Idaho National Engineering Laboratory

Managed by the U.S. Department of Energy

EG&G Idaho

Work performed under DOE Contract No. DE-AC07-76ID01570

EGG-SSRE-8972
May 1990

TECHNICAL REPORT

ESTIMATING HAZARD FUNCTIONS FOR REPAIRABLE COMPONENTS

Corwin L. Atwood

Prepared for the U.S. NUCLEAR REGULATORY COMMISSION

EGG-SSRE-8972

ESTIMATING HAZARD FUNCTIONS FOR REPAIRABLE COMPONENTS

Corwin L. Atwood

Published May 1990

Idaho National Engineering Laboratory
EG&G Idaho
Idaho Falls, ID 83415-3421

Prepared for the
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
Under DOE Contract No. DE-AC07-76IDO1570

ABSTRACT

This is a tutorial report, applying known formulas and tools in a way suitable for risk assessment. A parametric form is assumed for the hazard function of a set of identical components. The parameters are estimated, based on sequences of failure times when the components are restored to service (made as good as old) immediately after each failure. In certain circumstances, the failure counts are ancillary for the parameter that determines the shape of the hazard function; this suggests natural tools for diagnostic checks involving the individual parameters. General formulas are given for maximum likelihood estimators and approximate confidence regions for the parameters, yielding a confidence band for the hazard function. The results are applied to models where the hazard function is of linear, exponential, or Weibull form, and an example analysis of real data is presented.

KEY WORDS: Time-dependent failure rate, Non-homogeneous Poisson process, Poisson intensity, Exponential distribution, Exponential failure rate, Linear failure rate, Weibull distribution.

FIN No. A6389—Aging Components and Systems IV:
Risk Evaluation and Aging Phenomena

SUMMARY

This tutorial report presents a parametric framework for performing statistical inference on a hazard function, based on repairable data such as might be obtained from field experience rather than laboratory tests. This framework encompasses many possible forms for the hazard function, three of which are considered in some detail. The theory is neatest and the asymptotic approximations most successful when the hazard function has the form of a density in the exponential family. The results presented include formulas for maximum likelihood estimates (MLEs), tests and confidence regions, and asymptotic distributions. The confidence regions for the parameters are then translated into a confidence band for the hazard function. For the three examples considered in detail, a table gives all the building blocks needed to program the formulas on a computer; this table includes asymptotic approximations when they are necessary to maintain numerical accuracy. Diagnostic checks on the model assumptions are sketched.

The report gives an example analysis of real data. In this example, the methods are unable to discriminate among an exponential hazard function, a linear hazard function, and a Weibull hazard function. The MLE for the two parameters appears to have approximately a bivariate normal distribution under the exponential or Weibull hazard model, but not under the linear hazard model. If the analysis using approximate normality is carried out in any case, the results appear similar for all three models. If some model is preferred for theoretical or other reasons, the framework of this report indicates a way to use it.

ACKNOWLEDGMENTS

I am grateful to my colleagues Andrew J. Wolford and W. Scott Roesener for the stimulus and insights provided during the course of this work. I also thank Max Engelhardt for pointing out early related papers.

CONTENTS

TABLES

FIGURES

ESTIMATING HAZARD FUNCTIONS FOR REPAIRABLE COMPONENTS

1. INTRODUCTION

This report is concerned with the failure behavior of components. It is a tutorial report, applying previously known results in a way suitable for risk assessment. The model is defined in terms of the random variable T, the (first) failure time of a component. In many published articles, it is assumed that many components are tested until their first failure. The resulting failure times are used as data, and the properties of the distribution of T are then inferred. By contrast, this report deals with field data, not test data: it is assumed that each failed component is immediately restored to operability (made as good as old) and again placed in service. The data then consist of a sequence of failure times for each component.

A question of interest is whether the hazard function (or failure rate) is increasing, that is, whether the failures tend to occur more frequently as time goes on. This and related questions are investigated by postulating a parametric form for the distribution of T, and then performing the usual statistical inference about the parameters of the model, with special emphasis on the parameter(s) that determine whether the hazard function is increasing. The final goals of the inference are a point estimate and a confidence interval for the hazard function at any time t.

The general methods are applied in detail to three assumed parametric forms for the hazard function. A table gives all the formulas needed to implement the methods on a computer for these three models.

The outline of the report is as follows. Section 2 presents the assumptions and notation, and introduces three examples. Sections 3, 4, and 5 develop the likelihood formulas and equations for maximum likelihood estimators and tests/confidence intervals. Each of these three sections also discusses the application of the general results to the three examples. People who can appreciate theory without considering examples may skip the application portions. Section 6 outlines diagnostic checks, and Section 7 presents an analysis of data from motor-operated valves. Proofs are in Section 8.

2. MODEL FORMULATION

2.1 Basic Assumptions and Definitions

Assume that the failures of a component follow a time-dependent (or non-homogeneous) Poisson process. See, for example, Karr (1986) for a simple description, or Cox and Isham (1980) for a fuller introductory treatment. Alternatively, one can parallel the development from fundamental assumptions as given by Meyer (1970, Section 8.3) for the homogeneous case. The most important properties are the following: there is a nonnegative function $\lambda(t)$ defined for $t \geq 0$, with the probability of a failure in a short period $(t, t + \Delta t)$ asymptotically approaching $\lambda(t)\Delta t$ as $\Delta t \to 0$; the failure counts in non-overlapping time intervals are independent; and the number of failures occurring between 0 and t is a Poisson random variable with parameter $\Lambda(t)$, where

$$\Lambda(t) = \int_0^t \lambda(u)\, du \ .$$

Implicit in the independence property is the assumption that the component is restored to service immediately after any failure, with negligible repair time. In operational data, it is not uncommon to find that a component has failed several times in quick succession for the same reason. Presumably, the first repairs did not treat the true cause of the failure. This situation violates the independence property—the fact that a failure has occurred recently increases the chance that another failure will occur soon, because the problem may not have been really fixed. It may be difficult to force such data into the Poisson-process model: counting the failures as distinct ignores their apparent dependence, while counting them as a single failure may make the time to true repair far from negligible.

The function λ is called the *hazard function*, the *failure rate*, or the *intensity function* of the Poisson process, and Λ is the *cumulative hazard function*. Assume now that λ is continuous in t. It is related to the cumulative distribution function (c.d.f.) F of the time to first failure, and to the corresponding density function f by

$$\lambda(t) = f(t)/[1 - F(t)]$$

and

$$1 - F(t) = \exp[-\Lambda(t)] \ .$$

Any one of the three functions F, f, and λ uniquely determines the others. Note that because $F(t) \to 1$ as $t \to \infty$, it follows that

$$\lim_{t\to\infty} \Lambda(t) = \infty . \tag{1}$$

If $\lambda(t)$ is constant, as has been assumed for simplicity in many studies, the time to first failure has an exponential distribution. Often the concern is whether $\lambda(t)$ is increasing in t. It is therefore convenient to write λ in the form

$$\lambda(t) = \lambda_0 h(t;\beta). \tag{2}$$

Here, $\lambda_0 > 0$ is a constant multiplier and $h(t;\beta)$ determines the shape of $\lambda(t)$.

Because data generally come from more than one component, the following additional assumptions are made. The failures of one component are assumed to be independent of those of another component. All the components are assumed to have the same function h with the same value of β; that is, a proportional hazards model is assumed. Depending on the context, it may or may not be assumed that the different components have the same value of λ_0. Some simple regularity conditions on h, needed for asymptotic results, are discussed at the beginning of the section on confidence intervals and tests.

Sometimes there are gaps in the failure data. For example, the plant may have been shut down for an extended period, during which no component failures were possible, or the failure data may not have been collected for some period. This can be accommodated in the above framework by treating each component as two components, one observed before the gap and one after the gap, having the same installation date and, at the analyst's discretion, the same or possibly different values of λ_0.

2.2 One Notation for Two Types of Data

Types of Data

Failure data for a component can arise in a number of ways. Two simple ones to analyze are:

- A random number of failures in a fixed observation period (*time-censored* data)
- A fixed number of failures in a random observation period (*failure-censored* data).

The terms "time-censored" and "failure-censored" follow the analogous usage for tests that are terminated before all the items have failed (e.g. Nelson, 1982, Sec. 7.1). Time-censored data arise if there is a fixed time period when the component is watched or plant records are examined. During that time, the component is restored to service after each failure. Failure-censored data might arise if the component is repaired until a predetermined number of failures has occurred, at which time the component is removed from service and replaced by a new component. Both of these types of data result in tractable formulas for statistical inference.

In reality, the decision to repair or replace a component is based on a number of considerations, such as the availability of replacement components, the severity of the particular failure mode (including the difficulty and cost of repair), and any recent history of failures. These considerations are difficult to express in a simple mathematical model. Therefore, only the two types listed are analyzed here. In practice, one might simplify reality by treating failures that resulted in component replacement as if they were failure-censored.

Unified Notation

Let s_0 and s_1 denote the beginning and end of the component's observation period; s_0 does not necessarily coincide with the component's installation. Let n be the number of observed failures not counting any failure that results in replacement of the component. Let m be the total number of observed failures, including any failure that results in replacement. Let t_1, ..., t_m denote the ordered failure times. The two special cases then are

- Time-censored data: The observation period is from s_0 to a fixed time s_1. The random number of failures is n, and therefore m is random and equal to n.
- Failure-censored data: The number of failures is fixed at m, and n is therefore fixed at $m - 1$. The observation period starts at s_0 and ends at a random time s_1, with $s_1 = t_m$.

In general there are C components, indexed by j, and the quantities defined above are all indexed by j: s_{0j}, s_{1j}, n_j, m_j, and t_{ij}. In the formulas to be given, it is often convenient to define the midpoint $\bar{s}_j = (s_{0j} + s_{1j})/2$, and to define the range $r_j = (s_{1j} - s_{0j})$. This notation, sometimes with the subscript j suppressed, will be used without further comment.

Normally, time 0 is defined to be the installation time of the component. It may, however, be useful to center the data by measuring all times from some value in the middle of the observed time period(s). This can lead to negative failure times, allowed in the above formulation.

2.3 Examples

The methods of this report are applicable to a rather arbitrary hazard function, such as the ones discussed by Cox and Oakes (1984, Chapter 2). Three such examples of hazard functions are considered in this report. In each example, β is one-dimensional, the hazard function is increasing if $\beta > 0$, is constant if $\beta = 0$, and is decreasing if $\beta < 0$. The units of λ_0 are 1/time. The units of β depend on the example, but make $h(t;\beta)$ dimensionless in every case.

In some of the work presented below, the hazard function is treated as proportional to a density function. Therefore, models can be expected to be most tractable when the hazard function is of a standard form, such as a member of the exponential family. This is illustrated by the three examples of this report, with the linear hazard model consistently producing problems that the exponential and Weibull hazard models do not have. The differences result from the fact that $\log\lambda(t)$ is linear in β for the exponential and Weibull models, but not for the linear hazard model.

Various formulas and expressions are developed throughout this report. The forms that these expressions take in the example models are all collected in Table 1, given at the end of the report. To program the formulas for a computer, sometimes asymptotic approximations must be used to maintain numerical accuracy. These approximations are also given in Table 1. All the formulas of Table 1 were either derived or confirmed by using the symbolic computer program Mathematica (Wolfram, 1988).

Exponential Hazard Function

The hazard function is defined by

$\lambda(t) = \lambda_0 \exp(\beta t)$,

with β measured in units of 1/time. This example is considered in detail by Cox and Lewis (1966, Section 3.3). If β is negative, then λ does not integrate to ∞ and Equation (1) is not satisfied; therefore, λ is not a hazard function. This quirk is interesting, but is not important in practice. It is certainly possible for $\lambda(t)$ to have exponential form with negative β for t in the time period when data are observed, and to have some other form for other t, so that λ integrates to ∞. In this case, λ is a hazard function, and it is decreasing exponentially in the observed time period.

Table 1. Formulas for examples considered

Expression	Model: Exponential	Linear[a]	Weibull[b]
Constraints	None for t in finite interval	$-1/\max_{s_{1j}>0}(s_{1j})<\beta<-1/\min_{s_{0j}<0}(s_{0j})$	$\beta > -1$
$h(t)$ [Eq. (2)][c]	$\exp(\beta t)$	$1 + \beta t$	$(t/t_0)^{\beta}$
Cond. suff. stat for β	$\Sigma\Sigma T_{ij}$	$(\ldots, T_{ij}, \ldots)$	$\Sigma\Sigma \log T_{ij}$
$[\log h(t)]'$	t	$t/(1+\beta t)$	$\log(t/t_0)$
$[\log h(t)]''$	0	$-[t/(1+\beta t)]^2$	0
$\int [\log h(t)]'' h(t)$[d]	0	$-\{\log[(1+\beta s_1)/(1+\beta s_0)] - \beta r + \beta^2 r\bar{s}\}/\beta^3$	0
v [Eq. (3)][c]	$\exp(\beta s_0)[\exp(\beta r)-1]/\beta$	$r(1+\beta\bar{s})$	$t_0 C0^{b,e}/(\beta+1)$
Asymptotic[f]			
x, A	βr, $\exp(\beta s_0) r$		$\beta+1$, t_0
a_0	1		$D1^{b,e}$
a_1	1/2		$D2$
a_2	1/6		$D3$
v'	$\exp(\beta s_0)[\beta(s_1 e^{\beta r} - s_0) - (e^{\beta r}-1)]/\beta^2$	$r\bar{s}$	$t_0[C1^{b,e} - C0/(\beta+1)]/(\beta+1)$
Asymptotic[f]			
x, A	βr, $\exp(\beta s_0) r$		$\beta+1$, t_0
a_0	$\bar{s}$		$D2^{b,e}$
a_1	$s_0/2 + r/3$		$2D3$
a_2	$s_0/6 + r/8$		$3D4$
v''	$\exp(\beta s_0)[e^{\beta r}(1-\beta s_1)^2 - (1-\beta s_0)^2 + e^{\beta r} - 1]/\beta^3$	0	$t_0[C2 - 2C1/(\beta+1) + 2C0/(\beta+1)^2]/(\beta+1)^{b,e}$

Table 1. (continued)

Asymptotic[f]			
x, A	βr, $\exp(\beta s_0)r$		$\beta+1$, t_0
a_0	$s_0^2 + s_0 r + r^2/3$		$2D3$[b,e]
a_1	$s_0^2/2 + 2s_0 r/3 + r^2/4$		$6D4$
a_2	$s_0^2/6 + s_0 r/4 + r^2/10$		$12D5$
$[\log v]'$	$s_0 - 1/\beta + r/[1-\exp(-\beta r)]$	$\bar{s}/(1+\beta\bar{s})$	$C1/C0$[b,e] $- 1/(\beta+1)$
Asymptotic[f]			
x, A	βr, 1		$(\beta+1)$, $1/2$
a_0	$\bar{s}$		$\log s_0 + \log s_1$
a_1	$r/12$		$D1^2/6$[b,e]
a_2	0		0
a_3	$-r/720$		$-D1^4/360$
$[\log v]''$	$r^2 u/(1+a^2 u)$, $a = \beta r$, $u = (e^a + e^{-a} - 2 - a^2)/a^4$ $\approx (1/12)[1 + a^2/30 + a^4/1680]$	$-[\bar{s}/(1+\beta\bar{s})]^2$	$C2$[b,e]$/C0 - (C1/C0)^2 + 1/(\beta+1)^2$
$-\int[\log h(t)]'' h(t)/v + [\log v]''$	See individual terms	$\{-\beta r + (1+\beta\bar{s}) \times \log[(1+\beta s_1)/(1+\beta s_0)]\} / \{r\beta^3(1+\beta\bar{s})^2\}$	See individual terms
Asymptotic[f]			
x, A	See $[\log v]''$	β, $[r/(1+\beta\bar{s})]^2$	$(\beta+1)$, $D1^2/12$[b,e]
a_0		$1/12$	1
a_1		$-\bar{s}/6$	0
a_2		$(20\bar{s}^2 + r^2)/80$	$-D1^2/20$
a_3		$-(\bar{s}^3/3 + r^2\bar{s}/20)$ [g]	
a_4		$(560\bar{s}^4 + 168r^2\bar{s}^2 + 3r^4)/1344$	

Table 1. (continued)

$L'(0)/[I(0)]^{1/2}$	$\Sigma\Sigma(t_{ij}-\bar{s}_j)/[\Sigma n_j r_j^2/12]^{1/2}$	$\Sigma\Sigma(t_{ij}-\bar{s}_j)/[\Sigma n_j r_j^2/12]^{1/2}$	See text

---

a. If the data are centered at $t_{mid} = \Sigma r_j \bar{s}_j / \Sigma r_j$, then t_{ij}, s_{0j}, and s_{1j} must be replaced by $t_{ij} - t_{mid}$, $s_{0j} - t_{mid}$, and $s_{1j} - t_{mid}$, respectively, and Σv_j and its derivatives are replaced by 0.

b. For the Weibull failure rate model, any terms involving s_0 should be omitted if $s_0 = 0$. In this case, the asymptotic expressions are not needed.

c. Equation numbers refer to defining equations in text.

d. The integral is for t from s_0 to s_1.

e. The notation Ck is defined as $(s_1/t_0)^{\beta+1}[\log(s_1/t_0)]^k - (s_0/t_0)^{\beta+1}[\log(s_0/t_0)]^k$, for $k = 0, 1, 2$. The notation Dk is defined as $\{[\log(s_1/t_0)]^k - [\log(s_0/t_0)]^k\}/k!$, for $k = 1, 2, 3, 4, 5$.

f. The asymptotic approximation of the expression in the line immediately above is of the form $A\Sigma a_k x^k$. The next lines give the variable x and the values of A, a_0, a_1, ... The expression may be computed as $A(a_0 + a_1 x)$ if $a_2 x^2$ is numerically insignificant compared to a_0. For example, under the exponential failure rate model, the asymptotic approximation for v is

$v \approx \exp(\beta s_0) r[1 + (1/2)\beta r + (1/6)(\beta r)^2 + ...]$.

Therefore, v may be computed as $\exp(\beta s_0) r(1 + \beta r/2)$ if

$1 + (\beta r)^2/6 = 1$

to the limits of the machine accuracy.

g. On a machine where a number has approximately 16 significant digits (IBM PC double precision), for 5-digit accuracy in all cases, including cases when $\bar{s}$ is virtually zero, the expansion for the linear hazard model should be evaluated out to the β^4 term. If this term is negligible compared to a_0, the series through the β^3 term should be used to evaluate the expression.

The constant λ_0 is interpreted as the value of $\lambda(t)$ at time $t = 0$. This time 0 is customarily taken to be the component's installation time, but any other time is allowed in principle. Measuring t from a time other than the installation may make t negative, which is allowed. If each component has a different λ_{0j}, the hazard function of each component changes by the same relative amount in any specified time, but the hazard functions of the components are not equal. For example, the hazard function doubles every $(\log 2)/\beta$ time units, regardless of λ_{0j} and regardless of what time is assigned the value 0.

Linear Hazard Function

The hazard function is defined by

$$\lambda(t) = \lambda_0 + at = \lambda_0(1 + \beta t),$$

with β measured in units of 1/time. This distribution is mentioned by Johnson and Kotz (1970b). Salvia (1980) uses the model with test data, in which many components are tested until their first failures. Vesely (1987) uses the model with field data for which failures from aging (corresponding to the increasing portion of the hazard function) can be distinguished from failures from other causes (corresponding to the constant portion of the hazard function). The cases considered by Salvia and Vesely both turn out to be much simpler analytically than the cases considered in this report.

As with the exponential hazard model, it is possible that λ has the specified form for the time period for which data are observed, and some other form for other t. Therefore, it is possible for β to be negative. However, β must not be such that $\lambda(t)$ is negative in the observed time period. In fact, not even $\lambda(t) = 0$ is allowed, because $\log\lambda(t)$ is often used in the methods below. The details are complicated by the fact that it is sometimes convenient to center the data, leading to observed times expressed as negative values. Let s_{0j} and s_{1j} be the beginning and ending observation times for component j, following the unified notation defined above. To keep $\lambda(t)$ positive for all observed times, β must satisfy $\beta > -1/s_{1j}$ for all positive s_{1j}, and $\beta < -1/s_{0j}$ for all negative s_{0j}.

The constant λ_0 is the value of the hazard function at time $t = 0$. This time is the component's installation time, or the central time, depending on how time is measured. Note that the *relative* change in the hazard function approaches 0 as $t \to \infty$. For example when $\beta > 0$, the hazard function doubles from the value at $t = 0$ in $1/\beta$ time units, doubles again in the next $2/\beta$ time units, and so forth.

Weibull Hazard Function

The hazard function is defined by

$\lambda(t) = \lambda_0(t/t_0)^\beta$,

where $t_0 > 0$ is a normalizing time. It is common (Johnson and Kotz, 1970a, Cox and Oakes, 1984) to write the exponent as $c - 1$. The β notation is consistent with the other two examples because $\beta = 0$ corresponds to a constant failure rate. Both t and t_0 have units of time, and β is dimensionless. The constant λ_0 is measured in units of 1/time, and is the value of the failure rate at time $t = t_0$. Changing t_0 does not change the value of β, but does change the value of λ_0. For $\lambda(t)$ to be integrable at 0, β must satisfy the constraint $\beta > -1$. Negative times are not allowed. If $\beta > 0$, $\lambda(0)$ equals 0; if $\beta \leq 0$, $\lambda(0)$ is undefined.

The hazard function doubles between times t_1 and t_2 if $\log t_2 - \log t_1 = (\log 2)/\beta$. Because $\lambda(0)$ is either zero or undefined, the hazard function cannot double from the initial value.

3. LIKELIHOOD

3.1 Summary of Likelihood Formulas

In this section, the expressions for the likelihood are presented. All derivations and proofs are given in Section 8.

Let C denote the number of components. Define

$$H(t;\beta) = \int_0^t h(u;\beta)\, du$$

and

$$v_j(\beta) = H(s_{1j};\beta) - H(s_{0j};\beta) \quad . \tag{3}$$

Depending on whether the data are time- or failure-censored, v_j is fixed or is the realization of a random variable. The parameter β will sometimes not be shown.

The logarithm of the likelihood based on all the data is shown in Section 8 to be

$$L_{full}(\beta, \lambda_{01}, ..., \lambda_{0c}) = \sum_{j=1}^{c} \left[\sum_{i=1}^{m_j} \log h(t_{ij};\beta) + m_j \log\lambda_{0j} - \lambda_{0j} v_j(\beta) \right] . \quad (4)$$

This follows the unified notation established earlier, with the interpretation of m_j and s_{1j} depending on the way the data for the jth component were generated. The values of λ_{0j} may be distinct, or assumed to all be equal to a common λ_0. In the latter case, L_{full} depends only on β and λ_0, and can be written as

$$L_{full}(\beta, \lambda_0) = \sum_{j=1}^{c} \left[\sum_{i=1}^{m_j} \log h(t_{ij};\beta) + m_j \log\lambda_0 - \lambda_0 v_j(\beta) \right] . \quad (4')$$

Now consider the conditional distribution of the ordered failure times, conditional on the values of n_j or t_{mj}, whichever is random. The conditional log-likelihood is shown in Section 8 to be

$$L_{cond}(\beta) = \sum_{j=1}^{c} \left[\sum_{i=1}^{n_j} \log h(t_{ij};\beta) - n_j \log v_j(\beta) + \log(n_j!) \right] \quad (5)$$

$$= \sum_{j=1}^{c} \log\{(n_j!) \prod_{i=1}^{n_j} [h(t_{ij};\beta)/v_j(\beta)]\} . \quad (5')$$

From now on, the subscripts *full* and *cond* will be omitted, with the meaning being clear from the number of parameters given as arguments of L. It is crucial to note that the conditional log-likelihood (5) depends on β, but not on λ_0 or the λ_{0j}'s.

For component j, consider the term inside curly brackets in Expression (5′), and suppress the index j. The expression is the conditional joint density of the ordered failure times $(T_1, ..., T_n)$. Therefore, conditional on $N = n$ or $T_m = t_m$, the n unordered failure times T_i are independent and identically distributed (i.i.d.), each with density $h(t)/v$ on the interval $[s_0, s_1]$, and density 0 outside this interval. Therefore, inference for β can be performed in standard ways, based on observations that are conditionally independent, and conditionally identically distributed for each component. This can be done whether or not the components have a common value of λ_0.

Two other facts are needed to carry out inference for all the parameters. For time-censored data, N_j is Poisson($\lambda_{0j} v_j$). For failure-censored data, it is shown in Section 8 that $2\lambda_{0j} V_j$ has a $\chi^2(2m_j)$ distribution. The values of λ_{0j} may or may not be assumed to equal some common value.

3.2 Ancillarity

Suppose that there is a multidimensional paramenter (β, θ), and a sufficient statistic (X, Y). Y is said to be *ancillary* for β if the marginal distribution of Y does not depend on β. X is called *conditionally sufficient* for β if the conditional distribution of X given y does not depend on θ. When these conditions hold, inference for β should be based on the conditional likelihood of X given y. When maximum likelihood estimation is used, the same value for $\hat{\beta}$ is found whether the full likelihood or the conditional likelihood is used, but the appropriate variance of $\hat{\beta}$ is the conditional variance. See Kalbfleisch (1982) or Cox and Hinkley (1974, Sections 2.2viii and 4.8ii) for more information.

Return now to the setting of component failures, and consider time-censored data from C components, when either (1) the components are not assumed to have a common value of λ_0, or (2) the components have a common λ_0 and all the v_j's have a common value. In the examples of this report, case (2) can occur only if all the components are observed over the same period s_0 to s_1. For case (1), it is shown in Section 8 that (N_1, ..., N_C) is ancillary for β, and that the failure times T_{ij} form a conditionally sufficient statistic for β. (A lower dimensional conditionally sufficient statistic for β can be determined in some examples by examining the form of $\Sigma\Sigma\log h(T_{ij})$.) For case (2), the components may be pooled into a single super-component, and $N = \Sigma N_j$ is ancillary for β. In these cases, therefore, basing inference for β on Equation (5) is not only possible but best. In all other cases, basing inference for β on Equation (5) involves some loss of information.

3.3 Examples

The building blocks for the above formulas are all given in Table 1, at the end of this report. A few points are worth noting here: The exponential hazard model is worked out in some detail by Cox and Lewis (1966, Section 3.3). With this model, $\Sigma\Sigma\log h(T_{ij};\beta)$ equals $\beta\Sigma\Sigma T_{ij}$, and it follows that that $\Sigma\Sigma T_{ij}$ is conditionally sufficient for β. For the linear hazard function, $\Sigma\Sigma\log h(T_{ij};\beta)$ equals $\Sigma\Sigma\log(1 + \beta T_{ij})$, and there is no one-dimensional statistic that is conditionally sufficient for β. This is one of several problems with the linear hazard model, which will be mentioned in this report as they are encountered. For the Weibull hazard function, we have $\log h(T;\beta) = \beta\log(T/t_0)$. Therefore, $\Sigma\Sigma\log T_{ij}$ is conditionally sufficient for β.

4. MAXIMUM LIKELIHOOD ESTIMATION

4.1 Maximum Likelihood Estimation Based on the Conditional Likelihood

If $(N_1, ..., N_C)$ is ancillary for β, then inference for β should be based on the conditional log-likelihood given by Equation (5). Even in other cases, one could use this conditional log-likelihood at the cost of some loss of information. The maximum conditional likelihood equation is formed by setting the derivative of Expression (5) with respect to β equal to 0, resulting in:

$$L'(\beta) = \sum_{j=1}^{C} \sum_{i=1}^{n_j} \left\{ [\log h(t_{ij};\beta)]' - [\log v_j(\beta)]' \right\} = 0 \ . \tag{6}$$

Here, the prime denotes the derivative with respect to β. If β has dimension k, there are k such equations, each involving the partial derivative with respect to one component of β. The maximum likelihood estimate (MLE) $\hat{\beta}$ typically is found by numerical iteration to solve Equation (6). If any algebraic cancellation can be performed on the terms inside the curly brackets in Equation (6), then the order of evaluation should be as suggested by the bracketing, for numerical accuracy. If no algebraic cancellation can be performed, the evaluation may take advantage of the fact that $\Sigma_i[\log v_j]' = n_j[\log v_j]'$.

Suppose that no common value of λ_0 is assumed. The MLE of λ_{0j}, corresponding to the jth component, is $\hat{\lambda}_{0j} = m_j/v_j(\hat{\beta})$. This is shown directly from Equation (4) by maximizing $L(\hat{\beta}, \lambda_{01}, ..., \lambda_{0C})$ with respect to λ_{0j}. Suppose instead that a common value of λ_0 is assumed for all C components. Then it is shown similarly that $\hat{\lambda}_0 = \Sigma m_j/\Sigma v_j(\hat{\beta})$.

4.2 Maximum Likelihood Estimation Based on the Full Likelihood

Inference proceeds first by estimating λ_0, if a single common value is assumed, or by estimating the various λ_{0j}. Substitute the MLE(s) into the expression for the full log-likelihood, differentiate the resulting expression with respect to β, and find the MLE $\hat{\beta}$.

When no common λ_0 is assumed, the equation for $\hat{\beta}$ is

$$(\partial/\partial\beta)L(\beta, \hat{\lambda}_{01}, ..., \hat{\lambda}_{0C}) = \sum_{j=1}^{C} \sum_{i=1}^{m_j} \left\{ [\log h(t_{ij};\beta)]' - [\log v_j(\beta)]' \right\} = 0 \ . \tag{7}$$

This is identical to Equation (6), except that m appears in place of n. Therefore, use of either the conditional or the full likelihood yields the same MLE $\hat{\beta}$ from time-censored data; this agrees with the conclusion of the ancillarity argument given earlier. For failure-censored data, Equation (7) differs from Equation (6) by inclusion of the final failure times t_m and use of $m = n + 1$.

When a common λ_0 is assumed, the maximum likelihood equation for $\hat{\beta}$ is

$$\sum_{j=1}^{c} L_j'(\beta, \hat{\lambda}_0) = \sum_{j=1}^{c} \sum_{i=1}^{m_j} [\log h(t_{ij};\beta)]' - (\Sigma m_j)\, [\Sigma v_j'(\beta)]/[\Sigma v_j(\beta)] = 0 \quad . \tag{8}$$

This differs from Equation (6) in two ways: m_j is used instead of n_j, which makes a difference only with failure-censored data; and the portion involving v_j reverses the order of summation and multiplication and division.

4.3 Examples

All the expressions used in Equations (6) through (8) are presented in Table 1, for the three examples. A few points of interest are mentioned here. Typical features of all the models are discussed using the first example as an illustration.

Exponential Hazard Function

Consider first estimation based on the conditional likelihood. The maximum conditional likelihood equation for β is, from Equation (6) and the expressions given in Table 1,

$$\sum_{j=1}^{c} \sum_{i=1}^{n_j} (t_{ij} - s_{0j}) + \sum_{j=1}^{c} n_j/\beta - \sum_{j=1}^{c} n_j r_j/[1 - \exp(-\beta r_j)] = 0 \quad . \tag{9}$$

This agrees with the special case $C = 1$ and $s_0 = 0$ worked out by Cox and Lewis (1966). It must be solved numerically for $\hat{\beta}$. When β is near 0, the last two terms in Equation (9) are very large, although the difference is bounded. Therefore an asymptotic approximation should be used. From expressions given in Table 1, a first order approximation is

$$\sum_{j=1}^{c} \sum_{i=1}^{n_j} \left\{ (t_{ij} - s_{0j}) - (r_j/2)(1 + \beta r_j/6) \right\} = 0 \quad .$$

When β is small, this asymptotic approximation must be used to prevent complete loss of numerical significance; of course, when $\beta = 0$ the limiting value must be used. Note that $\hat{\beta}$ equals 0 when

$\Sigma\Sigma t_{ij} = \Sigma n_j \bar{s}_j$,

that is, when the sum of the (non-replacement) failure times equals the corresponding sum of the mid-points of the observation periods. This is intuitively consistent with the fact that when β equals 0, the conditional distribution of T_{ij} is uniform on (s_0, s_1). The MLE for λ_0 or for the λ_{0j}'s can be obtained in a direct way from the results given above.

Inference based on the full likelihood is similar, using Equation (7) or (8) and expressions given in Table 1.

Linear Hazard Function

It is straightforward to substitute the expressions for $h(t)$ and v_j into the general equations given above. For example, consider the conditional log-likelihood based on a single component. Its derivative is

$L'(\beta) = \Sigma t_i/(1 + \beta t_i) - n\,\bar{s}/(1 + \beta\bar{s})$.

It follows that the MLE $\hat{\beta}$, based on the conditional log-likelihood, equals zero if $\Sigma\Sigma t_{ij} = \Sigma n_j \bar{s}_j$, just as with the exponential hazard model. The following two points, however, deserve special notice:

The MLE $\hat{\beta}$ may be infinite. To see this, consider the expression for $L'(\beta)$ just given. If $t_i > \bar{s}$ for all i, then $L'(\beta)$ is positive for all β. There is no finite solution to the maximum likelihood equation. Thus, in cases when the evidence for an increasing failure rate is strongest, the rate of increase may not be estimable by maximum likelihood.

With time-censored data and a common λ_0 assumed, there is some advantage to centering the data. In this case $m_j \equiv n_j$, and the full log-likelihood is

$L(\beta, \lambda_0) = \Sigma n_j \log\lambda_0 + \Sigma\Sigma\log(1 + \beta t_{ij}) - \lambda_0\Sigma r_j - \lambda_0\beta\Sigma r_j\bar{s}_j$.

The last sum can be made to vanish by centering the data, that is, by measuring all times from

$t_{mid} = \Sigma r_j\bar{s}_j \,/\, \Sigma r_j$.

The log-likelihood then becomes

$L(\beta, \lambda_0) = \Sigma n_j \log\lambda_0 + \Sigma\Sigma\log[1 + \beta(t_{ij} - t_{mid})] - \lambda_0\Sigma r_j$.

In this formulation, λ_0 equals the value of $\lambda(t)$ at $t = t_{mid}$. If any value is assumed for β, ΣN_j is

Poisson($\lambda_0\Sigma r_j$), independent of β. Similarly, if any value is assumed for λ_0, $L(\beta,\lambda_0)$ is a function of λ_0 plus a function of β and the t_{ij}'s; therefore, inference for β is independent of λ_0. This ability to perform independent inference for β and λ_0 is a convenient property, which may be sufficient in the eyes of some analysts to justify centering the data.

Suppose that when the data are uncentered, there is no finite MLE $\hat{\beta}$. Centering the data is not a cure-all. When the data are centered, β is restricted to a finite range, as discussed in the introduction to the linear hazard model in Section 2. In this case, the MLE $\hat{\beta}$ is at an end point of the possible range; it is finite, but cannot be treated as asymptotically normal.

Weibull Hazard Function

In this case, $[\log h(t_{ij})]' = \log(t_{ij}/t_0)$. The remaining terms needed for Equations (6), (7), and (8) depend on whether s_{0j} is zero or nonzero, and are all given in Table 1.

There is a noteworthy simplification in Equations (6) and (7) when $s_{0j} = 0$ for all j, that is, when every component is observed from its time of installation. In this case, $[\log v]'$ equals $\log(s_1/t_0) - 1/(\beta + 1)$, and Equation (6) has the explicit solution

$$\hat{\beta} = -\Sigma n_j/\Sigma\Sigma\log(t_{ij}/s_{1j}) - 1 \ . \qquad (10)$$

The solution of Equation (7) replaces n_j by m_j. These are the only cases considered in this report for which the MLE $\hat{\beta}$ can be found without numerical iteration.

In this case, the value $\hat{\beta}$ satisfying Equation (6) equals 0 not when $\Sigma\Sigma t_{ij}$ equals $\Sigma n_j\bar{s}_j$, as in the other examples, but when

$$-\Sigma\Sigma\log(t_{ij}/s_{1j}) = \Sigma n_j \ .$$

This initially surprising fact has the following intuitive basis. For notational simplicity, consider a single component, suppress the index j, let $t_0 = 1$, and condition the observations on the value of n or s_1. To derive the conditional distribution of $-\log(T_i/s_1)$, begin with

$$P[-\log(T_i/s_1) > x] = P[T_i < s_1\exp(-x)] \ .$$

Following the discussion below Equation (5), T_i has conditional density $h(t)/v$; therefore, this probability equals

$$\{[s_1\exp(-x)]^{\beta+1}/(\beta + 1)\} \ / \ \{s_1^{\beta+1}/(\beta + 1)\} = \exp[-x(\beta + 1)] \ .$$

Therefore, the conditional distribution of $-\log(T_i/s_1)$ is exponential with mean $\mu = 1/(\beta + 1)$.

Equation (10) can be rewritten as

$-\Sigma\Sigma\log(t_{ij}/s_{1j}) \;/\; \Sigma n_j = 1/(\hat{\beta} + 1) = \hat{\mu}$,

that is, the MLE is based on equating the mean of $-\log(T_{ij}/s_{1j})$ to the sample mean. In particular, the case $\hat{\beta} = 0$ corresponds to $\hat{\mu} = 1$, that is, $-\Sigma\Sigma\log(t_{ij}/s_{1j}) \;/\; \Sigma n_j = 1$.

When the values of s_{0j} are not all zero, the expressions are more complicated, but the maximum likelihood equation is still equivalent to setting the mean of $\Sigma\Sigma\log T_{ij}$ equal to its sample mean.

5. CONFIDENCE REGIONS AND HYPOTHESIS TESTS

The standard regularity conditions, such as given by Cox and Hinkley (1974, Section 9.1) are assumed. The assumptions involving the parameter space, identifiability of the distributions, and existence of derivatives are all satisfied in the examples considered in this report. There is also an assumption involving the behavior of the third derivative of the log-likelihood as n goes to infinity. For field data, such an assumption is typically difficult to affirm or deny. Practitioners must always treat asymptotic approximations with care.

5.1 Inference Based on the Conditional Likelihood

The procedure described here might be used when β is the primary parameter of interest, or when $(N_1, ..., N_C)$ is ancillary for β. The presentation here assumes that β is one-dimensional. The generalizations to multidimensional β are straightforward. We remark in passing that when $\log h(t)$ is linear in one-dimensional β, as is the case for the exponential and Weibull models, then the one-sided tests given below are uniformly most powerful.

Inference for β

The derivative with respect to β of the conditional log-likelihood, $L'(\beta)$, is given by Equation (6). The information is

$I(\beta) = -E[L''(\beta)] = E\{\, [L'(\beta)]^2 \,\}$

$$= -E\left\{\sum_{j=1}^{c}\sum_{i=1}^{n_j}[\log h(t_{ij};\beta)]'' - \sum_{j=1}^{c} n_j\,[\log v_j(\beta)]''\right\}$$

$$= \Sigma_j n_j\{-\int [\log h(t;\beta)]'' h(t;\beta)\,dt/v_j(\beta) + [\log v_j(\beta)]''\}\ . \qquad (11)$$

If β is k-dimensional, $I(\beta)$ is the $k\times k$ matrix defined by taking all the mixed partial derivatives of L. Let β be the true value. Under the assumed regularity conditions, the expectation of $L'(\beta)$ is 0. and the variance (or covariance matrix for k-dimensional β) of $L'(\beta)$ is $I(\beta)$.

As a corollary to the Lindeberg-Feller Central Limit Theorem, Feller (1968, Section X.5) gives a sufficient condition for asymptotic normality of $L'(\beta)$. Rewrite Equation (6) as $L'(\beta) = \Sigma\Sigma X_k$. If there is a constant A such that $|X_k| < A$ for all k, and if (11) $\rightarrow \infty$, then

$$L'(\beta_0)\ /\ [I(\beta_0)]^{1/2} \qquad (12)$$

converges in distribution to normal(0,1). The assumptions must be verified for each example. Typically, the assumptions are satisfied if all the values of s_{0j} and s_{1j} are bounded by some constant, and if some fixed fraction of the r_j's is bounded away from 0. For the exponential hazard model, it is enough for the r_j's to be bounded by some constant and for a fixed fraction to be bounded away from 0. For the linear hazard model, it is necessary in addition for $1 + \beta s_{0j}$ and $1 + \beta s_{1j}$ to be uniformly bounded away from 0. Qualitatively, the approximation is best if the s_{0j}'s are approximately equal and if the s_{1j}'s are approximately equal. The approximation also is better if β and h are such that $[\log h(T_{ij};\beta)]'$ does not have a highly skewed distribution. If it is very important to know whether the normal approximation is adequate in some application, a simulation study should be performed.

An approximate confidence interval for β is the set of all β_0 such that the statistic (12) lies in the interval $(-c,\ c)$, where c is the appropriate number from a normal table; for example, $c = 1.96$ yields an approximate 95% confidence interval. Actually, this defines a confidence *region* for β. To show that the region is an interval rather than some more complicated set, one must show that Expression (12) is a monotone function of β_0. Monotonicity is difficult to show analytically. It can be checked numerically by a computer program in any example. In experience so far with real data, (12) has always been monotone for the exponential hazard model, but has not always been monotone with the linear hazard model when the confidence interval was unbounded, or for the Weibull hazard model near $\beta = -1$.

To test the hypothesis $\beta = \beta_0$ for some particular value β_0, the test statistic (12) can be used, and the hypothesis rejected if the test statistic is in an extreme tail of the normal distribution. In particular, the hypothesis $\beta = 0$ is often of interest; the test statistic (12) may then have an especially simple form, as discussed below for the examples.

Inference for λ_0

Once a value of β is assumed, it is easy to find a confidence interval for λ_0 or confidence intervals for the various λ_{0j}'s. The method is shown here when the components are assumed to have a single common λ_0.

For time-censored data, define $N = \Sigma N_j$ and $v = \Sigma v_j$ with v evaluated at the assumed value of β. Because N is Poisson($\lambda_0 \Sigma v_j$), a two-sided 100(1 − α)% confidence interval for λ_0 is given by Johnson and Kotz (1969, Section 6.2) as

$$\lambda_{0L} = \chi^2_{2n,\alpha/2}/(2v)$$
$$\lambda_{0U} = \chi^2_{2(n+1),1-\alpha/2}/(2v) \ . \qquad (13)$$

If instead the data are failure-censored, define $m = \Sigma m_j$ and $v = \Sigma v_j$ with v evaluated at the assumed value of β. Because $2\lambda_0 V$ has a $\chi^2(2m)$ distribution, a two-sided 100(1−α)% confidence interval for λ_0 is given by

$$\lambda_{0L} = \chi^2_{2m,\alpha/2}/(2v)$$
$$\lambda_{0U} = \chi^2_{2m,1-\alpha/2}/(2v) \ . \qquad (14)$$

Note that Formulas (13) and (14) agree except for the degrees of freedom.

A two-dimensional confidence region, with confidence coefficient approximately 100(1 − α)%, can be formed as follows. Form a 100(1 − α/2)% confidence region for β. At each β_0 in the confidence interval, evaluate v and form the resulting 100(1 − α/2)% confidence interval for λ_0. The approximation results from the use of a large-sample approximation for the confidence interval for β, and from the way the two individual confidence coefficients are combined to yield a joint confidence coefficient.

If β is treated as known and equal to $\hat{\beta}$, Equations (13) or (14) give an approximate confidence interval for λ_0. It is too short, however, because it does not account for the randomness of the estimator $\hat{\beta}$. If this interval for λ_0 depends strongly on the assumed value of β, a more exact confidence interval is obtained by taking the largest and smallest values of λ_0 in the two-dimensional region for (β, λ_0).

A conservative confidence interval for the hazard function $\lambda(t)$ is given by the largest and smallest values of $\lambda(t)$ attained in the two-dimensional confidence region for (β, λ_0).

5.2 Inference Based on the Full Likelihood

When all the model parameters are of interest, an analyst either could follow the procedure presented above, or could perform inference based on the full model as follows. The discussion assumes that all the components have a common λ_0. Formulas for λ_0 will be based on joint asymptotic normality. There are heuristic arguments for why parameterization in terms of $\rho = \log\lambda_0$ improves the normal approximation: for failure-censored data, this transformation replaces the scale parameter λ_0 by a location parameter; also, the log transformation of Equations (13) and (14) yields more nearly symmetrical intervals.

The log-likelihood $L(\beta, \lambda_0)$ is given by Equation (4′). The sample information matrix for $(\beta, \rho) \equiv (\beta, \log\lambda_0)$ is

$$SI(\beta, \log\lambda_0) = -\begin{bmatrix} (\partial^2/\partial\beta^2)L(\beta, \lambda_0) & (\partial^2/\partial\beta\partial\rho)L(\beta, \lambda_0) \\ (\partial^2/\partial\beta\, \partial\rho)L(\beta, \lambda_0) & (\partial^2/\partial\rho^2)L(\beta, \lambda_0) \end{bmatrix}$$

$$= \Sigma_j \begin{bmatrix} -\{\Sigma_i[\log h(t_{ij})]''\} + \lambda_0 v_j'' & \lambda_0 v_j' \\ \lambda_0 v_j' & m_j \end{bmatrix} . \qquad (15)$$

In some situations, evaluation of the above terms at $(\hat\beta, \hat\lambda_0)$ is made easier by using the identities $\Sigma m_j/\hat\lambda_0 = \Sigma v_j$ and $\Sigma\Sigma[\log h(t_{ij})]' = \hat\lambda_0 \Sigma v_j'$, with the second identity following from Equation (8) evaluated at $(\hat\beta, \hat\lambda_0)$.

The information matrix is then defined by

$I(\beta, \log\lambda_0) = E[SI(\beta, \log\lambda_0)]$.

The expectation is based on the randomness of T_{ij} and of either V_j or M_j. Depending on the form of

h, the analyst may choose to estimate the information matrix by $I(\hat{\beta}, \log\hat{\lambda}_0)$ or by $SI(\hat{\beta}, \log\hat{\lambda}_0)$; see Cox and Hinkley (1974, p. 302). In practice, expecially when V_j is random, it is much more convenient to use SI to estimate $I(\beta, \log\lambda_0)$.

Asymptotic inference is based on the fact that $(\hat{\beta}, \log\hat{\lambda}_0)$ is asymptotically normal with mean $(\beta, \log\lambda_0)$ and covariance matrix $I^{-1}(\beta, \log\lambda_0)$. This allows for approximate confidence intervals for β, for λ_0, and for functions of the two parameters, such as $\lambda(t)$. To do the last, write

$\log\hat{\lambda}(t) = \log\hat{\lambda}_0 + \log h(t;\hat{\beta})$.

Take the first-order Taylor expansion of $\log h(t;\hat{\beta})$ around $\hat{\beta} = \beta$. This yields the asymptotic distribution of $\log h(t;\hat{\beta})$, and its asymptotic covariance with $\log\hat{\lambda}_0$. Then $\log\hat{\lambda}(t)$ is asymptotically normal, with mean equal to the sum of the means, and variance equal to the sum of the variances plus twice the covariance. This may be used for t such that the Taylor approximation is adequate.

5.3 Examples

The building blocks for the formulas are all given in Table 1. Asymptotic approximations are also given, to be used when β is near 0 with an exponential or linear hazard function, and when β is near -1 with a Weibull hazard function. Special cases are now considered.

Exponential Hazard Function

To test $\beta = 0$, based on the conditional log-likelihood, the asymptotic formulas in Table 1 show that the test statistic (12) equals

$$\Sigma_j\left\{\Sigma_i t_{ij} - n_j\bar{s}_j\right\}\Big/[\Sigma_j n_j r_j^2/12]^{1/2} \qquad (16)$$

Here i goes from 1 to n_j. When there is just one component ($j = 1$), the statistic becomes

$[\Sigma_i t_i/n - \bar{s}]/[r/(12n)^{1/2}]$,

which has a simple intuitive interpretation. If the failure rate is constant ($\beta = 0$), the conditional distribution of the failure times for the component is uniform between s_0 and s_1. The test statistic is the average observed time minus the midpoint of the observation period, all divided by the standard deviation of an average of uniformly distributed variables. This test was first proposed by Laplace in 1773, according to Bartholemew (1955).

In this case, $\log\lambda(t) = \log\lambda_0 + \beta t$. Therefore, the asymptotic distribution of $\log\hat{\lambda}(t)$ follows neatly from the asymptotic distribution of $(\hat{\beta}, \log\hat{\lambda}_0)$.

Linear Hazard Function

Recall that time-censored data can be centered. This redefines the meaning of λ_0 and β, the function $h(t)$ becomes $1 + \beta(t - t_{mid})$, and $\Sigma v_j{}'$ equals 0. The sample information matrix (15) then becomes a diagonal matrix, and $\hat{\beta}$ and $\hat{\lambda}_0$ are asymptotically uncorrelated.

The test of $\beta = 0$, based on the conditional log-likelihood, can be built from the elements in Table 1. The statistic is given by Expression (16). That is, the natural large-sample test of constant failure rate is the same, whether an exponential or linear hazard model is postulated.

The asymptotic distribution of $\lambda(t)$ is obtained by making the approximation
$\log h(t;\hat{\beta}) \doteq \log(1 + \beta t) + (\hat{\beta} - \beta)t/(1 + \beta t)$.
The approximation may be used when the second term is small compared to 1. For practical use, the approximation is good enough if twice the standard deviation of $\hat{\beta}t/(1 + \beta t)$ is less than 0.1, and fair if this standard deviation is less than 0.5.

Weibull Hazard Function

The necessary expressions are given in Table 1. In this model, the test statistic (12) differs from Expression (16). When all the values of s_{0j} equal 0, the test statistic simplifies to

$$\{\Sigma\Sigma[\log(t_{ij}/s_{1j}) + 1]\} \;/\; (\Sigma n_j)^{1/2} \quad , \qquad (17)$$

with i going from 1 to n_j. Recall from the discussion of maximum likelihood estimation below Equation (10) that the conditional distribution of $-\log(T_{ij}/s_{1j})$ is exponential with mean and variance equal to $1/(\beta + 1)$, and that the MLE of $1/(\beta + 1)$ is the sample mean of the terms $-\log(t_{ij}/s_{1j})$. Therefore, the negative of the test statistic (17) can be written as the MLE of $1/(\beta + 1)$ standardized by the mean and variance when $\beta = 0$.

The estimated hazard function satisfies $\hat{\lambda}(t) = \log\hat{\lambda}_0 + \hat{\beta}\log(t/t_0)$, so the asymptotic normal distribution follows directly from the corresponding result for $(\hat{\beta}, \log\hat{\lambda}_0)$.

6. DIAGNOSTIC CHECKS

The methods presented above have assumed a common value of β for all components, perhaps a common value of λ_0, and a hazard function of the form $\lambda_0 h(t;\beta)$. Computations are often based on the assumption that asymptotic normality yields an adequate approximation. Diagnostic checks—both tests and plots—should be used to investigate the validity of these assumptions.

6.1 Common β

To see if a particular component, the kth say, has β significantly different from the other components, calculate the MLE based on the kth component only and on all components (pooled) except the kth. At this point there is no reason for confidence that the components have a common λ_0; therefore, use the MLE based on the conditional likelihood, which is independent of the value(s) of λ_0. The difference $\hat{\beta}_k - \hat{\beta}_{-k}$ has variance equal to the sum of the variances, and mean zero if all components have the same β. Therefore it yields a test, using asymptotic normality, of the hypothesis that the kth component has the same β as do the others. The C tests can be combined using the Bonferroni inequality to form an overall test of the hypothesis that the components have a common β. If any component has no nonreplacement failures, β cannot be estimated for that component, and fewer than C test statistics and confidence intervals can be calculated.

A single component may not have enough failures to justify asymptotic methods. In the extreme case when the kth component has only one non-replacement failure, a practical expedient is to treat β_{-k} as known, and test whether $\beta_k = \beta_{-k}$ based on the single observed failure time for the kth component. This test is based on the fact that the single failure has conditional density $h(t)/v_k$, with β set to β_{-k}.

In addition to the test for common β, a useful visual diagnostic is a plot of C confidence intervals for the parameter, placed side by side, with each interval based on the data from a single component.

6.2 Common λ_0

Suppose that the assumption of a common β is accepted, and consider how to test whether the components have a common λ_0. Treat β as known and equal to $\hat{\beta}$; this introduces an approximation into the tests for λ_0, but it does not a priori treat any component differently from any other. Consider now the kth component, pool all the components except the kth, and test whether λ_{0k} equals $\lambda_{0,-k}$. Assume for the moment the null hypothesis that the components have a common λ_0.

With time-censored data, the conditional distribution of N_k, conditional on the ancillary statistic Σn_j, is binomial(Σn_j, p_k), with $p_k = v_k(\beta)/\Sigma v_j(\beta)$. This yields a test of the hypothesis that λ_{0k} is the same as λ_0 for the other components. These tests may be combined with the Bonferroni inequality. Alternatively, if the failure counts are not too small, a χ^2 test may be used, based on the fact that $(N_1, ..., N_C)$ is multinomial $(\Sigma n_j, p_1, ..., p_C)$.

With failure-censored data, the distribution of $2\lambda_0 V_k(\beta)$ is $\chi^2(2m_k)$, and the sum of the observation periods for all components except the kth is likewise proportional to a χ^2 random variable. Therefore the ratio of V_k to the sum of such terms over all components except the kth is proportional to an F random variable. This yields a test of the hypothesis that λ_{0k} is the same as λ_0 for the other components. The tests may be combined with the Bonferroni inequality.

As when comparing the components for β, a side-by-side plot of confidence intervals for λ_{0j} provides useful visual diagnostic information.

6.3 Form of $h(t)$

To test whether h is of the assumed form, use the fact that for the jth component, conditional on the observed failure count n_j or on the final observation time s_{1j}, the T_{ij}'s are independent and for each component are identically distributed, with density proportional to h, as discussed below Equation (5′). Therefore, under the assumed model, the conditional probability that a random failure T occurs by time t is

$$\begin{aligned} P[T \le t] &= \sum_j P[T \le t \mid \text{failure is in component } j]\, P[\text{failure is in component } j] \\ &= \sum_j P[T \le t \mid \text{failure is in component } j]\, (n_j/\Sigma n_i) \;. \end{aligned}$$

with

$$\begin{aligned} P[T \le t \mid \text{failure is in component } j] &= [H(t) - H(s_{0j})]/v_j \quad &&\text{if } s_{0j} \le t \le s_{1j} \\ &= 0 &&\text{if } t < s_{0j} \\ &= 1 &&\text{if } t > s_{1j} \;. \end{aligned}$$

Tests for a hypothesized distribution may now be used, such as the Kolmogorov-Smirnov test or the

Anderson-Darling test.

Routine use of one of these tests gives a Type I error smaller than the nominal value: the test tends not to reject often enough. There are two reasons for this. One is the familiar reason that the estimated value of β must be used to evaluate H and v. The second reason arises if the components are observed over different time periods. The distribution used is conditional on the failure counts or final failure times, so the T_{ij}'s are not truly a random sample. As an extreme example, suppose that component 1 was observed for only the first year of its life and that it had n_1 failures, that component 2 was observed for only the second year of its life and that it had n_2 failures, and so forth. The conditional distribution then says that of Σn_j failures in the first C years, on the average n_i will occur in year i. The T_{ij}'s are a stratified sample from this distribution, and are therefore forced to fit the distribution rather well. They fit well regardless of the form of h, because the stratification does not involve the hypothesized h.

To avoid this difficulty, it is good to try to use components that are observed over the same time period; if a few components have a different observation window from all the others, try partitioning the data and performing the test on the two sets separately. In the extreme case given by the above example, the following method could be used. Find $\hat{\beta}$ using all the data, and treat it as known. Then for each of the C components perform a separate Kolmogorov-Smirnov test of H_0: $\beta = \hat{\beta}$. This yields p_1, ..., p_C, the attained significance levels or p-values. It is well-known that under H_0, a p-value is uniformly distributed on (0, 1), so that $-2\Sigma\log(p_j)$ has a $\chi^2(2C)$ distribution. Thus H_0 would be rejected at level α if $-2\Sigma\log(p_j) > \chi^2_{1-\alpha}(2C)$.

Two pictures may accompany the test. One is the plot of the above model-based c.d.f. overlayed with the empirical c.d.f.. The other is a Q-Q plot, as described, for example, by Snee and Pfeifer (1983). It plots the n observed failure times versus the inverse of the model-based c.d.f. evaluated at $1/(n+1)$, ..., $n/(n+1)$.

6.4 Adequacy of Asymptotic Normal Approximation

An MLE can be inspected to see if it is near the mid-point of a two-sided confidence interval; if not, the normal approximation may not be adequate. Also, a two-dimensional confidence region for $(\beta, \log\lambda_0)$ can be constructed from an interval for β and conditional intervals for λ_0 given β, as discussed below Equation (14). This can then be compared to the confidence ellipse based on the asymptotic joint normality of ($\hat{\beta}$, $\log\hat{\lambda}_0$). If the two regions are very different, approximate joint normality should be questioned.

7: EXAMPLE DATA ANALYSIS

A nuclear power plant for a commercial utility has 12 motor-operated valves in the auxiliary feedwater systems at the two units of the plant. Maintenance records covering about 10 years were examined, and the failure times for the valves were tabulated. The data are summarized in Table 2, and are given in more detail by Wolford et al. (1990). Three valves were replaced upon failure, and one was replaced for administrative reasons, leading to 16 valves shown in Table 2. The three valves that were replaced upon failure were regarded as failure-censored. The other 13 valves were regarded as time-censored. A Fortran program PHAZE (for Parametric HAZard Estimation) was written and used on a personal computer to analyze the data, following the methods of this report; the program is documented by Atwood (1990).

The valves were first compared to see if they have clearly different values of β. Figure 1 shows a side-by-side plot of the confidence intervals based on the individual components. It also shows the significance levels based on a comparison of $\hat{\beta}_k$ to $\hat{\beta}_{-k}$. The diamond in each confidence interval shows $\hat{\beta}_k$ while the square shows $\hat{\beta}_{-k}$. Note that there is no estimate or interval for components with no non-replacement failures. The overall significance level, based on the Bonferroni combination of the individual significance levels, is 1.0, confirming the pictorial impression that there is no real difference in β for the various components. The exponential hazard function was assumed for these calculations. The results were similar when the linear or Weibull hazard function was assumed. The only striking difference was that many of the MLEs and all of the upper confidence limits were infinite with the linear hazard function. A similar comparison of the components for λ_0 led to a conclusion that the components do not have greatly different values of λ_0. Therefore, the components were assumed to have a common value of β and of λ_0.

Tests of $\beta = 0$ were performed based on the test statistic (12), and the hypothesis was rejected in favor of $\beta > 0$. The test based on $\Sigma\Sigma t_{ij}$, when Expression (12) takes the form of Expression (16), rejected at one-sided level 0.021. The test based on $\Sigma\Sigma \log t_{ij}$, when Expression (12) is evaluated under the Weibull model, rejected at one-sided level 0.025.

Table 2. Summary of example data

Component	Nonrepl. Fails.	Observed Hrs.	Mean Failure Time (Normed)	Replaced on Fail.?	Initial Age (Hrs.)
MOV-1A	1	8.8584E+04	0.378		4.1448E+04
MOV-1B	1	8.8584E+04	0.086		4.1448E+04
MOV-1C	2	8.8584E+04	0.752		4.1448E+04
MOV-1D	7	8.8584E+04	0.743		4.1448E+04
MOV-1E	0	2.1840E+04		Y	4.1448E+04
MOV-1E(R)	3	6.6744E+04	0.498		0.0000
MOV-1F	3	4.3608E+04	0.568	Y	4.1448E+04
MOV-1F(R)	1	4.4976E+04	0.487		0.0000
MOV-2A	4	8.8584E+04	0.619		3.7824E+04
MOV-2B	5	8.8584E+04	0.567		3.7824E+04
MOV-2C	1	4.9728E+04	0.756	Y	3.7824E+04
MOV-2C(R)	1	3.8856E+04	0.866		0.0000E-01
MOV-2D	6	8.8584E+04	0.464		3.7824E+04
MOV-2E	0	2.2608E+04			3.7824E+04
MOV-2E(R)	2	6.5976E+04	0.698		0.0000
MOV-2F	7	8.8584E+04	0.593		3.7824E+04

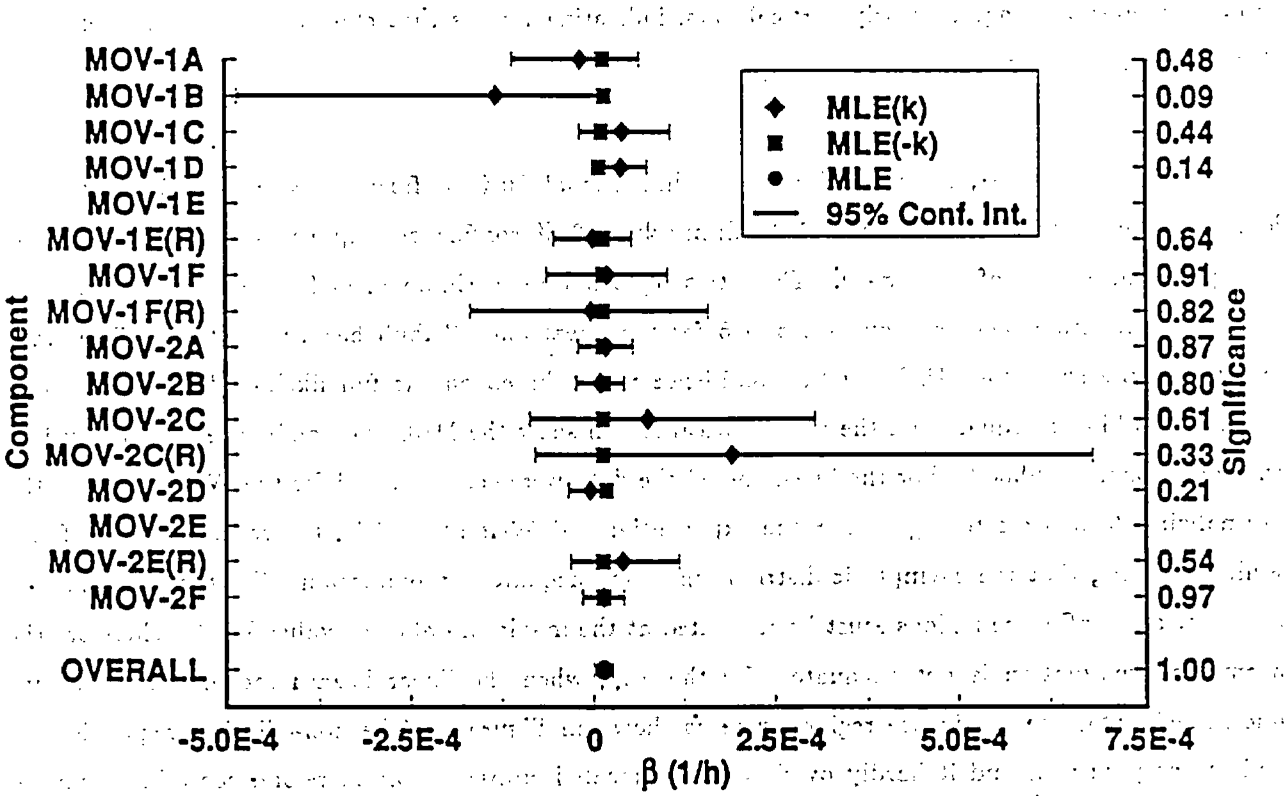

Figure 1. Component Comparisons for β, Exponential Hazard Model

To test the form of the model, the Kolmogorov-Smirnov test was performed, as described in Section 6.3. The test saw nothing wrong with any of the three models; the three significance levels were all greater than 0.8. To account for the partial stratification of the data, the components were partitioned into two groups, the twelve that were in place at the start of observation, and the four that were installed during the observation period. The overall MLE, based on the conditional likelihood for all the components, was used to estimate β. This value was treated as known in the two data sets, and the Kolmogorov-Smirnov test was used to test the fit of each data set to each of the three models. The three significance levels corresponding to the larger data set were calculated using asymptotic formulas and were all greater than 0.79; the significance levels corresponding to the smaller data set (seven failures) were not calculated exactly but were all substantially greater than 0.20. Even allowing for the fact that the hypothesized model had an estimated parameter, it seems that the data give no reason to question any of the three models.

Figure 2 shows the Q-Q plot of the full data set, based on the exponential hazard model. Q-Q plots based on the other models look similar. The only evident departure from the assumed model is shown by several strings of nearly vertical dots, indicating repairs that cluster in time. The effect of this clustering is ignored below.

For each model, an approximate two-dimensional 90% confidence region was found for (β, $\log\lambda_0$), as discussed below Equation (14). Similarly, a 90% confidence ellipse was found based on the asymptotic normality of ($\hat{\beta}$, $\log\hat{\lambda}_0$). These two regions are superimposed in Figure 3 for the exponential hazard function, and in Figures 4 and 5 for the linear and Weibull hazard functions. The circle and the ellipse show the MLE and the confidence region based on the full likelihood and asymptotic normality, while the square and the non-elliptical region show the MLE and confidence region based on the conditional likelihood. For the linear model the data were centered, and for the Weibull model the normalizing t_0 was set to t_{mid}. For the exponential and Weibull models, the regions overlap fairly well, suggesting that the asymptotic distribution is an adequate approximation. For the linear hazard function, the confidence regions must be truncated at the maximum allowed value for β. Therefore the normal approximation is not adequate. By the way, when the linear hazard model was used with uncentered data, the confidence regions were as shown in Figure 6. The non-elliptical region is thin and strongly curved, and it hardly overlaps the truncated ellipse at all; therefore, centering seems to improve the normal approximation, even though the approximation still is inadequate.

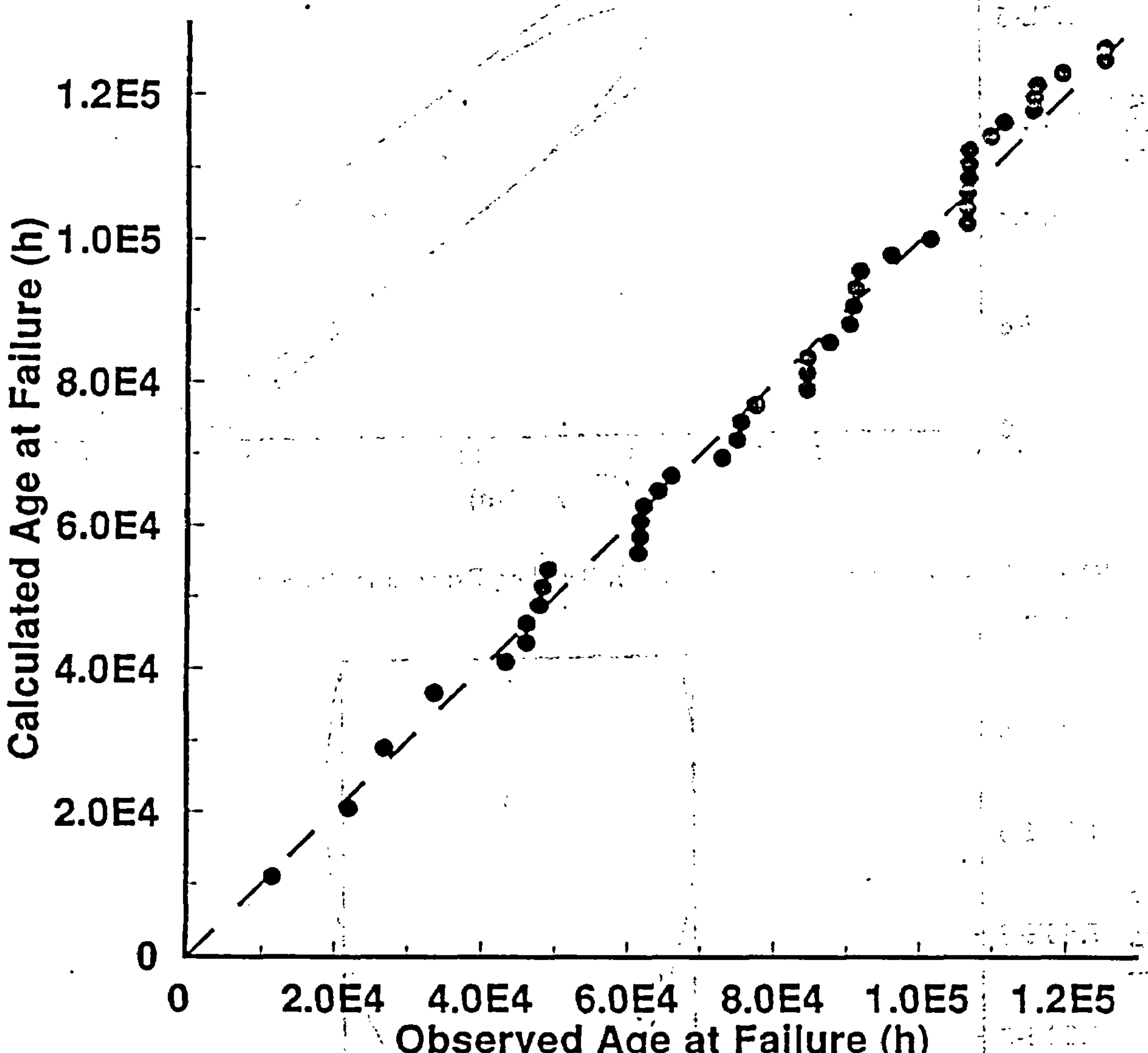

Figure 2. Q-Q Plot for Exponential Hazard Model

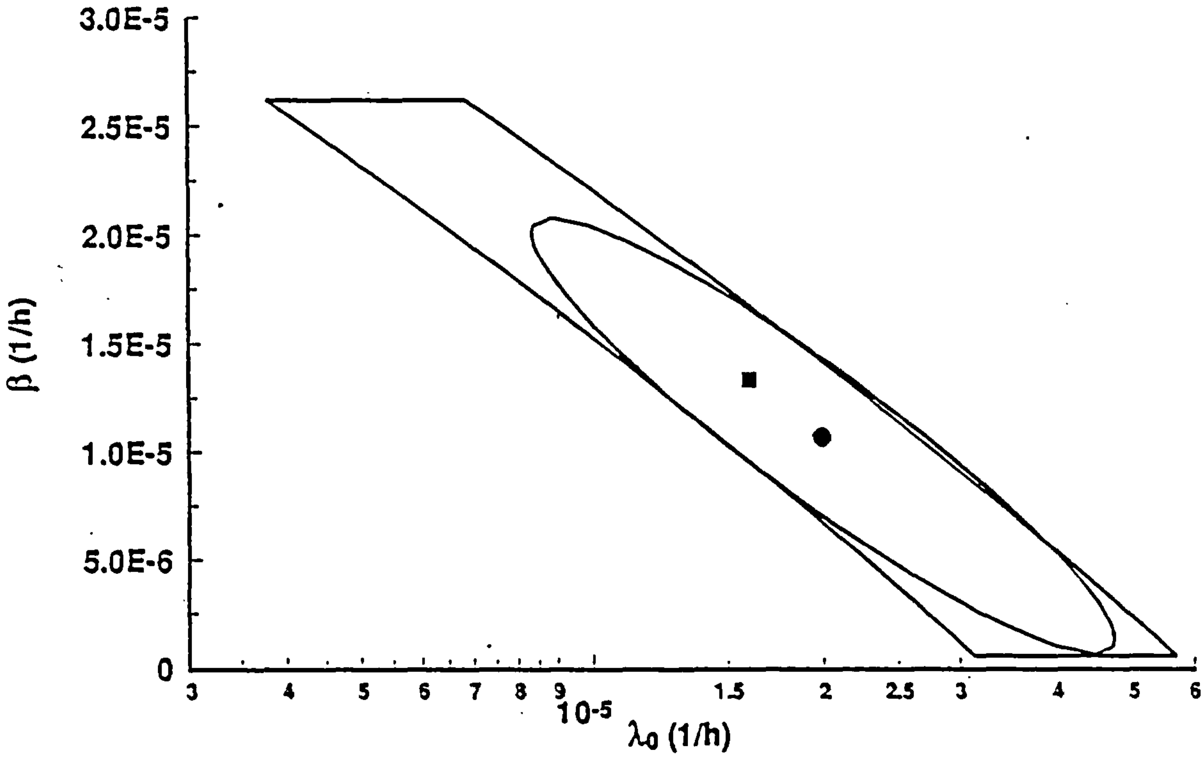

Figure 3. 90% Confidence Regions for (β, λ_0), Based on Exponential Hazard Model

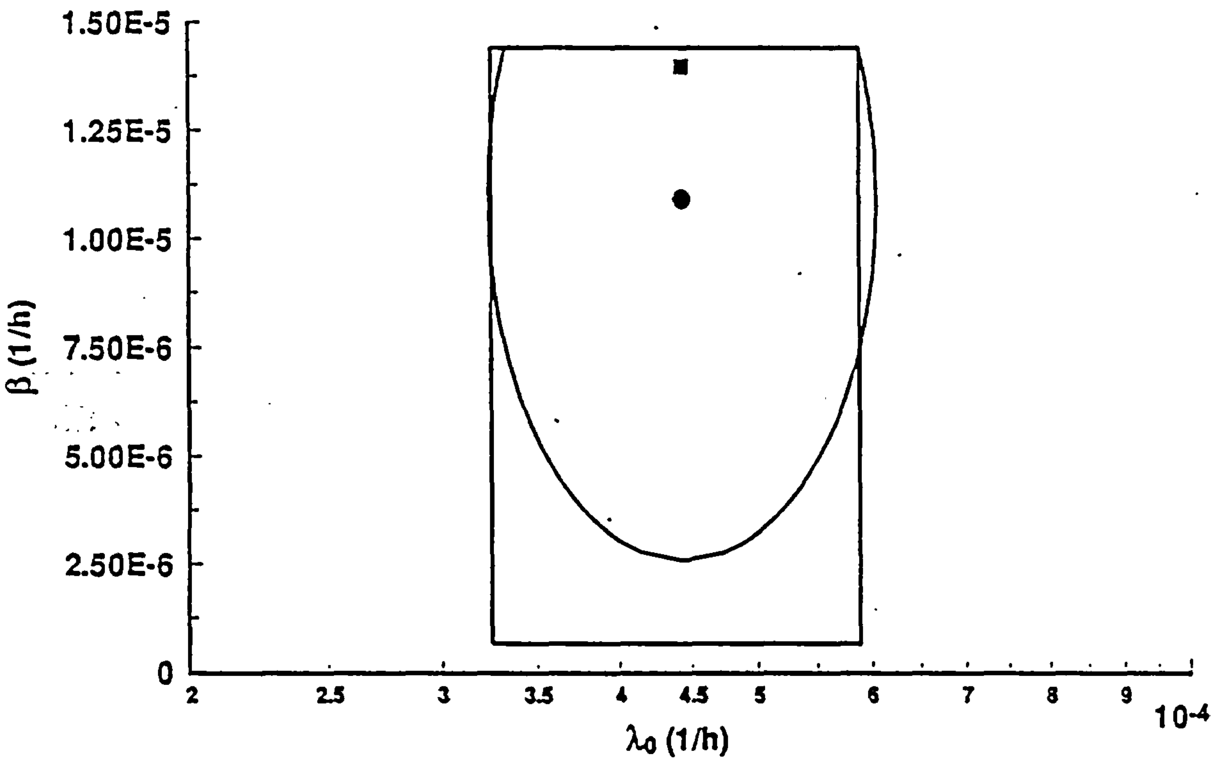

Figure 4. 90% Confidence Regions for (β, λ_0), Based on Linear Hazard Model, Centered Data

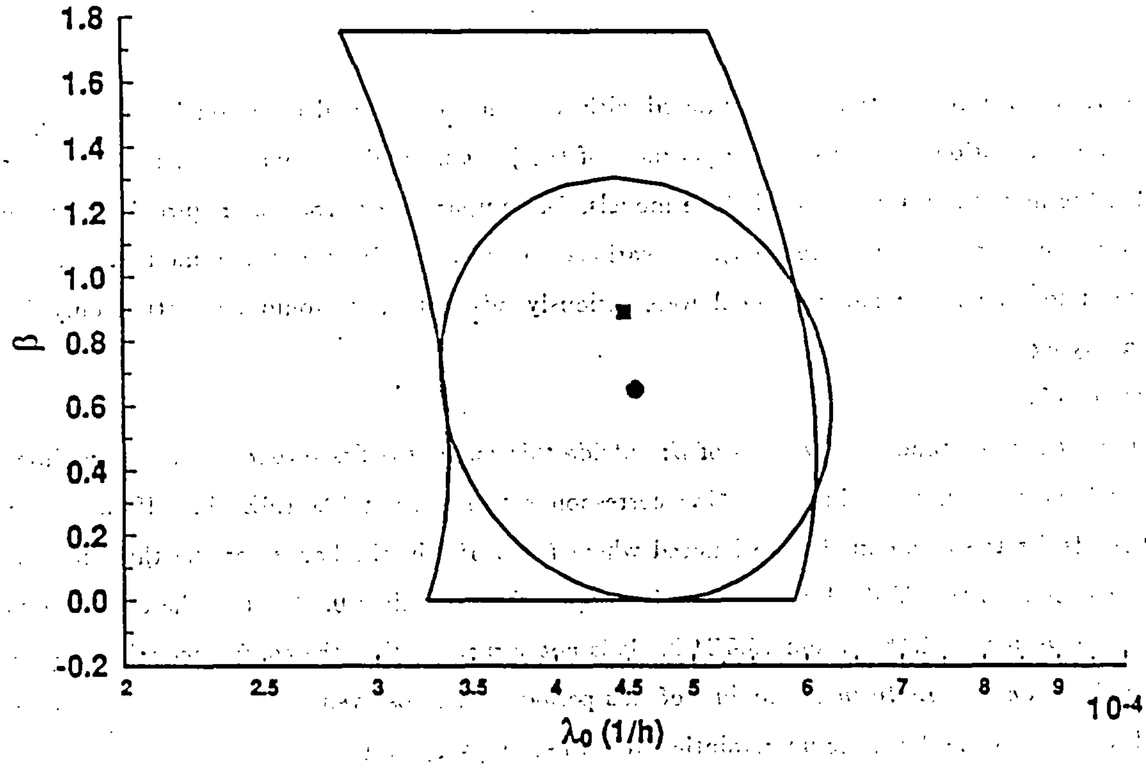

Figure 5. 90% Confidence Regions for (β, λ_0), Based on Weibull Hazard Model, Uncentered Data

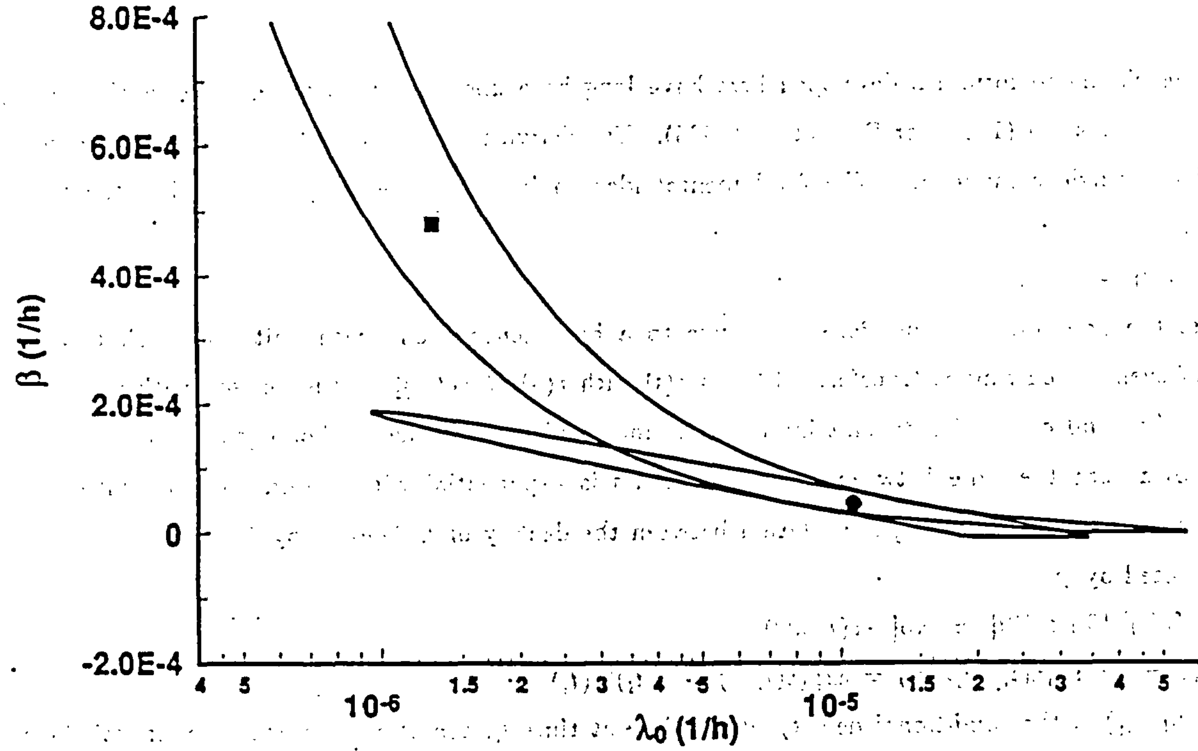

Figure 6. 90% Confidence Regions for (β, λ_0), Based on Linear Hazard Model, Uncentered Data

Finally, the hazard function was estimated with a confidence interval based on the asymptotic joint normal approximation. In spite of the poorness of the joint normal approximation for the linear hazard model, the method was used for all three models, for comparative purposes. Figure 7 shows the MLE and 90% confidence interval for $\lambda(t)$, at various values of t, for the three models. If the confidence band for the linear hazard model were seriously advocated, it would be plotted only for values of t satisfying

$$2\ sd\ t/(1+\hat{\beta}t) < 0.5,$$

where sd is the estimated standard deviation of $\hat{\beta}$; outside this range, the first-order Taylor approximation of $\log h(t;\beta)$ is inadequate. This restriction corresponds to requiring $t >$ 1.6E4 h. If the upper and lower bounds for the linear model are ignored where $t <$ 1.6E4 h, the bands for the three models look similar, except that the Weibull hazard function approaches 0 at time 0. Most of the components were observed between ages 4.1E4 h and 13.0E4 h. It is not surprising that the confidence intervals are narrowest [in the scale of $\log\lambda(t)$] in the middle of this period of the observed data. If the model were extrapolated far beyond the data, the uncertainties would become very large.

8. DERIVATIONS AND PROOFS

The likelihood formulas developed here have long been known; for example, see Equations (2.1) and (3.1) of Boswell (1966), or Bain et al. (1985). The derivations are sketched here for completeness. Consider a single component. The fundamental idea to be used repeatedly here is that the transformation

$$u(t) = \Lambda(t) - \Lambda(s_0)$$

converts the non-homogeneous Poisson process to a homogeneous one with unit rate. That is, the count of events occurring at transformed times $u(t)$ with $u(a) \leq u(t) \leq u(b)$ is Poisson with parameter $u(b) - u(a)$, and counts for disjoint intervals are independent. For such a homogeneous process, it is well known that the time between successive events is exponential with parameter 1.0. Likelihood formulas may be derived using the relation between the density of t, denoted by f, and the density of $u(t)$, denoted by g:

$$f(t) = g[u(t)]\ |\partial u(t)/\partial t| = \exp[-u(t)]\lambda(t)$$

$$f(t_i|t_{i-1}) = g[u(t_i)|u(t_{i-1})]\lambda(t_i) = \exp[u(t_{i-1}) - u(t_i)]\lambda(t_i)\ .$$

Here, $f(t_i|t_{i-1})$ is the conditional density of a failure at time t_i, conditional on the component's being operable (restored to service) at time t_{i-1}.

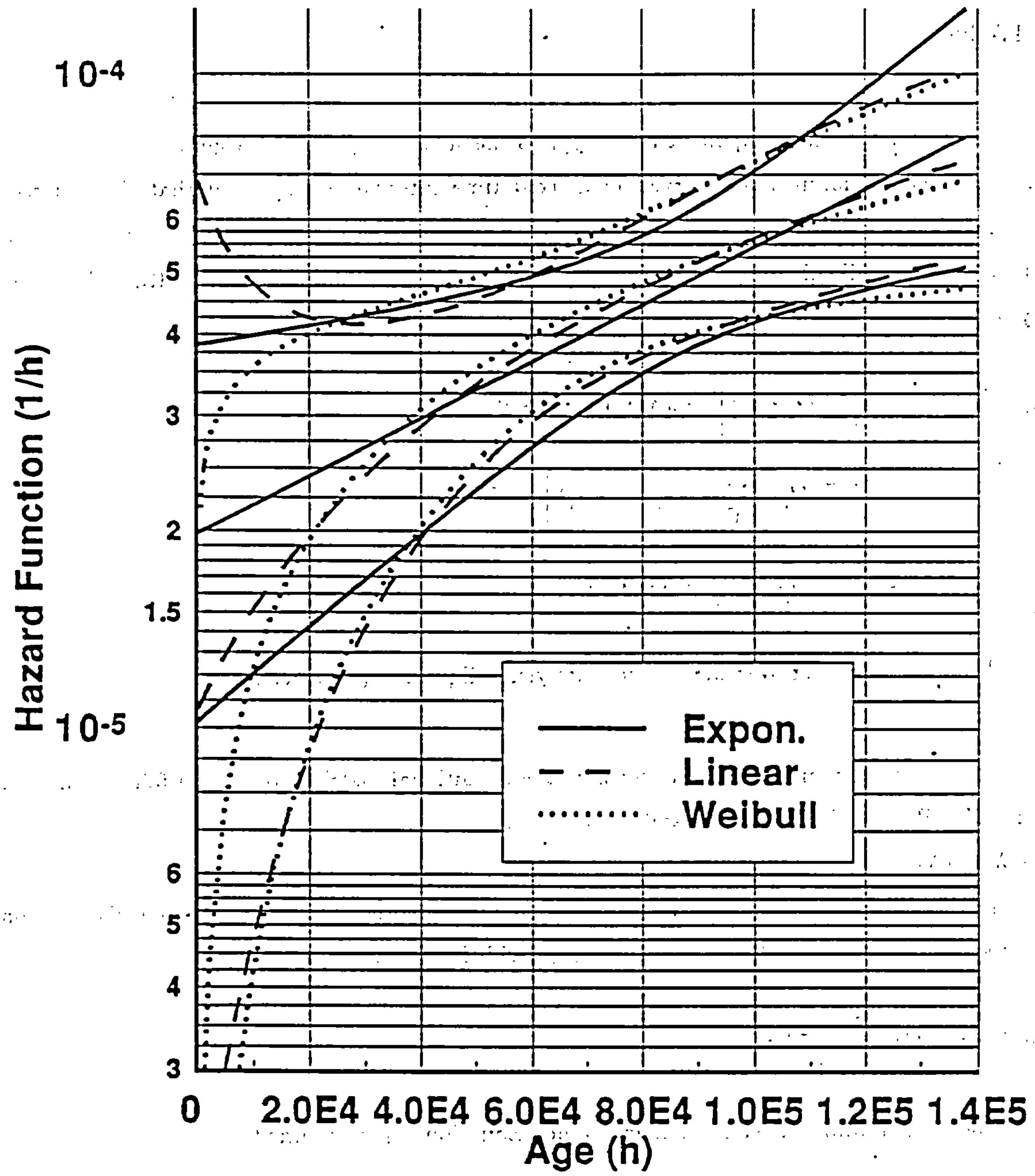

Figure 7. MLE and 90% Confidence Band for $\lambda(t)$, Based on Three Models

8.1 Derivation for Time-Censored Data

The Likelihood

Consider a single component and suppress the subscript j and the argument β. Suppose that a random number of failures is observed in a fixed time interval $[s_0, s_1]$, and that the ordered failure times are $t_1, \ldots, t_n$. In the formulas below, define $t_0 = s_0$ and $u_i = u(t_i)$. Note that $u(s_0) = 0$ and $u(s_1) = \lambda_0 v$. The likelihood is the joint density of the observed failure times, multiplied by the probability of no failures after t_n; that is,

$$
\begin{aligned}
l_{full}(\beta, \lambda_0) &= \left[\prod_1^n f(t_i | t_{i-1})\right] \exp[\Lambda(t_n) - \Lambda(s_1)] \\
&= \left[\prod_1^n \lambda(t_i)\right]\left[\prod_1^n \exp(u_{i-1} - u_i)\right] \exp[u_n - u(s_1)] \\
&= \lambda_0{}^n \left[\prod_1^n h(t_i)\right] \exp(-\lambda_0 v) \, , \qquad (18)
\end{aligned}
$$

Taking logs and summing over the components yields Equation (4), as claimed.

For a single component, consider now the conditional distribution of the failure times given n. Because N is Poisson$(\lambda_0 v)$, the probability of n failures is

$$\exp(-\lambda_0 v)\ (\lambda_0 v)^n \ / \ n! \ . \qquad (19)$$

Therefore the conditional likelihood, the likelihood corresponding to the conditional distribution of t_1, ..., t_n given n, is the quotient of Expression (18) divided by Expression (19):

$$l_{cond}(\beta) = \left[\prod_1^n h(t_i)\right] (v)^{-n} \ n! \ .$$

Taking logs and summing over components yields Equation (5), as claimed.

Ancillarity

Consider again a single component. The failure count N is ancillary for β. To see this, define $\mu = \lambda_0 v$. Reparameterize so that the parameters defining the model are μ and β. Then N is Poisson(μ), so the distribution of N involves only μ, not β. Given $N = n$, the unordered failure times T_i are i.i.d., each with density $h(t)/v$ on the interval $[s_0, s_1]$. This conditional density depends on β

only, not on μ. Therefore, N is ancillary for β and $(T_1, ..., T_n)$ is conditionally sufficient for β.

Suppose now that there are C components, $C > 1$, and that the components are not assumed to have a common value of λ_0. Then $(N_1, ..., N_C)$ forms a C-dimensional ancillary statistic for β. This is easily shown by a generalization of the above argument for a single component, parameterizing the model in terms of β and $(\mu_1, ..., \mu_C)$, with $\mu_j = \lambda_{0j} v_j$.

Similarly, suppose that there are C components with a common value of λ_0, and that v_j has the same value v for all the components, regardless of β. (Remark: In the three examples of this report, this can occur only if the components all have a common value of s_0 and s_1. To see this, set $v_j = v_k$ and $v_j' = v_k'$. Evaluate these quantities at $\beta = 0$ using the formulas of Table 1. It follows that $s_{0j} = s_{0k}$ and $s_{1j} = s_{1k}$; this is immediate for the exponential and linear hazard function, and can be shown with a little effort for the Weibull hazard function.) Now set $\mu = \lambda_0 v$ and note that $N = \Sigma N_j$ is Poisson$(C\mu)$. Consider the conditional log-likelihood analogous to Expression (5), only now conditional on n rather than on $(n_1, ..., n_C)$. It is equal to

$$\log\{(n!)C^n \prod_{j=1}^{C} \prod_{i=1}^{n_j} [h(t_{ij})/v]\} \ .$$

This is the log of the conditional density of the ordered failure times, with each time assigned at random to one of the C components. Therefore, the T_{ij}'s are conditionally i.i.d., each with conditional density $h(t)/v$ for $s_0 \le t \le s_1$. The components may therefore be pooled as a single super-component, and $N = \Sigma N_j$ is ancillary for β.

Finally, suppose that there are C components, $C > 1$, that the v_j's are not all equal, and that the components are assumed to have a common value of λ_0. There does not seem to be a reparameterization such that the distribution of $(N_1, ..., N_C)$ is independent of β. Therefore $(N_1, ..., N_C)$ does not appear to be ancillary. To show conclusively that $(N_1, ..., N_C)$ is not ancillary, we note that Equations (6) and (8) yield different values of $\hat{\beta}$.

8.2 Derivation for Failure-Censored Data

Now suppose that a single component is observed starting at time s_0, and that m failures are observed, with m fixed. The full likelihood is the joint density of the failure times:

$$l_{full}(\beta, \lambda_0) = \left[\prod_1^m f(t_i|t_{i-1})\right]$$

$$= \left[\prod_1^m \lambda(t_i)\right] \exp(u_0 - u_m)$$

$$= \lambda_0{}^m \left[\prod_1^m h(t_i)\right] \exp(-\lambda_0 v)]. \tag{20}$$

Taking logs and summing yields Equation (4).

To condition on the value t_m, the distribution of T_m must first be derived.

THEOREM. The time to the mth failure T_m has density

$$f_m(t_m) = w^{m-1} e^{-w} \lambda(t_m) / (m\text{-}1)! \tag{21}$$

where $w = \Lambda(t_m) - \Lambda(s_0)$, and $t_m \geq s_0$.

COROLLARY. Define $\lambda_0 V$ by $\Lambda(T_m) - \Lambda(s_0)$. Then $2\lambda_0 V$ has a $\chi^2(2m)$ distribution.

PROOF OF THEOREM. Here, $w = u(t_m)$, the mth transformed failure time. Because the transformed failure times correspond to a Poisson process with unit rate, it is well known that the mth transformed time has a gamma distribution. The asserted result follows. □

The conditional distribution of $(T_1, ..., T_m)$ given $T_m = t_m$ is (20) divided by (21). Take logs and sum over the components to show that $L_{cond}(\beta)$ is exactly equal to Expression (5).

9. REFERENCES

Corwin L. Atwood, *User's Guide to PHAZE, a Computer Program for Parametric Hazard Function Estimation*, EGG-SSRE-9017, 1990.

Lee J. Bain, Max Engelhardt, and F. T. Wright, "Tests for an Increasing Trend in the Intensity of a Poisson Process: A Power Study," *Journal of the American Statistical Association*, 80, 1985, pp. 419-422.

D. J. Bartholomew, Discussion of article by D. R. Cox, *Journal of the Royal Statistical Society, Series B*, 17, 1955, pp. 162-163.

M. T. Boswell, "Estimating and Testing Trend in a Stochastic Process of Poisson Type," *Annals of Mathematical Statistics*, 37, 1966, pp. 1564-1573.

D. R. Cox and D. V. Hinkley, *Theoretical Statistics*, London: Chapman and Hall, 1974.

D. R. Cox and Valerie Isham, *Point Processes*, London: Chapman and Hall, 1980.

D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events*, London: Chapman and Hall, 1966.

D. R. Cox and D. Oakes, *Analysis of Survival Data*, London: Chapman and Hall, 1984.

William Feller, *An Introduction to Probability Theory and Its Applications, Volume I*, Third Edition, New York: John Wiley & Sons, 1968.

Norman L. Johnson and Samuel Kotz, *Discrete Distributions*, New York: John Wiley & Sons, 1969.

Norman L. Johnson and Samuel Kotz, *Continuous Univariate Distributions* - 1, New York: John Wiley & Sons, 1970a.

Norman L. Johnson and Samuel Kotz, *Continuous Univariate Distributions* - 2, New York: John Wiley & Sons, 1970b.

John D. Kalbfleisch, "Ancillary Statistics," *Encyclopedia of Statistical Sciences*, Vol. 1, S. Kotz and N. L. Johnson, eds., New York: John Wiley & Sons, 1982.

Alan F. Karr, "Poisson Processes," *Encyclopedia of Statistical Sciences*, Vol. 7, S. Kotz and N. L. Johnson, eds., New York: John Wiley & Sons, 1986.

Paul L. Meyer, *Introductory Probability and Statistical Applications*, Reading MA: Addison-Wesley, 1970.

Wayne Nelson, *Applied Life Data Analysis*, New York: John Wiley & Sons, 1982.

Anthony A. Salvia, "Maximum Likelihood Estimation of Linear Failure Rate," *IEEE Transactions on Reliability*, R-29, 1980, pp. 49-50.

Ronald D. Snee and Charles G. Pfeifer, "Graphical Representation of Data," *Encyclopedia of Statistical Sciences*, Vol. 3, S. Kotz and N. L. Johnson, eds., New York: John Wiley & Sons, 1983.

William E. Vesely, *Risk Evaluations of Aging Phenomena: the Linear Aging Reliability Model and Its Extensions*, NUREG/CR-4769, EGG-2476, 1987.

Andrew J. Wolford, Corwin L. Atwood, and W. Scott Roesener, *Aging Risk Assessment Methodology: Demonstration Study on a PWR Auxiliary Feedwater System*, NUREG/CR-5378, EGG-2567, DRAFT, Rev. 1, 1990.

Stephen Wolfram, *Mathematica: A System for Doing Mathematics by Computer*, Redwood City, CA: Addison-Wesley Publishing Company, 1988.