Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques

Nan Laird; Donald Olivier

Journal of the American Statistical Association, Vol. 76, No. 374 (Jun., 1981), 231-240.

Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28198106%2976%3A374%3C231%3ACAOCSD%3E2.0.CO%3B2-N

Journal of the American Statistical Association is currently published by American Statistical Association.

Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques

NAN LAIRD and DONALD OLIVIER*

This paper unites two different fields, survival and contingency table analysis, in a single analytical framework based on the log-linear model. We demonstrate that many currently popular approaches to modeling survival data, including the approaches of Glasser (1967), Cox (1972), Breslow (1972, 1974), and Holford (1976), can be handled by using existing computer packages developed for the log-linear analysis of contingency table data. More important, we demonstrate that the log-linear modeling system used to characterize counted data structures directly characterizes survival data as well. Counted data methodologies for testing and estimation are also applicable here. Much of the theoretical basis for this work has been independently derived by Holford (1980) and Aitkin and Clayton (1980). The emphasis in this paper is not to develop new methodologies, but rather to present new uses and interpretations for already familiar methodologies.

KEY WORDS: Log-linear models; Life tables with covariates; Iterative proportional fitting.

1. INTRODUCTION

The usefulness of log-linear models for contingency table analysis is apparent from its current general popularity among both statisticians and applied researchers. Some of the more attractive features of this modeling system are the ease of model specification and reduction; its flexibility in treating both dependent and independent variables; and the fact that maximum likelihood estimates can be collectively characterized for an assortment of sampling distributions, including Poisson, multinomial, and product multinomial. Numerous computer packages for the computation of maximum likelihood estimates and associated tests are readily available. Most of these are based on the iterative proportional fitting (IPF) algorithm, although other maximization routines are used as well.

This paper assumes that the reader is generally familiar with log-linear models for contingency table analysis, associated maximum likelihood estimation and testing, and the use of IPF for computation. We will use notation originally introduced by Birch (1963), and extended by Bishop, Fienberg, and Holland (1975), who provide a comprehensive treatment of the models and methods; a newer book by Fienberg (1977) gives a briefer introduction.

The purpose of this paper is to point out that users of log-linear contingency table analysis can easily handle survival and life-table data as well. The existing log-linear models, tests, and estimates can be reinterpreted to provide a flexible modeling system and method of analysis for censored survival data with covariates. Some existing computer packages can be used without any modification to provide estimates and tests for survival data and life tables.

The theoretical basis for this claim rests on the existence of two simple relationships. First we show in Section 2 that log-linear models for the cell means of contingency tables with Poisson data are exactly equivalent to log-linear hazard models for survival data, when we specify (a) a piecewise exponential survival distribution and (b) categorical covariates. Second, we show in Section 3 that the likelihoods for these two cases, Poisson contingency table data and piecewise exponential survival data, are also equivalent. This last equivalence has the important implication that piecewise exponential survival data can be treated exactly as Poisson count data for the purpose of making likelihood based inferences. It is then easy to write expressions for maximum likelihood estimates and associated tests and to show how iterative proportional fitting may be used for computing them. Life-table analysis and Breslow's (1972, 1974) nonparametric version of Cox regression are discussed in Section 4. An example analyzing kidney graft survival data is also given.

The basic model we use was introduced in Holford (1976), who subsequently also discusses the Poisson representation of the likelihood and using IPF for computation (1980). We go beyond Holford (1980) by making more explicit the links with log-linear Poisson models, and by treating in some detail the scope and limitations on the use of IPF and its associated estimates, tests, and residuals. We provide a simple description of these results more accessible to readers familiar with log-linear contingency table analysis. Aitkin and Clayton (1980) also use the Poisson representation with Weibull and extreme-

* Nan Laird is Assistant Professor, Dept. of Biostatistics, School of Public Health, Harvard University, Boston MA 02115. Donald Olivier is with the Health Sciences Computing Facility, Harvard School of Public Health. This work was facilitated in part by National Science Foundation Grant MCS77-27119 and in part by the Health Sciences Computing Facility, Harvard School of Public Health. The calculations for this paper were carried out at the Health Sciences Computing Facility at the Harvard School of Public Health, using the LOGLIN computer package. Further information on this package can be obtained by writing to Donald Olivier. The authors thank Yvonne Bishop and Frederick Mosteller for helpful comments on an early draft, and the editors and referees for many suggestions that improved the clarity and style.

June 1981, Volume 76, Number 374
Applications Section

Table 1. Exposure in Months

	Aortic	Mitral	Total
Young	1,259	2,082	3,341
Old	1,417	1,647	3,064
Total	2,676	3,729	6,405

Table 2. Deaths

	Aortic	Mitral	Total
Young	4	1	5
Old	7	9	16
Total	11	10	21

value likelihoods, capitalizing on this representation by using GLIM for computation.

Before proceeding, we use a simple example to introduce the proposed method of analysis. The data are taken from a study by Cohn et al. (1981) on survival of patients after heart valve replacement operations. Patients are classified for this illustration by the type of valve replaced (aortic or mitral) and by age (under 55 versus 55 plus). A total of 109 patients were followed for periods ranging from 3 to 97 months. Table 1 gives the total exposure time in months in each category, with marginal totals, and Table 2 gives the number of deaths observed.

Defining risk as the number of deaths observed per number of person months of exposure, we obtain an overall risk of 21/6,405 = .0033 deaths/months of exposure. If we were concerned with only the single independent variable age, we would calculate separate risks for each age group as 5/3,341 = .0015 deaths/month for young patients and 16/3,064 = .0042 deaths/month for old patients. These three risks can be summarized by using a multiplicative model that says that risk in the ith age group is ra_i, where r is the overall risk, .0033, and $a_i = r_i/r$, where r_i is the risk in each age category (.0015 and .0042).

To take into account the effect of valve type, a simple extension of the multiplicative model is to assume that the risk of the ith age group and jth valve type is ra_iv_j where r is an overall risk, a_i is a factor that depends on age, and v_j is a factor that depends on valve type. This simple model will not in general exactly reproduce the observed risks for the four individual cells of the age-by-type table. It is possible, however, to choose a constant r and factors a_i and v_j so as to reproduce exactly the marginal risks—that is, the risks that would have been estimated for age class if valve type had been ignored, and correspondingly the risks for valve type, ignoring age.

Formal justification for marginal risk matching will be given in Section 3. Note, however, that if we were to assume that the number of deaths in each cell, d_{ij}, were independently distributed as Poisson with $E(d_{ij}) = E_{ij}ra_iv_j$, where E_{ij} is the exposure in cell ij, then the set of sufficient statistics would be d_{i+} and d_{+j}, where a "+" indicates summation over the subscript. Setting the sufficient statistics equal to their expectations, as required in exponential family settings, implies that estimates of r, a_i, and v_j must be chosen so that $d_{i+} = \hat{d}_{i+}$ and $d_{+j} = \hat{d}_{+j}$, where

$$\hat{d}_{ij} = E_{ij}\hat{r}\hat{a}_i\hat{v}_j. \tag{1.1}$$

It then follows that the estimated marginal risks, $\hat{d}_{i+}/E_{i+}$ and $\hat{d}_{+j}/E_{+j}$, will equal the observed marginal risks, d_{i+}/E_{i+} and d_{+j}/E_{+j}.

To estimate the parameters in this model, we construct the table of "expected deaths" $\{\hat{d}_{ij}\}$ whose cell values satisfy (1.1) and whose margins match the margins of $\{d_{ij}\}$. Note that the usual "expected" table for the multiplicative model $\{d_{i+}d_{+j}/d_{++}\}$ satisfies the margin matching, but does not satisfy (1.1). Since risk = deaths/exposure, and exposure is different in the four cells, the resulting table of risks will not be multiplicative, even though the table of expected deaths is.

To calculate our desired table, we will use an iterative process, basing our calculations on (1.1). We start with any values for ra_iv_j, say $r = .0033$ and $a_1 = a_2 = v_1 = v_2 = 1$, and calculate the expected number of deaths in each cell of the two-by-two table using (1.1). The result is shown in Table 3. Note that Table 3 is just the table of exposures (Table 1) multiplied by .0033.

Focusing first on age, we see that the expected number of deaths in the young category (11.02) is larger than the observed number (5), and the expected number in the old category (10.12) is smaller that the observed (16). These discrepancies can be rectified by modifying the values of a_1 to 5/11.02 and a_2 to 16/10.12. This gives new expected values, shown in Table 4.

We turn now to valve type. The expected numbers of deaths in the two categories (9.28 and 11.72) differ from the observed (11 and 10); we correct this by making $v_1 = 11/9.28$ and $v_2 = 10/11.72$, giving the new expected values shown in Table 5.

Adjusting the marginal values for valve types has altered those for age so that they again differ from the observed, although not by much. We correct the age

Table 3. Starting Point for Iterative Fitting

	Aortic	Mitral	Total
Young	4.15	6.87	11.02
Old	4.68	5.44	10.12
Total	8.83	12.31	

Table 4. After First Step of Iterative Fitting

	Aortic	Mitral	Total
Young	1.88	3.12	5.00
Old	7.40	8.60	16.00
Total	9.28	11.72	

marginals by multiplying a_1 by 5/4.89 and a_2 by 16/16.11, giving cumulated values of a_1 and a_2 of .46 and 1.56, and the expected values in Table 6.

The value-type marginals are off again (slightly); we could repeat the cycle if desired. The algorithm may be defined only in terms of the estimated cell values, $\hat{d}_{ij}$, as

$$\hat{d}_{ij}^{(0)} = E_{ij}$$

and

$$\hat{d}_{ij}^{(k)} = \hat{d}_{ij}^{(k-1)} d_{i+}/\hat{d}_{i+}^{(k-1)}$$

if k is odd, or

$$\hat{d}_{ij}^{(k)} = \hat{d}_{ij}^{(k-1)} d_{+j}/\hat{d}_{+j}^{(k-1)}$$

if k is even. As discussed in Section 3, cumulating the successive values of a_i and v_j in addition is useful for calculating their final estimates.

If we now divide each entry of Table 6 by the corresponding entry of Table 1, we obtain a set of estimated risks following the desired multiplicative model. Those familiar with survival analysis will recognize that the definition "risk = total deaths/total exposure" is appropriate only when the stochastic model for survival time is the exponential distribution, since this implies constant risk throughout the entire interval of observation. When the constant risk model is inappropriate, we use a simple modification that involves partitioning the time axis into intervals of constant risk.

Those familiar with log-linear contingency table analysis will recognize that the simple iterative process we have used is IPF (Deming and Stephan 1940). It can be readily generalized to fit any log-linear (i.e., multiplicative) model to a table with any number of dimensions and is widely used in fitting models to tables of counted data (Haberman 1974; Bishop, Fienberg, and Holland 1975; Fienberg 1977). It can be shown that the process always converges and yields the maximum likelihood parameter estimates for the given multiplicative model under a wide variety of sampling assumptions. Convergence is linear, in contrast to the quadratic convergence of Newton-Raphson and similar algorithms (Haberman 1976), but the algorithm is fast in most applications; its extreme simplicity and ability to handle models with large numbers of parameters compensate for this disadvantage.

2. LOG-LINEAR HAZARD MODELS FOR SURVIVAL DATA: EQUIVALENCE TO LOG-LINEAR MODELS FOR MEANS OF POISSON COUNT DATA

The term *survival data* is used in the context of measuring the duration of time until some event, for example,

Table 5. After Second Step of Iterative Fitting

	Aortic	Mitral	Total
Young	2.23	2.66	4.89
Old	8.77	7.34	16.11
Total	11.00	10.00	

Table 6. After Third Step of Iterative Fitting

	Aortic	Mitral	Total
Young	2.28	2.72	5.00
Old	8.71	7.29	16.00
Total	10.99	10.01	

death or recurrence of a disease. In such studies it is common to have only partial information on some individuals, called right censored, indicating that the event has not yet occurred before a given time. Thus the observed data in survival analysis often take the form of n independent observations, $(t_j, w_j, \mathbf{X}_j)$, $j = 1, \ldots, n$, where $t_j \in (0, \infty)$ is the time an individual is known to have survived before the event, $w_j = 1$ if the individual actually failed at t_j, $w_j = 0$ if the person was right censored at t_j, and $\mathbf{X}_j$ is a vector of known covariates.

In most areas of statistics, stochastic models for random variables are expressed in terms of densities and distributions. In survival (and reliability) analysis, more convenient, equivalent functions are the hazard and survivorship functions. If T denotes the random variable specifying time until the event of interest (and $\mathbf{X}$ denotes the associated set of fixed covariates), the survivorship function is defined as

$$S(t, \mathbf{X}) = \Pr(T > t \mid \mathbf{X}) = 1 - F(t, \mathbf{X}), \qquad (2.1)$$

where $F(t, \mathbf{X})$ is an ordinary distribution function for t given $\mathbf{X}$. The hazard function is

$$h(t, \mathbf{X}) = f(t, \mathbf{X})/S(t, \mathbf{X}), \qquad (2.2)$$

where $f(t, \mathbf{X})$ is the density associated with $F(t, \mathbf{X})$; that is,

$$f(T, \mathbf{X}) = \partial F(t, \mathbf{X})/\partial t. \qquad (2.3)$$

The hazard function is sometimes aptly called the force of mortality at t, since it is the instantaneous risk of failure at t, given survival to t. The exponential distribution, with $f(t) = \theta e^{-\theta t}$, $S(t) = e^{-\theta t}$, $t > 0$, $\theta > 0$, is often used to model accidents occurring randomly in time and has constant hazard θ.

We shall assume that the censoring mechanism is noninformative in the sense described by Lagakos (1979). Under noninformative censoring, the contribution of the jth individual to the likelihood is simply $h(t_j, \mathbf{X}_j)^{w_j} S(t_j, \mathbf{X}_j)$. As additional consequence of the noninformative censoring assumption is that in the case of parametric models, the asymptotic variances of maximum likelihood estimates can be obtained in the usual way, from the observed second derivative matrix of the log likelihood, evaluated at the parameter estimates. Lagakos (1979) discusses several models (one of these being the "random censorship" model) for the censoring mechanism that lead to noninformative censoring, and informative censoring models as well.

Perhaps the most appealing feature of the hazard function is that it allows a useful way of specifying the effect

of covariates on survival, which is now popular. The "proportional hazards model" introduced by Cox (1972) specifies

$$h(t, \mathbf{X}) = h_0(t)\, e^{\mathbf{X}^T\boldsymbol{\beta}}, \tag{2.4}$$

where $h_0(t)$ is the "underlying hazard function" and $\boldsymbol{\beta}$ is a column vector of unknown parameters specifying the effect of covariates. Here $h_0(t)$ can be chosen from any parametric family, such as exponential, Weibull, and so on, or it may be left unspecified, as in the Cox (1972) analysis and related nonparametric approaches such as Breslow's (1972, 1974).

The model (2.4) is termed "proportional hazards" for obvious reasons: the ratio of hazard functions for any two individuals with covariate vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ is simply $\exp\{(\mathbf{X}_1 - \mathbf{X}_2)^T\boldsymbol{\beta}\}$, independent of t. The fact that the ratio does not depend on t provides a convenient way of summarizing the effect of a covariate on survival. "Nonproportional hazards" models can be formulated by allowing $h_0(t)$ to depend on $\mathbf{X}$. "Time varying covariates" allow $\mathbf{X}$ to depend on t.

In this paper we use the proportional hazards model introduced by Holford (1976) for life-table analysis, where $h_0(t)$ is taken to be the hazard function associated with a piecewise exponential survival distribution. This model has also been used by Friedman (1980), who considers the asymptotic properties of maximum likelihood estimates. A piecewise exponential distribution is defined in the obvious way: we assume that the time axis can be partitioned into I mutually exclusive, exhaustive intervals, $\Omega_1, \ldots, \Omega_I$, such that the hazard function is constant within each interval. Letting h_i denote the constant hazard in Ω_i, $i = 1, \ldots, I$, we have

$$h(t, \mathbf{X}) = h_i e^{\mathbf{X}^T\boldsymbol{\beta}} \quad \text{for} \quad t \in \Omega_i. \tag{2.5}$$

As Holford (1976) points out, the two limiting cases ($I = 1$, $I \to \infty$) are of special interest. When $I = 1$, we have simple exponential survivals, and the model (2.5) is often called Glasser's (1967) model. At the other extreme, letting I become arbitrarily large so that each interval length becomes arbitrarily small yields a nonparametric model. This approach is essentially equivalent to the one outlined by Breslow (1972, 1974), which he shows yields an estimate of $\boldsymbol{\beta}$ equivalent to Cox's (1972) estimate (based on the partial likelihood) under certain conditions. This case is discussed further in Section 4.

A log-linear hazard model follows directly from (2.5):

$$\ln h(t, \mathbf{X}) = \ln h_i + \mathbf{X}^T\boldsymbol{\beta}, \quad t \in \Omega_i. \tag{2.6}$$

When the covariates specified by $\mathbf{X}$ are categorical, we can rewrite (2.6) by using ANOVA notation or, equivalently, the "u-term" notation often used in describing log-linear contingency table models. The advantage of the u-term notation is that it allows us to specify easily models with a broad range of complexity for any number of variables (covariates in this case), from the fully specified or "saturated" model, to a more parsimonious or "unsaturated" model.

We shall assume that $\mathbf{X}$ specifies the levels of p categorical covariates. We index the value of $\mathbf{X}$ by using the index set $(i_1, i_2, \ldots, i_p)$, where each covariate has l_m levels, $m = 1, \ldots, p$. We let i_0 index the time intervals, giving an augmented index set $(i_0, i_1, \ldots, i_p)$. We let $\theta_{i_0 i_1 \cdots i_p}$ denote the constant hazard in Ω_i appropriate for the set of covariates at levels $(i_1, \ldots, i_p)$; that is,

$$h(t, \mathbf{X}) = \theta_{i_0 i_1 \cdots i_p}, \quad t \in \Omega_i.$$

Then, employing the usual log-linear contingency table notation, we can rewrite (2.6) as

$$\begin{aligned}\ln\theta_{i_0 i_1 \cdots i_p} = {} & u + u_{0(i_0)} + u_{1(i_1)} + \cdots + u_{p(i_p)} \\ & + u_{12(i_1 i_2)} + \cdots + u_{12\cdots p(i_1 \cdots i_p)}.\end{aligned} \tag{2.7}$$

To avoid overparameterization, we assume $u_{0(+)} = \cdots = u_{12(i_1+)} = u_{12(+i_2)} = \cdots = 0$, as usual. Model (2.7) is fully saturated in the covariate effects; the $u_{m(i_m)}$ specify the "main effects" that the mth covariate has on survival, the $u_{mn(i_m i_n)}$ specify the joint effects on survival of the mth and nth covariates, and so on.

We formulate nonproportional hazard models by making the obvious generalization: in the completely saturated case

$$\begin{aligned}\ln\,\theta_{i_0 i_1 \cdots i_p} = {} & u + u_{0(i_0)} + \cdots + u_{01(i_0 i_1)} + \cdots \\ & + u_{012(i_0 i_1 i_2)} + \cdots + u_{01\cdots p(i_0 i_1 \cdots i_p)},\end{aligned} \tag{2.8}$$

where $u_{01(+i_1)} = u_{01(i_0+)} = \cdots = 0$. In this case the ratio of hazards $\theta_{i_0 i_1 \cdots p}/\theta_{i_0 i'_1 \cdots i'_p}$ for two sets of covariates $(i_1, \ldots, i_p)$ and $(i'_1, \ldots, i'_p)$ will vary over the different time intervals ($i_0 = 1, \ldots, I$), even though the covariate parameters are fixed.

We obtain unsaturated models, again in the obvious way, by setting higher order interaction terms to zero, the proportional hazards model of (2.7) being one special case. The log-linear model in (2.8) is like the log-linear model for the cell means of Poisson count data in that the total number of unconstrained parameters equals the total number of cells, IK, where

$$K = \sum_{m=1}^{p} l_m.$$

There is no constraint on the main effect "u," which here can be interpreted as the overall log-hazard rate. All subscripted u-terms may be omitted (in hierarchical fashion) if desired, and even the "null model," $\ln \theta_{i_0 i_1 \cdots i_p} = u$, may be of interest, since it implies simple exponential survivals with no effect of the covariates. Also, like Poisson models, the survival model allows only fixed covariates; all available degrees of freedom are used to specify the effects of the covariates on survival, or the shape of the hazard function.

Finally, we note that in contingency table analysis, there is an equivalence relationship between u-terms and odds ratios for multinomial, binomial, or product multinomial data. For Poisson count data, the u-terms specify the Poisson rates. For piecewise exponential survival

data with categorical covariates, the u-terms are used to model hazard functions.

3. POISSON REPRESENTATION OF SURVIVAL DATA LIKELIHOOD

It is simplest to begin by assuming exponentially distributed survivals and no covariates. Here $f(t) = \theta e^{-\theta t}$, $h(t) = \theta$, $S(t) = 1 - F(t) = e^{-\theta t}$. For a sample of n independent observations from $F(t)$ subject to right censoring, the likelihood is

$$l(\theta) = \prod_{j=1}^{n} \theta w_j e^{-\theta t_j} = \theta^d e^{-\theta T}, \tag{3.1}$$

where

$$d = \sum_{j=1}^{n} w_j; \quad \text{and} \quad T = \sum_{j=1}^{n} t_j.$$

Here d is the number of observed deaths or failures, and T is the total exposure, giving the total time the sample was known to be at risk of failure. It is easily seen that the maximum likelihood estimate of θ is the observed risk, d/T.

If we assume instead that d is Poisson, conditional on T with $E(d \mid T) = T\theta$, then the likelihood under this sampling model, $l_p(\theta)$, is

$$l_p(\theta) \propto (T\theta)^d e^{-T\theta}/d! \propto l(\theta).$$

Again, the maximum likelihood estimate of the Poisson rate θ is d/T. Since the likelihoods under these two different sampling models are proportional, one can use the likelihoods interchangeably for deriving maximum likelihood estimates, their asymptotic variances, and for calculating likelihood ratio statistics and their asymptotic sampling distributions. Thus for the purposes of making likelihood-based inferences about θ, one can use the exponential and Poisson sampling models interchangeably.

This result generalizes easily to piecewise exponentials, although the notation becomes more cumbersome. Here we need to represent t_j by $(t_{1j}, \ldots, t_{Ij})$, where $t_{ij} \in [0, \Delta_i]$ is the total time that individual j spent in Ω_i before failure and Δ_i is the length of Ω_i. Likewise, w_j becomes $(w_{1j}, \ldots, w_{Ij})$, where $w_{ij} = 1$ if individual j fails in Ω_i; otherwise, $w_{ij} = 0$ for all $i = 1, \ldots, I$. If $w_{ij} = 1$, then $t_{i'j} = 0$ for all $i' > i$. Likewise, $t_{i'j} = 0$ for all $i' > i$ if individual j is right censored in Ω_i. If $(\theta_1, \ldots, \theta_I)$ denotes the hazards in $\Omega_1, \ldots, \Omega_I$, the contribution of individual j to the likelihood is now

$$l_j = \prod_{i=1}^{I} \theta_i^{w_{ij}} e^{-\theta_i t_{ij}},$$

so the likelihood is simply

$$l(\boldsymbol{\theta}) = \prod_{j=1}^{n} l_j = \prod_{i=1}^{I} \theta_i^{d_i} e^{-\theta_i T_i},$$

where $d_i = \sum_{j=1}^{n} w_{ij}$ is the number of failures in Ω_i and $T_i = \sum_{j=1}^{n} t_{ij}$ is the total sample exposure in Ω_i. It follows that the same $l(\boldsymbol{\theta})$ will result if we assume that each d_i is an independent Poisson, conditional on T_i, with

$$E(d_i \mid T_i) = T_i\theta_i. \tag{3.2}$$

Note that we do not assume that (3.2) is true, but simply use it as a device for deriving the likelihood.

With covariates, we simply assume the log-linear hazard model presented in Section 2. Since individuals are independent, subgroups of individuals formed by stratifying on covariates will be independent. Letting $d_{i_0 i_1 \cdots i_p}$ denote the number of deaths in Ω_{i_0} among individuals with covariate levels $(i_1, \ldots, i_p)$, and letting $T_{i_0 i_1 \cdots i_p}$ denote the corresponding total exposure of that group in Ω_{i_0}, we can immediately write down the likelihood of $\boldsymbol{\theta}$, the $p + 1$-dimensional array of hazards $\theta_{i_0 i_1 \cdots i_p}$, as

$$l(\boldsymbol{\theta}) \propto \prod_{i_0 \cdots i_p} (m_{i_0 \cdots i_p})^{d_{i_0 \cdots i_p}} \exp(-m_{i_0 \cdots i_p}),$$

where

$$m_{i_0 \cdots i_p} = T_{i_0 \cdots i_p}\, \theta_{i_0 \cdots i_p}. \tag{3.3}$$

The resulting sampling model is merely a special case of the general log-linear model for Poisson means described in Sections 3.3.1 through 3.3.3 of Bishop, Fienberg, and Holland (1975), the only difference being the multiplicative cell constants, $T_{i_0 i_1 \cdots i_p}$. This feature is easily handled in characterizing maximum likelihood estimates, tests, and computational routines, by making use of the matrix of exposures, $\mathbf{T}$, as follows. Bishop et al. show that for the special case in which $T_{i_0 i_1 \cdots i_p} = 1$ for all cells, the maximum likelihood estimates of the cell means, $\hat{m}_{i_0 i_1 \cdots i_p}$, are characterized by the fact that certain marginals of $\hat{\mathbf{m}}$ (the $p + 1$-dimensional array of $\hat{m}_{i_0 \cdots i_p}$) must match the corresponding marginals of the array, $\mathbf{d}$, of observed cell counts $d_{i_0 \cdots i_p}$. The marginals matched are those specified by the nonzero u-terms in the model for $\ln \theta_{i_0 \cdots i_p}$. By differentiating $\ln l(\boldsymbol{\theta})$ with respect to the u-terms, using (3.3) and the appropriate version of (2.8) that specifies the θ's in terms of the u's, we find that the same marginals must be matched for $\hat{\mathbf{m}}$ with an arbitrary set of exposures, $\mathbf{T}$. In addition, however, the logarithms of the fitted cell counts must satisfy

$$\ln \hat{m}_{i_0 \cdots i_p} = \ln T_{i_0 \cdots i_p} + \ln \hat{\theta}_{i_0 \cdots i_p} \tag{3.4}$$

where $\ln \hat{\theta}_{i_0 \cdots i_p}$ is the sum of all the estimated u-terms specified as nonzero in the model.

This requirement can be satisfied, as in our introductory example, by using $\mathbf{T}$ as an initial estimate for $\hat{\mathbf{m}}$ and then using IPF to achieve the desired marginal matching of $\hat{\mathbf{m}}$ to $\mathbf{d}$. Thus any IPF computing routine that allows the user to specify an arbitrary matrix of starting values can be used to calculate maximum likelihood estimates, $\hat{\mathbf{m}}$, in the piecewise exponential sampling model. Existing IPF routines for contingency table analysis all use the same computational algorithm for obtaining tables of fitted counts, but may differ in the way that the parameters of the log-linear model (the u-terms) are calculated. In the simplest case, we can take logarithms of the fitted

counts and do mean removal (Tukey 1977) to get the estimated u-terms. This will not work when (3.4) holds, because of the additive constants, $\ln T_{i_0 \cdots i_p}$ (nor will it work if any cell has a zero estimate). Some IPF routines (for example, LOGLIN, produced by the Health Sciences Computing Facility at the Harvard School of Public Health) calculate the estimated u-terms by doing mean removal on the logs of the multipliers for each cell entry, cumulated at each step of the IPF routine, as in our introductory example. This method will produce the correct estimated u-terms for the piecewise exponential models, as well as in other cases in which a nonstandard start matrix is used.

Likelihood ratio tests as computed for Poisson count data will be valid, because of the proportionality of Poisson and piecewise exponential likelihoods. Likewise, Pearson goodness-of-fit and Freeman-Tukey deviates squared χ^2 tests, both based on the fitted $\hat{\mathbf{m}}$, should be valid asymptotically, but a formal demonstration of this would require investigation of the asymptotic sampling distribution of the $d_{i_0 \cdots i_p}$'s. The Freeman-Tukey cell deviates can be used to indicate model discrepancies in individual cells.

4. EXTENSIONS: LIFE TABLES AND NONPARAMETRIC REGRESSION ANALYSIS

As pointed out by Holford (1976), the piecewise exponential approach is a natural one for life-table analysis where the period of follow-up is divided into intervals, since a common assumption is that the hazard function is approximately constant within intervals. A problem with life-table data collected using the anniversary method of follow-up is that exact times of death or censoring are often not recorded. Holford also shows that assuming that all deaths and live withdrawals are observed for half the interval is not a bad assumption when censoring is uniform in the interval and the probability of surviving the interval is large.

The input for the start and data matrices for IPF come directly from quantities calculated in a standard actuarial life-table analysis. Using standard life-table notation (with X now indexing intervals), let n_X be the number alive at the beginning of interval X, d_X be the deaths in the interval, and w_X be the withdrawals. The data matrix is composed of the d_X column, and the start matrix is composed of the column of $n_X - (d_X + w_X)/2$. If interval widths are different, then $n_X - (d_X + w_X)/2$ must be multiplied by the interval width. Note that the columns of the start matrix will be almost equivalent to the adjusted n_X, $n_X' = n_X - w_X/2$, if the number of deaths is small and interval lengths are equal. The actuarial method estimates the conditional probabilities of death in an interval given survival to the beginning of an interval as $q_X = d_X/n_X'$, while our method estimates the hazard function for each interval.

Using the log-linear analysis in calculating life tables in a follow-up study is likely to be quite valuable when we have many groups and/or covariates and few individuals in each group. In such a circumstance, the analysis can be used to test for group differences in survival, and once a satisfactory model is found, the expected number of deaths (the fitted table) can be used as numerators in the standard actuarial analysis to give smoothed life tables for each group.

In his paper on regression models for life tables, Cox (1972) proposed an estimate for the parameter vector $\boldsymbol{\beta}$ in the proportional hazards model (2.4), based on maximizing what he termed the "conditional likelihood." To derive the conditional likelihood, Cox argued that conditional on a death occurring at time t_j and the set of people known to be alive at t_j (called the risk set, $R(t_j)$), the probability that the death happened to the jth person is just

$$p_j = h(t_j, \mathbf{X}_j) / \sum_{l \in R(t_j)} h(t_j, \mathbf{X}_l)$$

$$= e^{\mathbf{X}_j' \boldsymbol{\beta}} / \sum_{l \in R(t_j)} e^{\mathbf{X}_l' \boldsymbol{\beta}}.$$

Multiplying the p_j's for each observed death gives a likelihood depending only on $\boldsymbol{\beta}$, which we denote $L_C(\boldsymbol{\beta})$. Various discussants (Lindley 1972; Kalbfleisch and Prentice 1972; Breslow 1972) noted that $L_C(\boldsymbol{\beta})$ is not the claimed conditional likelihood, and Cox (1975) later justified its use as a "partial" likelihood.

In the case of categorical covariates, an estimate of $\boldsymbol{\beta}$ can be obtained by IPF. Breslow (1972) pointed out that Cox's estimate can be obtained by a nonparametric approach. He assumed that $h_0(t)$ is constant in a set of intervals whose endpoints are the set of observed death times and that each censored person is censored at the nearest preceding death. Writing down the complete likelihood $L_B(\boldsymbol{\beta}, \mathbf{h})$ for the sample under these assumptions, where $\mathbf{h}$ is the vector of constant hazards, Breslow showed that maximizing $L_C(\boldsymbol{\beta})$ gives the same $\hat{\boldsymbol{\beta}}$ as maximizing $L_B(\boldsymbol{\beta}, \mathbf{h})$ jointly for $\boldsymbol{\beta}$ and $\mathbf{h}$. An equivalent approach (Cox 1972; Holford 1976) is to assume $h(t)$ is zero everywhere except at the observed death times. This latter approach does not require moving censored observations to the nearest previous death, since the only information now of interest about the censored individuals is whether they are at risk for a given death.

Either case can be implemented using IPF, by taking the intervals of constant hazard to be defined by the distinct death times, so that in the absence of tied death times, $d_{i_0 + \cdots +} = 1$ for $i = 1, \ldots, I$, and $T_{i_0 i_1 \cdots i_p}$ is the number of people at risk for the ith death with covariates $(i_1 \cdots i_p)$. The $T_{i_0 i_1 \cdots i_p}$ can be multiplied by the length of the interval, Δ_i, to give exposure in the intervals between deaths. As long as the time effects are included in the model, the multiplication is unnecessary, since it merely acts like an arbitrary scale factor for the estimated $u_{0(i_0)}$'s. An even better approach would be not to move the censored observations and let $T_{i_0 i_1 \cdots i_p}$ denote the exact exposure in the interval.

When there are tied death times, $d_{i_0 i_1 \cdots i_p}$ is the number

of deaths occurring at the ith death time among the subgroup with covariate set $i_1 \cdots i_p$. The marginal cell $d_{i_0+\cdots+}$ is the number of deaths at the ith death time. As Breslow (1972) points out, his treatment of ties is the same as Peto's (1972) suggestion for handling ties. With the Cox approach, the likelihood must be modified to handle ties and the resulting estimate is no longer equivalent to Breslow's.

5. EXAMPLE

To illustrate the usefulness of our approach in survival analysis we consider its application to a set of data describing graft survival following kidney transplant operations. The data are preliminary data in life-table form from the Kidney Transplant Histocompatibility Study (Krakauer 1980) made available to us by Carl Cohen. Here we consider the effects of only two variables on survival—donor relationship and match grade. Donor relationship has two categories, cadaveric nonrelated (Cad), and living related donor (LRD). Match grade is the number of matched antigens out of a maximum of four matches—thus it varies from zero to four.

We construct our data and start matrices from 10 life tables, one for each donor relatedness/match grade combination, in the manner described in Section 4.1. The life tables and data and start matrices are given in Tables 7 and 8.

We begin our analysis by using the life-table time intervals and all five categories of match grade. The simplest model that provides a satisfactory fit is a proportional hazards model with main effects of donor relationship and match grade. The likelihood ratio statistic is 154.8 with 139 degrees of freedom. The estimated effects for this fit are:

Total Table: −7.47
Match Grade: .26 .23 .20 .00 −.69

Table 7. Life-Table Data on Graft Survival

	Match Grade														
	0			1			2			3			4		
Cad Time	E[a]	W[b]	D[c]	E	W	D	E	W	D	E	W	D	E	W	D
0–7d	242	0	13	386	0	16	393	0	20	130	0	8	18	0	0
7–15d	229	0	18	370	0	25	373	0	20	122	0	4	18	0	1
15–21d	211	0	14	345	0	12	353	0	18	118	0	8	17	0	0
21–30d	197	0	13	333	0	17	335	0	18	110	0	10	17	0	0
1–2m	184	0	21	316	0	43	317	0	36	100	0	14	17	0	1
2–3m	163	0	14	273	0	30	281	0	26	86	0	2	16	0	3
3–6m	149	0	23	243	0	35	255	0	36	84	0	10	13	0	2
6–9m	126	0	8	208	0	11	219	0	6	74	0	5	11	0	0
9–12m	118	0	3	197	0	16	213	0	10	69	0	1	11	0	0
1–1.5y	115	0	4	181	0	9	203	0	11	68	0	3	11	0	0
1.5–2y	111	2	9	172	6	6	192	8	7	65	2	4	11	0	1
2–2.5y	100	17	2	160	34	8	177	43	8	59	14	0	10	5	1
2.5–3y	81	21	3	118	26	4	126	31	6	45	11	0	4	1	1
3–3.5y	57	19	2	88	27	3	89	22	5	34	11	1	2	0	0
3.5–4y	36	19	1	58	33	1	62	27	1	22	11	0	2	0	0
>4y	16	16	0	24	23	1	34	34	0	11	9	2	2	2	0

	Match Grade														
	0			1			2			3			4		
LRD Time	E	W	D	E	W	D	E	W	D	E	W	D	E	W	D
0–7d	20	0	0	48	0	2	410	0	14	164	0	4	164	0	1
7–15d	20	0	1	46	0	0	396	0	7	160	0	3	163	0	2
15–21d	19	0	1	46	0	1	389	0	3	157	0	5	161	0	1
21–30d	18	0	0	45	0	2	386	0	4	152	0	7	160	0	1
1–2m	18	0	3	43	0	5	382	0	31	145	0	8	159	0	5
2–3m	15	0	1	38	0	3	351	0	21	137	0	1	154	0	1
3–6m	14	0	2	35	0	2	330	0	27	136	0	8	153	0	4
6–9m	12	0	0	33	0	0	303	0	14	128	0	3	149	0	1
9–12m	12	0	0	33	0	1	289	0	11	125	0	1	148	0	2
1–1.5y	12	0	0	32	0	2	278	0	9	124	0	0	146	0	0
1.5–2y	12	0	0	30	0	0	269	8	10	124	2	1	146	4	2
2–2.5y	12	2	0	30	4	0	251	53	7	121	17	2	140	23	1
2.5–3y	10	0	0	26	8	1	191	45	3	102	17	1	116	33	1
3–3.5y	10	5	0	17	3	0	143	45	2	84	26	1	82	27	1
3.5–4y	5	1	0	14	7	0	96	39	0	57	24	0	54	23	1
>4y	4	4	0	7	7	0	57	57	0	33	33	0	30	30	0

[a] E = Entered.
[b] W = Withdrawals.
[c] D = Deaths.

Table 8. Input Matrices for an IPF Program

DATA		Match					START		Match				
Type	Time	0	1	2	3	4	Type	Time	0	1	2	3	4
Cad	7d	13	16	20	8	0	Cad	7d	1,648.5	2,646.0	2,681.0	882.0	126.0
	15d	18	25	20	4	1		15d	1,760.0	2,860.0	2,904.0	960.0	140.0
	21d	14	12	18	8	0		21d	1,224.0	2,034.0	2,064.0	684.0	102.0
	1m	13	17	18	10	0		1m	1,714.5	2,920.5	2,934.0	945.0	153.0
	2m	21	43	36	14	1		2m	5,205.0	8,835.0	8,970.0	2,790.0	495.0
	3m	14	30	26	2	3		3m	4,680.0	7,740.0	8,040.0	2,550.0	435.0
	6m	23	35	36	10	2		6m	12,375.0	20,295.0	21,330.0	7,110.0	1,080.0
	9m	8	11	6	5	0		9m	10,980.0	18,225.0	19,440.0	6,435.0	990.0
	1y	3	16	10	1	0		1y	10,485.0	17,010.0	18,720.0	6,165.0	990.0
	1.5y	4	9	11	3	0		1.5y	20,340.0	31,770.0	35,550.0	11,970.0	1,980.0
	2y	9	6	7	4	1		2y	18,990.0	30,420.0	33,210.0	11,160.0	1,890.0
	2.5y	2	8	8	0	1		2.5y	16,290.0	25,020.0	27,270.0	9,360.0	1,260.0
	3y	3	4	6	0	1		3y	12,420.0	18,540.0	19,350.0	7,110.0	540.0
	3.5y	2	3	5	1	0		3.5y	8,370.0	13,140.0	13,590.0	5,040.0	360.0
	4y	1	1	1	0	0		4y	4,680.0	7,380.0	8,640.0	2,970.0	360.0
	>4y	0	1	0	2	0		>4y	1,440.0	2,160.0	3,060.0	990.0	180.0
LRD	7d	0	2	14	4	1	LRD	7d	140.0	329.0	2,821.0	1,134.0	1,144.5
	15d	1	0	7	3	2		15d	156.0	368.0	3,140.0	1,268.0	1,296.0
	21d	1	1	3	5	1		21d	111.0	273.0	2,325.0	927.0	963.0
	1m	0	2	4	7	1		1m	162.0	396.0	3,456.0	1,336.5	1,435.5
	2m	3	5	31	8	5		2m	495.0	1,215.0	10,995.0	4,230.0	4,695.0
	3m	1	3	21	1	1		3m	435.0	1,095.0	10,215.0	4,095.0	4,605.0
	6m	2	2	27	8	4		6m	1,170.0	3,060.0	28,485.0	11,880.0	13,590.0
	9m	0	0	14	3	1		9m	1,080.0	2,970.0	26,640.0	11,385.0	13,365.0
	1y	0	1	11	1	2		1y	1,080.0	2,925.0	25,515.0	11,205.0	13,230.0
	1.5y	0	2	9	0	0		1.5y	2,160.0	5,580.0	49,230.0	22,320.0	26,280.0
	2y	0	0	10	1	2		2y	2,160.0	5,400.0	46,800.0	22,050.0	25,740.0
	2.5y	0	0	7	2	1		2.5y	1,980.0	5,040.0	39,780.0	20,070.0	23,040.0
	3y	0	1	3	1	1		3y	1,800.0	3,870.0	30,060.0	16,740.0	17,820.0
	3.5y	0	0	2	1	1		3.5y	1,350.0	2,790.0	21,510.0	12,690.0	12,240.0
	4y	0	0	0	0	1		4y	810.0	1,890.0	13,770.0	8,100.0	7,560.0
	>4y	0	0	0	0	0		>4y	360.0	630.0	5,130.0	2,970.0	2,700.0

[a] The cell entries in the start matrix T are obtained by using the approach given in Section 4: $T = (E - (D + W)/2)\Delta$, where E, D, and W are given in Table 7 and Δ is the corresponding interval in length in days. We assume that the last interval is six months in duration.

Donor Type: .34 −.34
Time: 2.04 2.00 2.08 1.85 1.57 1.18 .57 −.47 −.49 −1.31 −1.21 −1.36 −1.44 −1.39 −2.21 −1.42

These estimates indicate lower risk of loss for better-matched kidneys and for living related donors, and in the period immediately following the transplant a relatively high risk, which diminishes with time. The estimates suggest that the effect of match grade is not linear, and in fact that there is little difference between match grades 0, 1, and 2. In addition, the time effects suggest there is no need to retain so many categories for time. On the basis of these estimates we decided to combine the 0, 1, and 2 categories of match grade, the first four time categories, the fifth and sixth time categories, the eighth and ninth time categories, and the last seven time categories. These divisions correspond to intervals of 0–1 month, 1–3 months, 3–6 months, 6 months–1 year, and greater than 1 year. The original data and start matrices were thus reduced to dimension $3 \times 2 \times 5$.

Table 9 summarizes the likelihood ratio statistics for various possible models fit to the collapsed table. With the collapsed table, model B, a proportional hazards model with nonadditive effects for type and match grade, is now the simplest model with adequate fit. The nonproportional hazards models (C, D, E, and F) do not seem to improve the fit, but the interaction between donor type and match grade is significant.

The estimated effects for model B (with main and interaction effects added together) are

Total Table: −7.12
Time: 1.56 .98 .15 −.91 −1.78

Match X Donor:

	Cad	LRD
0-2	.62	.01
3	.53	−.42
4	.42	−1.15

The ratio of estimated hazard functions for any two match-by-donor groups is obtained by exponentiating the difference of their estimated effects. Thus for cadaveric transplants there is little effect of match type, the ratio of hazard functions for the 0–2 group versus the 4 group being $\exp\{.62 - .42\} = 1.22$. By contrast, the same ratio of hazard functions for living related donors is 3.19.

To illustrate the effect of donor type and match grade on duration of function, we have plotted the log of the estimated survival function versus time for all six groups in Figure 1. Here the log-survival function is calculated according to the formula

$$\ln S(t) = -e^{\hat{u}^*}\left[\sum_{j=0}^{i-1}(t_{j+1} - t_j)e^{\hat{u}_{0(j)}} + (t - t_i)e^{\hat{u}_{0(i)}}\right], \quad t_i < t \le t_{i+1}$$

where $t_0 = 0$, $t_1 = 1$ month, $t_2 = 3$ months, and so on, $u_{0(1)} = 1.56$, $u_{0(2)} = .98$, and so on, and $\hat{u}^*$ is the total table effect plus the appropriate group effect. The linearity of log survival in the periods 0–6 months and greater than 6 months suggests that two or three time intervals might suffice. To illustrate how well these curves fit the data, Figure 1 also shows the empirical life-table curves for the two most extreme groups, the 0 match cadaveric transplants and the 4 match living relateds.

When model B was fit to the original uncollapsed table, the likelihood ratio statistic was 146.6 with 135 degrees of freedom. Testing for a match grade by donor type interaction on the uncollapsed table gives a χ^2 of 8.2, on four degrees of freedom, which is to be compared with 7.5 on two degrees of freedom on the collapsed table. The estimated match by donor effects for model B on the uncollapsed data are

	Cad	LRD
0	.53	−.07
1	.51	−.11
2	.43	−.13
3	.40	−.56
4	.28	−1.28

Thus the trends found on the collapsed and uncollapsed data are essentially the same, only their statistical significance is altered, presumably because of the extra unnecessary parameters being fit on the uncollapsed data. The use of significance testing in this example is of course exploratory and their p values are not to be taken literally.

Table 9. Summary of Fitted Models for Graft Survival

	Model[a]	Likelihood Ratio	Degrees of Freedom	P Value
A	1/2/3	35.5	22	.034
B	1 2/3	28.0	20	.11
	A-B	7.5	2	.023
C	1 3/2	23.5	14	.053
	A-C	12.0	8	.15
D	1/2 3	29.9	18	.038
	A-D	5.6	4	.23
E	1 2/1 3	16.1	12	.19
	B-E	11.9	8	.16
F	1 2/2 3	22.4	16	.13
	B-F	5.6	4	.23

[a] The notation (1/2/3) denotes a model with main effects for all three variables and no interactions. A model with an interaction between variables 1 and 2 and a main effect for variable 3 is (1 2/3), etc. Variable 1 represents match grade 0-2/3/4; variable 2, donor type Cad/LRD; variable 3, time interval 1m/3m/6m/1yr/>1yr.

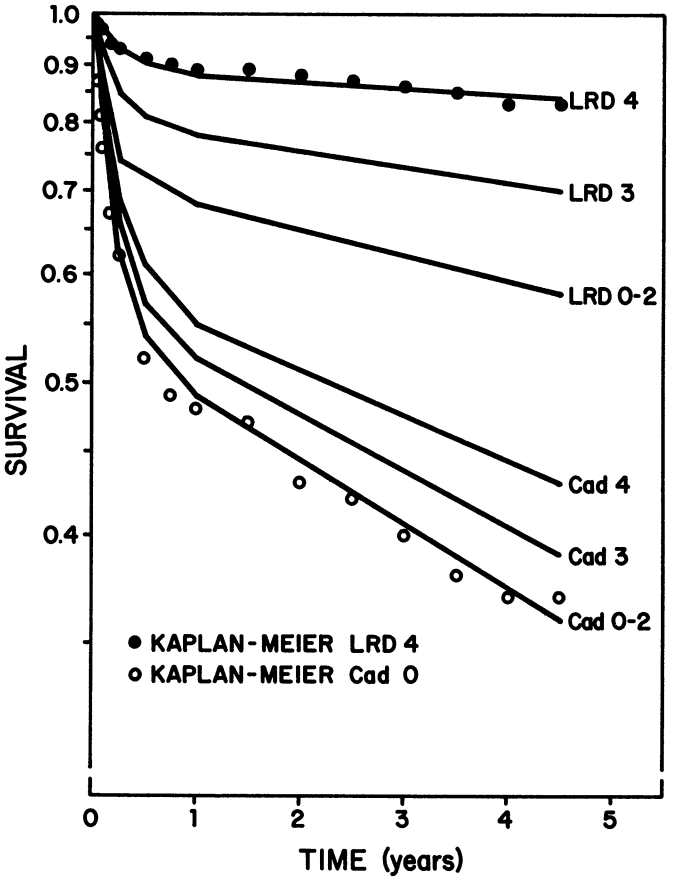

Figure 1. Actual and Fitted Graft Survival Curves

6. SUMMARY AND CONCLUSIONS

This paper demonstrates how model fitting, estimation, and testing methods developed for log-linear contingency table analysis can be used to analyze survival and life-table data. Several features of this approach make it attractive for handling survival analysis.

From a practical standpoint, many computing facilities already have log-linear contingency table packages based on IPF routines that can be used without modification. The advantages of using IPF over straightforward numerical maximization are that high-parameter models, including interactions, can be easily fit without added complexity and that very large data sets are readily handled by reduction to sufficient statistics. Some modifications could make these packages even more attractive. For example, making more efficient use of storage allocation when a modified Cox approach is used would make this method more feasible with IPF. Other desirable features would include the ability to calculate data and start matrices from life-table and survival data and to compute quantities used routinely in survival analysis, such as hazard functions, survivorship functions, and life tables. In addition, methods for handling ordered categories, such as those described in Fienberg (1977), would be useful.

From a data modeling standpoint, piecewise exponential models are very flexible. If nothing is known or as-

sumed about the underlying survival distribution, then an essentially nonparametric analysis can be implemented by making the time intervals sufficiently small. The time effects estimates can be inspected for clues as to an appropriate model, as in the example in Section 5. Future developments in this area should include more formal methods for model simplification, based on fitting a time curve to the estimated time effects.

From an analytical standpoint, it is both intuitively appealing and conceptually expedient to include survival analysis in the general counted data framework. The most obvious benefit is that the whole class of log-linear models used to characterize counted data structures carries over directly to characterize survival data. Competing risk analyses and time varying covariates are easily handled in this general framework.

[*Received March 1980. Revised January 1981.*]

REFERENCES

AITKIN, M., and CLAYTON, D. (1980), "The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data Using GLIM," *Journal of the Royal Statistical Society*, Ser. C, 29, 156–163.

BIRCH, M.W. (1963), "Maximum Likelihood in Three-Way Contingency Tables," *Journal of the Royal Statistical Society*, Ser. B, 25, 220–233.

BISHOP, Y.M.M., FIENBERG, S.E., and HOLLAND, P.W. (1975), *Discrete Multivariate Analysis: Theory and Practices*, Cambridge, Mass.: M.I.T. Press.

BRESLOW, N.E. (1972), Discussion of Cox (1972), *Journal of the Royal Statistical Society*, Ser. B, 34, 187–220.

——— (1974), "Covariance Analysis of Censored Survival Data," *Biometrics*, 30, 89–99.

COHN, L.H., MUDGE, G.H., PRATTER, F., and COLLINS, J.J., JR. (1981), "Five- to Eight-Year Follow-up of Patients Undergoing Porcine Bioprosthetic Valve Replacement," *New England Journal of Medicine*, 304, 258–262.

COX, D.R. (1972), "Regression Models and Life Tables" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 34, 187–220.

COX, D.R. (1975), "Partial Likelihood," *Biometrika*, 62, 269–276.

DEMING, W.E., and STEPHAN, F.F. (1940), "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known," *Annals of Mathematical Statistics*, 11, 427–444.

FIENBERG, S.E. (1977), *The Analysis of Cross Classified Data*, Cambridge, Mass.: M.I.T. Press.

FRIEDMAN, M. (1980), "Piecewise Exponential Models for Survival Data With Covariates," unpublished paper.

GLASSER, M. (1967), "Exponential Survival With Covariance," *Journal of the American Statistical Association*, 62, 561–568.

HABERMAN, S.J. (1974), *The Analysis of Frequency Data*, Chicago: University of Chicago Press.

HOLFORD, T.R. (1976), "Life Tables With Concomitant Information," *Biometrics*, 32, 587–597.

——— (1980), "The Analysis of Rates and Survivorship Using Log-linear Models," *Biometrics*, 36, 299–306.

KALBFLEISCH, J., and PRENTICE, R.L. (1972), Discussion of Cox (1972), *Journal of the Royal Statistical Society*, Ser. B, 34, 187–220.

KRAKAUER, H. (1980), "Kidney Transplant Histocompatibility Study: Analysis of the Data. Phase I, an Overview," Publication No. NIH 80-2614 and -2615, National Institutes of Health, Bethesda, Md.

LAGAKOS, S.W. (1979), "General Right Censoring and Its Impact on the Analysis of Survival Data," *Biometrics*, 35, 139–156.

LINDLEY, D.V. (1972), Discussion of Cox (1972), *Journal of the Royal Statistical Society*, Ser. B, 34, 187–220.

PETO, R. (1972), Discussion of Cox (1972), *Journal of the Royal Statistical Society*, Ser. B, 34, 187–220.

TUKEY, J.W. (1977), *Exploratory Data Analysis*, Reading, Mass.: Addison-Wesley.