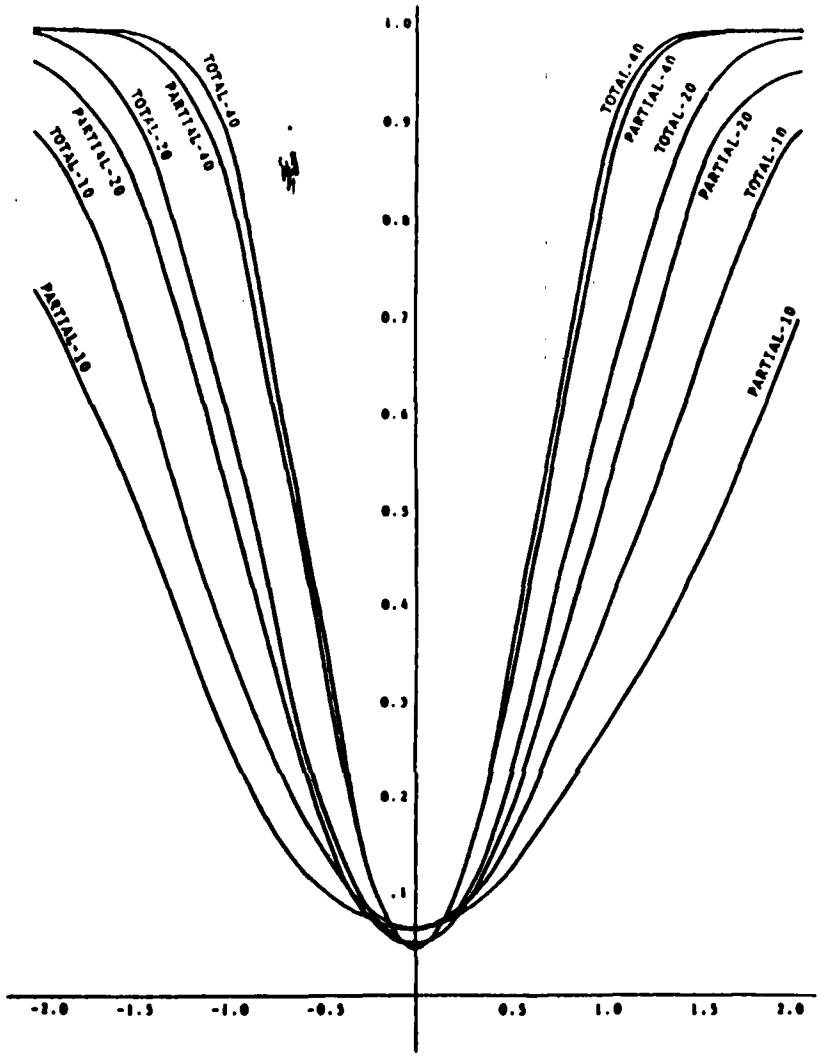
Percentage Occurrence of Covariate Combinations

and Probability of Censoring When $\beta = \tilde{\beta}$



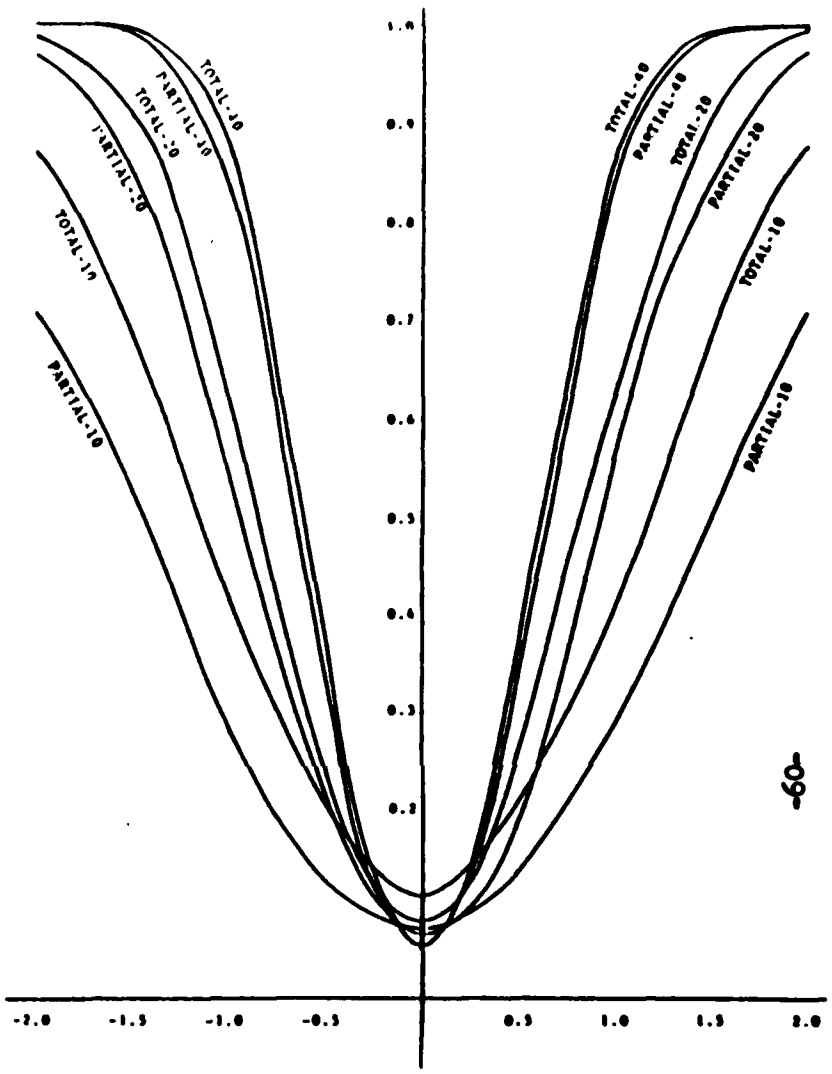
Kaplan-Meier estimate of the experimental censoring distribution

Figure 1.



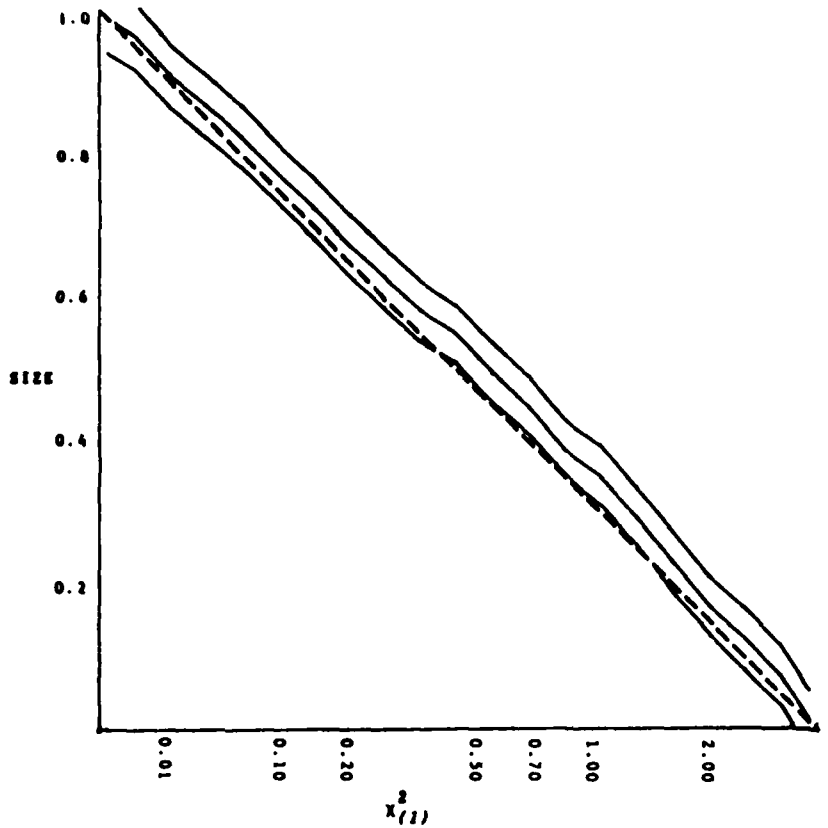
Power Curves - BLOGL Test - H_0:θ=θ - Exponential Model

Figure 2.



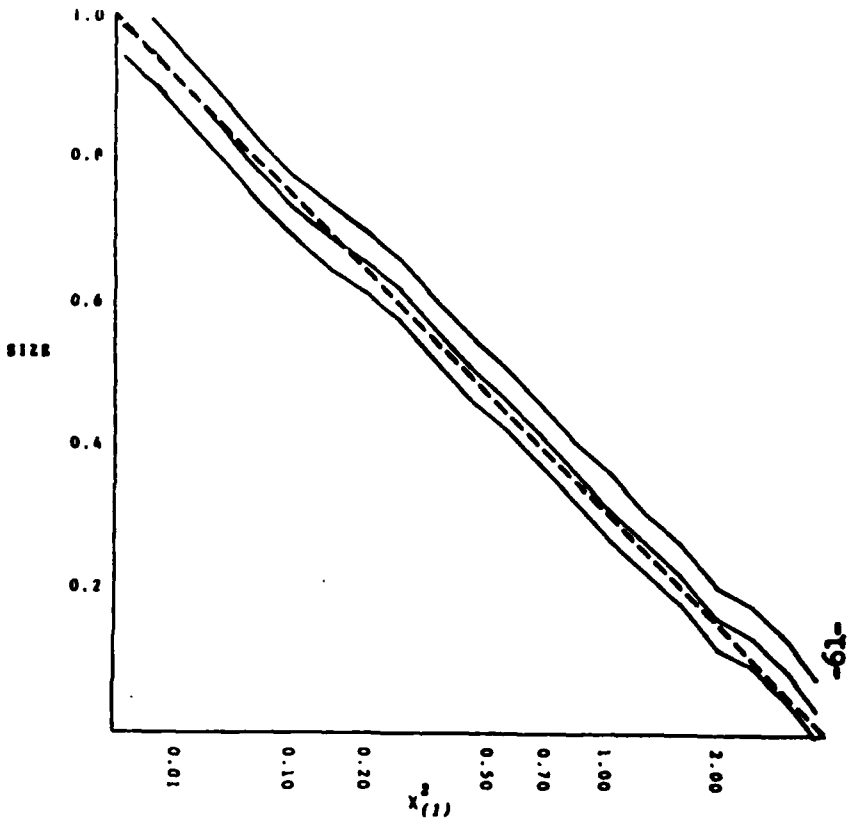
Power Curves - BLOGL Test - H_0:θ=θ - Weibull Model

Figure 3.



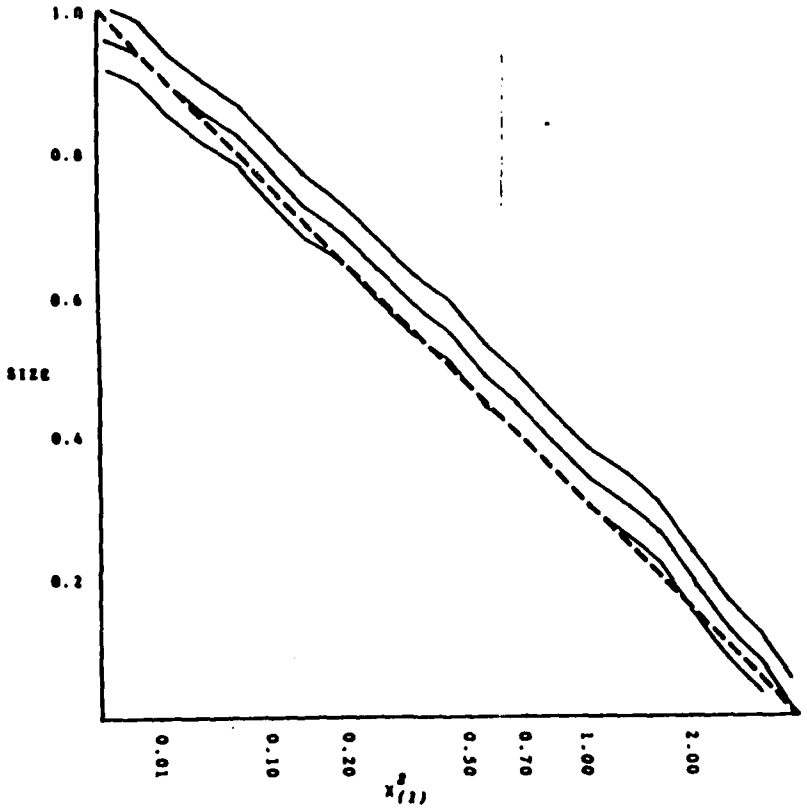
$X^2_{(1)}$ vs. Size: Empirical Alongside Exact DLOGL Test, Partial Likelihood, 1000 Samples of Size 10

Figure 4.



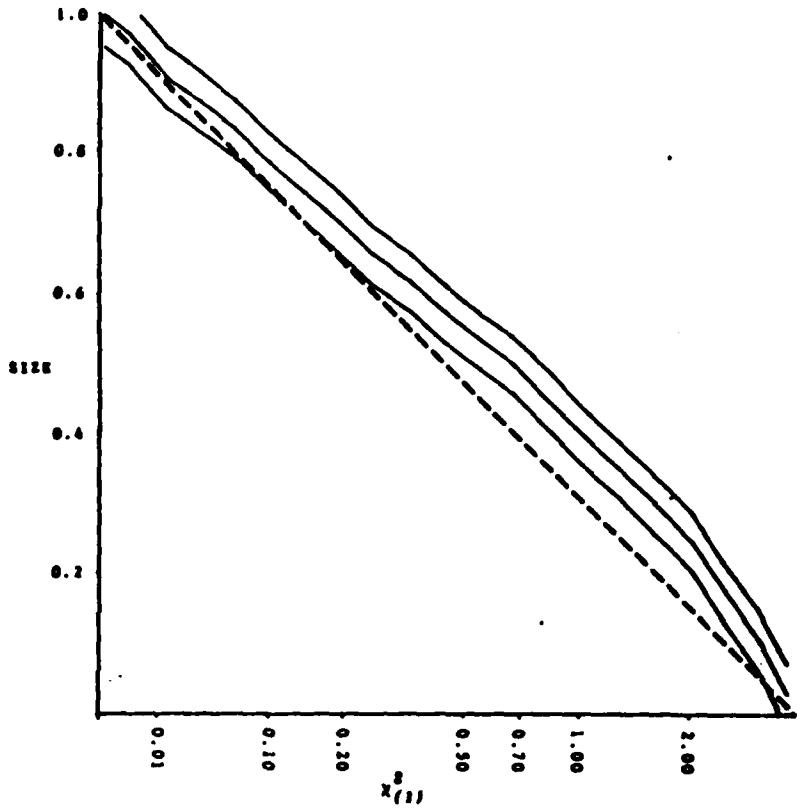
$X^2_{(1)}$ vs. Size: Empirical Alongside Exact DLOGL Test, Exponential MTLE, 1000 Samples of Size 10

Figure 5.



$X^2_{(1)}$ vs. Size: Empirical Alongside Exact BLOCL Test, Partial Likelihood, 1000 Samples of Size 10

Figure 6.



$X^2_{(1)}$ vs. Size: Empirical Alongside Exact BLOCL Test, Weibull MTLE, 1000 Samples of Size 10

Figure 7.


ATE
LME